22/4/2010
Understanding variations in sports participation

Understanding variations in sports participation

Technical Report

Contents

Executive Summary

Background and approach

Sport England has commissioned a series of robust quantitative models aimed at better understanding the factors which account for variations in sports participation, and thereby identify the levers most amenable to public policy intervention.

The objectives for this study included:

- A strengthened theoretical framework for understanding variations in participation in sport.
- Robust quantitative models that test the impact of various inputs, and activities on participation, the nature and strength of relationships between a range of outputs and the intended outcome.
- Illustration of those factors that on their own, or in combination, make the best public policy 'buy' to grow and sustain community participation in sport.

Sheffield Hallam first built a model in 2007 to understand variations in participation rates as driven by demographic factors such as age and income*. The differences between actual and expected observed in the original Sheffield model clearly showed that whilst demographic factors were important there are other factors (perhaps more amenable to intervention) that affect participation rates. This project sought to identify and better understand some of these factors.

As part of this project, we have updated the Sheffield Hallam model to estimate Local Authority participation using data from more recent waves of the Active People Survey.

The principal thrust of the modelling work has been Mindshare's NI8 model, which extends the Sheffield Hallam model model by taking into account, in addition to demographics, additional quantifiable information about an individual's surroundings, such as weather, local authority interventions, and access to facilities.

In addition to the NI8 model, the second part of the analysis seeks to understand the variations in participation rates between 11 different sports:

- Athletics
- Tennis
- Football
- Rugby Union
- Rugby League
- Squash
- Badminton
- Swimming
- Cycling
- Cricket
- Golf

The main criteria for selecting these sports were that they should have a high enough level of participation to provide a large enough sample for modelling within the Active People survey and that they were of strong interest to Sport England and its partners.

In the case of all 11 sports, we have first built a selection model. This model identifies those who engage in at least some sport – taken to be at least 1 session over the last four weeks that is within the definition used for the 1 million sport indicator. Individual sport specific models have then been developed to understand the drivers of frequency of participation in the individual sports within this population.

Key findings from the Mindshare NI8 model

Demographic factors

The model has confirmed and reinforced the importance of demographic factors in driving participation in sport and variations between participation rates between local authorities, i.e.:

- Age: the likelihood of meeting the NI8 criteria declines with age, although it does so much more sharply for men than women
- Income: higher household income has a significantly positive impact on the probability of reaching the NI8 criteria, and the selection model criteria used for the sports specific model of one session of sport over the last four weeks
- Education: across the NI8 model, as well as the selection and sports specific models, there is a consistently emerging pattern that those who have attained a higher-education qualification, degree or otherwise are more likely to participate in sport
- Children in household: individuals with children in the household are less likely to meet the NI8 criteria
- Population density: the modeling has found that higher population density increases the likelihood of reaching the NI8 criteria

Cultural engagement

The model provides evidence to support the hypothesis that people who engage more in cultural activities or civic life engage more in sport, as the number of cultural events that an individual has attended in the last year has a significant impact on the likelihood of reaching the NI8 criteria. Those who have attended three or more events are more likely to reach the criteria than those who have attended fewer or no events.

Lottery funding

The model confirms our hypothesis that areas that have received higher levels of Sport England lottery funding have higher participation rates. Within the NI8 model, higher Sport England lottery award amounts within 10km of a respondent lead to that respondent being more likely to participate in sport to an NI8 level (at least 12 sessions in last 4 weeks).

It is possible that this is because the lottery funding data analysed reflects long-term investment in an area, as the lottery grant award data we have used stretches back to 1995. Furthermore, given that in order to receive lottery funding a local area will likely have to demonstrate a long-term commitment to invest in the frameworks, people and programmes that support sport in the area, lottery funding itself may be representative of a broader long-term commitment to sport in that area.

Factors we were unable to test

In addition, there were a number of factors that we were unable to provide evidence to support due to the limitations of the data available or limitations within the modeling techniques.

- Total spend on sport in an area: we do not have access to data that records total expenditure in sport over the long-term in an area across the public and private sectors

- Quality of facilities and local government sport provision: while we were able to test distance to facilities and accredited clubs, we do not have data to assess the relative quality of facilities in one area versus another

- Events and competitions: we were unable to prove the hypothesis that greater access to competitions and events increases participation. In addition to the availability of comprehensive events data, this is due to an issue of causality – whilst those training for an event may train more than those who are not, it may also be the case that those who participate more then choose to participate in events.

Key findings from the sports specific models

Gender gap

In each sports model, we have tested whether gender is a significant driver which increases the ***frequency*** of participation in that sport. We have found that gender impacts on the frequency of participation in football, athletics, rugby league, cycling, badminton, golf, squash and cricket.

Asian ethnicity

We found that, within the sports selection model (which is used as the first measure of whether someone is "active" or not, based on the criteria of one 30 minute session of sport in the past 4 weeks), Asians are less likely to be active as they get older than people from other ethnicities.

Of the 11 sports we tested, we found that for athletics and rugby union Asian people are significantly *less* likely to participate than other ethnicities. However, for badminton and cricket, Asian people are *more* likely to participate.

Club sports

We found that for the team based sports of rugby league, rugby union, cricket and football, individuals who belong to a club tend to play more often than non-club members. This is likely to be a combination of training sessions and matches/competition. The effect is more pronounced for rugby league, rugby union and football than for cricket, which may suggest that cricket has less of a training element than other sports.

However, given that club membership is by far the biggest driver of frequency of participation for these sports, we have provided additional mezzanine models which seek to explain the drivers of club membership for each of these sports.

The effects of the drivers within the "mezzanine" model can then be interpreted as the impact of that variable on either increasing or decreasing the probability that an individual is a member of a club in that particular sport. The difference between the coefficients on frequency of participation between the two models account for the bias introduced due to the endogeneity of club membership.

Consistent across all four of the team sports are the drivers of male gender and non-gym membership which make individuals more likely to be club members. For football, rugby union and rugby league, probability of being a club member declines with age; however for cricket, older individuals are more likely to be a club member.

Family sports

We found that in contrast with the NI8 model, for cricket and swimming, frequency of participation increases with the number of children in the household. This is consistent with the hypothesis of swimming as a family friendly sport that is a complement, rather than substitute, for time with the family. In the case of cricket, this effect was unexpected and may merit further investigation.

Furthermore, for tennis and badminton, having older children rather than younger children increases the frequency of participation, which suggests that these sports are well adapted to parents and older children playing together as a family.

1. Background and Approach

1.1. Objectives and key measures

The objectives for the analysis of understanding variations in sports participation can be split into three specific parts:

1. Quantification of the drivers of reaching the 3x30 sport and NI8 measures of sports participation,
2. A predicted NI8 participation rate by LA, and
3. The results of testing a series of hypotheses and whether the factors have been found to have a significant impact on sports participation.

These have involved looking at different definitions of participation in sport according to the performance indicator which is most relevant and the requirements for the analysis.

The 3x30 sport indicator is the target measure for Sport England and is defined as the percentage of the population taking part in at least moderate sport for at least 30 minutes duration on at least 3 days a week. This has been used to understand the drivers of participation in sport at a national level.

The NI8 measure currently used as a target for local authorities differs from the 3x30 sport indicator in the range of activities it includes. It is also based on the Active People Survey and is defined as the percentage of the population taking part in at least moderate intensity sport or active recreation for at least 30 minutes duration on at least 3 days a week. The inclusion of recreational walking and cycling within N!8 is the main difference between the two indicators.

1.2. Exploring hypotheses about drivers of sports participation

At the outset of the project, a meeting was held with Sport England research and policy people that sought to capture as many of the hypotheses around the drivers of sports participation as possible. Mindshare, The Futures Company and Sport England together developed these hypotheses. As a result of the meeting, a framework was put together for classifying hypotheses:

Decision Influencer	Have I got time available to play?	Have I got the energy to play?	Do I know how to play?	Am I interested in playing?	Have I got what I need to play?
Individual					
Community					
Wider					

We have used the meeting and framework to identify a number of testable hypotheses which we have gone on to test in either the model to understand variations in participation between local authorities or to understand differences in participation between sports. The completed framework with a summary of the hypotheses is included for reference in the appendix to this report.

1.3. Causality

A key area that has been considered in the modeling is that of causality. There are many cases where the direction of causality is not obvious when considering the drivers of sports participation.

For instance, whilst it may be that a respondent's satisfaction with sporting facilities drives sports participation, it is also possible that respondents who do more sport are more satisfied with their sports facilities due to doing more sport. Where there have been contentious areas around causality, we have not included the variable in the model.

2. The Dataset

As the basis for the modeling, we have built a modeling dataset which comprises information from the latest available Active People Survey (July 2008 – July 2009) and combined this with additional datasets from other different sources.

For each individual, we have used the postcode that they provided as the basis for calculating the distance to various facilities and other geo-located features, including Active Place facilities and Clubmark clubs. Where it has not been possible to obtain information on the location of facilities, we have used any information available on the LA that the facility or feature is in.

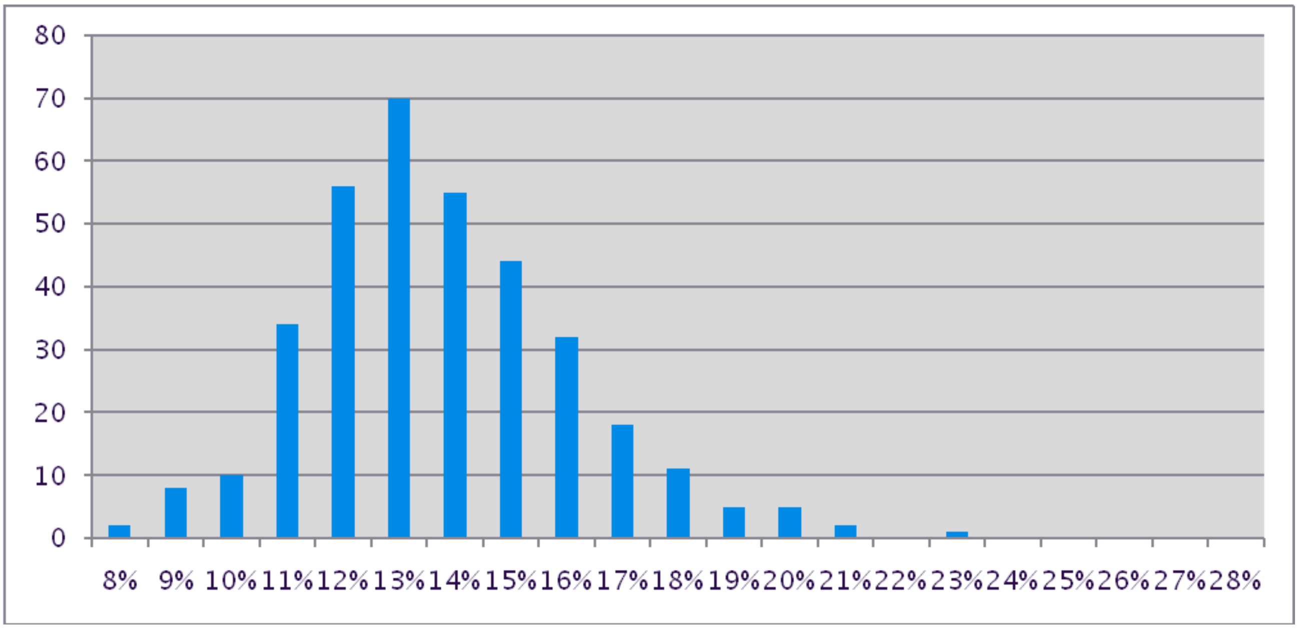
2.1. Active People Survey

The Active People Survey has been used as the basis for our analysis. The most recent data available at the time of the analysis was July 2008 – July 2009. The survey contains information for 190,899 respondents across Local Authorities in England.

Of the full sample,21,658 respondents (11.3% of the total) did not report postcode information. We have analysed the distribution of unreported postcodes across Local Authorities, since if there were certain authorities where a much higher proportion of respondents had not reported postcodes, this may lead to a bias being present in the dataset. As shown in Figure 1 below, the distribution of unreported postcodes is centred around 13%, with a close to Normal distribution. We have therefore concluded that unreported postcodes should not have a significant impact on the results from the modeling.

Figure 1 Frequency of Unreported Postcodes

% of respondents within LA not reporting postcode



2.2. Additional datasets

In order to enable us to test some of the hypotheses about drivers of participation not covered by the Active People survey, we incorporated additional datasets into the models where possible.

2.2.1. Active Places

We wanted to test how the proximity of sports facilities affected levels of individual sports participation.

The Active Places Dataset contains information on 31,308 facilities within England, across 16 different type of facility. For each respondent who has reported postcode information, we have calculated the distance to the nearest Active Places facility. In addition, we have also calculated the distance to the nearest Active Places facility for a number of sports which are relevant to the individual sports modeling, such as Active Places Rugby League pitch and Active Places swimming pool.

2.2.2. Clubmark

We wanted to test how proximity to clubs and the quality of local club networks affected sports participation.

Clubmark was introduced in 2002 by Sport England to:

- Ensure that accrediting partners apply core common criteria consistent with good practice and that minimum operating standards are delivered through all club development and accreditation schemes.
- To empower parent(s)/carer(s) when choosing a club for their children.
- To ensure that Clubmark accredited clubs are recognised through a common approach to branding.
- To provide a focus around which all organisations involved in sport can come together to support good practice in sports clubs working with children and young people.

The Clubmark dataset that has been used to build our modeling dataset contains information on 5,613 clubs that have achieved the Clubmark accreditation. This covers 49 different sports.

The Clubmark dataset includes the postcode of the club location. For each APS respondent that reported their postcode, the distance to the nearest Clubmark accredited club was calculated for each sport of interest.

In addition, the Clubmark dataset also contains the Local Authority (LA) that the club is located in. The number of Clubmark clubs located in each LA was also calculated and applied to all individuals in that authority.

2.2.3. SportsMark and ActiveMark

We wanted to test how the quality of local school PE provision affected adult sports participation.

The SportsMark and ActiveMark dataset contains information on schools that have been awarded either or both awards in 2008. In total, 18,454 schools are covered by the dataset, with 2,423 having achieved the SportsMark award and 16,520 having achieved the ActiveMark. Within the two sets there are 429 schools that achieved both.

For each individual, we have calculated the number of ActiveMark and SportsMark schools that are within their Local Authority.

2.2.4. Quality Assurance List

We wanted to test how the quality of local sports facilities affected levels of individual sports participation.

The Quality Assurance List contains information on 1,270 sports facilities that have attained one of the following quality assurances: Green Flag, CharterMark, ISO 9001:2000 or Quest. The dataset also contains information on the local authority that the facility is located in.

For each individual, we have calculated the number of facilities of each type of quality assurance within the local authority that they live in.

2.2.5. Sport England Lottery Grants

We wanted to test whether there was a relationship between the level of Sport England lottery funding in a local area and levels of individual sports participation.

The Sport England Lottery grant dataset has information on Sport England lottery grants that have been invested since 1995. Information on each grant includes the location of the grant when the grant is location-specific; the sport being supported; the award amount; and the total project cost.

For each respondent who has reported their postcode, we have calculated the Sports England Lottery Award amounts within 1, 2, 5, 10 and 20km for each of the 11 sports of interest and also the total amount across all sports. We have also calculated the total project cost in each case.

2.2.6. Big Lottery Fund Grants

We wanted to test whether there was a relationship between the level of other lottery funding in a local area and levels of individual sports participation.

We have used the grants information from the Big Lottery Fund website[1] to calculate the total amount of grants given by Local Authority.[2] We have applied this value to all individuals in the authority.

2.2.7. Sports Colleges

We wanted to test whether proximity to a specialist sports college affected adult sports participation.

We have identified schools that have been awarded specialist status as a Sports College. In total, there are 447 Sports Colleges in England.

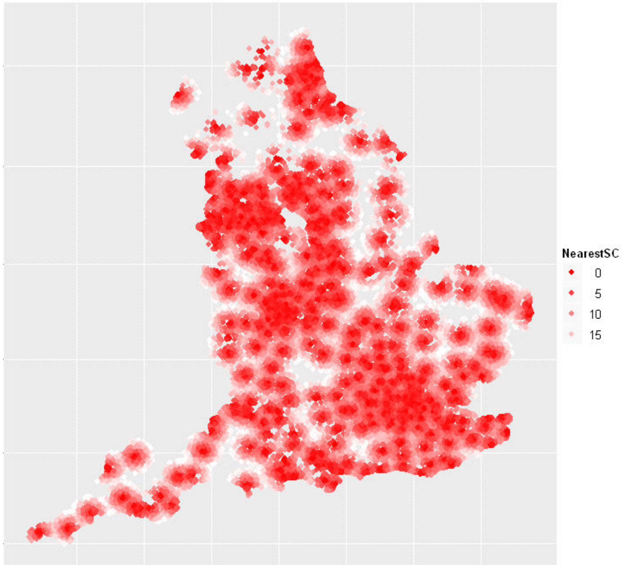
For each individual with postcode information we have calculated the distance to the nearest Sports College (in kilometres). In addition, we have also calculated the number of sports colleges within 1, 2, 5 and 10km of each individual.

---

[1] Based on using the search engine at http://www.biglotteryfundgrants.org.uk:8080/grant-search/gs_001.xsql#__utma%3D1.463642987.1274355601.1274355601.1274355601.1%3B%2B__utmz%3D1.1274355601.1.1.utmccn%3D%28referral%29|utmcsr%3Dbiglotteryfund.org.uk|utmcct%3D%2F|utmcmd%3Dreferral%3B%2B

[2] In total, there have been 20, 973 grants provided, with a total value of £3.3 billion.

Distance to nearest Sports College



2.2.8. School Sports Survey 07/08

We wanted to test how the quality of local school sport provision affected adult sports participation.

The School Sports Survey collects information on levels of participation in Physical Education and school sport amongst schools in the School Sports Partnership Programme. In 2007/8, the survey included over 21,000 schools.

The survey contains information by Local Authority against a number of measures, including:

i. Percentage of pupils who participated in at least two hours of high quality PE and out of hours school sport in a typical week

ii. Percentage of pupils involved in inter-school competition during this academic year

iii. Percentage of pupils participating in one or more community sports, dance or multi-skill clubs with links to the school during this academic year - analysis by Local Authority

For each individual, we have linked the above three values for both 2007 and 2008 onto the dataset based on the local authority in which the respondent lives.

2.2.9. Local Authority Population Density

We wanted to test whether local population density affected levels of individual sports participation.

We have used data from the 2001 UK census to calculate the population density for each Local Authority (expressed in number of persons per hectare). For each individual, we have identified the population density of their local authority. In addition, we have also identified the size of the local authority in hectares.

2.2.10. Local Authority Performance

We wanted to test whether the assessed quality of local government (including cultural ervices) affected levels of individual sports participation.

We have used Comprehensive Performance Assessment scores published by the Audit Commission as a measure of Local Authority performance.[3] Each LA is given an overall CPA score, which is split into five categories ranging from zero to four stars. The CPA score brings together assessment scores for use of resources, service assessments and corporate assessment. We have used data for 2008, which was published in March 2009.

For each individual, we have linked the CPA star category in the LA of the respondent to the dataset, along with the individual scores for Corporate Assessments and Use of Resources. Within Use of Resources, we also included information on scores for Culture and Children and Young People.

2.2.11. Local Authority Spend Levels

We wanted to test whether the level of local government expenditure on sport affected levels of individual sports participation.

We have used information on LA spend levels for the Financial Year 2007/08. This dataset splits spend into a number of categories.

For each individual we have calculated the following spends within their local authority:

i. Sports development and community recreation

ii. Sports and recreation facilities including golf courses

iii. Arts development and support

2.2.12. Local Authority Stretch Targets

We wanted to test whether the presence of local government improvement targets affected levels of individual sports participation.

As a measure of performance, local authorities choose a set of National Indicators against which they are assessed. Eighty two local authorities have chosen NI8 as one of these national indicators and a further 15 have chosen sport and recreation as a 'local target'. For each individual, we have identified whether the individual lives in a local authority that has one of these two targets.

2.2.13. Local Authority GCSE Performance

We wanted to test whether local levels of educational attainment affected levels of individual sports participation.

We have collected data on the average GCSE score performance across each local authority for 2009[4]. The score is based on the number of GCSEs achieved and the grading within them. Higher scores indicate a better GCSE performance.

For each individual, we have calculated the average GCSE score in their LA.

2.2.14. Local Authority Obesity amongst pupils

We wanted to test whether there was a relationship between levels of obesity and sports participation.

[3] CPA – The Harder Test Local government National report, National Audit Commission, March 2009

[4] Based on data collected from the Department for Children, Schools and Families' Achievement and attainment tables 2009, Available at http://www.dcsf.gov.uk/performancetables/schools_09.shtml

We have used data published from the National Child Measurement Programme for the 2007/08 school year to identify the prevalence of underweight, overweight and obese children in each LA. The dataset includes the proportion of both reception class and Year 6 pupils who are overweight and obese within the LA.

For each individual, we have calculated the proportions of underweight, overweight and obese children in both school years in their LA.

2.2.15. Lakes

We wanted to test whether proximity to inland water affected levels of individual sports participation.

UKLakes.net is a database derived from digital map data and holds a wide range of environmental information about lakes and lochs. In total, it contains the locations of more than 14,167 lakes in the UK.

For each individual where we have postcode information, we have calculated both the distance to the nearest body of open water and also the number of bodies of open water within 1, 2, 5, 10 and 20km.

2.2.16. Temperature and Rainfall Data

We wanted to test whether patterns of weather affected levels of individual sports participation.

We have collated monthly information on temperature and rainfall from 192 weather stations in the UK over the survey period.

For each individual, we have identified the location of the nearest weather station and subsequently the average temperature and total rainfall over the month of interview. Note that average temperature is across the full 24 hours of each day rather than during daytime hours.

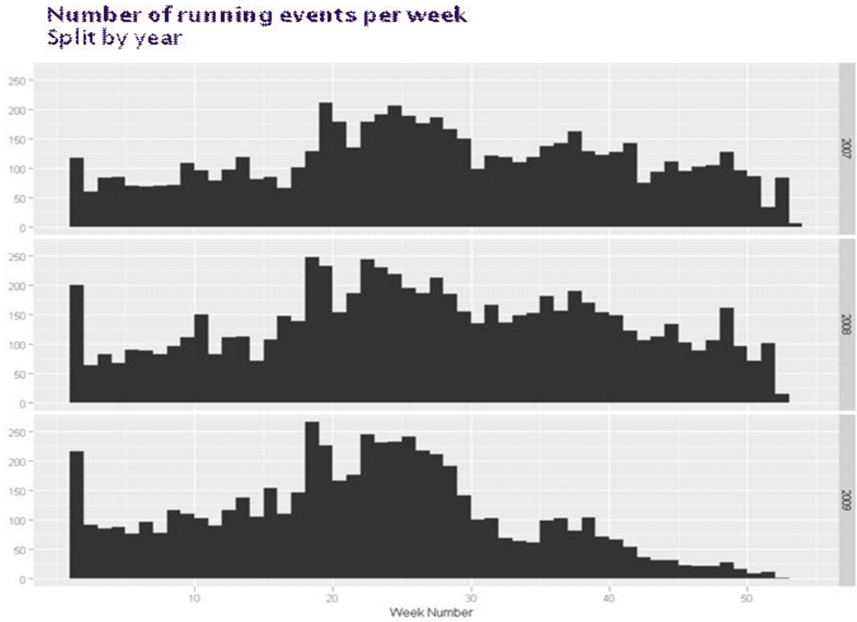
2.2.17. Running Events

We wanted to test whether the availability of competitions and events affected levels of individual sports participation.

We have compiled a database of running events that have occurred during the survey period (June 2008 – May 2009) using search engines to identify events and where possible, making use of running-related websites to identify events that have occurred previously. We have recorded details of the distance, location and date of the event.

The most consistent measure of location across events was the county in which the event occurred. For each individual we have then calculated the number of events that occurred in the month of interview in the respondent's county, along with the events that occurred up to 2 months before and 2 months after the interview. The intention was to capture events which may have caused the respondent to run more, because they were training for an event that had occurred or was about to occur. Figure 2 below shows the number of events each week between June 2007 and May 2009.

Figure 2 Number of Running Events per Week



3. Understanding variations in sports participation between Local Authorities – Model of Demographically Adjusted Participation Rates

3.1. Background

The Sport Industry Research Centre at Sheffield Hallam University built a model to understand variations in sports participation using APS1 in 2007 entitled Active People: The Model of Demographically Adjusted Participation Rates ("the Sheffield Hallam model").

The rationale behind building a model to understand variations in participation was, and continues to be, that variations in actual participation rates between local authorities are driven in part by variations in the demographics of the authorities. Without taking account of such variations in demographics the fair comparison of the relative *performance* of local authorities is difficult.

By modeling these factors, such as age and income, a participation rate can be predicted for each local authority which is based only on the population demographics and not on all the other factors that affect sports participation including the interventions of the authority.

Comparing the actual rate of participation with the predicted rate for a local authority then provides a measure of over or underperformance for the authority. Any difference between the actual and predicted will be driven by factors not included in the model, such as local authority interventions and weather conditions.

The original modeling used 'Percentage of the population taking part in at least moderate intensity sport and active recreation for at least 30 minutes duration on at least 3 days a week' (KPI 1) as the measure of sports participation.

As part of this project, we have used the variables included in the Sheffield Hallam Model to estimate NI8 participation by Local Authority using the APS2 and APS3 datasets combined. We have re-estimated the coefficients on the variables based on the different dataset.

3.2. Methodology

As far as possible, we have made use of the same set of variables that were used for the Sheffield Hallam model.

3.2.1. Change in Dependent Variable

One key difference in re-estimating the original model is the use of the NI8 measure of sports participation rather than the KPI1 participation variable that was used in the original modeling. The change is due to NI8 now being considered the most relevant measure of participation for local government. The key difference between the measures is the inclusion of 5 low intensity sports (yoga, pilates, indoor and outdoor bowls, archery and croquet) for individuals aged 65 and over.

3.2.2. Modeling Sub-Sample

A key factor when using the APS dataset for modeling purposes is the difference between the survey sample and the modeling sample. The modeling sample is a subset of the survey sample for the following key reasons:

1. As the APS is survey based, individuals can choose to refuse to answer questions or to quit the interview if they wish to. From a modeling perspective, this leads to missing values for

certain individuals for certain variables. For example, an individual may have reported their age, but not their income level.

2. At the same time, questions often have “Don’t Know” as a possible answer. From a modeling perspective, if individuals have responded “Don’t Know” then the interpretation of the results for other levels of that variable (namely “Yes” and “No”) are not as clear as they otherwise would be.

   At the same time, excluding such individuals from the modeling sample can reduce its size markedly.

Within the dataset, respondents have answered "Don't Know", "Refused" and "Respondent Quits Interview" to a number of questions. The following criteria have been applied to the APS2&3 dataset, which have reduced the survey sample to the modeling sample:

a. Respondents who did not provide income information have been removed from the dataset,

b. Respondents who did not provide an answer to Education Level, Number of Children in Household or House ownership or age finished full time education have also been removed from the dataset.

The net result of this is that the sample size is reduced from 385,272 to 251,022 which has been predominantly due to the income criteria. Note that these are "or" conditions, so if any of them are not met then that respondent is excluded from the modeling.

In addition to filtering the survey sample to the modeling sample, we have adjusted the age finished full time education so that it is interpreted as a continuous variable. This has been done in the following way: any respondent who answered "14 or less" to the question has been recoded as 14 and any respondent reporting an answer of "21 or over” has been recoded as 21.

3.2.3. Model Estimation

As described in the Sheffield Hallam report, the model has been estimated with two different weightings:

a. We have used the annual weighting to estimate the logit model results detailed in section 1.2
b. We have used the local authority weighting to estimate the coefficients used for predicting local authority participation rates.

3.2.4. Participation Rate Estimation

One output from the model is a predicted probability for each respondent as to whether or not, based on the factors included in the model, he or she will meet the NI8 criteria. This probability varies between 0 (will not meet the criteria) to 1 (meets the criteria).

To calculate predicted participation by LA, we have weighted each individual’s predicted probability of participation weighted by their LA weight. As a formula, this is:

Sum (Predicted Probability * LA Weight)/Sum (LA Weight).

The same method has been applied to calculate actual participation:

Sum (Participation * LA Weight)/Sum (LA Weight)

As noted above, a subset of the full dataset has been used for the modeling – the ***modeling sub-sample***. As we have used this sub-sample to estimate actual participation rates, these actual rates will vary from those reported based on the full dataset due to variations in the make-up of the sub-sample versus the full sample.

From a modeling perspective, the sub-sample predicted and actual participation rates should be compared when identifying Local Authority under and over-performance.

3.3. Differences to original methodology

As far as possible, we have replicated the methodology that was used by Sheffield Hallam when originally modeling the drivers of sports participation.

The reasons for the difference will be where we have had to make assumptions about the Sheffield Hallam modeling methodology, namely:

1. The restrictions described above where individuals have not provided information regarding questions including income level, which reduces the sample size;
2. Removing the access to local facilities variable from the modeling;
3. A different dependent variable has been used when re-estimating the model (NI8 vs. KPI 1); and
4. The estimation technique used to aggregate the individual level results to local authority level to calculate the predicted probabilities from the modeling to estimate participation at LA level (weighted by individual's LA weight) may differ to that used in the original modeling work.

   The Model of Demographically Adjusted Participation Rates Report does not detail how Sheffield Hallam have performed this calculation and there are several approaches to making this type of aggregation.[5]

3.4. NI8 Results

We have followed Sheffield Hallam's approach in estimating our models:

i. A national model with respondents weighted by annual weight, and

ii. A model with respondents weighted by local authority weight.

Section 1.2 provides the detailed model results.

Appendix 2 provides the detailed predicted versus actual results, split by Local Authority.

Key differences that should be noted from re-estimating the model are that three of the regions that were identified as significantly impacting on participation in the original Sheffield Hallam

[5] A typical alternative to our approach is to have a "cut off" value which gives each individual a 1 or 0 predicted participation depending on whether or not they reach the cut off criteria (e.g. a respondent with predicted probability of 0.55 would have a "1" and a respondent with predicted probability of 0.45 would have a "0"). One reason that we have not done this is that the majority of predicted probabilities from the model are less than 0.5 which is the standard "cut off" value used. If we did use this cut off criteria, the LA predicted participation rates are very low.

model are no longer significant at the 95% level. This is likely to be due to the change in the time period for the modeling, along with the change in the dependent variable.

Statistically significant changes in estimates of coefficients

In order to identify statistically significant changes in the impact of the coefficients, it is necessary to look at both the coefficient of the variable and also the standard error of the coefficient.

We have tested whether the changes in coefficients are significant by considering whether there is overlap in the 90% confidence interval around the two estimated coefficients (calculated as the coefficient ± 1.645*standard error). Please note that as the standard errors were not provided in the original report provided by Sheffield Hallam, we have re-estimated the coefficients using the APS1 dataset.

The following variables have shown statistically significant changes between the APS1 and APS2+3. Where a change has been found, the sign of that change is given.

Table 1 – Statistically significant changes between APS1 and APS2+3 model

Variable	Direction of change
Number of adults in household	Less negative
Age Band 25-34	Less positive
Age Band 35-44	Less positive
Age finished full time education	Less positive
Student full-time	Less positive
Age Band 55-64	More negative
Age Band 75-84	More negative
Age Band 85+	More negative
Male	More positive
Ethnicity-Ethnic White	More positive
Higher Education (Degree Equivalent)	More positive
Income £45,800-£51,999	More positive
One child in Household	Positive to negative
Two children in Household	Positive to negative
Three children in Household	Positive to negative
1st oldest child's age (multiple children households.)	Positive to negative

4. Understanding variations in sports participation between Local Authorities – Mindshare model

4.1. Background

Previous modeling made use exclusively of the Active People Survey information and focused primarily on the demographic characteristics of respondents. Modeling the impact of demographic factors only goes so far in explaining the local variations we see in sports participation rates. By including additional information in the model about an individual's surroundings, such as weather, local authority interventions and access to facilities, hypotheses around how these factors impact on sports participation can be tested.

4.2. Dataset

We have used an APS dataset that includes the last quarter of APS2 and the first three quarters of APS3 as the basis for our analysis. In addition to the responses from the APS survey itself, we have merged the datasets described in section 1 above into the APS dataset. As such, we have based our modeling sample on those who have reported their postcode and applied the same modeling sub-sample requirements as detailed in section 3.2.2 above.

4.3. Additional variables created from APS data for modeling

In addition to the variables that have been merged onto the APS dataset from other sources, a number of additional variables have been created by making use of the APS dataset itself. Details of these additional variables are provided below:

Variable Name	Description	Calculation
Simpson Ethnic Diversity Index	A measure of the diversity of the local authority. The index measures the probability that two individuals randomly selected from a LA will belong to the same ethnicity.	$\frac{\sum n(n-1)}{N(N-1)}$
Own Ethnic Percentage in LA	Proportion of the APS respondents within the LA who share the same ethnicity as the respondent.	$\frac{\sum n}{N}$
Own Income Percentage in LA	Proportion of the APS respondents within the LA who are within the same income band as the respondent.	$\frac{\sum n}{N}$

4.4. Methodology

We have modelled the drivers that influence whether an individual will reach the NI8 criteria, which is defined as at least 12 sessions over the last 4 weeks of sport and active recreation included in NI8 definition. As the dependent variable is a dummy, taking the value 1 if the individual has met the criteria and 0 otherwise, we have used a Logit modeling approach to take account of this.

The expected participation rates generated from a logit model are designed to have a minimum value of zero and a maximum of one. This makes it ideal for binary variables such as participation

(0 stands for non participation and 1 for participation). Values outside this range are meaningless. Linear OLS models are not restricted in this range and may return results outside the 0-1 domain.

Secondly, a non-linear logit model does not make the unrealistic 'constant returns' assumption embodied in linear models. In the OLS model a 1% change in income would return the same change in participation independently of the starting level of income. So it does not make any difference if we change by 1% an annual income of £6,000 or an annual income of £60,000 resulting in misspecification of the expected participation model.

As far as possible, we have used the most straightforward, yet statistically valid technique for our modelling.

4.4.1. Quantification of the drivers of reaching the NI8 measure of sports participation

The estimates of the coefficients from the modelling are provided in Section 1.4 below.

It should be noted that the coefficient estimates from a Logit model are the change in the log-odds ratio due to a one-unit change in the variable. Variables where the coefficient is positive increase the probability that the individual is likely to meet the NI8 criteria if they have more of that variable and negative coefficients point to the opposite. Tables 1 and 2 below summarise the drivers in the model.

Table 2 – Impact of variables in NI8 model

Variable	Impact
Social club membership	Positive
Attended cultural events over the last year	Positive
Region: East Midlands, North East, North West, South East, South West, Yorkshire	Positive
A-Levels	Positive
5 or more GCSEs	Positive
Higher education at degree level	Positive
Average temperature	Positive
Total rainfall	Positive
Income Level	Positive
Own ethnicity in area	Positive
White ethnicity	Positive
Attend cultural events	Positive
Single adult household	Positive
Male	Positive
National lottery grants awarded within 10kms	Positive
Lakes within 10kms	Positive
Own home outright	Positive
Number of children in household	Negative
Population density in local area	Negative
Live in council housing	Negative
Number of children in household	Negative
Car Van Available	Positive
Age	Negative
Illness	Negative
Four or more adults in household	Negative

Table 3 – Quantification of factors driving probability of reaching NI8 criteria[6]

<table>
<tr><th></th><th>Estimate</th><th>Std. Error</th><th>z</th><th>Pr(>|z|)</th><th>odds-ratio</th><th colspan="2">More/Less likely</th></tr>
<tr><td>(Intercept)</td><td>-2.890</td><td>0.099</td><td>-29.13</td><td>0.000</td><td>0.06</td><td></td><td></td></tr>
<tr><td>SocialClubMemberAdj</td><td>0.959</td><td>0.022</td><td>43.68</td><td>0.000</td><td>2.61</td><td></td><td>||||||||||</td></tr>
<tr><td>culturalevent1</td><td>0.034</td><td>0.030</td><td>1.13</td><td>0.259</td><td>1.03</td><td></td><td>|||</td></tr>
<tr><td>culturalevent2</td><td>0.013</td><td>0.026</td><td>0.50</td><td>0.620</td><td>1.01</td><td></td><td>|</td></tr>
<tr><td>culturalevent3</td><td>0.209</td><td>0.018</td><td>11.76</td><td>0.000</td><td>1.23</td><td></td><td>|||||||</td></tr>
<tr><td>SE_RegionEast Midlands</td><td>0.053</td><td>0.034</td><td>1.57</td><td>0.116</td><td>1.05</td><td></td><td>||||</td></tr>
<tr><td>SE_RegionLondon</td><td>-0.043</td><td>0.041</td><td>-1.05</td><td>0.292</td><td>0.96</td><td>|||</td><td></td></tr>
<tr><td>SE_RegionNorth East</td><td>0.120</td><td>0.041</td><td>2.96</td><td>0.003</td><td>1.13</td><td></td><td>|||||</td></tr>
<tr><td>SE_RegionNorth West</td><td>0.084</td><td>0.032</td><td>2.66</td><td>0.008</td><td>1.09</td><td></td><td>|||||</td></tr>
<tr><td>SE_RegionSouth East</td><td>0.028</td><td>0.029</td><td>0.96</td><td>0.339</td><td>1.03</td><td></td><td>||</td></tr>
<tr><td>SE_RegionSouth West</td><td>0.124</td><td>0.033</td><td>3.81</td><td>0.000</td><td>1.13</td><td></td><td>|||||</td></tr>
<tr><td>SE_RegionWest Midlands</td><td>-0.001</td><td>0.034</td><td>-0.02</td><td>0.980</td><td>1.00</td><td>|</td><td></td></tr>
<tr><td>SE_RegionYorkshire</td><td>0.125</td><td>0.033</td><td>3.82</td><td>0.000</td><td>1.13</td><td></td><td>|||||</td></tr>
<tr><td>awardamount10</td><td>0.000</td><td>0.000</td><td>3.62</td><td>0.000</td><td>1.00</td><td></td><td>|||</td></tr>
<tr><td>AverageTemp</td><td>0.019</td><td>0.002</td><td>11.59</td><td>0.000</td><td>1.02</td><td></td><td>|||||</td></tr>
<tr><td>TotalRainfallAdj</td><td>0.004</td><td>0.002</td><td>1.76</td><td>0.078</td><td>1.00</td><td></td><td>|</td></tr>
<tr><td>d6heduc1</td><td>0.307</td><td>0.024</td><td>12.94</td><td>0.000</td><td>1.36</td><td></td><td>|||||||</td></tr>
<tr><td>d6heduc2</td><td>0.326</td><td>0.030</td><td>11.02</td><td>0.000</td><td>1.39</td><td></td><td>||||||||</td></tr>
<tr><td>d6alevels</td><td>0.235</td><td>0.025</td><td>9.33</td><td>0.000</td><td>1.26</td><td></td><td>|||||||</td></tr>
<tr><td>d6gcse</td><td>0.166</td><td>0.027</td><td>6.09</td><td>0.000</td><td>1.18</td><td></td><td>||||||</td></tr>
<tr><td>Male</td><td>1.780</td><td>0.105</td><td>16.96</td><td>0.000</td><td>5.93</td><td></td><td>|||||||||||||</td></tr>
<tr><td>poly(RespondentAge,2,raw=TRUE)1</td><td>0.026</td><td>0.004</td><td>6.78</td><td>0.000</td><td>1.03</td><td></td><td>||||||||</td></tr>
<tr><td>poly(RespondentAge,2,raw=TRUE)2</td><td>-0.001</td><td>0.000</td><td>-11.98</td><td>0.000</td><td>1.00</td><td>||||||||</td><td></td></tr>
<tr><td>ethwethnic</td><td>0.162</td><td>0.079</td><td>2.06</td><td>0.039</td><td>1.18</td><td></td><td>||||||</td></tr>
<tr><td>OwnEthnicPct</td><td>0.269</td><td>0.100</td><td>2.70</td><td>0.007</td><td>1.31</td><td></td><td>||||</td></tr>
<tr><td>d23_bands_7\24315,600 to \24320,799</td><td>0.104</td><td>0.032</td><td>3.23</td><td>0.001</td><td>1.11</td><td></td><td>|||||</td></tr>
<tr><td>d23_bands_7\24320,800 to \24325,999</td><td>0.121</td><td>0.035</td><td>3.49</td><td>0.000</td><td>1.13</td><td></td><td>|||||</td></tr>
<tr><td>d23_bands_7\24326,000 to \24331,199</td><td>0.233</td><td>0.033</td><td>7.17</td><td>0.000</td><td>1.26</td><td></td><td>|||||||</td></tr>
<tr><td>d23_bands_7\24331,200 to \24336,399</td><td>0.303</td><td>0.035</td><td>8.62</td><td>0.000</td><td>1.35</td><td></td><td>|||||||</td></tr>
<tr><td>d23_bands_7\24336,400 to \24351,999</td><td>0.367</td><td>0.030</td><td>12.35</td><td>0.000</td><td>1.44</td><td></td><td>||||||||</td></tr>
<tr><td>d23_bands_7\24352,000 or more</td><td>0.645</td><td>0.031</td><td>21.03</td><td>0.000</td><td>1.91</td><td></td><td>|||||||||</td></tr>
<tr><td>LAPopulationDensity</td><td>0.002</td><td>0.001</td><td>1.95</td><td>0.051</td><td>1.00</td><td></td><td>|||</td></tr>
<tr><td>I(CarVanAvailable==1)TRUE</td><td>0.323</td><td>0.038</td><td>8.52</td><td>0.000</td><td>1.38</td><td></td><td>||||||||</td></tr>
<tr><td>NumAdultsHousehold==1TRUE</td><td>0.183</td><td>0.022</td><td>8.48</td><td>0.000</td><td>1.20</td><td></td><td>||||||</td></tr>
<tr><td>illness1</td><td>-0.347</td><td>0.022</td><td>-16.11</td><td>0.000</td><td>0.71</td><td>|||||||</td><td></td></tr>
<tr><td>d7own</td><td>0.199</td><td>0.020</td><td>10.05</td><td>0.000</td><td>1.22</td><td></td><td>||||||</td></tr>
<tr><td>d7council</td><td>-0.158</td><td>0.039</td><td>-4.02</td><td>0.000</td><td>0.85</td><td>||||||</td><td></td></tr>
<tr><td>NumChildHouseholdAdj1</td><td>-0.107</td><td>0.022</td><td>-4.95</td><td>0.000</td><td>0.90</td><td>|||||</td><td></td></tr>
<tr><td>NumChildHouseholdAdj2</td><td>-0.163</td><td>0.023</td><td>-6.95</td><td>0.000</td><td>0.85</td><td>||||||</td><td></td></tr>
<tr><td>NumChildHouseholdAdj3</td><td>-0.064</td><td>0.041</td><td>-1.58</td><td>0.114</td><td>0.94</td><td>||||</td><td></td></tr>
<tr><td>NumChildHouseholdAdj4 or more</td><td>-0.424</td><td>0.083</td><td>-5.09</td><td>0.000</td><td>0.65</td><td>|||||||</td><td></td></tr>
<tr><td>lakeswithin10</td><td>0.000</td><td>0.000</td><td>1.70</td><td>0.089</td><td>1.00</td><td></td><td>|</td></tr>
<tr><td>Male:poly(RespondentAge,2,raw=TRUE)1</td><td>-0.073</td><td>0.005</td><td>-14.85</td><td>0.000</td><td>0.93</td><td>|||||||||</td><td></td></tr>
<tr><td>Male:poly(RespondentAge,2,raw=TRUE)2</td><td>0.001</td><td>0.000</td><td>12.84</td><td>0.000</td><td>1.00</td><td></td><td>|||||||||||</td></tr>
<tr><td>LAPopulationDensity:I(CarVanAvailable==1)TRUE</td><td>-0.005</td><td>0.001</td><td>-5.98</td><td>0.000</td><td>1.00</td><td>||||</td><td></td></tr>
</table>

[6]Note: the bar chart is approximate – the sizes of bars are proportional to the odds ratios raised to the power of one for dummy variables or the odds ratios raised to the power of one standard deviation for continuous variables

4.4.2. Interpretation of Outright Home Ownership within modeling

Within both the NI8 modeling and also the sports specific modeling, outright home ownership has been found to be a significant, typically positive, driver of sports participation measured in terms of both the probability of reaching the NI8 criteria and also the probability of being active and the frequency of sport participation.

There are reasons to think that, by itself, outright home ownership may have a positive impact, such as the additional disposable income available from not having mortgage or rent payment contributions. However, at the same time, there are likely to be additional elements of a respondent's affluence and life stage that are also being captured by the estimates of the impact of the variable (and outright home ownership may be a proxy for these things).

4.4.3. Interpretation of Car Availability impact within modeling

Within the modeling, car/van availability has been found to have a positive effect on the probability of achieving the NI8 criteria. At the same time, individuals who live in more dense Local Authorities are more likely to reach the criteria than individuals who live in sparser populated authorities.

However, an interaction also exists between Population Density and Car Ownership. That is to say that the impact of car ownership varies by the population density of the area where an individual lives.

In rural areas, car ownership has a positive impact on the probability of reaching NI8 criteria. In urban areas, it has a negative impact – car owners are less likely to meet the target.

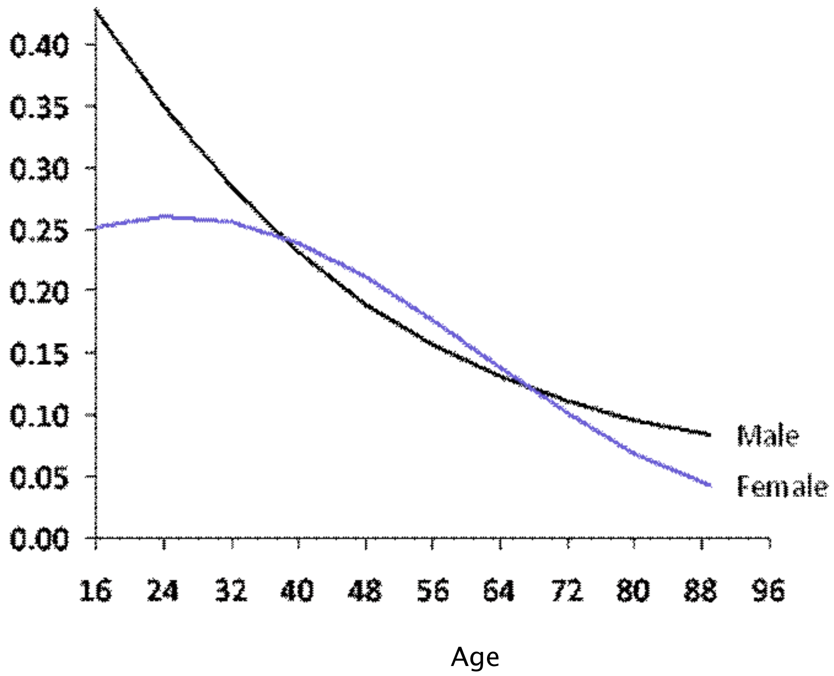
This is likely to be due to the better transport links and proximity to sport facilities in urban areas making it easier for an individual to participate in sport without having access to a car. In urban areas, car ownership may also be correlated with other lifestage and lifestyle effects which are linked to reduced participation in sport.

4.4.4. Interpretation of Respondent Age impact within modelling.

Within the modeling, we have used a quadratic form to model the impact that age has on the probability of reaching the NI8 criteria. We have also tested whether the impact of age varies between males and females and found it to differ.

Figure 3 below shows how the probability of achieving the NI8 criteria changes with age, holding everything else equal. Whilst the probability of participation for men is significantly higher for men at young ages, it falls much more rapidly, dropping below the females' probability in the forties and fifties, before rising above it again in the late sixties.

Figure 3 – Probability of achieving NI8 criteria, split by gender

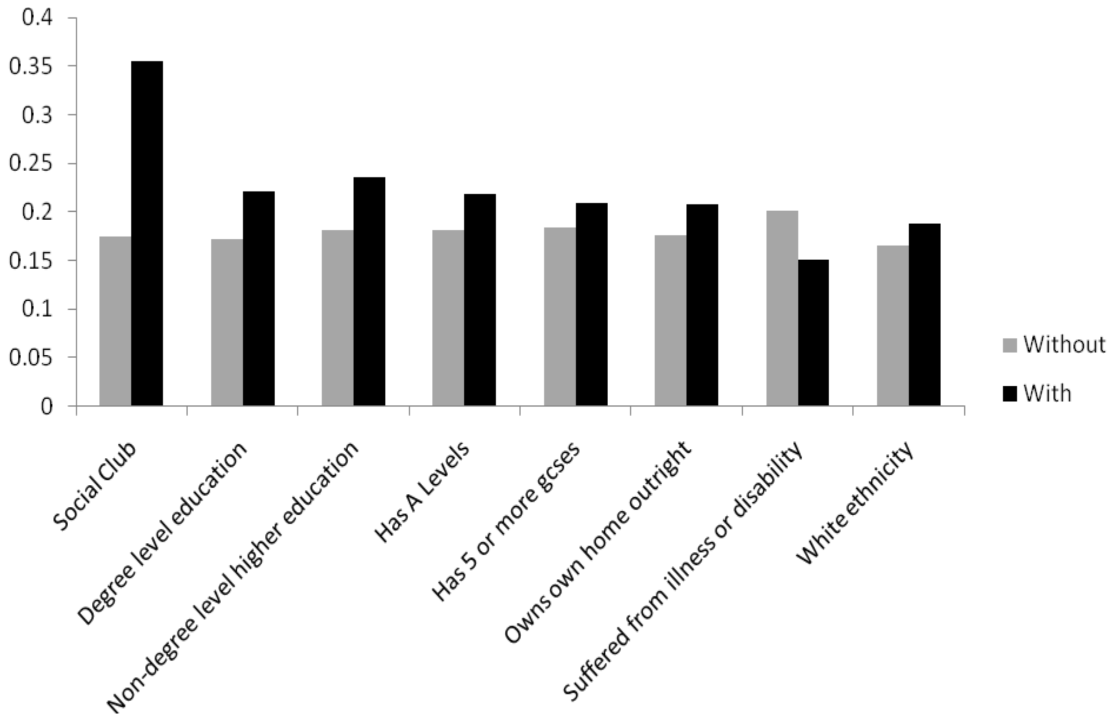


4.4.5. Interpretation of coefficients

To enable further interpretation of the model, we have estimated the impact that each of the drivers has on reaching the NI8 criteria, holding all other variables at their average and estimating the average participation rate at different levels of the variable.

In the case of dummy variables, such as gender, this is the average predicted participation rate where the individual does and does not meet the dummy, i.e. Male versus Female. The figure below provides estimates of the probability of achieving the NI8 criteria when each dummy variable is met versus not being met. For instance, the values for Male indicate the expected probabilities of Male (dummy=1) versus Female (dummy=0), holding everything else constant.

Figure 4 - Probability of reaching NI8 criteria based on whether or not dummy variable is achieved



4.4.6. Results from Hypothesis Testing

As described above, we have used a hypothesis-based approach to building each of our models. The completed framework with a summary of the hypotheses is included for reference in the appendix to this report,

The following paragraphs summarise how the results from the models relate to each hypothesis. Instances where this study has not identified a significant correlation between a variable and sports participation should not be interpreted as a conclusive evidence that these factors are of no importance. Other factors relating to the limitations and availability of data will have prevented comprehensive and conclusive testing in some areas.

a. Have I got what I need?

Me

Higher incomes will increase participation due to more resources being available

We have found within the NI8 modeling and also within the modeling of individual sports that higher household income has a significantly positive impact on both the probability of reaching the NI8 criteria and on the probability of being an active individual in the selection model.

Within the individual sports modeling, household income has been found to have a positive impact on the frequency of participation in golf.

Team sport is harder to organise so has lower participation than individual sports

Participation rates tend to be lower in team sports than individual sports like swimming and cycling. However, this is not a testable hypothesis within the modeling framework that we have used.

Amount of coastline (and open water) in local area (and access to this) drives participation in water sports

Within the NI8 model, the number of lakes within 10km of the individual has been found to be significant as a positive driver of being more likely to reach the NI8 criteria. It should be noted that part of this effect may capture other elements of the natural environment such as open spaces around lakes which could be used for activites like running,.as well as the impact on water sports

People will switch to cheaper sports when their economic position worsens

The time period over which the model has been developed is too short to be able to isolate the impact of macroeconomic effects and data for individual respondents relates to one point in time. However, there is some evidence to suggest that some sports may be more sensitive to changes in people's economic circumstances than others. In particular, income is a strong driver of frequency of participation in golf suggesting that it may be subject to such an effect.

Within the NI8 and 1 million indicator modeling, the quarter of interview has not been found to have a significant impact on the probability of reaching the criteria. It has been found to be a driver in some of the sports models, including tennis. However, this is more likely to be reflective of the seasonality of tennis rather than economic conditions.

b. Community and Local Institutions

Higher population density gives you greater critical mass around which to organize participation. Lower density populations give you more space for outdoor activity. Rural areas for some sports have lower participation rates due to lack of facility provision.

The modeling has tended to find that higher population density increases the likelihood of reaching the NI8 criteria, but also has a positive impact on the frequency of participation in tennis amongst people who have taken part in at least one session of sport in the last four weeks.

However, at the same time it has a negative impact on frequency of participation in golf, squash and badminton – that is to say that, all else being equal, respondents living in more rural local authorities tend to participate more frequently in these sports.

It may be hypothesised that there is a more complex mechanism at work where lifestyle choice, community and demand in rural areas encourage sports such as badminton.

Local authorities that invest more in sport have higher rates of participation

We have tested the total and per capita amount of spend by local authorities in FY07/08 across sports related categories, specifically spend within the respondent's Local Authority on Sports Development, Sports Facilities, Museums and Galleries and Arts Development. Within the model these spends have not been found to have a significant impact on the probability that an individual will achieve the NI8 criteria.

Schools with accreditation in sports generally or sports-specific accreditations lead to long-term higher levels of participation for their students / local population

Within the modeling, we have tested whether the distance to the nearest Sports College and number of sports colleges within 1, 2, 5 and 10km has an impact on the probability of reaching the NI8 criteria and also on frequency of participation in individual sports. In both cases, it was not found to be a significant driver of participation.

We have also tested whether the number of ActiveMark and SportsMark accredited schools within the respondents local authority has an impact on participation and have not found it to have an impact.

Local authorities with higher CPA scores have higher participation rates. Some local authorities that are not good at engaging with their communities in sport will also be worse at engaging generally with the community at a broader level

Within the NI8 modeling, we did not find that CPA star category in the LA of the respondent impacted on the probability of achieving the NI8 criteria. In addition, we also tested whether the individual scores for Corporate Assessments, Use of Resources and Culture had an impact on the probability of participation, but did not find these to have a significant impact on the probability of reaching the NI8 criteria.

Local authorities with a strong commitment to improvement of cultural / sporting services (e.g. stretch target that includes NI8, etc.) will have higher participation rates

We have tested whether individuals who live in local authorities that have an NI8 stretch target tend to be more likely to reach the NI8 criteria of participation. However, we have not found it to be a significant driver of increased probability of participation.

Areas where SSPs have higher levels of PE, offer the greatest range of activities and have the greatest number of club links have higher participation rates

Within the NI8 model, we have tested whether the percentage of pupils who participated in at least two hours of high quality PE and out-of-hours school sport in a typical week have higher participation rates. However, this was not found to be a significant driver of the probability of reaching the NI8 criteria.

It should be noted that whilst some information on SSPs was available, this did not include information on the location of the schools, making it not possible to geo-locate the school and tie the information to individual respondents.

The availability of facilities (quality facilities) in an area will have an impact on sports participation

In each sport model, we have tested the distance to both facilities (as recorded by the Active Places database) and accredited clubs (as recorded by the Clubmark database). In general, we have not found that overall sports participation (as captured by the NI8 measure) is impacted by the availability of facilities or accredited clubs. It should be noted the number of running clubs within 20km has been found to be a significant driver of increased frequency within the athletics model.

Within the NI8 modeling, we have also tested whether the number of Quest, Green Flag, ISO 9001:2000 or Charter Mark accredited facilities within the respondent's local authority impact on the probability of a respondent reaching the NI8 criteria. In all cases, we have not found the number of accredited facilities to be a significant driver of changing the probability of reaching the criteria.

One reason to think that this is the case is that whilst the availability of facilities increases the *opportunity* to participate in sport, looking at actual participation there will not be sufficient difference between individuals who do participate and those who don't based on the distance to sports facilities.

Areas receiving higher levels of lottery funding (SE lottery funding) for sports projects have higher participation rates

We have tested overall SE Lottery funding within the NI8 model and also specific sports funding within each of the 11 sports. Within the NI8 model, SE lottery amounts within 10km of a respondent lead to that respondent being more likely to participate in sport to an NI8 level (at least 12 sessions in last 4 weeks).

c. Am I interested in playing?

Me

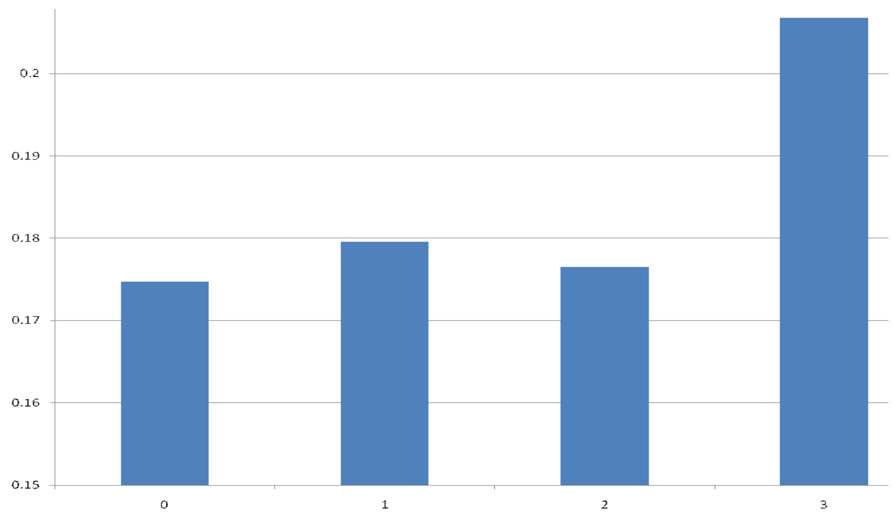
People who continue in higher education may have higher participation rates over long-term

Across the NI8 model, selection and outcome models described above, there is a consistently emerging trend that those who have attained a higher-education qualification, degree or otherwise are more likely to participate in sport, and across many sports do it more frequently having controlled for all other factors.

People who engage more in cultural activities or civic life engage more in sport

The number of cultural events that the individual has attended in the last year has a significant impact on the likelihood of reaching the NI8 criteria. Those who have attended three or more events are more likely to reach the criteria than those who have attended fewer or no events. The variable that we have used when testing this (and in all cases where Cultural Events are included within the sport specific modeling is based on DCMS Q4 "How many events have you attended? (all that have attended any event in last 12 months)". Figure 5 below shows the probability of reaching the NI8 criteria based on attendance at cultural events, holding all else equal.

Figure 5 - Probability of reaching NI8 criteria based on attendance at cultural events



Attendance at cultural events has also had an impact on frequency of participation within a selection of the sports models. However, there have been mixed results, with attendance having a positive effect on frequency of swimming and rugby union participation, but a negative impact on both cycling and football. This may be due to a time trade-off, where cultural engagement leaves less time available for sports participation.

Sports with more satisfied participants have higher levels of participation

Although the original Sheffield Hallam model included the respondent's satisfaction with local sporting provision within the list of drivers of participation, as described above, we have omitted this variable from our modeling due to issues of causality, since people who participate in sport are likely to rate sporting provision differently from people who do not participate in sport.

It should be noted that this modeling does not make use of the Satisfaction Survey dataset (due to the lack of sufficiently granular geo-locational data) which provides insight into participant satisfaction.

People who watch more sport also participate more in sport

We have been unable to test this hypothesis as the APS survey does not include information on the respondent's television viewing habits.

It should be noted that the Taking Part survey does include data about media consumption and modeling of this data elsewhere does suggest a relationship.

Sports where participation can be training for another sport have higher participation rates

From the APS, it was possible to identify which sports an individual participated in and the overlaps between different sports. However, there is not a definite causal direction in these overlaps – whilst an individual may participate in two sports, to identify causality requires something more than just correlation.

It should be noted that in more recent versions of the Active People Survey (which were not available at the time of modeling), a question has been added to the survey for those respondents reporting participation in weight training. The question enquires as to whether those respondents do it as preparation for another sport. With the addition of this question, it is possible that future modeling may be able to address this hypothesis.

Community and Local Institutions

A more diverse community may result in there being less social cohesion, and hence lower overall rates of participation within the community

We have used three measures of diversity when considering the impact on participation rates:

a. the proportion of other respondents in the Local Authority who share the ethnicity of the individual,
b. the proportion of other respondents in the Local Authority who are in the same household income band as the individual, and
c. the Simpson Ethnic Diversity Index, which gives the probability that two individuals randomly chosen from a Local Authority will be of the same ethnicity

Within the NI8 modeling, it has been found that the proportion of the local authority that is the same ethnicity as the respondent has a positive impact on the probability of reaching the NI8 criteria. This variable has been calculated in the same way for all respondents, however its impact is mainly on respondents from ethnic minorities since its value does not vary as much for white (majority) respondents.

It was not found that the Simpson Ethnic Diversity Index nor the proportion of respondents within the same income band as the respondent had a significant impact on the probability of reaching the NI8 criteria.

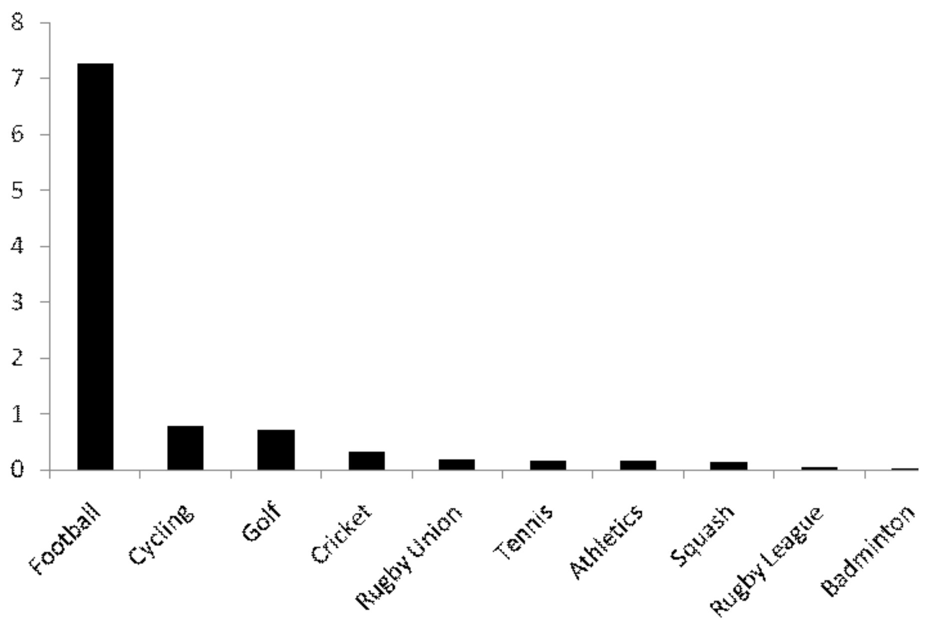
Local authorities that consistently have high performing academic scores will have higher levels of participation

The average GCSE score of the local authority has been used to test this hypothesis. Within the NI8 model, it was not found to be a significant driver of the probability of reaching the NI8 criteria. It has been found to be significant as a driver of respondents being more likely to be active within the selection model (doing at least one "1 million" indicator session over the last 4 weeks). Given that the criteria required to achieve the required level of activity (only one session over last four weeks), this suggests that average GCSE score only has an impact on the likelihood of being active rather than reaching a higher level of sport participation.

The size of the gap between gender participation rates is a driver of overall participation rates

In each sports model, we have tested whether gender is a significant driver which increases the frequency of participation in that sport. We have found that gender impacts on football, athletics, rugby league, cycling, badminton, golf, squash and cricket. The chart below shows the variations in the impact that being male has on each of these sports in terms of incremental days of the sport over 28 days. Football is the sport where being male has the biggest impact.

Figure 6 – Incremental days of participation due to being male, based on Outcome Models



It should be noted that gender impacts on all models. In the case of swimming, which is not included in the figure, it has been found that changes in frequency of participation due to age differ between men and women.

We have not tested whether the gap between gender participation rates impacts the frequency of participation as the models have not been set up in such a way as to test this.

Sports where patterns of participation are more equitable (e.g. they appear to appeal to a wider demographic group) have higher participation rates overall

We have tested whether factors such as household income and type of occupation impact on the frequency of participation across each of the sports. Higher household income has been found to have a positive impact on Rugby League, Rugby Union and Golf. Occupation type has been found to have an impact on tennis and football, with those in professional occupations more likely to take part in tennis and less likely to take part in football.

There are a number of sports where neither household income nor occupation has an impact on participation: athletics, cycling, swimming, badminton, squash and cricket. Of these sports, athletics is the only one where ethnicity does not impact on the frequency of participation – it may be hypothesised that these sports are more “democratic” and accessible to a wider variety of individuals, particularly in the case of athletics.

Greater access to competitive opportunities (leagues, tournaments, events) will increase participation

Participation in organised events and higher participation in the related sport are closely linked. However, the direction of causality is not obvious – whilst those training for an event may train more than those who are not, it may also be the case that those who participate more then choose to participate in events.

Within the athletics model, it was tested whether the number of running events in the respondent’s county in the months around the month of interview had an impact on the frequency of participation in athletics. However, it was not found to be significant driver.

National popularity (linked to elite success and / or major events) at the time will increase participation in a particular sport. More TV Screening of a sport will increase participation in a particular sport.

We have not been able to directly test whether popularity of a sport at the time will increase participation.

As described above, we have tested whether the quarter of interview has an impact on the probability of participation to reach the NI8 or 1 million indicator criteria. As a proxy for changes in the popularity we have tested whether the quarter of interview has an impact on frequency of participation within each of the sports models. An impact has been measured within the tennis model. However, this is likely to be reflecting a combination of factors – both the interest in tennis events around Wimbledon, but also the seasonality in playing tennis.

5. Understanding variations in participation between sports

5.1. Background

In addition to understanding the variations in participation rates between Local Authorities, the second part of the analysis seeks to understand the variations in participation rates between 11 different sports.

We have used the APS2Q4+APS3Q1-Q3 dataset as the basis for the modeling. The same sub-sample that was used for the NI8 modeling (described in Section 3) has also been used for modeling the individual sports.

Each regression identifies the factors that determine sports participation and the strength of its influence,

5.2. Sports Analysed

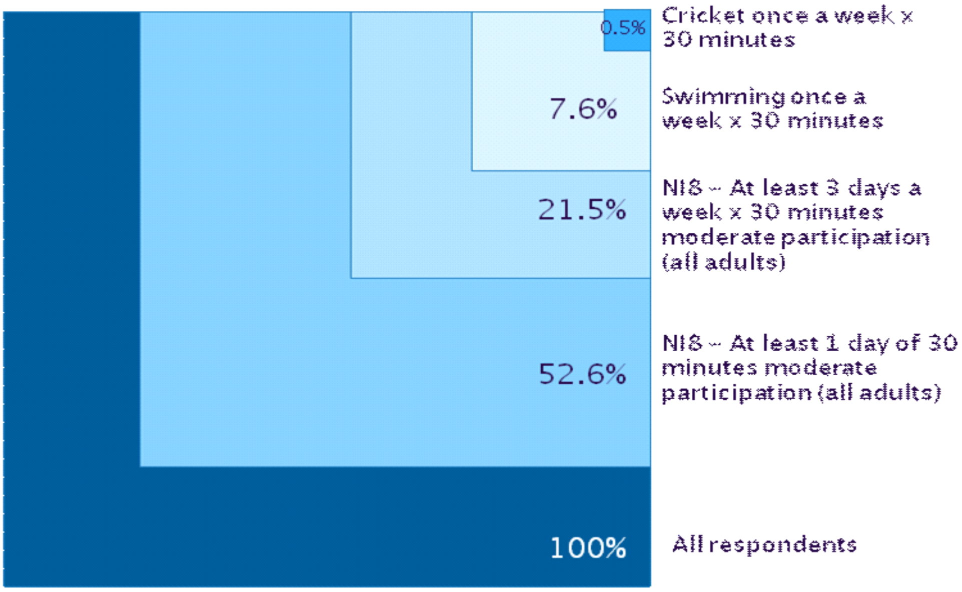
We have analysed the following eleven sports:

1. Athletics
2. Tennis
3. Football
4. Rugby Union
5. Rugby League
6. Squash
7. Badminton
8. Swimming
9. Cycling
10. Cricket
11. Golf

5.3. Modeling Technique

Within the modeling of individual sports, participation rates are typically much lower than for a measure of overall sports participation. Within the 11 sports that we have analysed, participation rates (at least once a week) vary from 7.6% in Swimming to 0.5% in the case of cricket.

Figure 7 – Analysis of sample based on proportion achieving criteria listed



In the region of 90-95% of respondents will not have participated in a particular sport. This leads to the dataset being heavily censored at 0 within each sport as it is not possible for respondent to do a negative amount of days of sport. This censoring in the dataset leads to a selection bias if the model is estimated using a typical OLS technique.

Heckman (1976) proposed a two-stage estimation procedure using the Inverse Mills' Ratio to take account of the selection bias. In a first step, a regression for observing a positive outcome of the dependent variable is modelled with a probit model ("the selection model"). The second stage is then an OLS model of frequency which uses only a subset of the modeling sample, based on whether the respondent has achieved the positive outcome in the first stage ("the outcome model").

The estimated parameters from the selection model are used to calculate the inverse Mills ratio, which is then included as an additional explanatory variable in the OLS estimation.

In the case of all 11 sports, we have first built a selection model. This model identifies those who engage in at least some sport – taken to be at least 1 session over the last four weeks that is within the definition used for the 1 million sport indicator. As the dependent variable in this model is a binary choice: those who have attained the criterion and those who have not, we have used a probit rather than a standard OLS model as the latter would lead to bias results.

Figure 8 Overview of Factors

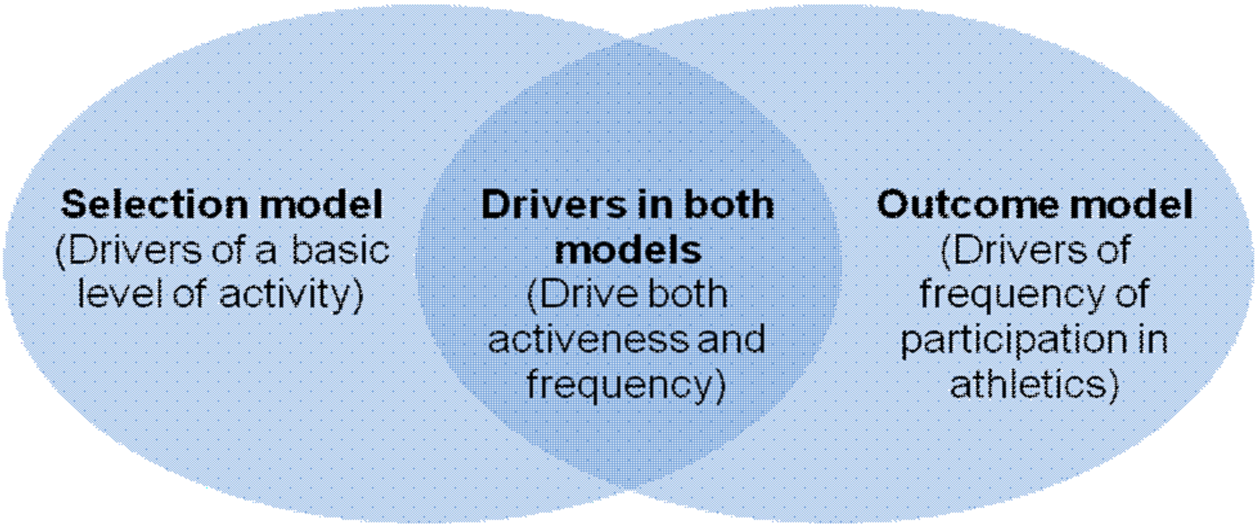
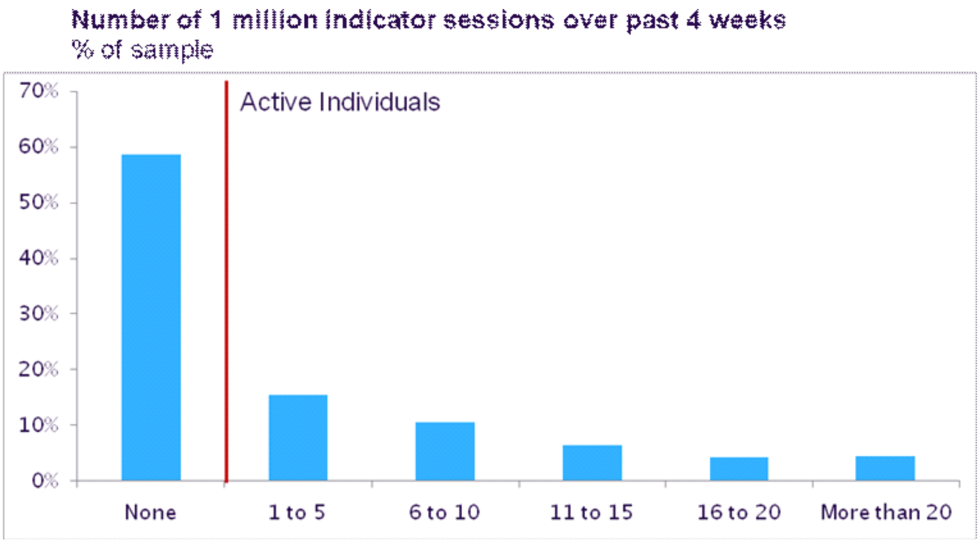


Figure 9



Relative to the NI8 modeling described above, where the NI8 criterion of at least 12 sessions over the last four weeks was used as the binary choice variable, the variable is significantly more lenient in terms of attainment (not because of what is included in the '1 million indicator' which is narrower than NI8 but because of a relaxation of the frequency threshold from three sessions a week (12 in 4 weeks) to one in 4 weeks.

The second stage of the modeling involved building a separate model for each sport to understand what the drivers of frequency of participation are. We have used frequency rather than a discrete dummy variable due to the additional detail that the frequency data provides and the application that this has to policy. An Ordinary Least Squares (OLS) modeling technique has been used to estimate the model coefficients.

5.4. Test of suitability of selection model methodology

One test for the use of the two step selection model methodology is that the coefficient on the Inverse Mills Ratio (IMR) in the outcome model is significantly different from zero. For each of the 11 individual sports models, the IMR was found to be significant.[7]

5.5. Interpretation of Results

5.5.1. Selection Model Coefficients

The estimated coefficients from the selection model are interpreted in the same way as from any other probit model – that is that it is the z-score for that particular variable. For instance, the coefficient on the "Male" variable is the change in the z-score due to being male. For a continuous variable, the coefficient is the change in z-score due to a 1 unit increase in the variable.

To aid the interpretation of the coefficients, we have evaluated the probability of achieving the criteria at different levels of each variable (both continuous and discrete). Each calculation is based on holding the value of all other variables in the equation at their mean values.

These results are similar, but not identical to the results of the NI8 model described earlier in this report,

a. Demographics

Respondents are more likely to be active where:

a. Respondents who live in households with higher incomes than those in households with lower incomes,
b. Those who attend cultural events (taken as three or more) than those who attend fewer or no cultural events,
c. Those who have attained a degree-equivalent qualification,
d. Those who have attained a non-degree higher education qualification,
e. Those who have attained an A-Level qualification
f. Males are more likely to be active than females.
g. Respondents where there are children in the household. The impact is greatest where there are two children in the household. There is a smaller, but still positive effect with three children.

Respondents are less likely to be active where:

a. Respondents live in council owned accommodation,
b. Respondents suffer from a limiting long term illness,
c. Respondents have access to a car or van are less likely to be active than those who do not have access to a car.

Ethnicity also has a role in the likelihood of being active: white individuals are more likely to be active, all else being equal. At the same time, Asians are less likely to be active as they get older than other ethnicities.

b. Local Environment

The higher the proportion of individuals in the local authority who share the same ethnicity as the individual increases the probability that an individual will be active.

[7] The Inverse Mills Ratio is the ratio of the probability density function over the cumulative distribution function of a distribution.

Also, individuals who live in local authorities that have a higher deprivation index are less likely to be active than those who live in areas with a lower index. At the same time, those who live in Local Authorities with a higher average GCSE score are more likely to be active than individuals who live in Local Authorities with lower scores. Although significant here, these variables did not show a significant impact in the NI8 model.

Temperature also impacts on the likelihood of being active, with those who were interviewed where the weather was warmer being more likely to be active than those who were interviewed where the temperature was cooler. It should be noted that the average temperature in the month of interview is a function of both the location of the individual and also the month of interview.

5.5.2. Outcome Model Coefficients

In the case of the interpretation of the estimated coefficients from the outcome model, variables fall into 2 groups:

1. Variables that are only a driver of frequency, and
2. Variables that are both a driver of frequency for particular sport (the outcome model), but also a driver of participation in sport more generally (the selection model).

In the case of (1), the coefficient is interpreted as the change in the number of days in a 4 week period that an individual will participate in a particular sport, assuming the individual is active (has taken part in at least one 30 minute session of sport in the last four weeks).

In the case of (2), the interpretation of the coefficient is slightly more complicated. Taken alone, the coefficient is the change in frequency for only active people. However, as the variable has also had an impact on the probability that the individual is active, the overall effect may be ambiguous where the coefficients are of opposite signs in the selection and outcome equations.

We have used the following equation to calculate the combined effect from the two models:

$$gamma = beta - (alpha * rho * sigma * IMRdelta)$$

Where gamma is the combined effect on frequency, beta is the coefficient on the variable in the outcome model, alpha is the coefficient in the selection model, rho is the estimated correlation coefficient between the error term of the selection equation and the outcome equation, sigma is the standard error of the error terms of the outcome equation and IMRdelta is the δs calculated from the inverse Mills Ratios and the results of the 1st step probit estimation.

Where necessary, for variables where the functional form differs between the two equations alpha or beta has been treated as being 0 and the resultant coefficients used to construct the necessary function to understand the overall impact that the variable has on participation.

For each sport described below, we have provided a commentary of the findings from the modeling, along with a bar chart showing the outcome model and combined effect (where necessary) for all categorical variables. Where the outcome model effect and combined effect are the same (i.e. where the variable only appears in the outcome model) we have reported only the outcome model effect.

The table of coefficients provides a full list of variables within the model. It should be noted that all levels of a categorical variable have typically been included in the table, although some may not be statistically significant. Where the t-value is less than 1.645, the coefficient on the variable is not statistically significant from zero (at the 90% confidence level).

5.5.3. Membership of clubs within team sports

Within the team sports, namely Football, Cricket, Rugby League and Rugby Union, an individual often needs to be part of a team to be able to participate at a sport. Therefore, club membership is an important driver of participation and it would be expected to have a positive impact on the frequency of participation within the team sports.

At the same time, many of the factors that are likely to have an impact on the frequency of participation would also be expected to have an impact on the probability that an individual is a club member. This means that club membership is an endogenous variable, that is, influenced by other variables within the model. This potentially leads to bias in the estimates of the effects of other variables within the model.

We have therefore taken the approach of building an additional "mezzanine" model for these four sports which seeks to explain the drivers of being a club member in each of the sports. The predicted probability of being a club member (the dependent variable from the mezzanine model) is then used as an explanatory driver in the frequency of sport participation model. The effects of the drivers within the "mezzanine" model can be interpreted as the impact of that variable on either increasing or decreasing the probability that an individual is a member of a club in that particular sport.

A probit modeling technique has been used for the mezzanine model. The Inverse Mills Ratio from the selection model has also been included in this model, which is based on the active people subset.

For each of the team sports, we have built two models of the effects of drivers of frequency of participation – one which includes the club membership variable and the other which instead includes the predicted probability of club membership. The difference between the coefficients on frequency of participation between the two models will include the bias due to the endogeneity of club membership.

a. Athletics

Athletics is primarily composed of running – around 95% of all athletics is road or cross country running. On average, people who participate in sport (defined as at least one session to 1 million indicator definition over the last 4 weeks) will do 1.33 sessions of running over the course of 4 weeks.

Figure 10

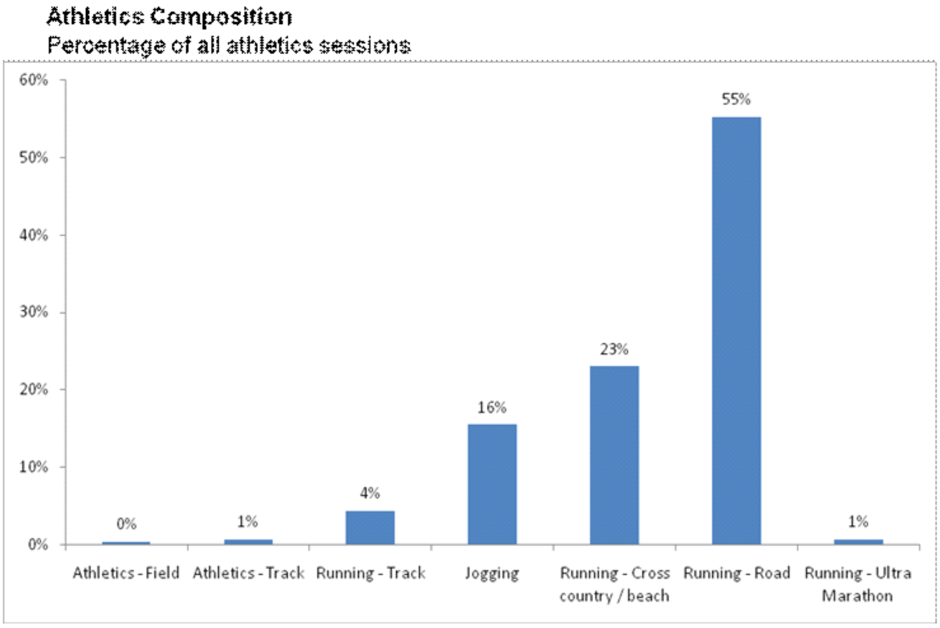
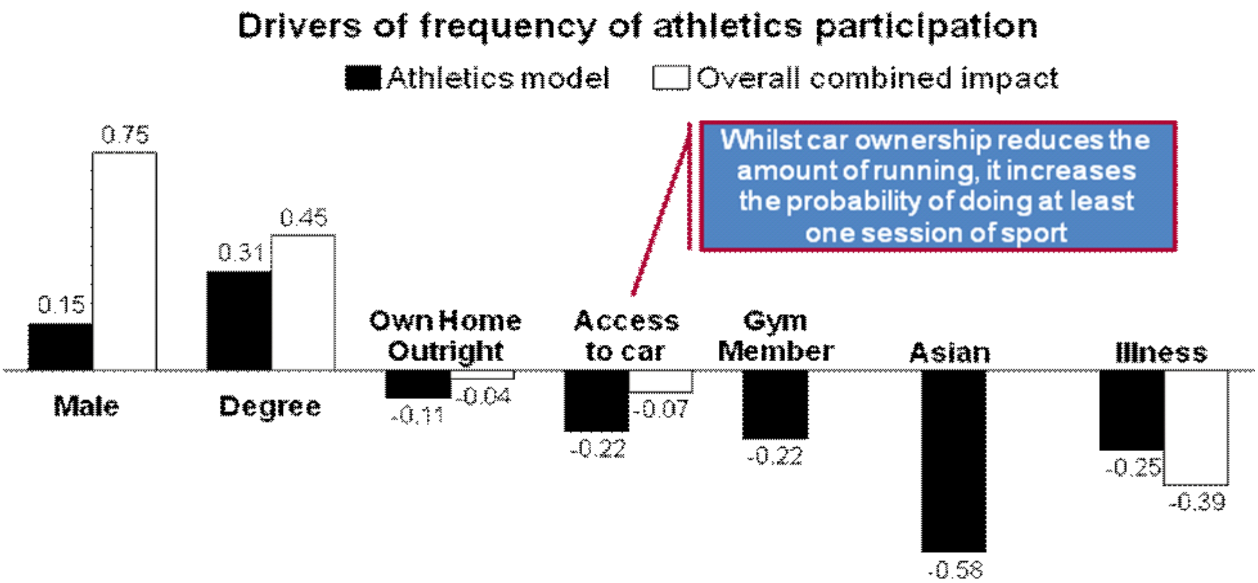


Figure 11 – Drivers of incremental days of Athletics



The key drivers of increased participation in athletics are set out in Table 4 and summarised in the following paragraphs:

Asian people tend to participate in athletics significantly less often than other ethnicities.

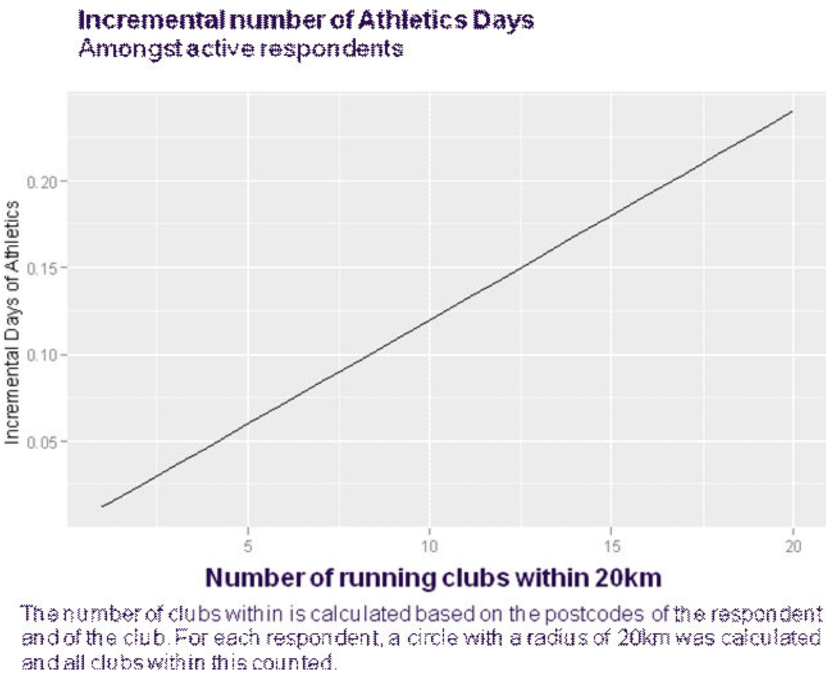
The age of the respondent is a significant driver of frequency of participation, with athletics frequency amongst active individuals peaking in the thirties before declining. However, when combining this result with how age influences whether an individual is active, the overall impact of the two is that running frequency tends to decline with age. This result corresponds to findings from *Social Interactions and the Demand for Sport – An Economic Analysis.*[8]

Gym membership has a negative impact on running frequency, with members running 0.29 days less than non-members. This suggests a substitution effect, with active individuals using gym as an alternative way of keeping fit to athletics.

Higher education is an important positive driver of frequency – above and beyond the effect that it has on determining whether an individual participates in sport. Combining the two effects, it is estimated that higher education to degree level increases the number of days of athletics by 0.45 days over four weeks.

The club network is a particularly important area within the range of possible interventions. From the modeling, the number of clubs within 20km of an individual has a significant positive influence on participation. Figure 12 visualises the incremental days of running due to the number of running clubs within 20km of the respondent.

Figure 12



[8] Downward and Riordan, *Social Interactions and the Demand for Sport – An Economic Analysis. Available for download at https://uhra.herts.ac.uk/dspace/bitstream/2299/2868/1/902959.pdf*

Table 4 – Athletic Model Coefficients

	Estimate	t value
(Intercept)	1.7887	9.8
Attained a degree-level qualification	0.3221	8.06
Respondent Age	0.0196	3.17
Respondent Age ^ 2	-0.0003	-5.21
Male	0.1639	4.38
Log (Deprivation Index)	-0.0933	-2.73
Gym Member	-0.2157	-5.58
Own Home Outright	-0.101	-2.26
Car Van Available	-0.19	-2.97
Number of athletics clubs within 20km	0.0126	4.2
Asian Ethnicity	-0.598	-5.48
Respondent has limiting long lasting illness	-0.2708	-5.37

a. Tennis

The key drivers of increased participation in tennis set out in Table 5 are:

- having a degree level education and/or A levels
- working in a professional occupation
- living in a Local Authority with a denser population
- having older children rather than younger children

Seasonality also matters, with active individuals who were interviewed during the early summer (Apr – June) playing more tennis than those in late summer and during the winter months.

People that participate in sport who live in more deprived area are less likely to play tennis as are those who are members of a gym.

Regionally, those living in the South East and South West tend to play tennis more frequently (all else being equal) and those living in the North East and Yorkshire less often (in both cases relative to the East region).

Figure 13

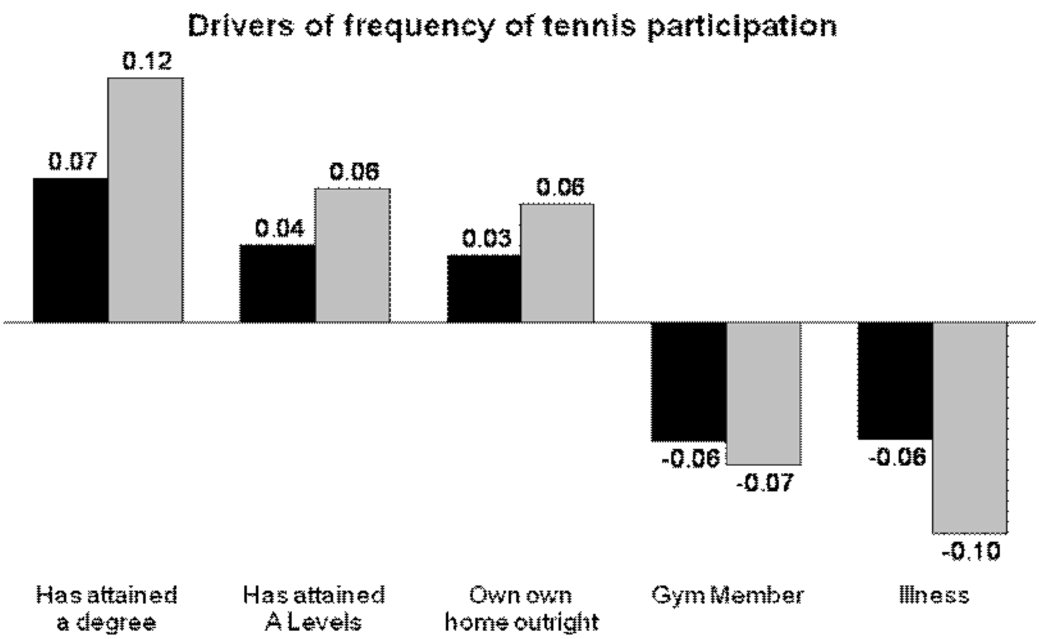


Figure 14 – Incremental days of tennis by SE Region[9]

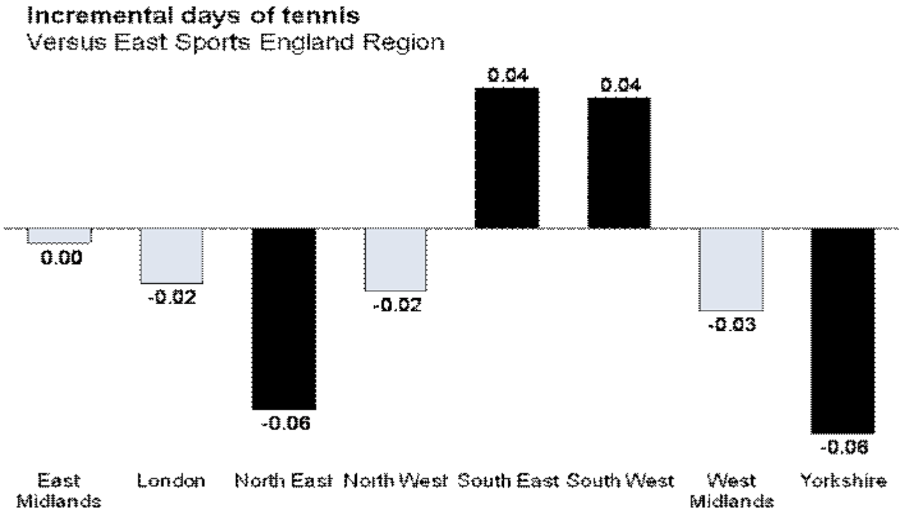
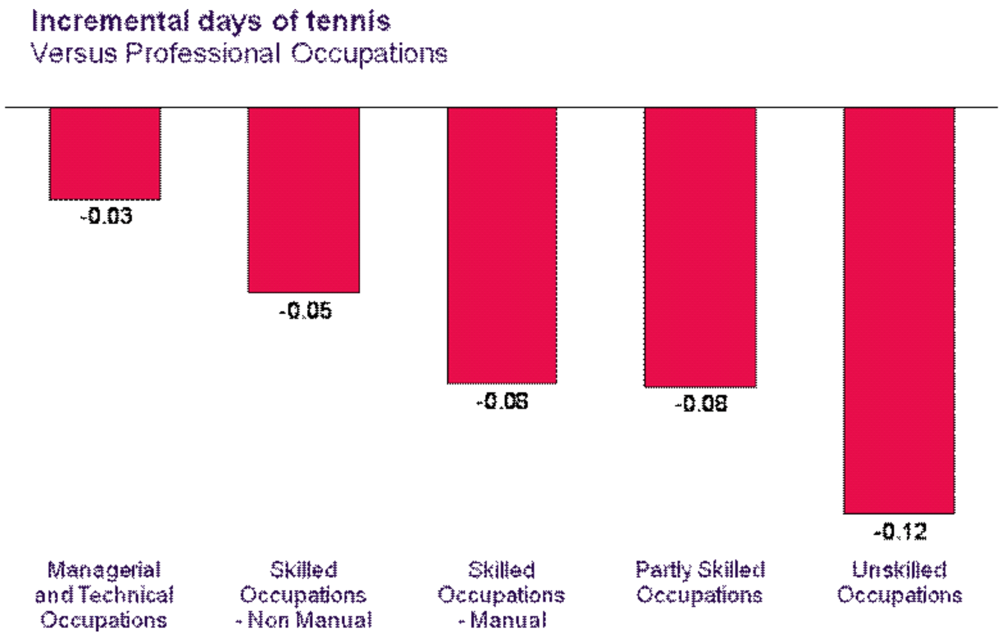


Figure 15



[9] Bars in grey represent regions where frequency of participation is not statistically different to the East Region.

Figure 16

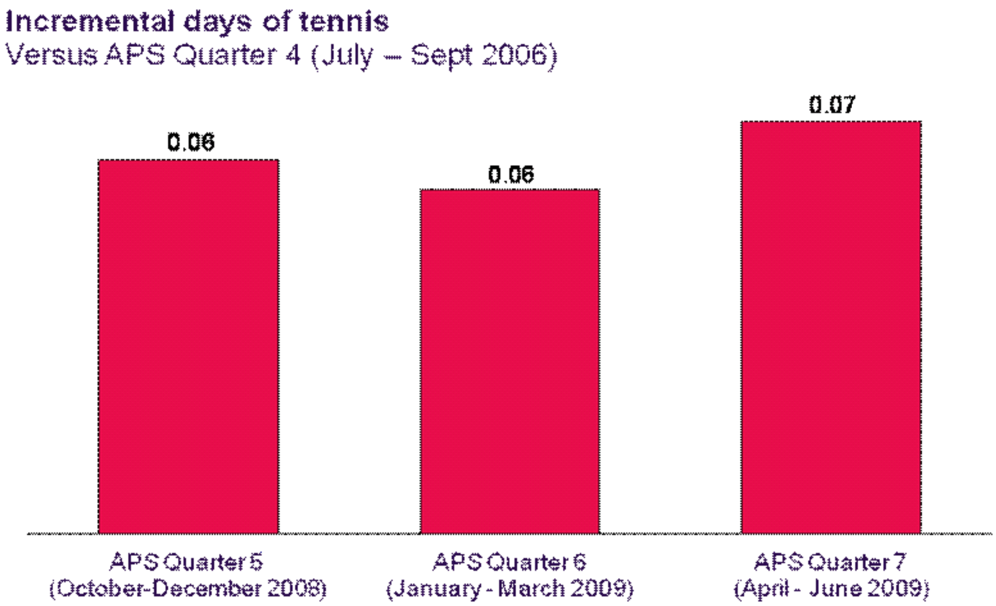


Figure 17

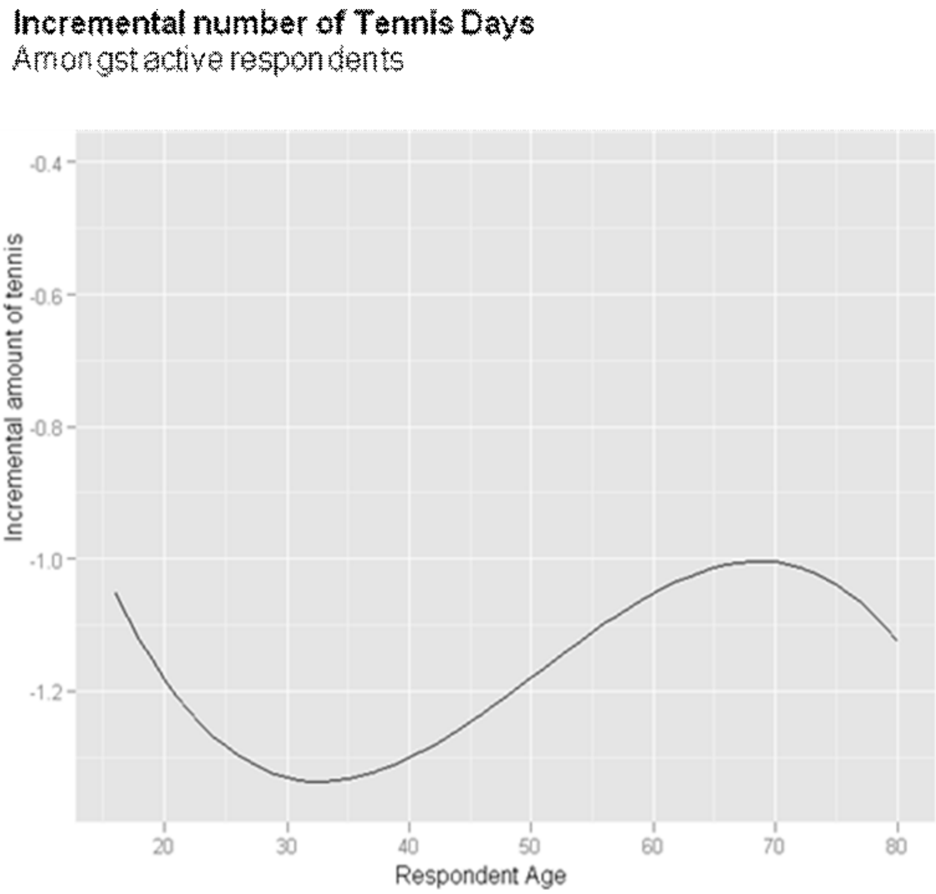


Table 5 – Tennis Model Coefficients

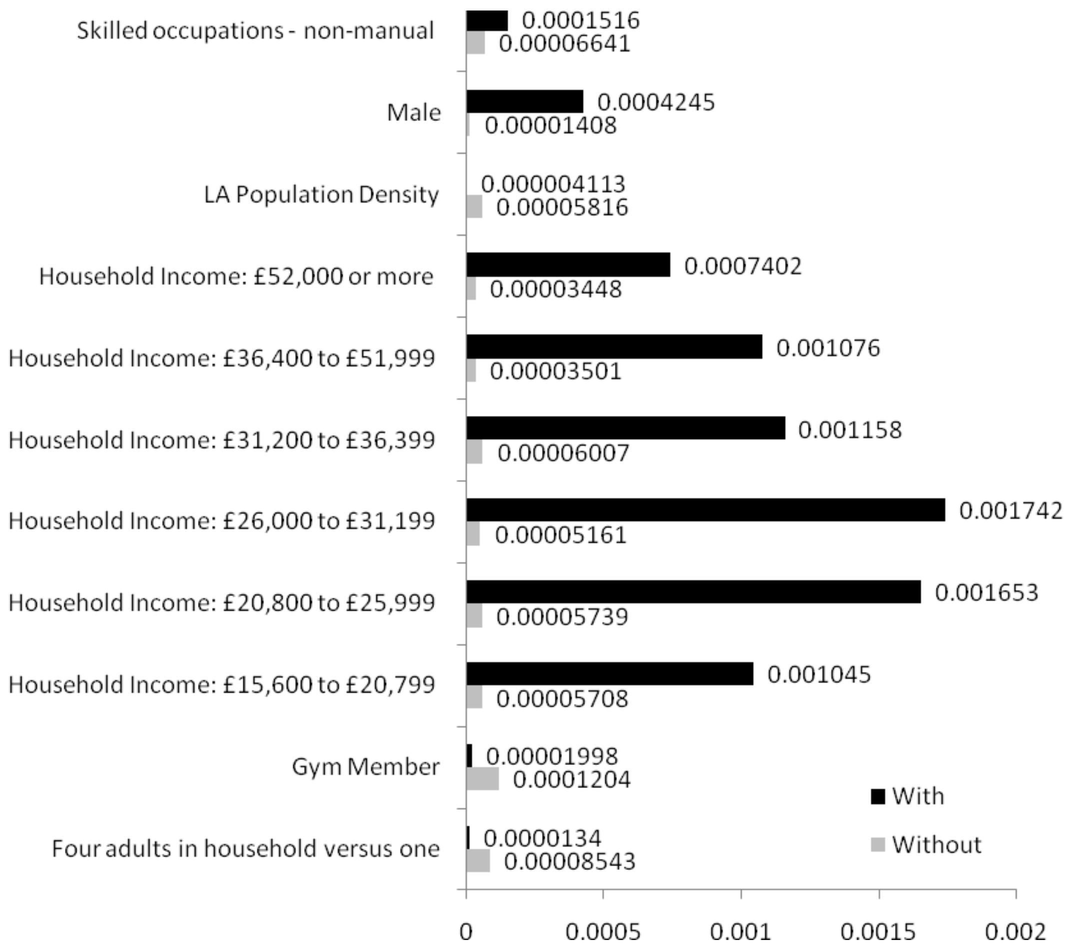
	Estimate	t-value
(Intercept)	1.6856	11.97
Age of oldest child in household	0.0085	5.27
1st oldest child's age (multiple children households)	0.0099	5.32
Own home outright	0.0327	1.88
Log(Deprivation Index)	-0.0591	-3.16
Gym Member	-0.0565	-3.83
Age	-0.0971	-11.13
Age ^ 2	0.0022	11.38
Age ^ 3	0	-10.89
Average Temperature	-0.0023	-0.29
Average Temperature ^ 2	0.0008	2.05
Managerial and Technical occupations	-0.0268	-1.16
Skilled occupations - manual	-0.08	-3.02
Skilled occupations - non-manual	-0.0539	-1.99
Partly skilled occupations	-0.0812	-2.35
Unskilled occupations	-0.1176	-2.06
Interview Quarter 5	0.0648	2.03
Interview Quarter 6	0.0595	1.6
Interview Quarter 7	0.0714	3.36
Attained degree-level qualification	0.0692	4.15
Attained A-Levels	0.0374	2.07
SE Region: East Midlands	-0.0044	-0.17
SE Region: London	-0.0176	-0.5
SE Region: North East	-0.0571	-1.72
SE Region: North West	-0.02	-0.76
SE Region: South East	0.0444	1.94
SE Region: South West	0.0441	1.66
SE Region: West Midlands	-0.0262	-0.99
SE Region: Yorkshire	-0.0644	-1.89
LA Population Density	0.0021	3.89
Respondent has limiting long lasting illness	-0.0558	-2.94

b. Rugby League

Drivers of Club Membership

Household income and being male both have a positive impact on the probability of an individual being a member of a rugby league club (income having the most positive impact around the middle income brackets). In addition, those who work in skilled non-manual occupations are more likely to be members. The probability of being a member of a club falls as age increases (which is in line with both football and rugby union club membership), as does being a member of a gym.

Figure 18 – Probability of being member of rugby league club[10]



Drivers of Frequency of Participation

The key drivers of rugby league participation are region: those who live in Yorkshire and the North West tend to play more often than elsewhere. Men also play more often than females, with frequency of participation falling as age increases. Those who are in households with higher incomes tend to play less often than those with lower incomes.

The distance to Clubmark rugby league clubs and Active Places relating to rugby league were also tested in our model. However and more so than other sports, the locations of both Clubmark clubs

[10] Age values are at 32 years for without and 56 years for with, LA Population Density values are at 29 for without and 101 for with.

and Active Places is likely to be demand driven and closely related to areas closely associated with traditionally playing rugby league. It was not unexpected then that both were found to be insignificant when tested.

Impact of using Mezzanine Club Membership probability variable

The coefficients remain relatively unchanged using the probability variable instead of the dummy variable within the model, all retain the same sign and the magnitudes remain relatively similar.

Figure 19 – Incremental Days of rugby league amongst Active and All People

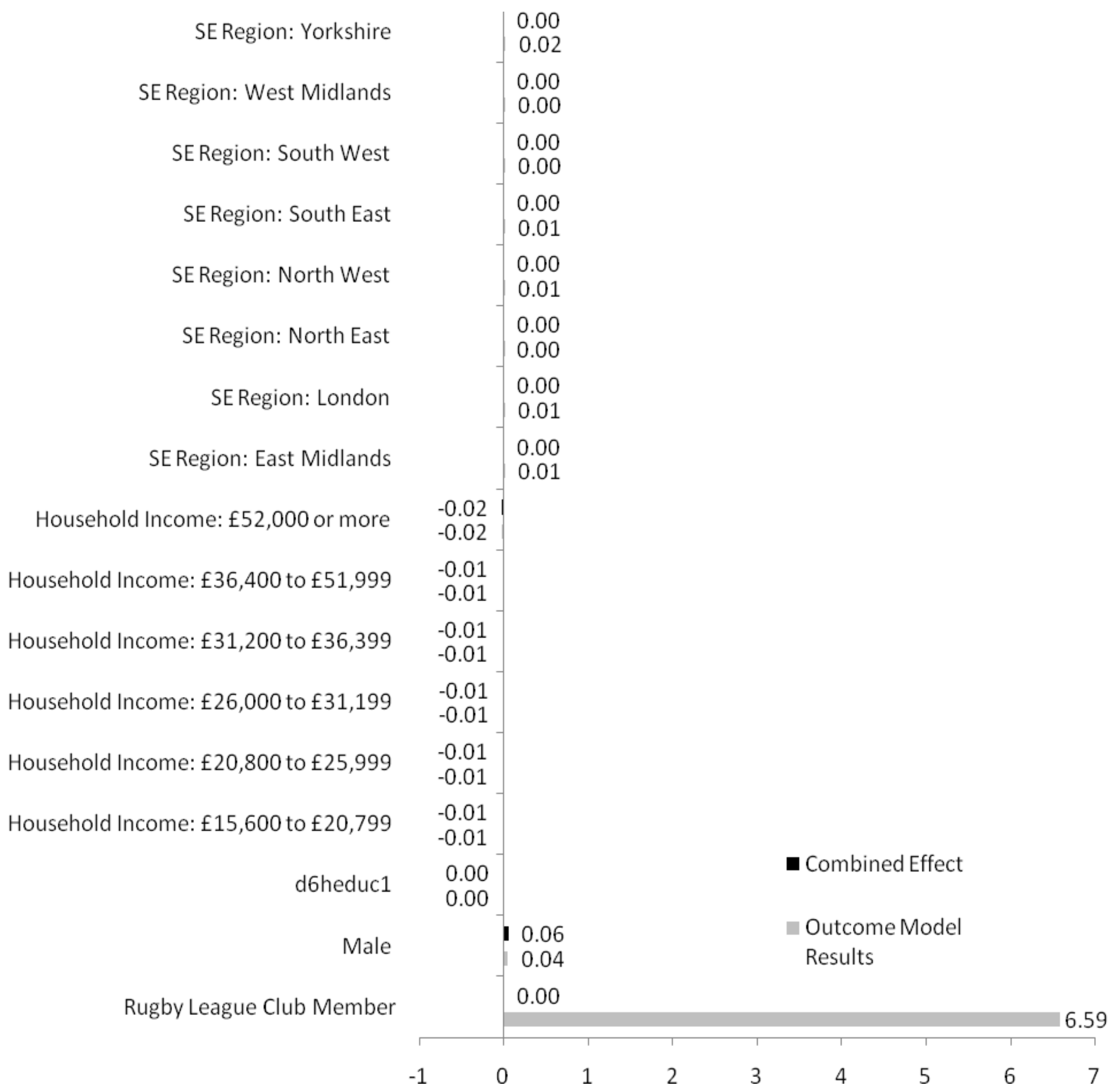


Table 6 – Rugby League Model Coefficients

	Estimate	t value
(Intercept)	0.022	2.18
Rugby league club member	6.5891	177.82
Male : respondent age	-0.0008	-5.42
Male	0.0444	5.62
Degree-level higher education attained	-0.0035	-1.25
Household Income: £15,600 to £20,799	-0.0148	-2.9
Household Income: £20,800 to £25,999	-0.0116	-2.1
Household Income: £26,000 to £31,199	-0.008	-1.5
Household Income: £31,200 to £36,399	-0.0121	-2.1
Household Income: £36,400 to £51,999	-0.013	-2.61
Household Income: £52,000 or more	-0.0212	-3.92
SE Region: East Midlands	0.0066	1.29
SE Region: London	0.0064	1.23
SE Region: North East	0.0011	0.18
SE Region: North West	0.0103	2.15
SE Region: South East	0.0116	2.62
SE Region: South West	0.0048	0.94
SE Region: West Midlands	0.0042	0.83
SE Region: Yorkshire	0.0219	3.37

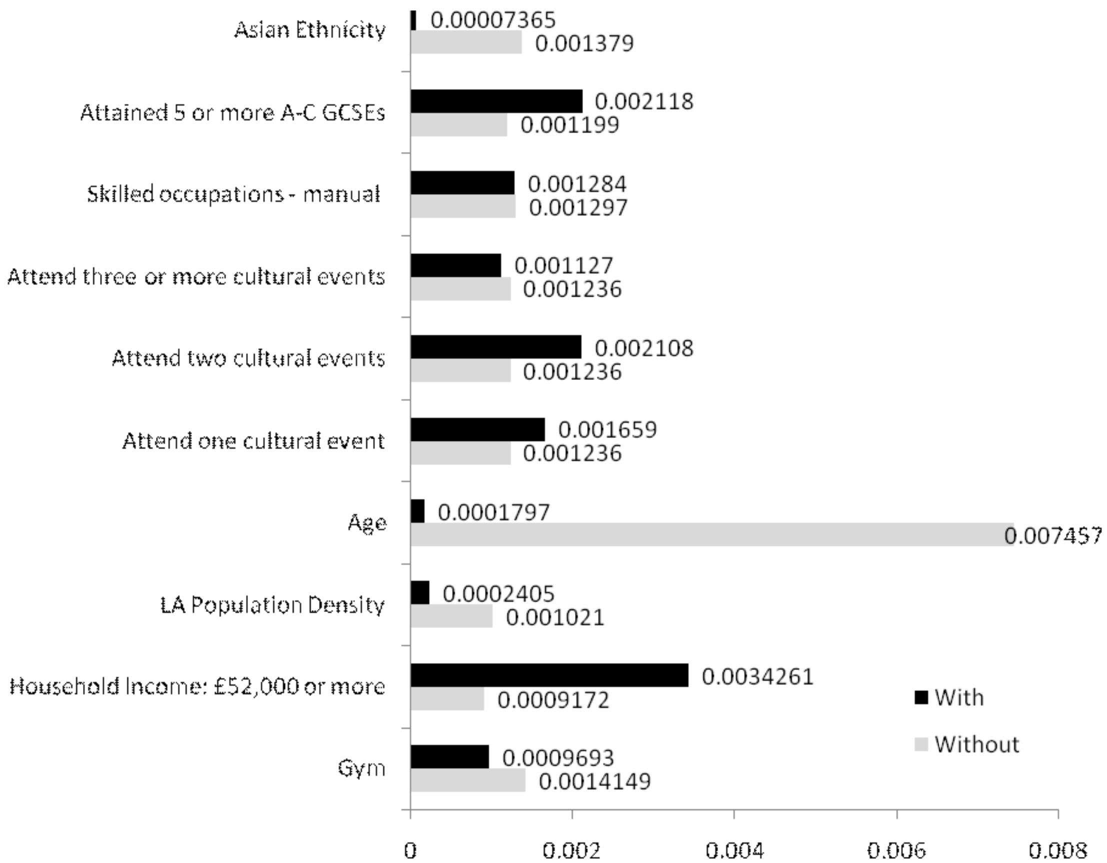
c. Rugby Union

Drivers of Club Membership

Attendance at cultural events and being male are the main positive drivers of being a member of a rugby union club. Those who live in areas where the population is less dense are more likely to be members.

In line with rugby league, age and gym membership both have a negative impact on the probability of being a member. In addition, respondents who are Asian are less likely to be members, as are those who work in skilled manual occupations.

Figure 20 – Probability of being member of rugby union club[11]



Drivers of Frequency of Participation

Key drivers of participation include the level of education that the individual has achieved, with individuals who have achieved some level of higher education playing less frequently than others. As would be expected, male active individuals participate more frequently than females. Frequency of participation falls as age increases for men.

[11] Age values are at 32 years for without and 56 years for with, LA Population Density values are at 29 for without and 101 for with.

Active respondents who are Asian participate less frequently than other ethnicities. In line with football, the number of adults in the household has an impact on the frequency of participation, with households containing three or four adults playing more frequently.

Temperature also has an impact with frequency of participation falling amongst active respondents as temperature increases – this effect is likely to be capturing the seasonality of the sport.

Impact of using Mezzanine Club Membership probability variable

Using the Club Membership probability variable within the modeling has had an impact on the majority of the variables. Key changes to coefficients include the impact of being male falling significantly and the negative impact of attaining higher education reducing in magnitude. These differences between Figures 20 and 21 represent the bias introduced with the club membership variable and the impact of both should be considered as drivers of participation.

Figure 21 – Incremental Days of rugby union amongst Active and All People

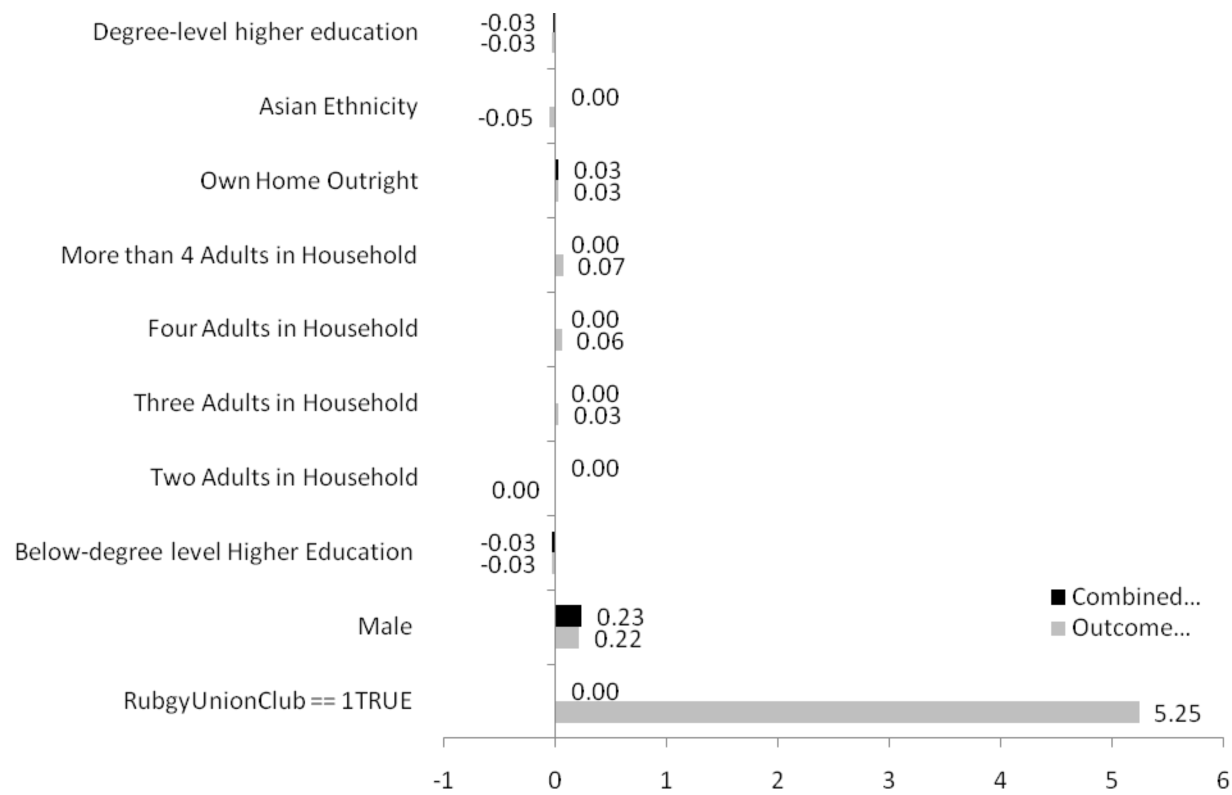


Table 7 – Rugby Union Model Coefficients

	Estimate	t value
(Intercept)	0.0396	1.7
RubgyUnionClub == 1TRUE	5.2462	154.73
SE Lottery Awards for rugby union within 10km	0	1.87
Male	0.2151	11.8
Below-degree level Higher Education	-0.0293	-3.04
Log (Deprivation Index)	-0.0026	-0.47
Two Adults in Household	-0.0008	-0.12
Three Adults in Household	0.0281	2.75
Four Adults in Household	0.0559	3.89
More than 4 Adults in Household	0.0661	2.41
Own Home Outright	0.0277	4.16
Asian Ethnicity	-0.0499	-2.69
Average Temperature	-0.0007	-1.05
Male : RespondentAge	-0.0042	-11.64
Degree-level higher education	-0.0294	-4.22

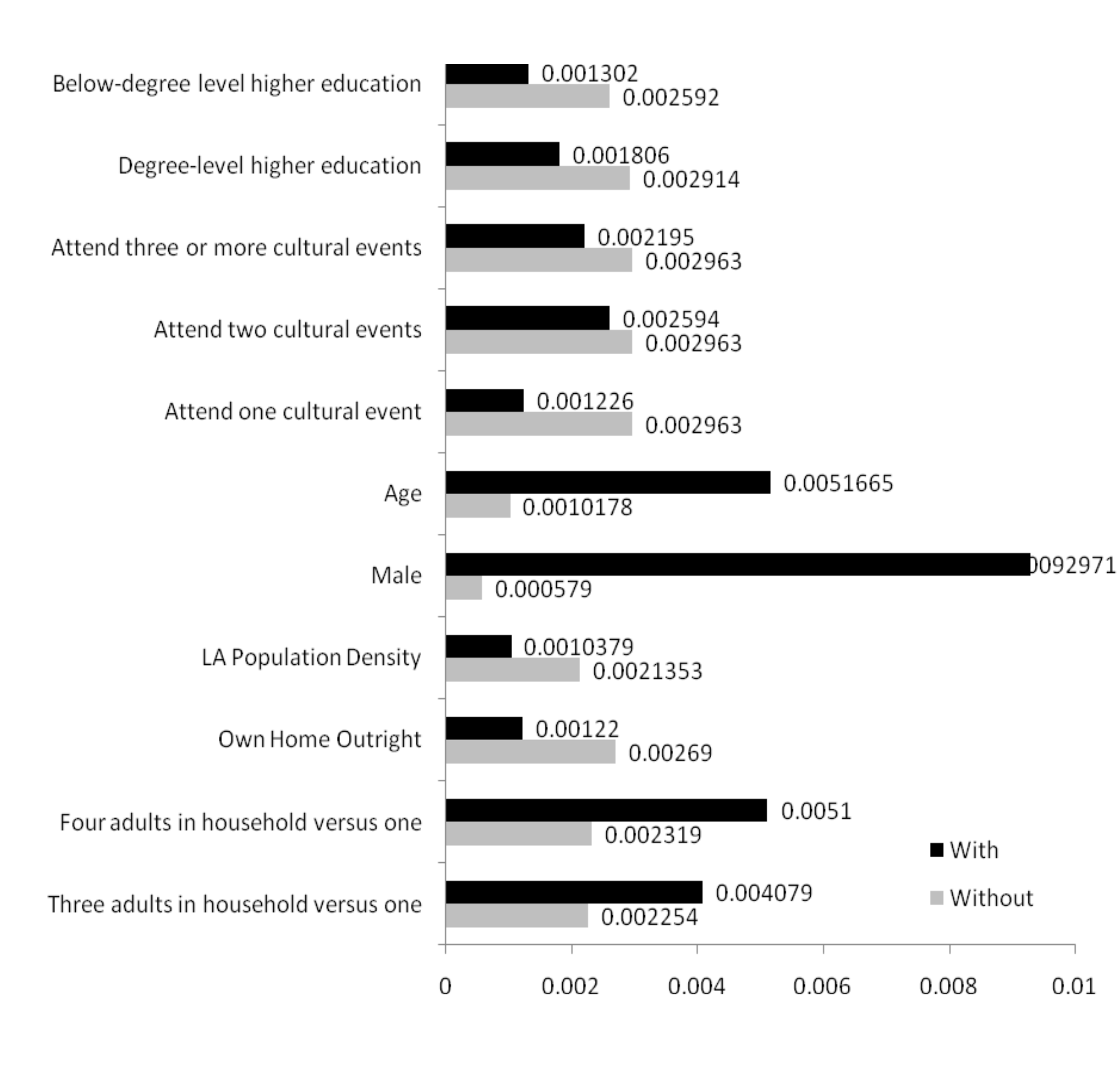
d. Cricket

Drivers of Club Membership

In contrast with the other team membership sports, age has a positive impact on the probability of being a member of a cricket club – those who are older are more likely to be a member. In addition, those who live in rural areas are more likely to be members (which is in line with rugby union club membership). Respondents who live in households with more adults are also more likely to be members.

In line with the other club sports, gym membership has a negative impact on probability of being a member.

Figure 22 – Probability of being member of Cricket Club[12]



Drivers of Frequency of Participation

Cricket has the lowest participation rate of the 11 sports modelled, with only 0.5% of all respondents having played cricket at least once a week.

As with rugby league, as cricket is a team sport, some type of club membership is typically required in order to play cricket. Those who are a member of a club tend to play on average once more a week than those who are not members of a club.

Factors that were found to increase the amount of cricket being played by active individuals were:

- Asian ethnicity
- Male
- Having children in the household (the impact of this increases up to three children before decreasing for four or more children, although this is still positive)

[12] Age values are at 32 years for without and 56 years for with, LA Population Density values are at 29 for without and 101 for with.

Impact of using Mezzanine Club Membership probability variable

The coefficients remain relatively unchanged using the probability variable instead of the dummy variable within the model, all retain the same sign and the magnitudes remain relatively similar.

Figure 23 – Incremental Days of Cricket amongst Active and All People

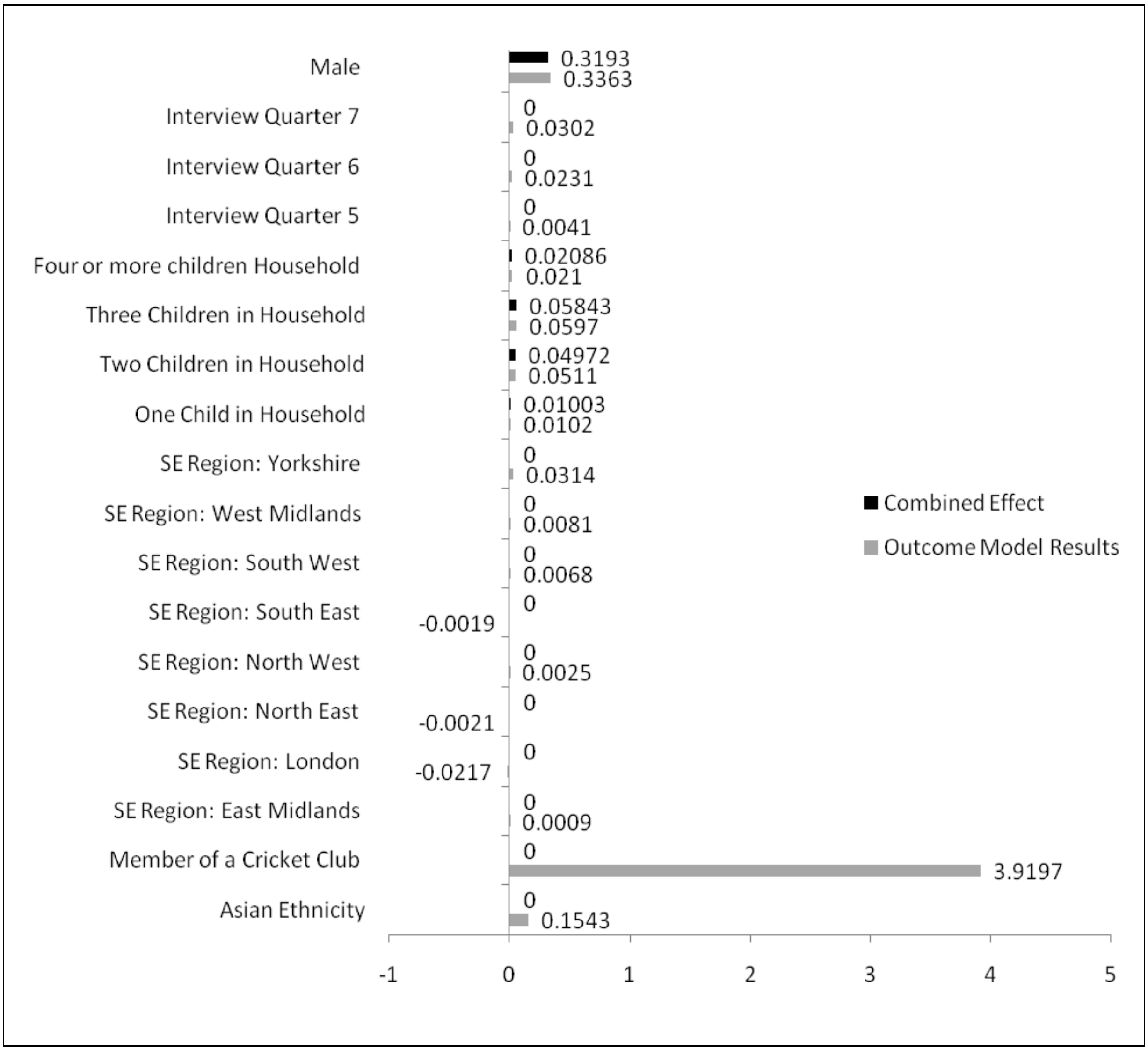


Table 8 – Cricket Model Coefficients

	Estimate	t value
(Intercept)	-0.0685	-2.63
Asian Ethnicity	0.1543	8.74
Member of a Cricket Club	3.9197	126
SE Region: East Midlands	0.0009	0.08
SE Region: London	-0.0217	-1.88
SE Region: North East	-0.0021	-0.16
SE Region: North West	0.0025	0.24
SE Region: South East	-0.0019	-0.2
SE Region: South West	0.0068	0.6
SE Region: West Midlands	0.0081	0.73
SE Region: Yorkshire	0.0314	2.21
One Child in Household	0.0102	1.3
Two Children in Household	0.0511	6.42
Three Children in Household	0.0597	4.23
Four or more children Household	0.021	0.76
Interview Quarter 5	0.0041	0.3
Interview Quarter 6	0.0231	1.45
Interview Quarter 7	0.0302	3.29
Male	0.3363	10.75
Respondent Age	-0.0117	-8.7
Respondent Age ^ 2	0.0001	7.08
Average Temperature	-0.0042	-1.27
Average Temperature ^ 2	0.0006	3.4
Total Rainfall	0.0037	3.95

e. Swimming

Swimming is the highest participation sports within the 11 modeled, with 7.6% of all individuals having gone swimming at least once a week in the last four weeks.

In line with cricket, swimming frequency increases with the number of children in the household – this is consistent with the hypothesis of swimming as a family friendly sport that is a complement, rather than substitute, for time with the family.

Those who are gym members tend to go swimming more often than individuals who are not gym members. This may be in part due to health clubs often having swimming pools associated with them, thus increasing the accessibility to swimming.

Also included in the model is the distance to the nearest Active Places swimming pool. Those who live closer to a pool tend to swim more often than those who live further away. This is consistent with the hypothesis that living closer to a facility will increase participation due to the reduced travelling times.

Ethnicity also matters to swimming frequency. Active individuals who are white tend to swim more often than other ethnicities, whilst those who are black tend to swim less often. Combining these two effects, an individual who is white does an extra session of swimming than an individual who is black every four weeks.

Attendance at cultural events also features in the model, with those attending three or more cultural events swimming more frequently than those who attend fewer events. Those who own their own home outright also tend to go swimming more often than those who do not. Both of these effects may capture the life stage of the respondent to some extent.

There is also some regional variation in the frequency of swimming – individuals outside of the East, South East and London tend to swim more often than respondents in those regions. The North East is the highest of these regions, closely followed by the South West and Yorkshire.

Figure 24 – Incremental swimming sessions over four weeks split by SE Region versus East Region[13]

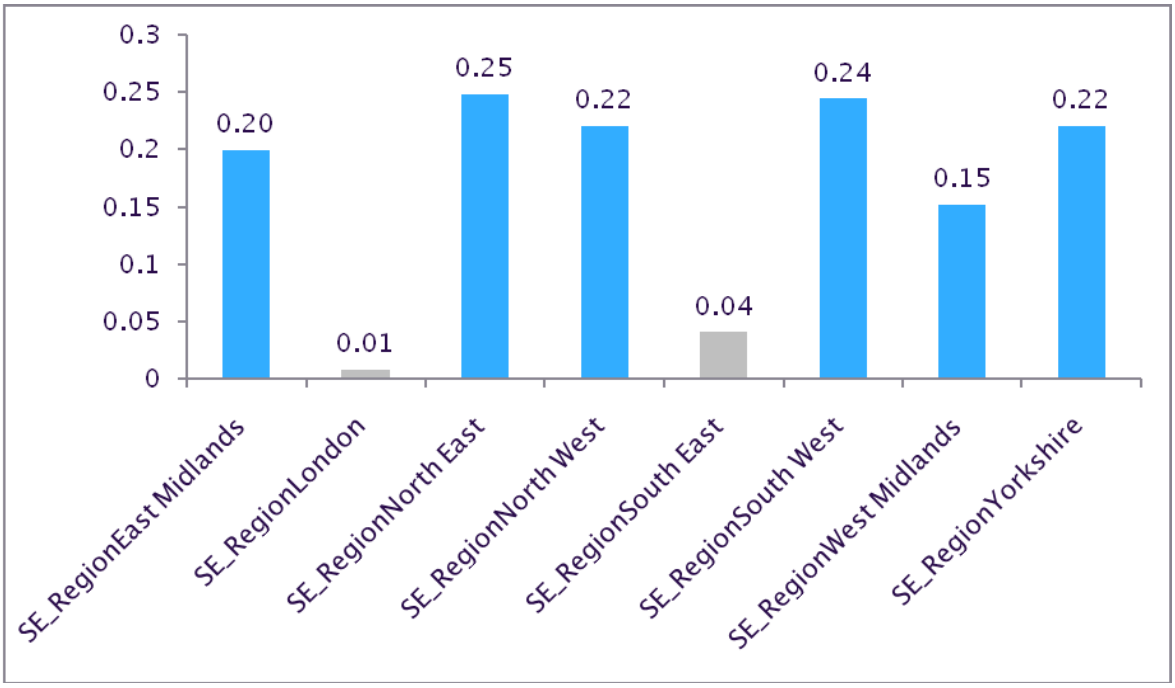
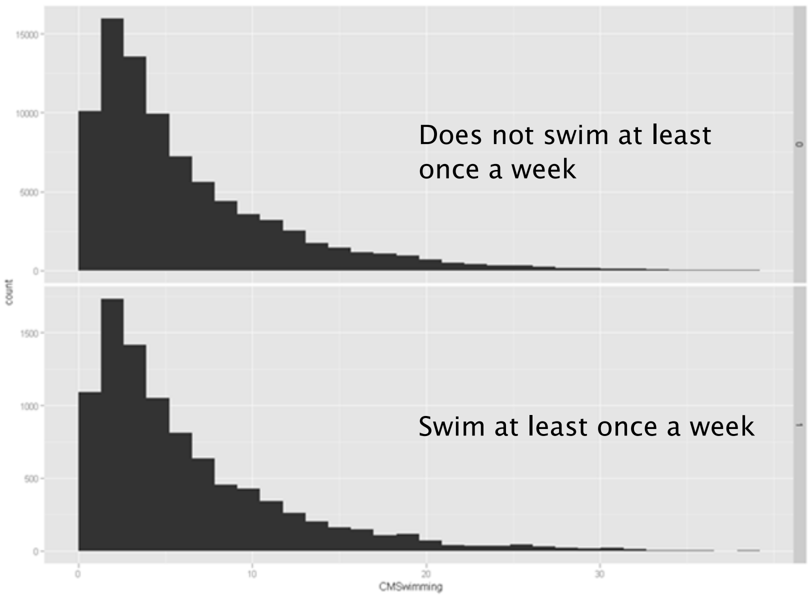


Figure 25 – Distribution of Distance to Nearest Active Places Pool



[13] Bars in light grey represent SE Regions that are not statistically significantly different from the

Figure 26 – Drivers of Incremental Days of Swimming

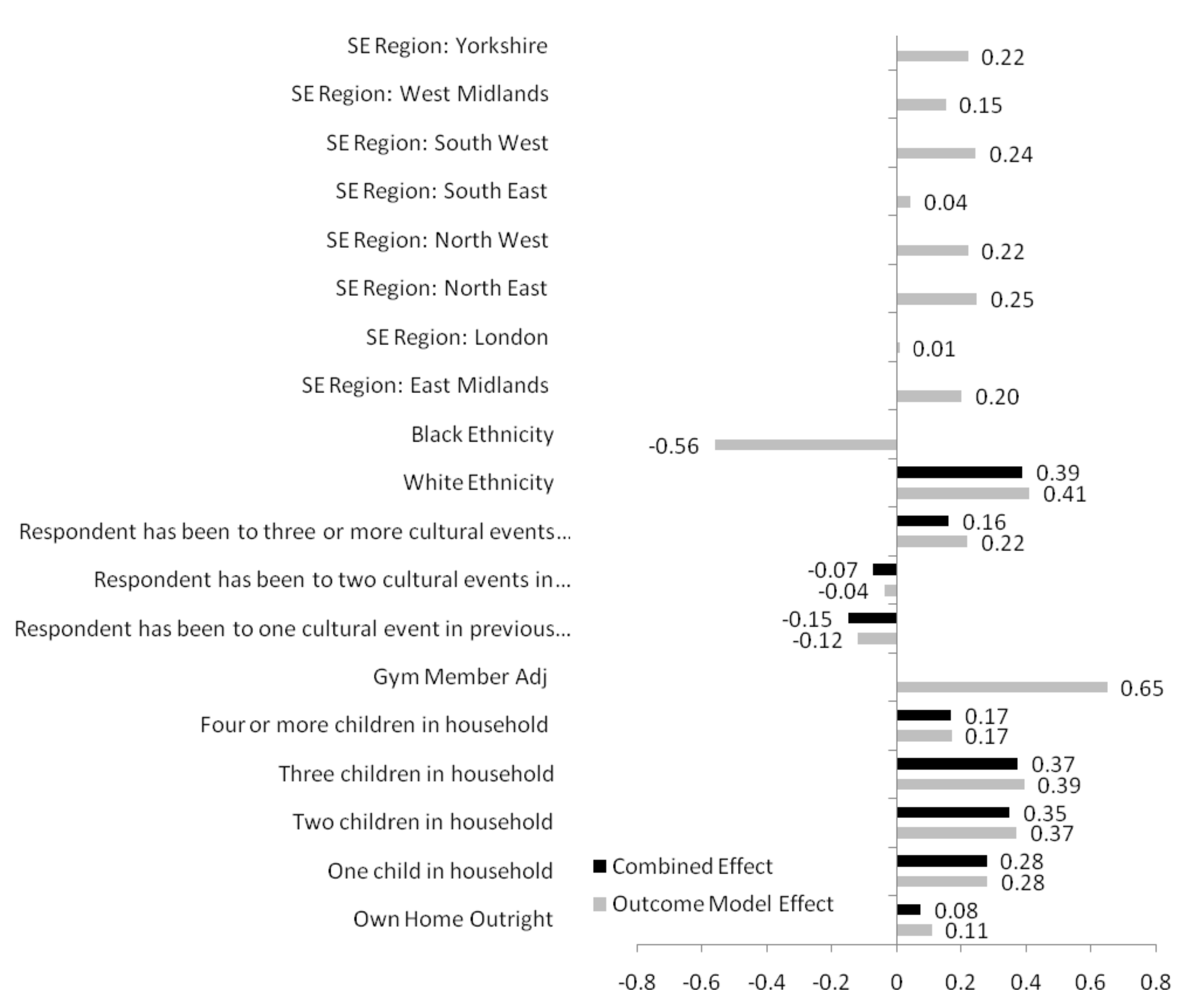


Table 9 – Swimming Model Co-Efficient Estimates

	Estimate	t-value
(Intercept)	1.0659	7.9
Own Home Outright	0.109	2.59
One child in household	0.2807	5.53
Two children in household	0.3713	7.27
Three children in household	0.3939	4.38
Four or more children in household	0.1701	0.97
Gym Member Adj	0.6492	16.12
Male: Age	-0.0439	-15.43
Male: Age ^ 2	0.0004	8.98
Respondent has been to one cultural event in previous 12 months	-0.1198	-1.71
Respondent has been to two cultural events in previous 12 months	-0.0389	-0.66
Respondent has been to three or more cultural events in previous 12 months	0.2176	5.25
White Ethnicity	0.4103	4.64
Black Ethnicity	-0.5604	-3.54
SE Region: East Midlands	0.1997	2.82
SE Region: London	0.0084	0.11
SE Region: North East	0.248	2.92
SE Region: North West	0.2203	3.31
SE Region: South East	0.0411	0.67
SE Region: South West	0.2444	3.42
SE Region: West Midlands	0.152	2.17
SE Region: Yorkshire	0.2205	2.45
Distance to nearest Active Place Swimming Pool	-0.0335	-3.94

f. Cycling

Along with swimming, football and athletics, cycling is one of the highest participation sports included within the modeling. In line with the participation results published by Sport England, we have used the definition of cycling that requires a threshold of 30 minutes at moderate intensity. Cycling includes recreational and competitive cycling but excludes any cycling which is exclusively for travel purposes only.

The availability of a car or van in the household has a significantly negative effect on active individuals, with those with access doing more than a day less cycling than those without access to a car.

There are also particular regions where cycling is less frequent: respondents living in the West Midlands, North East, North West and London all cycle less frequently than other parts of England. This is likely to be due to a combination of factors including topography, cycling facilities and tradition.

Those who are gym members also tend to cycle less frequently, as do respondents of Asian ethnicity. There is also evidence of a trade off with cultural events, as attendance at such events causes a lower frequency of participation.

Figure 27 – Drivers of Incremental Cycling Days

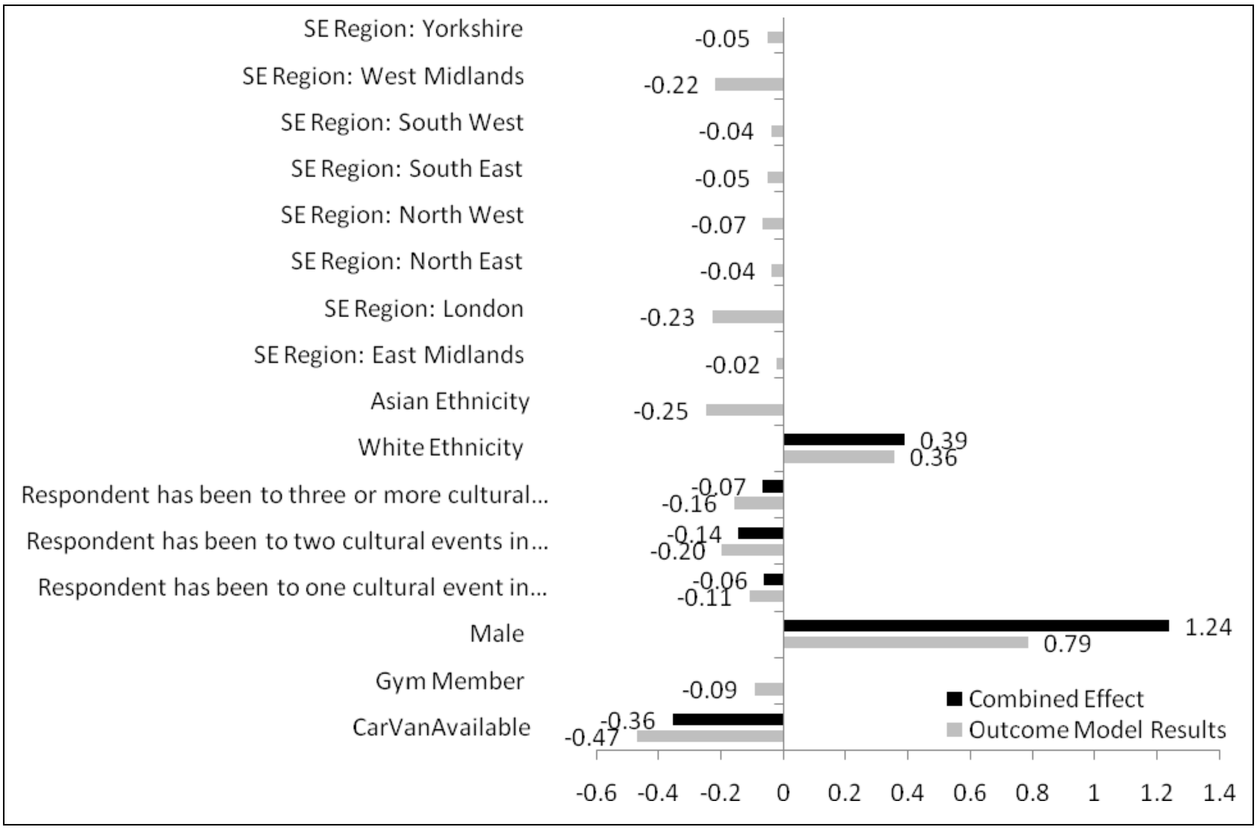


Table 10 - Cycling Model Co-Efficient Estimates

	Estimate	t value
(Intercept)	0.4062	2.3
LAPopulationDensity : CarVanAvailable	-0.0019	-1.73
CarVanAvailable	-0.4701	-7.22
Gym Member	-0.092	-2.53
Male	0.7869	22.24
Respondent Age	0.054	9.47
Respondent Age ^ 2	-0.0006	-10.06
Respondent has been to one cultural event in previous 12 months	-0.1095	-1.73
Respondent has been to two cultural events in previous 12 months	-0.1961	-3.64
Respondent has been to three or more cultural events in previous 12 months	-0.1582	-3.84
White Ethnicity	0.3566	3.97
Asian Ethnicity	-0.2484	-1.92
SE Region: East Midlands	-0.0219	-0.34
SE Region: London	-0.2277	-2.97
SE Region: North East	-0.0383	-0.5
SE Region: North West	-0.0666	-1.11
SE Region: South East	-0.0508	-0.92
SE Region: South West	-0.0392	-0.61
SE Region: West Midlands	-0.2186	-3.45
SE Region: Yorkshire	-0.0508	-0.63

g. Badminton

Within the racquet sports analysed, badminton participation is the only one where both active Asian and Chinese respondents are more likely to participate. Active males are more likely to play badminton more frequently than active females, everything else being equal.

Furthermore, respondents who live in areas which are less ethnically diverse, as measured by the Simpson Ethnic Diversity Index, are more likely to participate than those who live in more diverse areas.

Active individuals who live in local authorities with denser populations participate in badminton less frequently than those in live in sparsely populated local authorities. At the same time, living closer to an Active Places sports hall leads to an increase in the frequency of participation.

Other key drivers of participation tie in with the importance of life stage, namely the age of the oldest child in the household (which is consistent with a similar finding within the tennis model), which increases the frequency of participation and the number of children in the household which reduces the frequency of participation. Active people who own their own home participate more frequently in badminton.

Figure 28 – Drivers of Incremental Days of Badminton

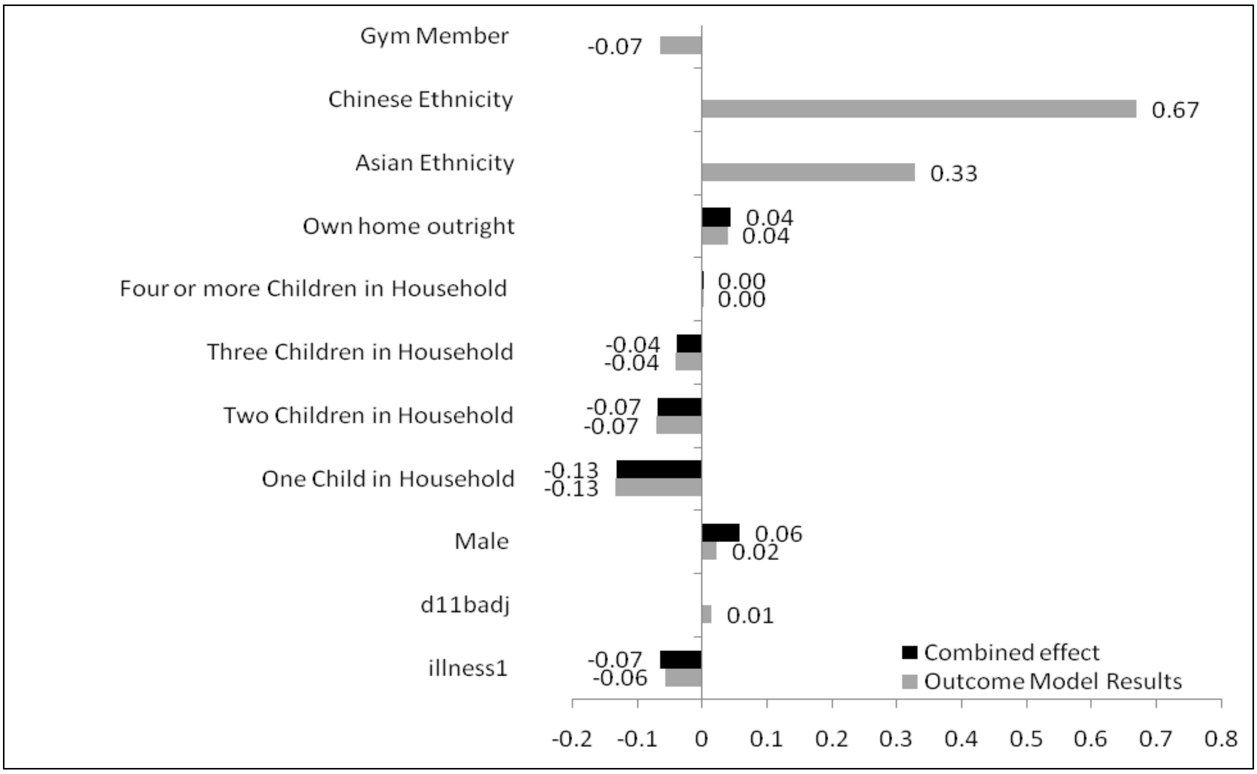


Table 11 – Badminton Model Results

	Estimate	t value
(Intercept)	-0.0054	-0.05
illness1	-0.0578	-3.84
LA Population Density	-0.0015	-5.21
Simpson Diversity Index	0.2099	1.92
childage1	0.0144	5.67
Male	0.0225	2.03
Respondent Age	0.0064	3.47
Respondent Age ^ 2	-0.0001	-3.68
One Child in Household	-0.1333	-5.21
Two Children in Household	-0.0712	-4.67
Three Children in Household	-0.0415	-1.56
Four or more Children in Household	0.0002	0
Own home outright	0.0391	2.85
Asian Ethnicity	0.3274	9.84
Chinese Ethnicity	0.6688	5.92
Distance to nearest Active Place Sports Hall	-0.0131	-2.59
Gym Member	-0.0656	-5.51

h. Squash

As with badminton, active individuals who live in sparser populated local authorities tend to play squash more frequently than those who live in denser local authorities. Another common feature is that active males play squash more frequently than active females.

In addition, the time of year has an impact on frequency of participation, with individuals playing more often during the autumn and winter. This is likely to be due to squash being an indoor sport and possibly a substitute for alternative outdoor sports that are less attractive during the winter months.

In contrast with badminton, active individuals who are gym members participate more frequently than individuals who are not members. We would speculate that this is likely to be due to squash courts often being found with gyms, whilst the relationship between gyms and badminton courts is much weaker.

As with badminton, the number of children in the household has a negative impact on frequency of participation.

Figure 29 – Drivers of Incremental Days of Squash

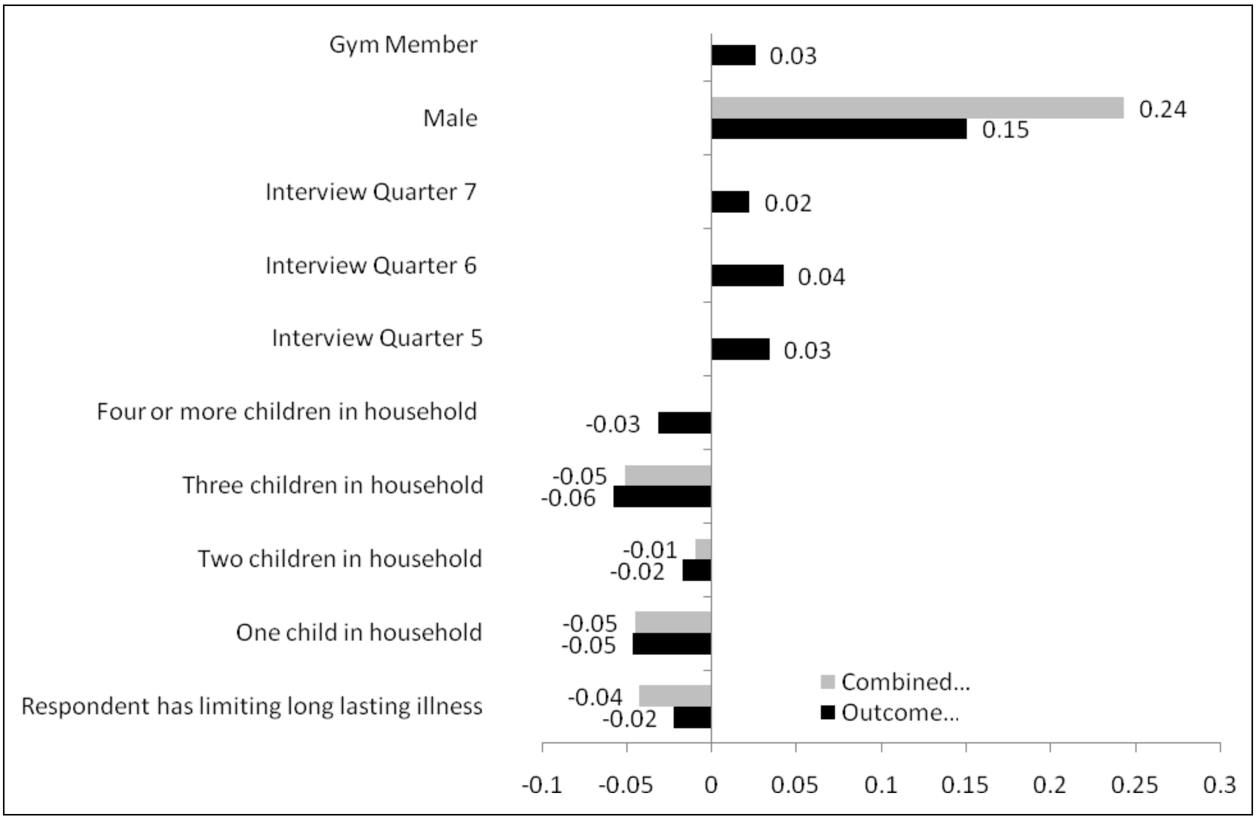


Table 12 – Squash Model Coefficients

	Estimate	t-value
(Intercept)	-0.0079	-0.2
Respondent has limiting long lasting illness	-0.0225	-1.68
LA Population Density	-0.0007	-3.14
One child in household	-0.0463	-3.46
Two children in household	-0.0175	-1.3
Three children in household	-0.0581	-2.47
Four or more children in household	-0.0316	-0.69
Interview Quarter 5	0.0339	2.67
Interview Quarter 6	0.0422	3.26
Interview Quarter 7	0.0221	1.8
Male	0.1502	15.38
Respondent Age	0.008	4.89
Respondent Age ^ 2	-0.0001	-5.06
Gym Member	0.0255	2.41

i. Golf

Along with squash and badminton, active individuals who live in local authorities with lower population density play golf more often than individuals who live in denser authorities, holding everything else constant. This is likely due to the greater proximity to golf courses for these individuals.

Rugby union, rugby league and golf are the only sports where household income level impacts on the frequency of participation. In the case of rugby league, higher household income levels had a negative impact on frequency. However, in the case of golf, active individuals who live in households with higher income tend to play golf more often than active individuals who live in households with lower incomes.

In addition, individuals who are in managerial and skilled occupations tend to play golf more often than other occupations. Those who own their own home outright are also likely to play more often than others. Active male individuals play golf more frequently than active females.

The number of children in the household has a significant negative impact on the frequency of golf playing amongst active individuals. At the same time, being a member of a gym also has a negative impact.

Temperature also has an impact on frequency of golf playing, with participation maximised at an average temperature of around 18 degrees Celsius. This is likely to be due to the outdoor nature of the sport.

Active individuals who are white tend to play golf more often than other ethnicities, whilst frequency of participation tends to increase with age.

Figure 30 – Drivers of Incremental Golf Days

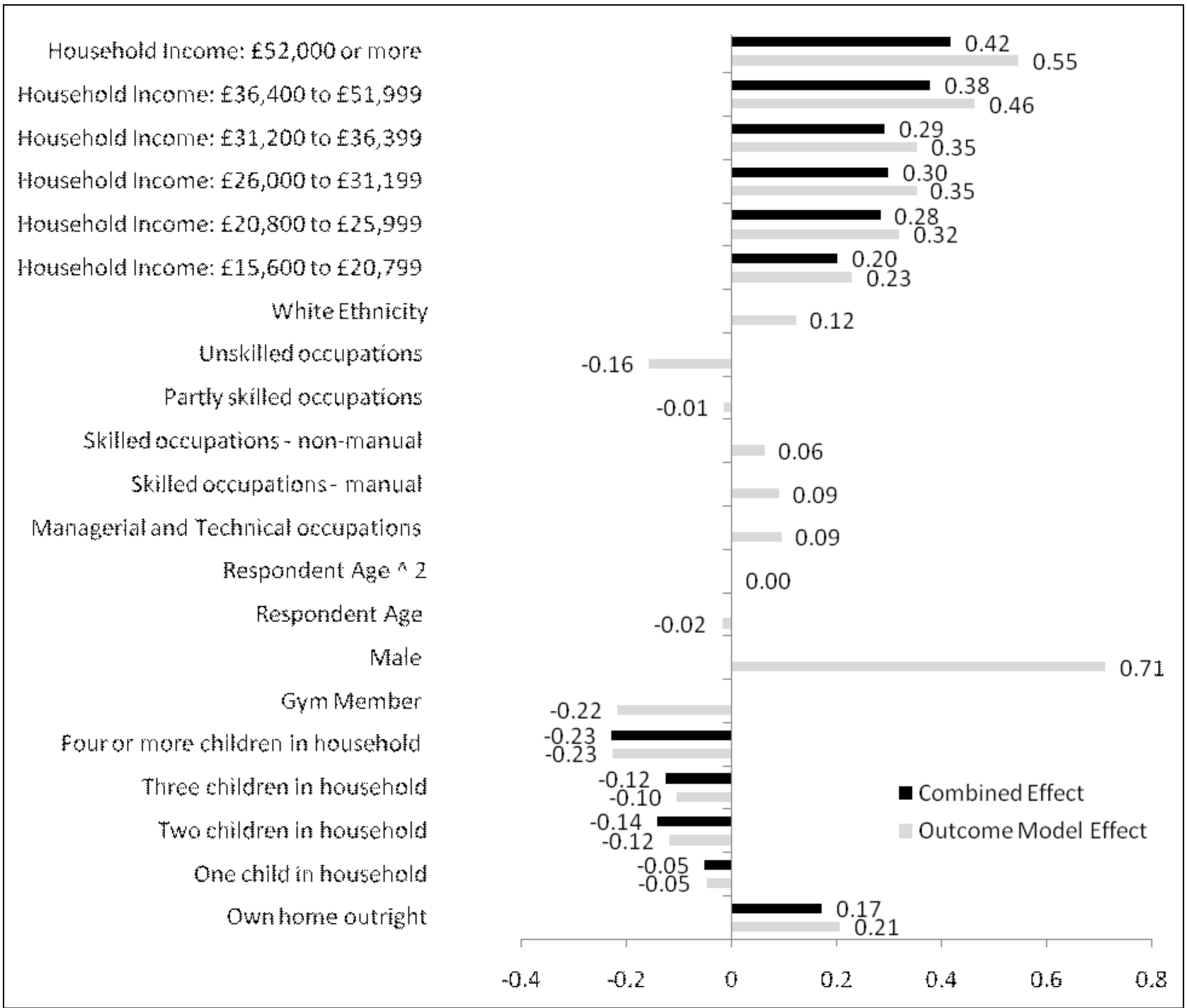


Table 13 – Golf Model Coefficients

	Estimate	t-value
(Intercept)	-0.9811	-8.83
LA Population Density	-0.0039	-8.49
Own home outright	0.2058	7.93
One child in household	-0.0478	-1.71
Two children in household	-0.1176	-4.16
Three children in household	-0.1034	-2.1
Four or more children in household	-0.2251	-2.35
Gym Member	-0.2172	-9.85
Average Temperature	0.048	4.96
Average Temperature ^ 2	-0.0013	-2.64
Male	0.7111	33.98
Respondent Age	-0.0182	-4.97
Respondent Age ^ 2	0.0004	10.14
Managerial and Technical occupations	0.0947	2.78
Skilled occupations – manual	0.0899	2.31
Skilled occupations – non-manual	0.0624	1.57
Partly skilled occupations	-0.0143	-0.28
Unskilled occupations	-0.1571	-1.86
White Ethnicity	0.1221	2.87
Household Income: £15,600 to £20,799	0.2289	5.91
Household Income: £20,800 to £25,999	0.318	7.56
Household Income: £26,000 to £31,199	0.3518	8.59
Household Income: £31,200 to £36,399	0.3533	7.86
Household Income: £36,400 to £51,999	0.463	11.67
Household Income: £52,000 or more	0.5452	12.25

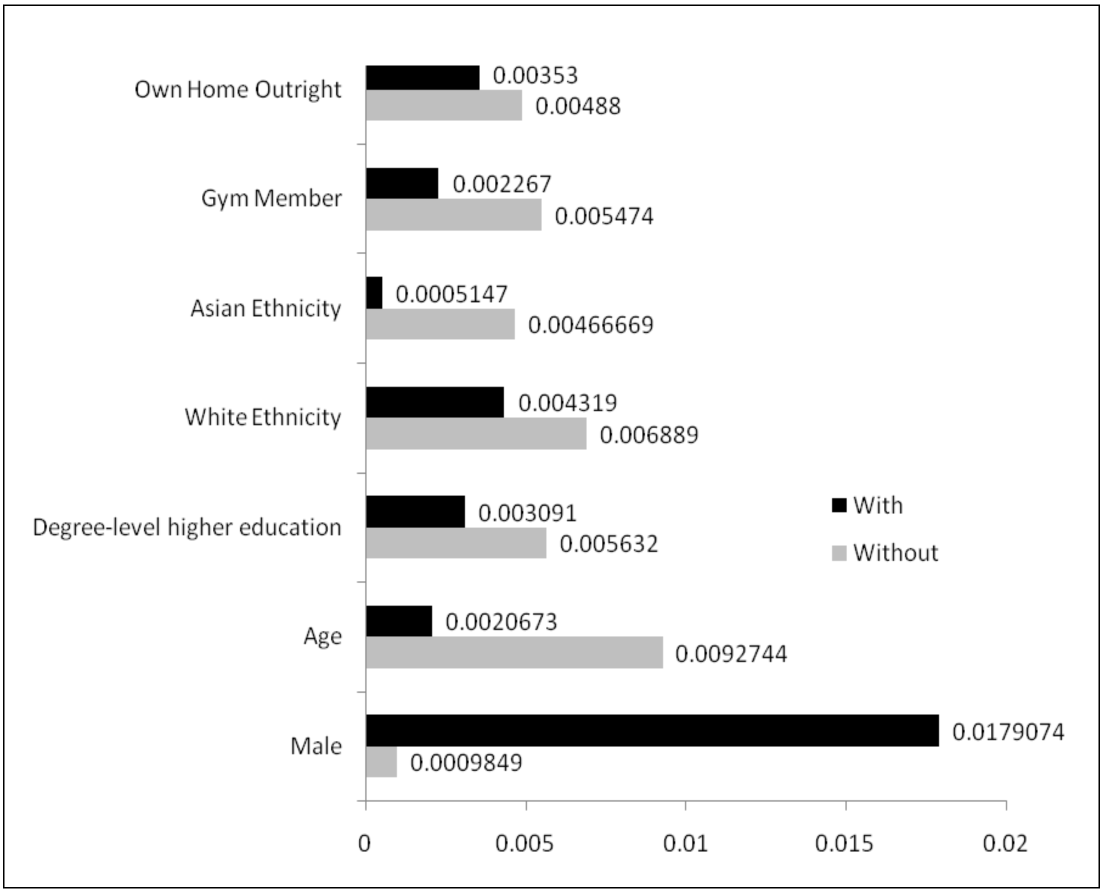
j. Football

Drivers of Club Membership

Gender is again a key positive driver of being a member of a club. In addition, respondents who are younger are more likely to be a member of a club (which is in line with rugby league and rugby union).

Those who are gym members are less likely, as are home owners and who have attained a degree-level qualification. Both Asian and White respondents are also less likely to be a member of a football club.

Figure 31 – Probability of being member of Football Club[14]



Drivers of Frequency of Participation

In comparison to athletics, where individuals who had attained a degree level education participated more frequently than other individuals, the opposite is true for football; individuals with such a qualification, a degree equivalent qualification or 5 or more GCSEs play football less often than other active individuals.

Active males are a lot more likely to participate than females, playing seven more football sessions over 28 days compared to females. In addition, those who are members of a football club tend to play almost 5 more sessions than those who are not members – this is likely to be a combination of matches and football training sessions.

[14] Age values are at 32 years for without and 56 years for with, LA Population Density values are at 29 for without and 101 for with.

Attendance at cultural events and gym membership both have a significant negative effect on the frequency of participation. Those who live in council-owned accommodation, or who live in households with a higher number of adults tend to play more frequently.

Impact of using Mezzanine Club Membership probability variable

The coefficients remain relatively unchanged using the probability variable instead of the dummy variable within the model, all retain the same sign and the magnitudes remain relatively similar.

Figure 32 – Drivers of Incremental Days of Football

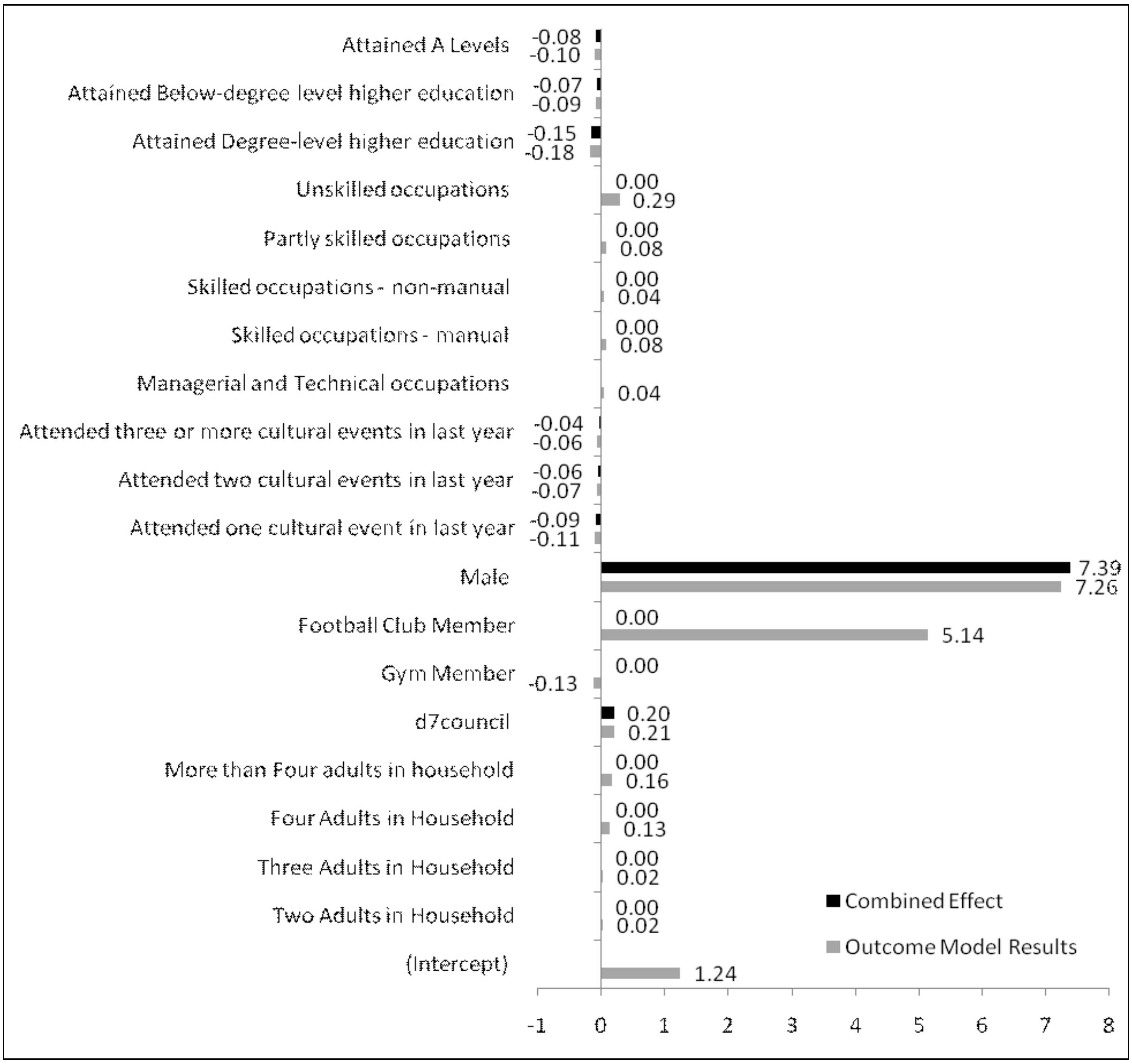


Table 14 – Football Outcome Model Coefficients

	Estimate	t value
(Intercept)	1.2368	4.74
NumAdultsHouseholdAdj2	0.0166	0.8
NumAdultsHouseholdAdj3	0.0192	0.58
NumAdultsHouseholdAdj4	0.1315	2.83
NumAdultsHouseholdAdjMore than 4	0.1642	1.89
d7council	0.2079	4.21
GymMemberAdj	-0.1331	-6.31
FootballClub == 1TRUE	5.1403	75.46
Male	7.259	21.74
poly(RespondentAge, 3, raw = TRUE)1	-0.0474	-2.67
poly(RespondentAge, 3, raw = TRUE)2	0.0008	2.15
poly(RespondentAge, 3, raw = TRUE)3	0	-1.79
culturalevent1	-0.1053	-2.86
culturalevent2	-0.0739	-2.36
culturalevent3	-0.0634	-2.69
sec4hII: Managerial and Technical occupations	0.0441	1.34
sec4hIIIM: Skilled occupations - manual	0.083	2.2
sec4hIIIN: Skilled occupations - non-manual	0.0373	0.97
sec4hIV: Partly skilled occupations	0.0792	1.61
sec4hV: Unskilled occupations	0.2876	3.53
AP_Grass_Football	-0.0209	-2.81
d6heduc1	-0.1838	-6.89
d6heduc2	-0.0886	-2.68
d6alevels	-0.1023	-3.71
Male:poly(RespondentAge, 3, raw = TRUE)1	-0.3218	-13.65
Male:poly(RespondentAge, 3, raw = TRUE)2	0.0049	9.46
Male:poly(RespondentAge, 3, raw = TRUE)3	0	-7.13

1. Appendix

1.1. Hypotheses for exploration

Decision Influencer	Have I got time available to play?	Have I got the energy to play?	Do I know how to play?	Am I interested in playing?	Have I got what I need?
Me	Those who work full time will have lower participation rates due to less free time Weather will have a more significant impact on outdoor sports	Sports of a higher intensity will have lower participation rates amongst older age groups Areas with higher obesity levels will tend to have lower sport participation rates	Early experience of sport makes a big difference in long-term participation Proximity to HE/FE colleges will impact on sports participation	People who continue in higher education may have higher participation rates over long-term People who engage in sport more generally engage more in cultural activities or civic life	Amount of coastline (and open water) in local area (and access to this) drives participation in water sports Higher incomes will increase participation due to more resources being available
Community & Local Institutions	Areas with poor public transport connections will have lower participation levels Greater distance from home or work to gym or facility leads to lower participation		Schools with accreditation in sports generally or sports-specific accreditations lead to long-term higher levels of participation for their students Density of clubs in an area matters more than quality	A more diverse community may result in there being less social cohesion, and hence lower overall rates of participation within the community Local authorities that consistently have high performing academic scores in scores will have higher levels of participation The size of the gap between gender participation rates is a driver of overall participation rates The availability of gyms in an area will have an impact on sports participation	Higher population density gives you greater critical mass around which to organize participation Lower density populations give you more space for outdoor activity Rural areas for some sports have lower participation rates due to lack of facility Local authorities that invest more in sport have higher rates of participation More funding into coaching/ infrastructure does not result in much higher overall participation rates. Some local authorities that are not good at engaging with their communities in sport will also be worse at engaging generally with the community at a broader level Local Authorities with a stretch target that includes NI8, etc. Will have a higher participation rate
National Institutions				National popularity at the time will increase participation in a particular sport More TV Screening of a sport will increase participation in a particular sport Social marketing campaigns have increased sport participation	County Sport Partnerships are not making a difference to sport participation Section 106 planning agreements and CASC tax breaks will increase participation in those areas

Factors that are not easily influenced by public policy

Factors that are influenced by non-SE public policy

Factors that are influenced by SE policy

Factors that are influenced by private and third sector

Codebook for variables used

Variable Name	Description of variable
age16_24	Respondent is aged 16-24
age25_34	Respondent is aged 25-34
age35_44	Respondent is aged 35-44
age55_64	Respondent is aged 55-64
age65_74	Respondent is aged 65-74
age75_84	Respondent is aged 75-84
age85plus	Respondent is aged 85+
AP_Grass_Football	Distance to nearest Active Places grass football pitch
AP_SportsHall_Main	Distance to nearest Active Places sports hall
awardamount10	Sport England Lottery Award Amounts within 10km of respondent
culturalevent1	Respondent has been to one cultural event in previous 12 months
culturalevent2	Respondent has been to two cultural events in previous 12 months
culturalevent3	Respondent has been to three or more cultural events in previous 12 months
CarVanAvailable	Respondent has access to a car or van
athleticsclubs20	The number of athletics clubs within 20km of respondent
CricketClub	Respondent is a member of a cricket club
numberchild1	One child in Household
numberchild2	Two children in Household
numberchild3	Three children in Household
childage1	Child's age (single child household)
childage1	Age of oldest child in household
childage_1	Age of oldest child in household
childage_2	2nd oldest child's age(multiple children households)
d13adj	Number of cars/vans in household
d19retired	Retired-Not working
d19stufull	Student full-time
d19unemp1	Unemployed less than 12 months
d19wpart	Working part-time
Gender	Gender
d23_bands_7£15,600 to £20,799	Household income: £15,600 to £20,799
d23_bands_7£20,800 to £25,999	Household income: £20,800 to £25,999
d23_bands_7£26,000 to £31,199	Household income: £26,000 to £31,199
d23_bands_7£31,200 to £36,399	Household income: £31,200 to £36,399
d23_bands_7£36,400 to £51,999	Household income: £36,400 to £51,999
d23_bands_7£52,000 or more	Household income: £52,000 or more
d23incm1	Respondent has household income level up to £5,199
d23incm10	Respondent has household income level £45,800-£51,999
d23incm11	Respondent has household income level £52,000 or more
d23incm2	Respondent has household income level £5,200-£10,399
d23incm3	Respondent has household income level £10,400-£15,599
d23incm4	Respondent has household income level £15,600-

Variable Name	Description of variable
	£20,799
d23incm7	Respondent has household income level £31,200-£36,399
d23incm8	Respondent has household income level £36,400-£41,599
d23incm9	Respondent has household income level £41,600-£45,799
d5adj	Age respondent finished full time education
d6alevels	Respondent has attained A Level qualification as highest qualification
d6gcse5	Respondent has attained GCSEs (5 or more) as highest qualification
d6heduc1	Respondent has attained a degree-equivalent qualification as highest qualification
d6heduc2	Respondent has attained a other higher education qualification below degree level as highest qualification
d7own	Respondent owns home outright
ethasian	The respondent is Asian
ethblack	Respondent is black
ethchinese	Respondent is Chinese
ethwethnic	Respondent is white
FootballClub	Respondent is member of a football club
GymMemberAdj	Respondent is member of a health club
I(Male * RespondentAge)	Interaction term between Age and Gender
illness1	Respondent has limiting long lasting illness
Imdlnadj	LN(index of deprivation)
imdlnadj	Log of deprivation index in Local Authority (2007)
Int_Quarter5	Respondent interviewed in APS Quarter 5
Int_Quarter6	Respondent interviewed in APS Quarter 6
Int_Quarter7	Respondent interviewed in APS Quarter 7
lakeswithin10	Number of open water spaces within 10km of respondent
LAPopulationDensity	The population density of the Local Authority that the individual lives in
Male	Respondent is male
NumChildHouseholdAdj1	One child in household
NumChildHouseholdAdj2	Two children in household
NumChildHouseholdAdj3	Three children in household
NumChildHouseholdAdj4 or more	Four or more children in household
OwnEthnicPct	Percentage of respondents in Local Authority that have same ethnicity as respondent
poly(AverageTemp, 2, raw = TRUE)1	Average temperature in month of interview at weather station nearest to respondent
poly(AverageTemp, 2, raw = TRUE)2	(Average temperature)^2
poly(RespondentAge, 3, raw = TRUE)1	Age of respondent
poly(RespondentAge, 3, raw = TRUE)2	(Age of respondent)^2
poly(RespondentAge, 3, raw = TRUE)3	(Age of respondent)^3
q19adj	The respondent has done voluntary sports work in the past 12 months
q21adj	Level of Satisfaction with sports provision in LA
Regeast	Region-East

Variable Name	Description of variable
Reglondon	Region-London
Regneast	Region-North East
Regseast	Region-South East
Regswest	Region-South West
Regwmids	Region-West Midlands
RubgyUnionClub	Respondent is a member of a rugby union club
RugbyLeagueClub	Respondent is a member of a rugby league club
rugbyunion_awardamount10	SE Lottery Award amounts related to rugby union within 10km of respondent
s3adj	Number of adults in household
SE_RegionEast Midlands	Respondent lives in East Midlands SE Region
SE_RegionLondon	Respondent lives in London SE Region
SE_RegionNorth East	Respondent lives in North East SE Region
SE_RegionNorth West	Respondent lives in North West SE Region
SE_RegionSouth East	Respondent lives in South East SE Region
SE_RegionSouth West	Respondent lives in South West SE Region
SE_RegionWest Midlands	Respondent lives in West Midlands SE Region
SE_RegionYorkshire	Respondent lives in Yorkshire SE Region
sec4hII: Managerial and Technical occupations	Managerial/technical occupations
sec4hIIIM: Skilled occupations - manual	Skilled occupations - manual
sec4hIIIN: Skilled occupations - non-manual	Skilled occupations - non-manual
sec4hIV: Partly skilled occupations	Partly skilled occupations
sec4hV: Unskilled occupations	Unskilled occupations
sec4manag	Managerial/technical occupations
sec4prof	Professional occupations
sec4skill1	Skilled occupations - Non Manual
SimpsonDiversityIndex	Simpson Ethnic Diversity Index for Local Authority that individual lives in
SocialClubMemberAdj	Respondent is a member of a Social Club

1.2. Guide to interpreting detailed model results

	Description
Parameter	The name of the factor (e.g. Male). In cases like income, where there are a number of categories, a number of dummy variables are used, each representing a different category – in each case a respondent is attributed either a 1 (if they meet the category criteria) or 0. Where categorical variables have been used, one level of the variable is omitted to negate perfect multi-collinearity.
Estimate	This is the impact of a one unit change in the parameter, e.g. age increasing by one year. In both logistic and probit regression models, linear changes in the independent variables produce nonlinear changes in the probability of success. In logit equations (in this case the NI8 models), this is the impact on the log-odds ratio. In the case of probit models (in this case the selection model), this is the change in the z-score, which can then be evaluated using the normal distribution. In standard OLS models (in this case the outcome models), this is the change in the number of days of participation over 4 weeks.
Error	The standard error of the estimate at the 95% level. This is a measure of the range of possible values that the estimate is likely to take. It is used to calculate a test of significance.
Chi-Square	A measure of significance. The Chi-Square test statistic is the squared ratio of the Estimate to the Standard Error of the respective predictor. The Chi-Square value follows a central Chi-Square distribution which is used to test the hypothesis that the Estimate is not equal to zero.
Pr > ChiSq	A measure of significance. The probability that a particular **Chi-Square** test statistic is as extreme as, or more so, than what has been observed under the null hypothesis.
t value	A measure of whether the estimate of the parameter is statistically significant (different from zero). It is the ratio of the estimated coefficient to its estimated standard error. If the null hypothesis is to be rejected, the "t-stat" must be larger (in absolute value) than the critical point on the t- distribution. For a 90% confidence, the critical value is 1.645, anything greater than this indicates that the estimate is statistically significantly different from zero.
Pr (> |t|)	An alternative way of displaying the t-value.

R Squared	Measures the percent of variation in the "dependent" variable that can be accounted for or "explained" by the "independent" variables.
Adjusted R Squared	The adjusted R Squared is a measure of the noise around the regression line, correcting for over fitting, where the number of parameters is close to the number of observations. This correction is negligible if the number of observations is a large multiple of the number of variables in the equation.
Sigma	The standard error of the error terms of the outcome equation.
Rho	The estimated correlation coefficient between the error term of the selection equation and the outcome equation
Inverse Mills Ratio	The ratio of the probability density function over the cumulative distribution function of a distribution.

1.3. Sheffield Hallam NI8 Model Results (based on APS2+3 sample)

Parameter	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	-1.5082	0.1833	67.7048	<.0001
d13adj	0.0485	0.00524	85.6895	<.0001
q19adj	0.6768	0.0172	1539.6391	<.0001
q21_adj	0.3273	0.0109	904.7274	<.0001
s3adj	-0.0783	0.00657	141.8927	<.0001
Male	0.1617	0.011	216.4946	<.0001
age16_24	0.6738	0.0218	958.9783	<.0001
age25_34	0.433	0.0169	656.2955	<.0001
age35_44	0.2542	0.0163	243.9562	<.0001
age55_64	-0.3627	0.0198	335.6047	<.0001
age65_74	-0.4862	0.0322	228.4754	<.0001
age75_84	-1.1491	0.041	785.7023	<.0001
age85plus	-2.1396	0.1409	230.5945	<.0001
d5adj	0.0127	0.00329	14.943	0.0001
ethwethnic	0.0887	0.0288	9.487	0.0021
ethasian	-0.3628	0.0405	80.1099	<.0001
ethblack	-0.3285	0.0479	46.9325	<.0001
ethchinese	-0.5816	0.1129	26.5508	<.0001
d6heduc1	0.3448	0.0199	299.8336	<.0001
d6heduc2	0.2843	0.0217	172.4502	<.0001
d6alevels	0.2401	0.0185	168.863	<.0001
d6gcse	0.1626	0.0191	72.2424	<.0001
d7own	0.1917	0.0139	189.3158	<.0001
d7council	-0.1346	0.0262	26.4217	<.0001
numberchild1	-0.4233	0.0256	273.728	<.0001
numberchild2	-0.3531	0.0282	157.0806	<.0001
numberchild3	-0.3128	0.0375	69.5142	<.0001
d19wpart	0.1249	0.0163	58.4394	<.0001
d19stufull	0.1848	0.0424	18.9621	<.0001
d19retired	0.3069	0.0262	137.0485	<.0001
regneast	0.0233	0.0248	0.8818	0.3477
regwmids	-0.0796	0.0189	17.7012	<.0001
regeast	-0.0998	0.0188	28.1649	<.0001
regswest	0.0213	0.0184	1.3328	0.2483
regseast	-0.0895	0.0171	27.4591	<.0001
reglondon	-0.0894	0.0179	24.862	<.0001
d19unempl1	0.1748	0.0354	24.3751	<.0001
childage1	0.0343	0.00254	181.6968	<.0001
d11old1	-0.0157	0.0052	9.1311	0.0025
childage_2	0.0379	0.00566	44.828	<.0001
sec4prof	0.0576	0.0216	7.1437	0.0075
sec4manag	0.0854	0.0133	41.1286	<.0001
sec4skill1	0.0575	0.0161	12.8152	0.0003

Parameter	Estimate	Error	Chi-Square	Pr > ChiSq
d23incm1	-0.2529	0.0402	39.6749	<.0001
d23incm2	-0.228	0.0277	67.594	<.0001
d23incm3	-0.2468	0.0232	113.3625	<.0001
d23incm4	-0.101	0.0207	23.7399	<.0001
d23incm7	0.1301	0.0211	37.9541	<.000
d23incm8	0.1357	0.0221	37.731	<.0001
d23incm9	0.1933	0.0232	69.2621	<.0001
d23incm10	0.3281	0.0225	212.5678	<.0001
d23incm11	0.4945	0.017	841.5169	<.0001
illness1	-0.3776	0.0147	659.0421	<.0001
lmDep	-0.0455	0.0127	12.8527	0.0003

1.4. Mindshare NI8 Model Results (based on APS2Q4 + APS3Q1-Q3)

	Estimate	Std. Error	z	Pr(>|z|)
(Intercept)	-2.89E+00	9.92E-02	-29.13	2.00E-16
SocialClubMemberAdj	9.59E-01	2.20E-02	43.68	2.00E-16
culturalevent1	3.39E-02	3.01E-02	1.13	0.25926
culturalevent2	1.29E-02	2.60E-02	0.5	0.6201
culturalevent3	2.09E-01	1.78E-02	11.76	2.00E-16
SE_RegionEast Midlands	5.34E-02	3.40E-02	1.57	0.11647
SE_RegionLondon	-4.31E-02	4.09E-02	-1.05	0.2924
SE_RegionNorth East	1.20E-01	4.06E-02	2.96	0.00311
SE_RegionNorth West	8.43E-02	3.17E-02	2.66	0.00785
SE_RegionSouth East	2.77E-02	2.90E-02	0.96	0.33916
SE_RegionSouth West	1.24E-01	3.25E-02	3.81	0.00014
SE_RegionWest Midlands	-8.38E-04	3.38E-02	-0.02	0.98025
SE_RegionYorkshire	1.25E-01	3.26E-02	3.82	0.00013
awardamount10	1.09E-07	3.03E-08	3.62	0.0003
AverageTemp	1.89E-02	1.63E-03	11.59	2.00E-16
TotalRainfallAdj	4.28E-03	2.43E-03	1.76	0.07798
d6heduc1	3.07E-01	2.37E-02	12.94	2.00E-16
d6heduc2	3.26E-01	2.96E-02	11.02	2.00E-16
d6alevels	2.35E-01	2.51E-02	9.33	2.00E-16
d6gcse	1.66E-01	2.72E-02	6.09	1.10E-09
Male	1.78E+00	1.05E-01	16.96	2.00E-16
poly(RespondentAge,2,raw=TRUE)1	2.59E-02	3.82E-03	6.78	1.20E-11
poly(RespondentAge,2,raw=TRUE)2	-5.11E-04	4.27E-05	-11.98	2.00E-16
ethwethnic	1.62E-01	7.86E-02	2.06	0.03915
OwnEthnicPct	2.69E-01	9.98E-02	2.7	0.00703
d23_bands_7\24315,600 to \24320,799	1.04E-01	3.23E-02	3.23	0.00124
d23_bands_7\24320,800 to \24325,999	1.21E-01	3.47E-02	3.49	0.00048
d23_bands_7\24326,000 to \24331,199	2.33E-01	3.25E-02	7.17	7.50E-13
d23_bands_7\24331,200 to \24336,399	3.03E-01	3.51E-02	8.62	2.00E-16
d23_bands_7\24336,400 to \24351,999	3.67E-01	2.97E-02	12.35	2.00E-16
d23_bands_7\24352,000 or more	6.45E-01	3.07E-02	21.03	2.00E-16
LAPopulationDensity	1.77E-03	9.06E-04	1.95	0.05133
I(CarVanAvailable==1)TRUE	3.23E-01	3.80E-02	8.52	2.00E-16
NumAdultsHousehold==1TRUE	1.83E-01	2.16E-02	8.48	2.00E-16
illness1	-3.47E-01	2.15E-02	-16.11	2.00E-16
d7own	1.99E-01	1.98E-02	10.05	2.00E-16
d7council	-1.58E-01	3.92E-02	-4.02	5.80E-05
NumChildHouseholdAdj1	-1.07E-01	2.17E-02	-4.95	7.50E-07
NumChildHouseholdAdj2	-1.63E-01	2.34E-02	-6.95	3.70E-12
NumChildHouseholdAdj3	-6.40E-02	4.05E-02	-1.58	0.11381
NumChildHouseholdAdj4 or more	-4.24E-01	8.33E-02	-5.09	3.60E-07
lakeswithin10	4.85E-04	2.85E-04	1.7	0.08863
Male:poly(RespondentAge,2,raw=TRUE)1	-7.27E-02	4.90E-03	-14.85	2.00E-16
Male:poly(RespondentAge,2,raw=TRUE)2	6.84E-04	5.33E-05	12.84	2.00E-16
LAPopulationDensity:I(CarVanAvailable==1)TRUE	-4.74E-03	7.91E-04	-5.98	2.20E-09

1.5. Individual Sports Model Results

1.5.1. Selection Model Results

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-3.73E-01	7.64E-02	-4.89	1.00E-06
ethwethnic	7.66E-02	4.19E-02	1.83	0.06727
OwnEthnicPct	7.42E-02	4.27E-02	1.74	0.08233
I(ethasian*RespondentAge)	-3.15E-03	8.25E-04	-3.81	0.00014
avggcsescore	2.99E-04	1.03E-04	2.89	0.00386
AverageTemp	1.58E-02	8.78E-04	17.95	2.00E-16
d6heduc1	2.41E-01	1.22E-02	19.77	2.00E-16
d6heduc2	1.74E-01	1.51E-02	11.47	2.00E-16
d6alevels	1.36E-01	1.31E-02	10.36	2.00E-16
d6gcse	8.26E-02	1.41E-02	5.85	4.90E-09
Male	1.03E+00	7.08E-02	14.51	2.00E-16
poly(RespondentAge,2,raw=TRUE)1	-2.37E-02	2.03E-03	-11.66	2.00E-16
poly(RespondentAge,2,raw=TRUE)2	8.74E-05	2.05E-05	4.27	2.00E-05
imdlnadj	-3.50E-02	8.75E-03	-4	6.40E-05
culturalevent1	1.11E-01	1.66E-02	6.68	2.40E-11
culturalevent2	1.21E-01	1.39E-02	8.7	2.00E-16
culturalevent3	2.09E-01	9.58E-03	21.79	2.00E-16
d23_bands_7£15,600to20799	9.91E-02	1.47E-02	6.75	1.50E-11
d23_bands_7£20,800to25999	1.26E-01	1.64E-02	7.67	1.70E-14
d23_bands_7£26,000to31199	1.97E-01	1.60E-02	12.28	2.00E-16
d23_bands_7£31,200to36399	2.25E-01	1.81E-02	12.43	2.00E-16
d23_bands_7£36,400to51999	3.07E-01	1.50E-02	20.45	2.00E-16
d23_bands_7£52,000ormore	4.60E-01	1.61E-02	28.67	2.00E-16
I(NumAdultsHousehold	-2.88E-02	1.93E-02	-1.49	0.13667
CarVanAvailable	2.61E-01	1.31E-02	20	2.00E-16
I(NumAdultsHousehold	1.29E-01	9.78E-03	13.16	2.00E-16
illness1	-2.24E-01	1.01E-02	-22.22	2.00E-16
d7own	1.23E-01	1.06E-02	11.57	2.00E-16
d7council	-9.05E-02	1.90E-02	-4.76	2.00E-06
NumChildHouseholdAdj1	1.07E-02	1.29E-02	0.83	0.40548
NumChildHouseholdAdj2	8.45E-02	1.34E-02	6.33	2.40E-10
NumChildHouseholdAdj3	7.56E-02	2.30E-02	3.29	0.00101
NumChildHouseholdAdj4ormore	1.01E-02	4.27E-02	0.24	0.81366
Male:poly(RespondentAge,2,raw=TRUE)1	-2.78E-02	2.96E-03	-9.41	2.00E-16
Male:poly(RespondentAge,2,raw=TRUE)2	1.97E-04	2.92E-05	6.76	1.40E-11

1.5.2. Athletics Outcome Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	1.7887	0.1824	9.8	0
d6heduc1	0.3221	0.04	8.06	0
poly(RespondentAge, 2, raw = TRUE)1	0.0196	0.0062	3.17	0.0015
poly(RespondentAge, 2, raw = TRUE)2	-0.0003	0.0001	-5.21	0
Male	0.1639	0.0374	4.38	0
imdlnadj	-0.0933	0.0342	-2.73	0.0064
GymMemberAdj	-0.2157	0.0387	-5.58	0
d7own	-0.101	0.0447	-2.26	0.0239
CarVanAvailable	-0.19	0.0639	-2.97	0.0029
athleticsclubs20	0.0126	0.003	4.2	0
ethasian	-0.598	0.1091	-5.48	0
illness1	-0.2708	0.0505	-5.37	0
R Squared	0.025			
Adjusted R Squared	0.025			
Sigma	3.518			
Rho	-0.219			
Inverse Mills Ratio	-0.7705	0.1213	-6.35	0

1.5.3. Tennis Outcome Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	1.6856	0.1408	11.97	0
childage_1	0.0085	0.0016	5.27	0
childage1	0.0099	0.0019	5.32	0
d7own	0.0327	0.0175	1.88	0.0606
imdlnadj	-0.0591	0.0187	-3.16	0.0016
GymMemberAdj	-0.0565	0.0148	-3.83	0.0001
poly(RespondentAge, 3, raw = TRUE)1	-0.0971	0.0087	-11.13	0
poly(RespondentAge, 3, raw = TRUE)2	0.0022	0.0002	11.38	0
poly(RespondentAge, 3, raw = TRUE)3	0	0	-10.89	0
poly(AverageTemp, 2, raw = TRUE)1	-0.0023	0.0078	-0.29	0.7682
poly(AverageTemp, 2, raw = TRUE)2	0.0008	0.0004	2.05	0.0399
sec4hII: Managerial and Technical occupations	-0.0268	0.0231	-1.16	0.2458
sec4hIIIM: Skilled occupations - manual	-0.08	0.0265	-3.02	0.0025
sec4hIIIN: Skilled occupations - non-manual	-0.0539	0.0271	-1.99	0.0465
sec4hIV: Partly skilled occupations	-0.0812	0.0346	-2.35	0.0189
sec4hV: Unskilled occupations	-0.1176	0.0571	-2.06	0.0393
Int_Quarter5	0.0648	0.0319	2.03	0.042
Int_Quarter6	0.0595	0.0372	1.6	0.1097
Int_Quarter7	0.0714	0.0213	3.36	0.0008
d6heduc1	0.0692	0.0167	4.15	0
d6alevels	0.0374	0.0181	2.07	0.0381
SE_RegionEast Midlands	-0.0044	0.0262	-0.17	0.8658
SE_RegionLondon	-0.0176	0.035	-0.5	0.615
SE_RegionNorth East	-0.0571	0.0332	-1.72	0.086
SE_RegionNorth West	-0.02	0.0263	-0.76	0.4462
SE_RegionSouth East	0.0444	0.023	1.94	0.0529
SE_RegionSouth West	0.0441	0.0265	1.66	0.0965
SE_RegionWest Midlands	-0.0262	0.0266	-0.99	0.3232
SE_RegionYorkshire	-0.0644	0.034	-1.89	0.0586
LAPopulationDensity	0.0021	0.0005	3.89	0.0001
illness1	-0.0558	0.019	-2.94	0.0033
R Squared	0.015			
Adjusted R Squared	0.014			
Sigma	1.3479			
Rho	-0.2211			
Inverse Mills Ratio	-0.298	0.0403	-7.39	0

1.5.4. Rugby League Model Results (with Club Membership Variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	0.0303	0.0132	2.3	0.0217
Male : Respondent Age	-0.0018	0.0002	-8.71	0
Male	0.0981	0.0103	9.51	0
Attained degree-level higher education qualification	-0.0078	0.0037	-2.12	0.0343
Household Income: £15,600 to £20,799	-0.0126	0.0067	-1.88	0.0604
Household Income: £20,800 to £25,999	-0.0056	0.0072	-0.78	0.4373
Household Income: £26,000 to £31,199	-0.0036	0.0069	-0.52	0.605
Household Income: £31,200 to £36,399	-0.0142	0.0076	-1.87	0.0612
Household Income: £36,400 to £51,999	-0.0128	0.0065	-1.97	0.0489
Household Income: £52,000 or more	-0.0255	0.0071	-3.62	0.0003
SE Region: East Midlands	0.0042	0.0067	0.63	0.5262
SE Region: London	0.0028	0.0068	0.42	0.6774
SE Region: North East	0.0007	0.008	0.09	0.9279
SE Region: North West	0.0209	0.0063	3.33	0.0009
SE Region: South East	0.0103	0.0058	1.78	0.0754
SE Region: South West	0.0046	0.0067	0.69	0.4905
SE Region: West Midlands	0.0067	0.0066	1	0.315
SE Region: Yorkshire	0.0546	0.0085	6.43	0
R Squared	0.007			
Adjusted R Squared	0.007			
Sigma	0.34194			
Rho	-0.07354			
Inverse Mills Ratio	-0.0251	0.0098	-2.56	0.0104

1.5.5. Swimming Model Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	1.0659	0.135	7.9	0
d7own	0.109	0.042	2.59	0.0095
NumChildHouseholdAdj1	0.2807	0.0507	5.53	0
NumChildHouseholdAdj2	0.3713	0.0511	7.27	0
NumChildHouseholdAdj3	0.3939	0.09	4.38	0
NumChildHouseholdAdjMore than 4	0.1701	0.1759	0.97	0.3335
GymMemberAdj	0.6492	0.0403	16.12	0
I(Male * poly(RespondentAge, 2, raw = TRUE))1	-0.0439	0.0028	-15.43	0
I(Male * poly(RespondentAge, 2, raw = TRUE))2	0.0004	0	8.98	0
culturalevent1	-0.1198	0.0701	-1.71	0.0876
culturalevent2	-0.0389	0.0589	-0.66	0.5092
culturalevent3	0.2176	0.0415	5.25	0
ethwethnic	0.4103	0.0884	4.64	0
ethblack	-0.5604	0.1581	-3.54	0.0004
SE_RegionEast Midlands	0.1997	0.0707	2.82	0.0048
SE_RegionLondon	0.0084	0.0732	0.11	0.909
SE_RegionNorth East	0.248	0.0848	2.92	0.0035
SE_RegionNorth West	0.2203	0.0666	3.31	0.0009
SE_RegionSouth East	0.0411	0.0616	0.67	0.5044
SE_RegionSouth West	0.2444	0.0714	3.42	0.0006
SE_RegionWest Midlands	0.152	0.0701	2.17	0.0301
SE_RegionYorkshire	0.2205	0.0899	2.45	0.0142
AP_Swimming_Main	-0.0335	0.0085	-3.94	0.0001

R Squared	0.035			
Adjusted R Squared	0.035			
Sigma	3.6293			
Rho	0.1121			
Inverse Mills Ratio	0.4068	0.0765	5.32	0

1.5.6. Cycling Outcome Model Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	0.4062	0.1764	2.3	0.0213
I(LAPopulationDensity * (CarVanAvailable == 1))	-0.0019	0.0011	-1.73	0.084
CarVanAvailable	-0.4701	0.0651	-7.22	0
GymMemberAdj	-0.092	0.0363	-2.53	0.0113
Male	0.7869	0.0354	22.24	0
poly(RespondentAge, 2, RAW=TRUE)1	0.054	0.0057	9.47	0
poly(RespondentAge, 2, RAW=TRUE)2	-0.0006	0.0001	-10.06	0
culturalevent1	-0.1095	0.0633	-1.73	0.0838
culturalevent2	-0.1961	0.0538	-3.64	0.0003
culturalevent3	-0.1582	0.0412	-3.84	0.0001
ethwethnic	0.3566	0.0898	3.97	0.0001
ethasian	-0.2484	0.1296	-1.92	0.0552
SE_RegionEast Midlands	-0.0219	0.0636	-0.34	0.7302
SE_RegionLondon	-0.2277	0.0767	-2.97	0.003
SE_RegionNorth East	-0.0383	0.0762	-0.5	0.6149
SE_RegionNorth West	-0.0666	0.06	-1.11	0.2673
SE_RegionSouth East	-0.0508	0.0554	-0.92	0.3595
SE_RegionSouth West	-0.0392	0.0641	-0.61	0.5412
SE_RegionWest Midlands	-0.2186	0.0633	-3.45	0.0006
SE_RegionYorkshire	-0.0508	0.0809	-0.63	0.5301

R Squared	0.026			
Adjusted R Squared	0.025			
Sigma	3.29			
Rho	-0.198			
Inverse Mills Ratio	-0.653	0.1033	-6.32	0

1.5.7. Football Outcome Model Results (with Club Membership Variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	1.1415	0.2579	4.43	0
NumAdultsHouseholdAdj2	0.0159	0.0206	0.77	0.4414
NumAdultsHouseholdAdj3	0.0276	0.0327	0.84	0.3987
NumAdultsHouseholdAdj4	0.1453	0.0459	3.17	0.0015
NumAdultsHouseholdAdjMore than 4	0.1651	0.0859	1.92	0.0547
d7council	0.2085	0.0488	4.27	0
GymMemberAdj	-0.1322	0.0208	-6.35	0
FootballClub	4.8576	0.0584	83.17	0
Male	7.2594	0.3298	22.01	0
poly(RespondentAge, 3, raw = TRUE)1	-0.0409	0.0175	-2.33	0.0197
poly(RespondentAge, 3, raw = TRUE)2	0.0007	0.0004	1.84	0.0664
poly(RespondentAge, 3, raw = TRUE)3	0	0	-1.5	0.1334
culturalevent1	-0.0998	0.0364	-2.74	0.0061
culturalevent2	-0.0813	0.0309	-2.63	0.0084
culturalevent3	-0.0647	0.0233	-2.78	0.0054
sec4hII: Managerial and Technical occupations	0.0299	0.0325	0.92	0.3572
sec4hIIIM: Skilled occupations - manual	0.0723	0.0373	1.94	0.0528
sec4hIIIN: Skilled occupations - non-manual	0.0228	0.0381	0.6	0.55
sec4hIV: Partly skilled occupations	0.0715	0.0486	1.47	0.1416
sec4hV: Unskilled occupations	0.2804	0.0805	3.49	0.0005
AP_Grass_Football	-0.0198	0.0074	-2.7	0.007
d6heduc1	-0.1808	0.0264	-6.86	0
d6heduc2	-0.0825	0.0327	-2.53	0.0115
d6alevels	-0.0951	0.0273	-3.49	0.0005
Male:poly(RespondentAge, 3, raw = TRUE)1	-0.3259	0.0233	-14	0
Male:poly(RespondentAge, 3, raw = TRUE)2	0.0051	0.0005	9.81	0
Male:poly(RespondentAge, 3, raw = TRUE)3	0	0	-7.48	0
R Squared	0.285			
Adjusted R Squared	0.285			
Sigma	1.8724			
Rho	-0.1021			
Inverse Mills Ratio	-0.1912	0.0638	-3	0.0027

1.5.8. Football Outcome Model Results (with Mezzanine Club Membership Variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	0.8833	0.279	3.17	0.0015
NumAdultsHouseholdAdj2	0.0123	0.0221	0.56	0.5774
NumAdultsHouseholdAdj3	0.0148	0.035	0.42	0.6719
NumAdultsHouseholdAdj4	0.1	0.0494	2.03	0.0428
NumAdultsHouseholdAdjMore than 4	0.1559	0.0922	1.69	0.0906
d7council	0.2041	0.0523	3.9	0.0001
GymMemberAdj	-0.0707	0.0237	-2.98	0.0028
Football.club.prob	12.4147	0.9118	13.62	0
Male	5.2285	0.4306	12.14	0
poly(RespondentAge, 3, raw = TRUE)1	-0.0346	0.0188	-1.84	0.0662
poly(RespondentAge, 3, raw = TRUE)2	0.0006	0.0004	1.48	0.1383
poly(RespondentAge, 3, raw = TRUE)3	0	0	-1.25	0.2109
culturalevent1	-0.1204	0.0391	-3.08	0.0021
culturalevent2	-0.0624	0.0331	-1.88	0.0595
culturalevent3	-0.0405	0.0251	-1.61	0.1065
sec4hII: Managerial and Technical occupations	0.0309	0.0348	0.89	0.3752
sec4hIIIM: Skilled occupations - manual	0.0606	0.0401	1.51	0.1307
sec4hIIIN: Skilled occupations - non-manual	0.0152	0.041	0.37	0.7103
sec4hIV: Partly skilled occupations	0.0671	0.0521	1.29	0.1985
sec4hV: Unskilled occupations	0.2805	0.0863	3.25	0.0011
AP_Grass_Football	-0.0235	0.0079	-2.98	0.0029
d6heduc1	-0.1227	0.0292	-4.2	0
d6heduc2	-0.0663	0.0351	-1.89	0.0591
d6alevels	-0.0978	0.0292	-3.35	0.0008
Male:poly(RespondentAge, 3, raw = TRUE)1	-0.2315	0.0271	-8.55	0
Male:poly(RespondentAge, 3, raw = TRUE)2	0.0035	0.0006	6.16	0
Male:poly(RespondentAge, 3, raw = TRUE)3	0	0	-4.7	0
Inverse Mills Ratio	-0.1074	0.0691	-1.56	0.1198

1.5.9. Badminton Outcome Model Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	-0.0054	0.1004	-0.05	0.9571
illness1	-0.0578	0.015	-3.84	0.0001
LAPopulationDensity	-0.0015	0.0003	-5.21	0
SimpsonDiversityIndex	0.2099	0.109	1.92	0.0543
childage1	0.0144	0.0025	5.67	0
Male	0.0225	0.0111	2.03	0.0423
poly(RespondentAge, 2, raw = TRUE)1	0.0064	0.0019	3.47	0.0005
poly(RespondentAge, 2, raw = TRUE)2	-0.0001	0	-3.68	0.0002
NumChildHouseholdAdj1	-0.1333	0.0256	-5.21	0
NumChildHouseholdAdj2	-0.0712	0.0152	-4.67	0
NumChildHouseholdAdj3	-0.0415	0.0266	-1.56	0.1181
NumChildHouseholdAdjMore than 4	0.0002	0.0518	0	0.9964
d7own	0.0391	0.0137	2.85	0.0043
ethasian	0.3274	0.0333	9.84	0
ethchinese	0.6688	0.1129	5.92	0
AP_SportsHall_Main	-0.0131	0.0051	-2.59	0.0096
GymMemberAdj	-0.0656	0.0119	-5.51	0
R Squared	0.007			
Adjusted R Squared	0.007			
Sigma	1.0625			
Rho	-0.0473			
imrData$IMR1	-0.0502	0.0285	-1.76	0.0782

1.5.10. Golf Outcome Model Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	-0.9811	0.1111	-8.83	0
LAPopulationDensity	-0.0039	0.0005	-8.49	0
d7own	0.2058	0.0259	7.93	0
NumChildHouseholdAdj1	-0.0478	0.0279	-1.71	0.0868
NumChildHouseholdAdj2	-0.1176	0.0283	-4.16	0
NumChildHouseholdAdj3	-0.1034	0.0492	-2.1	0.0358
NumChildHouseholdAdjMore than 4	-0.2251	0.0958	-2.35	0.0188
GymMemberAdj	-0.2172	0.0221	-9.85	0
poly(AverageTemp, 2, raw = TRUE)1	0.048	0.0097	4.96	0
poly(AverageTemp, 2, raw = TRUE)2	-0.0013	0.0005	-2.64	0.0082
Male	0.7111	0.0209	33.98	0
poly(RespondentAge, 2, raw = TRUE)1	-0.0182	0.0037	-4.97	0
poly(RespondentAge, 2, raw = TRUE)2	0.0004	0	10.14	0
sec4hII: Managerial and Technical occupations	0.0947	0.0341	2.78	0.0054
sec4hIIIM: Skilled occupations - manual	0.0899	0.039	2.31	0.0212
sec4hIIIN: Skilled occupations - non-manual	0.0624	0.0398	1.57	0.117
sec4hIV: Partly skilled occupations	-0.0143	0.0513	-0.28	0.7805
sec4hV: Unskilled occupations	-0.1571	0.0846	-1.86	0.0632
ethwethnic	0.1221	0.0425	2.87	0.0041
d23_bands_7£15,600 to £20,799	0.2289	0.0387	5.91	0
d23_bands_7£20,800 to £25,999	0.318	0.0421	7.56	0
d23_bands_7£26,000 to £31,199	0.3518	0.041	8.59	0
d23_bands_7£31,200 to £36,399	0.3533	0.045	7.86	0
d23_bands_7£36,400 to £51,999	0.463	0.0397	11.67	0
d23_bands_7£52,000 or more	0.5452	0.0445	12.25	0

R Squared	0.082			
Adjusted R Squared	0.082			
Sigma	1.9903			
Rho	0.2061			
Inverse Mills Ratio	0.4102	0.0721	5.69	0

1.5.11. Squash Outcome Model Results

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	-0.0079	0.0386	-0.2	0.8382
illness1	-0.0225	0.0134	-1.68	0.0924
LAPopulationDensity	-0.0007	0.0002	-3.14	0.0017
NumChildHouseholdAdj1	-0.0463	0.0134	-3.46	0.0005
NumChildHouseholdAdj2	-0.0175	0.0134	-1.3	0.194
NumChildHouseholdAdj3	-0.0581	0.0236	-2.47	0.0136
NumChildHouseholdAdjMore than 4	-0.0316	0.046	-0.69	0.4915
Int_Quarter5	0.0339	0.0127	2.67	0.0077
Int_Quarter6	0.0422	0.013	3.26	0.0011
Int_Quarter7	0.0221	0.0123	1.8	0.072
Male	0.1502	0.0098	15.38	0
poly(RespondentAge, 2, raw = TRUE)1	0.008	0.0016	4.89	0
poly(RespondentAge, 2, raw = TRUE)2	-0.0001	0	-5.06	0
GymMemberAdj	0.0255	0.0106	2.41	0.0158
R Squared	0.011			
Adjusted R Squared	0.011			
Sigma	0.949			
Rho	-0.142			
Inverse Mills Ratio	-0.1348	0.025	-5.4	0

1.5.12. Cricket Outcome Model Results (with Cricket Club Membership Variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	-0.0685	0.0261	-2.63	0.0086
ethasian	0.1543	0.0176	8.74	0
CricketClub == 1TRUE	3.9197	0.0311	126	0
SE_RegionEast Midlands	0.0009	0.0112	0.08	0.9349
SE_RegionLondon	-0.0217	0.0115	-1.88	0.0597
SE_RegionNorth East	-0.0021	0.0135	-0.16	0.8761
SE_RegionNorth West	0.0025	0.0105	0.24	0.8128
SE_RegionSouth East	-0.0019	0.0098	-0.2	0.8435
SE_RegionSouth West	0.0068	0.0113	0.6	0.5508
SE_RegionWest Midlands	0.0081	0.0111	0.73	0.4658
SE_RegionYorkshire	0.0314	0.0142	2.21	0.0272
NumChildHouseholdAdj1	0.0102	0.0079	1.3	0.1941
NumChildHouseholdAdj2	0.0511	0.008	6.42	0
NumChildHouseholdAdj3	0.0597	0.0141	4.23	0
NumChildHouseholdAdjMore than 4	0.021	0.0277	0.76	0.448
Int_Quarter5	0.0041	0.0137	0.3	0.7625
Int_Quarter6	0.0231	0.016	1.45	0.1472
Int_Quarter7	0.0302	0.0092	3.29	0.001
Male	0.3363	0.0313	10.75	0
Male : Respondent Age	-0.0117	0.0013	-8.7	0
Male : Respondent Age ^ 2	0.0001	0	7.08	0
Average Temp	-0.0042	0.0033	-1.27	0.203
Average Temp ^ 2	0.0006	0.0002	3.4	0.0007
TotalRainfallAdj	0.0037	0.0009	3.95	0.0001
Inverse Mills Ratio	0.0246	0.0119	2.06	0.0393

1.5.13. Cricket Outcome Model Results (with Mezzanine Club Membership Variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	-0.0839	0.0305	-2.75	0.0059
ethasian	0.1531	0.0205	7.46	0
Cricket.club.prob	5.4726	0.5563	9.84	0
SE_RegionEast Midlands	0.0028	0.013	0.21	0.8322
SE_RegionLondon	-0.0113	0.0135	-0.83	0.4054
SE_RegionNorth East	-0.0056	0.0157	-0.36	0.7196
SE_RegionNorth West	0.0007	0.0122	0.06	0.9543
SE_RegionSouth East	0.0028	0.0113	0.24	0.8068
SE_RegionSouth West	0.0201	0.0132	1.53	0.1269
SE_RegionWest Midlands	0.0167	0.0129	1.3	0.1935
SE_RegionYorkshire	0.0301	0.0165	1.82	0.0686
NumChildHouseholdAdj1	0.017	0.0091	1.87	0.0621
NumChildHouseholdAdj2	0.0484	0.0093	5.23	0
NumChildHouseholdAdj3	0.0582	0.0164	3.54	0.0004
NumChildHouseholdAdjMore than 4	0.0011	0.0322	0.03	0.9726
Int_Quarter5	-0.0109	0.0159	-0.69	0.4916
Int_Quarter6	0.0166	0.0186	0.89	0.3715
Int_Quarter7	0.0342	0.0107	3.2	0.0014
Male	0.2759	0.0409	6.74	0
I(Male * poly(RespondentAge, 2, raw = TRUE))1	-0.0102	0.0016	-6.39	0
I(Male * poly(RespondentAge, 2, raw = TRUE))2	0.0001	0	5.37	0
poly(AverageTemp, 2, raw = TRUE)1	-0.0046	0.0039	-1.19	0.2335
poly(AverageTemp, 2, raw = TRUE)2	0.0006	0.0002	3.11	0.0019
TotalRainfallAdj	0.0043	0.0011	3.99	0.0001
Inverse Mills Ratio	0.0381	0.0143	2.67	0.0077

1.5.14. Rugby Union Model Results (with Club Membership variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	0.0396	0.0233	1.7	0.0885
RubgyUnionClub == 1TRUE	5.2462	0.0339	154.73	0
rugbyunion_awardamount10	0	0	1.87	0.0616
Male	0.2151	0.0182	11.8	0
d6heduc2	-0.0293	0.0096	-3.04	0.0024
imdlnadj	-0.0026	0.0057	-0.47	0.6403
NumAdultsHouseholdAdj2	-0.0008	0.0065	-0.12	0.907
NumAdultsHouseholdAdj3	0.0281	0.0102	2.75	0.006
NumAdultsHouseholdAdj4	0.0559	0.0144	3.89	0.0001
NumAdultsHouseholdAdjMore than 4	0.0661	0.0274	2.41	0.0159
d7own	0.0277	0.0067	4.16	0
ethasian	-0.0499	0.0186	-2.69	0.0072
AverageTemp	-0.0007	0.0006	-1.05	0.2951
I(Male * RespondentAge)	-0.0042	0.0004	-11.64	0
d6heduc1	-0.0294	0.007	-4.22	0
Inverse Mills Ratio	-0.0228	0.0144	-1.58	0.1133

1.5.15. Rugby Union Model Results (with Mezzanine Club Membership variable)

	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	-0.0088	0.029	-0.3	0.7616
RU.club.prob	11.6966	0.44	26.58	0
rugbyunion_awardamount10	0	0	2.74	0.0061
Male	-0.0547	0.0289	-1.89	0.0587
d6heduc2	-0.0254	0.0118	-2.14	0.0321
imdlnadj	0.0063	0.007	0.9	0.3677
NumAdultsHouseholdAdj2	-0.0142	0.0081	-1.76	0.0788
NumAdultsHouseholdAdj3	-0.0034	0.0127	-0.27	0.7876
NumAdultsHouseholdAdj4	0.0376	0.0177	2.12	0.034
NumAdultsHouseholdAdjMore than 4	0.0117	0.0339	0.34	0.7308
d7own	0.0224	0.0082	2.74	0.0062
ethasian	0.0268	0.0234	1.15	0.2521
AverageTemp	-0.0009	0.0008	-1.16	0.2465
I(Male * RespondentAge)	0	0.0005	0.03	0.9742
d6heduc1	-0.0111	0.0087	-1.28	0.2007
Inverse Mills Ratio	0.0042	0.0179	0.24	0.8124

1.5.16. Cricket Club Membership Model Results

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	-2.6619	0.2909	-9.15	0
NumAdultsHouseholdAdj2	0.0327	0.0589	0.55	0.5792
NumAdultsHouseholdAdj3	0.1931	0.0772	2.5	0.0124
NumAdultsHouseholdAdj4	0.2603	0.0996	2.61	0.0089
NumAdultsHouseholdAdjMore than 4	0.5154	0.1416	3.64	0.0003
d7council	-0.2109	0.1838	-1.15	0.2514
d7own	-0.1213	0.0619	-1.96	0.0501
GymMemberAdj	-0.2507	0.0625	-4.01	0.0001
d23_bands_7£15,600 to £20,799	0.0193	0.1169	0.17	0.8685
d23_bands_7£20,800 to £25,999	-0.0162	0.1213	-0.13	0.8939
d23_bands_7£26,000 to £31,199	-0.0617	0.1175	-0.53	0.5995
d23_bands_7£31,200 to £36,399	-0.1451	0.1274	-1.14	0.2547
d23_bands_7£36,400 to £51,999	-0.1025	0.1189	-0.86	0.3884
d23_bands_7£52,000 or more	-0.2089	0.1361	-1.54	0.1246
LAPopulationDensity	-0.0031	0.0013	-2.41	0.016
Male	0.8953	0.0898	9.97	0
poly(RespondentAge, 2, raw = TRUE)1	0.0211	0.0103	2.04	0.041
poly(RespondentAge, 2, raw = TRUE)2	-0.0001	0.0001	-1	0.3149
culturalevent1	-0.2775	0.1039	-2.67	0.0075
culturalevent2	-0.0422	0.0747	-0.57	0.572
culturalevent3	-0.0958	0.0602	-1.59	0.1119
sec4hII: Managerial and Technical occupations	0.0678	0.078	0.87	0.3842
sec4hIIIM: Skilled occupations - manual	-0.0243	0.0917	-0.27	0.7906
sec4hIIIN: Skilled occupations - non-manual	0.1184	0.0923	1.28	0.1997
sec4hIV: Partly skilled occupations	0.0632	0.1277	0.49	0.6207
sec4hV: Unskilled occupations	-0.1034	0.2347	-0.44	0.6594
d6heduc1	-0.1482	0.0836	-1.77	0.0761
d6heduc2	-0.2125	0.0988	-2.15	0.0315
d6alevels	-0.0715	0.0798	-0.9	0.3705
d6gcse	-0.0623	0.0861	-0.72	0.4693
ethwethnic	0.0991	0.1554	0.64	0.5239
ethasian	0.2846	0.1917	1.48	0.1377
Inverse Mills Ratio	-1.3209	0.2858	-4.62	0

1.5.17. Football Club Membership Model Results

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	-1.5032	0.22	-6.83	0
NumAdultsHouseholdAdj2	0.0319	0.0451	0.71	0.479
NumAdultsHouseholdAdj3	0.0459	0.0607	0.76	0.4493
NumAdultsHouseholdAdj4	0.1168	0.075	1.56	0.1194
NumAdultsHouseholdAdjMore than 4	0.1357	0.12	1.13	0.258
d7council	-0.0151	0.0861	-0.18	0.8609
d7own	-0.1099	0.0523	-2.1	0.0358
GymMemberAdj	-0.2938	0.0477	-6.16	0
d23_bands_7£15,600 to £20,799	0.0654	0.0833	0.79	0.4323
d23_bands_7£20,800 to £25,999	0.1048	0.0855	1.23	0.2205
d23_bands_7£26,000 to £31,199	0.133	0.0831	1.6	0.1093
d23_bands_7£31,200 to £36,399	0.0708	0.0897	0.79	0.4304
d23_bands_7£36,400 to £51,999	0.1141	0.0848	1.35	0.1786
d23_bands_7£52,000 or more	0.0364	0.0978	0.37	0.7097
LAPopulationDensity	-0.0013	0.0009	-1.52	0.1284
Male	0.9958	0.0624	15.95	0
poly(RespondentAge, 2, raw = TRUE)1	-0.0213	0.0089	-2.39	0.0169
poly(RespondentAge, 2, raw = TRUE)2	-0.0001	0.0001	-1.02	0.3076
culturalevent1	0.0585	0.06	0.98	0.3295
culturalevent2	-0.0157	0.0563	-0.28	0.7797
culturalevent3	-0.0654	0.0452	-1.45	0.1478
sec4hII: Managerial and Technical occupations	0.0367	0.0649	0.57	0.5718
sec4hIIIM: Skilled occupations - manual	0.0607	0.0717	0.85	0.3974
sec4hIIIN: Skilled occupations - non-manual	0.0793	0.074	1.07	0.2836
sec4hIV: Partly skilled occupations	0.0472	0.0934	0.5	0.6137
sec4hV: Unskilled occupations	0.0087	0.1421	0.06	0.9514
d6heduc1	-0.2005	0.0621	-3.23	0.0012
d6heduc2	-0.0795	0.0708	-1.12	0.2614
d6alevels	-0.0228	0.0563	-0.41	0.6849
d6gcse	-0.0386	0.0593	-0.65	0.5146
ethwethnic	-0.1637	0.0775	-2.11	0.0346
ethasian	-0.6833	0.1441	-4.74	0
Inverse Mills Ratio	-0.1473	0.2	-0.74	0.4614

1.5.18. Rugby Union Club Membership Results

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	-2.5364	0.3179	-7.98	0
NumAdultsHouseholdAdj2	0.1076	0.0687	1.57	0.1173
NumAdultsHouseholdAdj3	0.1087	0.089	1.22	0.2217
NumAdultsHouseholdAdj4	-0.0025	0.1143	-0.02	0.9825
NumAdultsHouseholdAdjMore than 4	0.115	0.1691	0.68	0.4964
d7council	-0.1095	0.1504	-0.73	0.4666
d7own	-0.0452	0.071	-0.64	0.5246
GymMemberAdj	-0.1137	0.0613	-1.86	0.0636
d23_bands_7£15,600 to £20,799	-0.0016	0.1391	-0.01	0.9908
d23_bands_7£20,800 to £25,999	0.125	0.1364	0.92	0.3594
d23_bands_7£26,000 to £31,199	0.2095	0.1298	1.61	0.1066
d23_bands_7£31,200 to £36,399	0.1188	0.1411	0.84	0.3998
d23_bands_7£36,400 to £51,999	0.207	0.1319	1.57	0.1165
d23_bands_7£52,000 or more	0.4142	0.1476	2.81	0.005
LAPopulationDensity	-0.0056	0.0014	-3.89	0.0001
Male	1.1948	0.1079	11.07	0
poly(RespondentAge, 2, raw = TRUE)1	-0.0466	0.0116	-4.03	0.0001
poly(RespondentAge, 2, raw = TRUE)2	0.0002	0.0001	1.71	0.0865
culturalevent1	0.0903	0.086	1.05	0.2939
culturalevent2	0.1655	0.0733	2.26	0.0239
culturalevent3	-0.0276	0.0643	-0.43	0.6684
sec4hII: Managerial and Technical occupations	0.0021	0.0798	0.03	0.9791
sec4hIIIM: Skilled occupations - manual	-0.1954	0.0961	-2.03	0.042
sec4hIIIN: Skilled occupations - non-manual	-0.1274	0.0995	-1.28	0.2002
sec4hIV: Partly skilled occupations	0.0148	0.1274	0.12	0.9074
sec4hV: Unskilled occupations	-0.2869	0.2436	-1.18	0.239
d6heduc1	0.0211	0.0932	0.23	0.8206
d6heduc2	0.1422	0.1024	1.39	0.1652
d6alevels	0.1094	0.0867	1.26	0.207
d6gcse	0.1758	0.0884	1.99	0.0467
ethwethnic	0.1489	0.1419	1.05	0.2939
ethasian	-0.802	0.3576	-2.24	0.0249
Inverse Mills Ratio	0.363	0.2935	1.24	0.2161

1.5.19. Rugby League Club Model Membership Results

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	-2.8908	0.7742	-3.73	0.0002
NumAdultsHouseholdAdj2	-0.1411	0.1315	-1.07	0.2831
NumAdultsHouseholdAdj3	-0.2999	0.1836	-1.63	0.1024
NumAdultsHouseholdAdj4	-0.4452	0.2395	-1.86	0.0631
NumAdultsHouseholdAdjMore than 4	-0.1633	0.3231	-0.51	0.6131
d7council	0.2877	0.217	1.33	0.1849
d7own	-0.0722	0.1596	-0.45	0.6508
GymMemberAdj	-0.4322	0.1937	-2.23	0.0257
d23_bands_7£15,600 to £20,799	0.7908	0.3896	2.03	0.0424
d23_bands_7£20,800 to £25,999	0.9387	0.3926	2.39	0.0168
d23_bands_7£26,000 to £31,199	0.9839	0.3938	2.5	0.0125
d23_bands_7£31,200 to £36,399	0.8262	0.4159	1.99	0.0469
d23_bands_7£36,400 to £51,999	0.935	0.4061	2.3	0.0213
d23_bands_7£52,000 or more	0.8442	0.4368	1.93	0.0533
LAPopulationDensity	-0.006	0.0035	-1.72	0.0854
Male	0.8567	0.2265	3.78	0.0002
poly(RespondentAge, 2, raw = TRUE)1	-0.092	0.0238	-3.87	0.0001
poly(RespondentAge, 2, raw = TRUE)2	0.0007	0.0002	3.07	0.0022
culturalevent1	0.0769	0.1808	0.43	0.6708
culturalevent2	0.1186	0.161	0.74	0.4613
culturalevent3	0.0131	0.1378	0.09	0.9243
sec4hII: Managerial and Technical occupations	0.4087	0.309	1.32	0.186
sec4hIIIM: Skilled occupations - manual	0.4786	0.3203	1.49	0.1351
sec4hIIIN: Skilled occupations - non-manual	0.5506	0.3217	1.71	0.0869
sec4hIV: Partly skilled occupations	0.4675	0.3695	1.27	0.2057
sec4hV: Unskilled occupations	0.6965	0.4201	1.66	0.0973
d6heduc1	-0.0426	0.183	-0.23	0.8158
d6heduc2	-0.05	0.2127	-0.24	0.8141
d6alevels	-0.1348	0.1645	-0.82	0.4127
d6gcse	-0.1602	0.1695	-0.94	0.3447
ethwethnic	0.3574	0.3615	0.99	0.3229
ethasian	-2.9946	101.6993	-0.03	0.9765
Inverse Mills Ratio	0.1644	0.6445	0.26	0.7987