



department for
**culture, media
and sport**

Models of sporting and cultural activity

Analysis of the Taking Part Survey

Geoff Inglis and Joel Williams, TNS-BMRB

August 2010

Our aim is to improve the quality of life for all through cultural and sporting activities, support the pursuit of excellence, and champion the tourism, creative and leisure industries.

Contents

Executive Summary	2
Chapter 1: Introduction	5
Chapter 2: Modelling recreational physical activity and sport	10
Chapter 3: Modelling arts activity	17
Chapter 4: Modelling museum and gallery visits	26
Chapter 5: Modelling heritage site visits.....	36
Chapter 6: Modelling library visits	46
Endnotes	53

Executive Summary

Background and aims of the research

This report presents findings from multivariate analysis of the 2007/08 Taking Part survey on the predictors of participation in culture and sport.

The analysis was commissioned by the Department for Culture, Media and Sport and undertaken by TNS-BMRB.

The research was driven by one overarching research question:

What are the predictors of participation in culture and sport?

By using a combination of multivariate research techniques (logistic regression and multiple classification analysis), the study identifies the sort of characteristics and behaviours associated with participation in five sectors: sport, arts, heritage, museums and galleries, and libraries.

Summary of findings

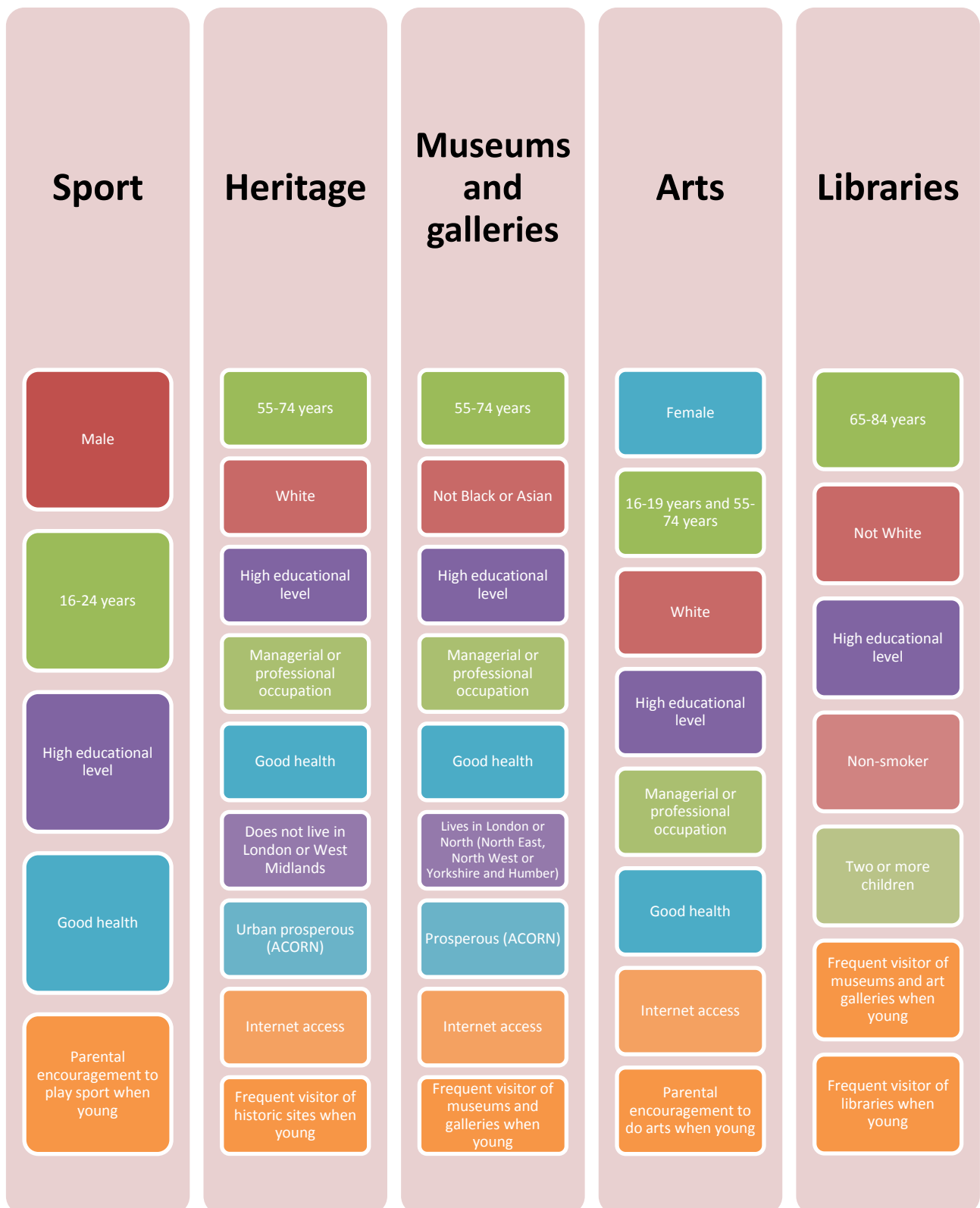
The modelling for this study was conducted in two stages. Headline findings from each stage are summarised in Charts A and B.

Among background characteristics, age and education predict participation in all sectors; health and ethnicity predict participation in all sectors except libraries.

The direction of influence is generally the same across all sectors, but there are exceptions. For example, youth is a strong predictor of sports participation while being older is a predictor of participation in the other sectors. The relationship between ethnicity and participation is not straightforward either. While Black and Asian people are less likely to participate in heritage, museum and galleries and the arts, they are more likely than White people to attend libraries

Participating in a sector as a child also consistently predicts adult participation in that sector, i.e. people who frequently visited libraries as a child were more likely to have attended a library in the last 12 months. Similarly, people who were encouraged to play sport as a child were more likely to participate in sport as an adult (Chart A).

Chart A: Predictors of participation: summary of background characteristics

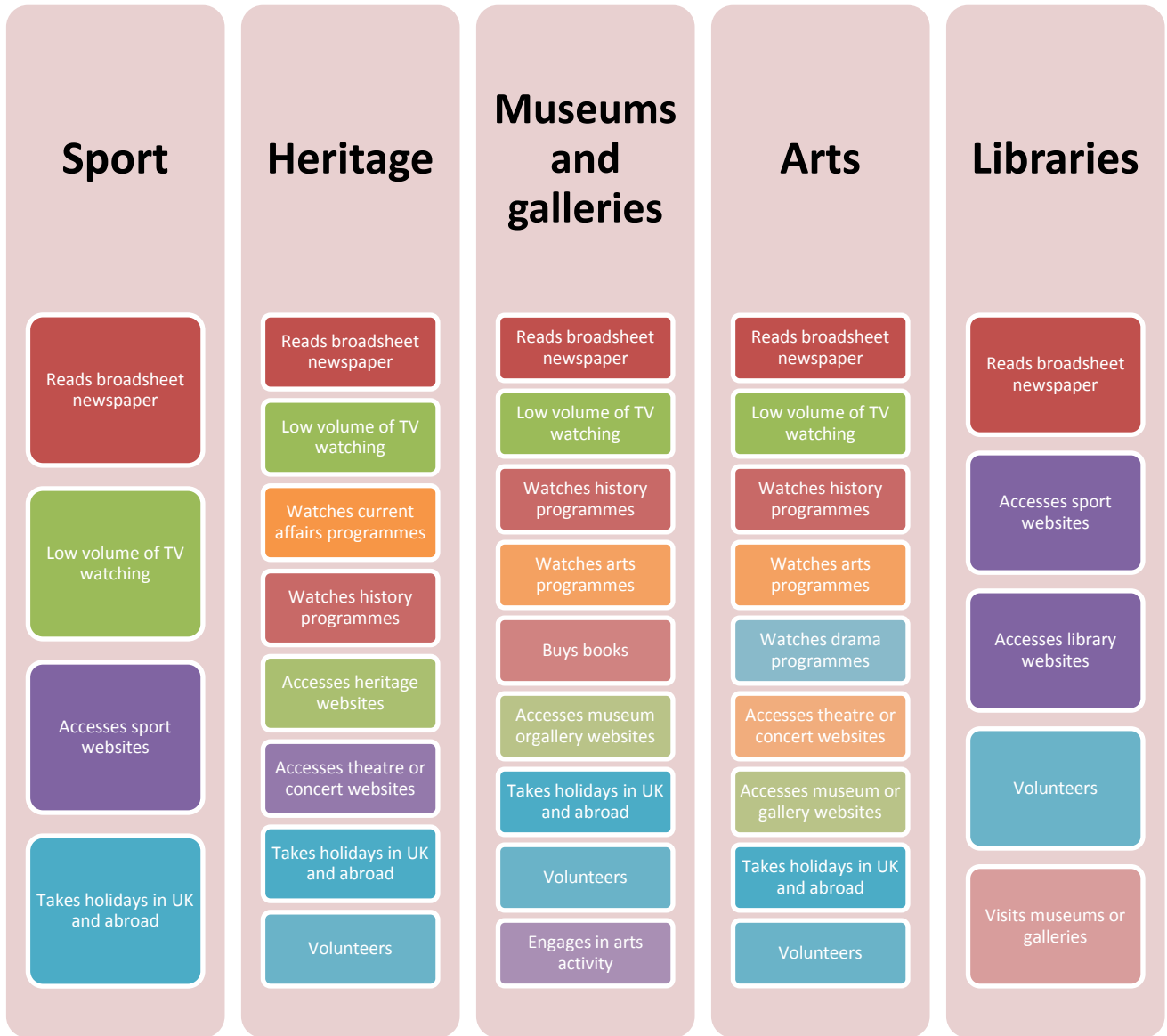


At the second stage, current activities were modelled to see if these also had an impact on participation in culture and sport.

Among these current activities, newspaper readership – reading a broadsheet newspaper in particular – predicts participation in all sectors; not watching very much TV and holiday

activity both predict participation in all sectors except libraries while volunteering predicts participation in all sectors except sport. Not surprisingly, a wider interest in a sector also tends to predict participation in it – for example, accessing sports websites predicts sports participation while watching history shows predicts attendance at heritage sites (Chart B).

Chart B: Predictors of participation: summary of ‘current activities’ models



Chapter 1: Introduction

This report presents findings from analysis of the Taking Part survey conducted by TNS-BMRB for the Department for Culture, Media and Sport (DCMS) on the predictors of participation in culture and sport.

TNS-BMRB produced models for participation across five sport and culture sectors, each broadly representing one element of PSA21's Indicator 6ⁱ:

1. Recreational physical activity and sport (participated in 30 minutes of moderate intensity level sport and active recreation on three or more days in the past week)
2. Arts activity (engaged in the arts at least three times in the past 12 months)
3. Museum and gallery visits (attended a museum or gallery at least once in the past 12 months)
4. Heritage site visits (visited at least two historic environment sites in the past 12 months)
5. Library visits (used a public library service at least once in the past 12 months)

The analysis used the 2007/08 Taking Part dataset and was commissioned to gain a better understanding of the sort of behaviours and socio-demographic characteristics which predict participation. .

While not an exact fit for PSA21 performance assessment, each model provides a good summary of the main predictors of participation for that sector.

Analytic approach

TNS-BMRB used two multivariate analysis techniques to produce the models: logistic regression and multiple classification analysis (MCA).

Logistic regression is the industry standard method for modelling the relationship between a selection of 'predictor' variables and a single binary dependent variable. All the PSA21 variables are binary: either an individual reports a sufficient level of activity in sector x or he/she does not.

However, the output data from logistic regression is not simple to interpret so TNS-BMRB used multiple classification analysis to present the findingsⁱⁱ.

The MCA method shows how the proportion of people participating would change if the 'real' distributionⁱⁱⁱ of a particular predictor variable is replaced with an imaginary distribution in which the whole population has a specified characteristic. For example, if everyone was educated to degree level rather than to the different levels we observe, the proportion of people participating would go up^{iv}.

TNS-BMRB analysts distinguished between two types of predictor variables: (1) background characteristics, and (2) current activities.

'Background characteristics' includes demographic variables like sex, age, health and educational level as well as data about past experiences, demographic variables over which the individual has some choice (e.g. occupational type) and 'structural' factors such as location and even internet access.

As a group these variables can be highly predictive of sporting and cultural activity but their population distributions tend to change very slowly and are resistant to most direct interventions.

In contrast, behaviour that is less fundamental and more a product of real choice can mutate quite rapidly. These variables have been grouped under the banner 'current activities'. These activities include TV viewing behaviour, propensity to engage in voluntary work, holiday choices and activity in other sport and cultural sectors.

Taking Part includes a variety of questions covering both background characteristics and current activities. Rather than generate models in which both types of variable are mixed together, TNS-BMRB elected to produce two-stage models.

The first stage model was based only on background characteristics. Using logistic regression, a 'probability value' was generated for each sample case. This gives the base likelihood of meeting PSA activity criteria, given the individual's background characteristics.

TNS-BMRB then divided the sample into three roughly equal-sized groups, labelling them as having relatively 'high', 'middle' or 'low' base likelihoods.

Second-stage models were based only on current activities and were generated separately for each of the three base likelihood groups.

In this way, TNS-BMRB was able to show how the relationship between the target PSA activity and other current activities differs between the base likelihood groups.

While qualitative measures are extremely useful when modelling behaviour, allowing us to differentiate between demographically similar individuals, the models do not include attitudinal questions. Consequently, the models described in this paper are fairly 'weak'. Models are considered weak if they lack explanatory power. In other words, much of the observed variety in activity levels cannot be explained with reference to the other data collected by the survey. While this is a fairly common occurrence in social research it is nevertheless disappointing.

The remainder of this paper details the models for each sector: sport, museums and galleries, libraries, the arts and heritage. Each chapter follows a set pattern to make comparison between the models simple.

Please note that the magnitude of an effect is an estimate subject to sampling error (as well as other sources of error). Some of the observed effects may be the consequence of sampling error rather than reflecting a true effect. In particular, effect sizes can be unstable if the base sample size is small. TNS-BMRB has chosen to not report effect sizes based on samples <50 but to include effect sizes based on samples of 50+ even if the probability of 'spuriousness' exceeds the standard level of 5%^v.

Overall model summary

All five models have a lot in common. Table 1.1 shows the 16 background characteristics that appear across the five predictive models. Two variables – educational level and age cohort – appear in all five models and two others – general health and ethnic group – appear in four.

The direction of influence is usually the same across all five models but there are exceptions. For example, sport and recreational physical activity is associated with youth but other activities are associated most strongly with late middle age. Another interesting example is ethnic group. For the most part Black and Asian people are less likely than others to engage in cultural activities but the exception is library use.

Table 1.1 Summary of ‘background characteristics’ models

VARIABLE	MODEL				
	Sport	Heritage	Museums	Arts	Libraries
Sex	* (male)			* (female)	
Age cohort	* (16-24)	* (55-74)	* (55-74)	* (16-19, 55-74)	* (65-84)
Ethnic group		* (White)	* (not Black or Asian)	* (white)	* (not white)
Educational level	* (high)	* (high)	* (high)	* (high)	* (high)
Occupation		* (managerial / professional)	* (managerial / professional)	* (managerial / professional)	
General health	* (good)	* (good)	* (good)	* (good)	
Smoking status					* (non-smoker)
Region		* (not London or West Midlands)	* (North East, North West and Yorkshire and the Humber or London)		
Housing segment (ACORN)		* (urban prosperous)	* (prosperous)		
Internet access		* (yes)	* (yes)	* (yes)	
Number of children					* (2+)
Level of parental encouragement to play sport when young	* (high)				
Level of parental encouragement to play musical instruments, act, dance, sing etc.				* (high)	
Frequency of visiting museums/art galleries when young			* (frequent)		* (frequent)
Frequency of visiting historic sites when young		* (frequent)			
Frequency of visiting libraries when young					* (frequent)

Table 1.2 covers the second-stage models based on current activities and was generated separately for each of the three base likelihood groups. For the most part the base likelihood groups differ only in the *magnitude* of the penalties and rewards associated with particular behaviours, not the direction of influence.

One variable – newspaper readership - appears in all five models and three others – volume of TV viewing, holiday activity and voluntary work – appear in four.

Table 1.2 Summary of 'current activities' models

VARIABLE	MODEL				
	Sport	Heritage	Museums	Arts	Libraries
Newspaper readership	* (b/sheet)	* (b/sheet)	* (b/sheet)	* (b/sheet)	* (b/sheet)
Volume of TV viewing	* (low)	* (low)	* (low)	* (low)	
Watching current affairs programmes		* (yes)			
Watching history programmes		* (yes)	* (yes)	* (yes)	
Watching arts programmes			* (yes)	* (yes)	
Watching drama programmes				* (yes)	
Buying books			* (yes)		
Reading for pleasure					* (yes)
Accessing sport websites	* (yes)				
Accessing historical websites		* (yes)			
Accessing theatre/concert websites		* (yes)		* (yes)	
Accessing museum/gallery websites			* (yes)	* (yes)	
Accessing library websites					* (yes)
Holiday activity	* (UK & abroad)	* (UK & abroad or UK only)	* (UK & abroad)	* (UK & abroad)	
Voluntary work		* (yes)	* (yes)	* (yes)	* (yes)
Arts activity (PSA)			* (high)		
Visiting					* (yes)

**museums/galleries
(PSA)**

--

--

--

--

--

Chapter 2: Modelling recreational physical activity and sport

Background characteristics

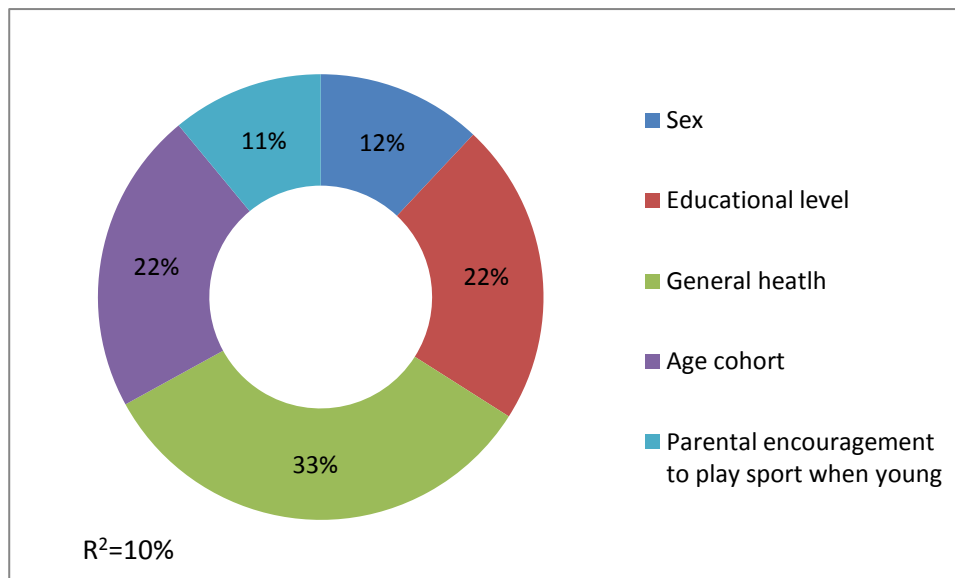
The model for recreational physical activity and sport was the weakest. It is estimated that the stage 1 'background characteristics' model explains only 10% of the observed variety in activity levels^{vi}.

Only five background characteristics proved significantly^{vii} predictive:

- General health
- Age cohort
- Educational level
- Sex
- Parental encouragement to play sport when young

Chart 2.1 shows the relative weight of each of these variables in terms of independent predictive power.

Chart 2.1 Predictors of sports participation – background characteristics



As shown in Chart 2.1, the predictors of sports participation were: very good general health, youth, a high educational level, and strong parental encouragement to play sport when young. Men were also more likely than women to participate, even when holding the other characteristics constant.

The predictive power of the model was not improved by adding interaction terms. An interaction term would be effective if, for example, the model was weaker for older women than it is for younger women (i.e. the influence of sex varies with age).

However, that was not the case here or, indeed, for any other model described in this paper^{viii}.

The three most *positive* characteristics were:

- Very good health (+10 percentage points (PP))
- Aged under 25 (+14PP if aged 16-19, +8PP if aged 20-24)
- Educated to degree level (+7PP)

The scores in brackets show the expected increase in the proportion participating in 30 minutes of moderate intensity level sport and active recreation on three or more days in the past week if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'reward'.

The three most *negative* characteristics are:

- Fair health or worse (-8PP if 'fair', -12PP if 'bad', -13PP if 'very bad')
- Aged >74 (-8PP if aged 75-84, -11PP if aged 85+)
- No qualifications (-5PP)

The scores in brackets show the expected decrease in the proportion participating in 30 minutes of moderate intensity level sport and active recreation on three or more days in the past week if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty'.

Charts 2.2 to 2.6 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Chart 2.2 Gain in probability – general health

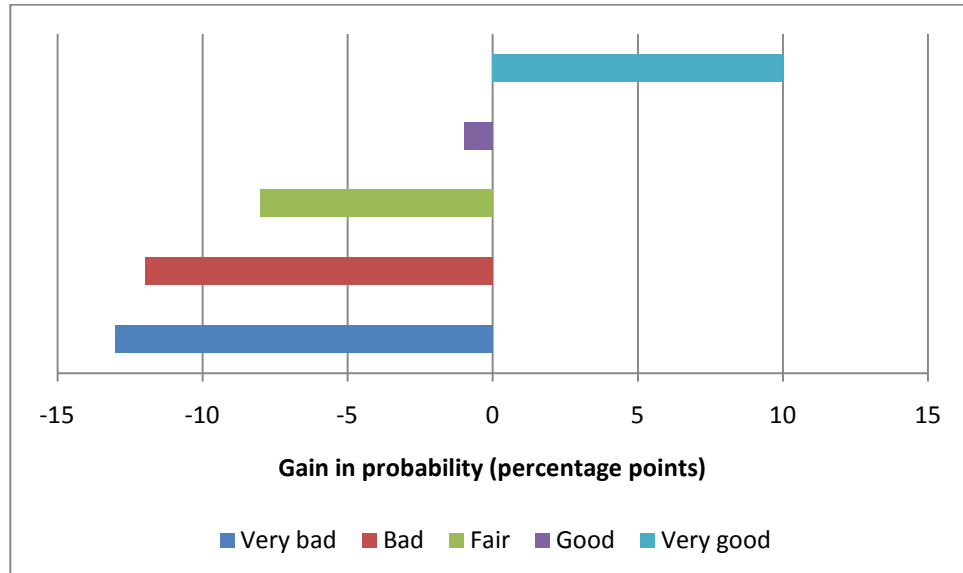


Chart 2.3 Gain in probability – age cohort

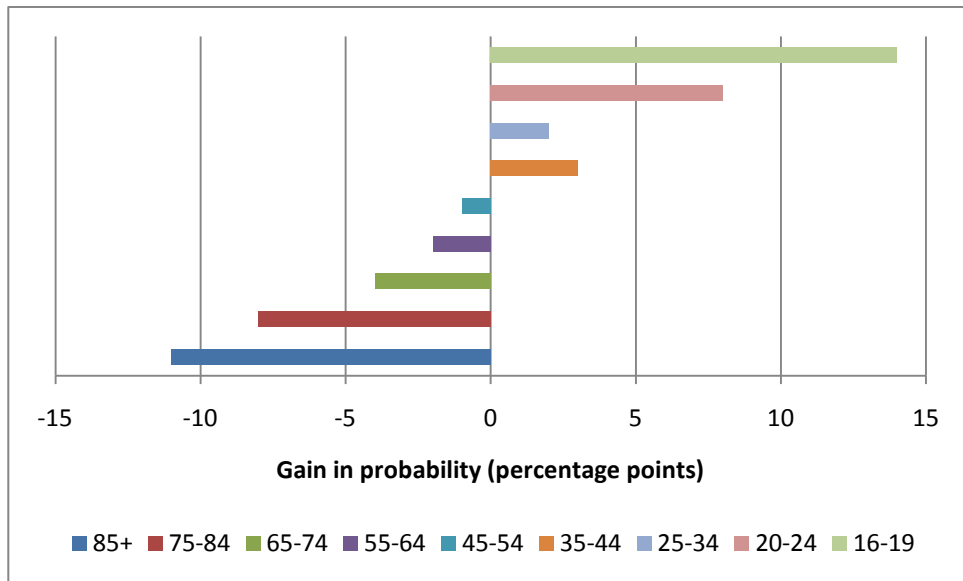


Chart 2.4 Gain in probability – educational level

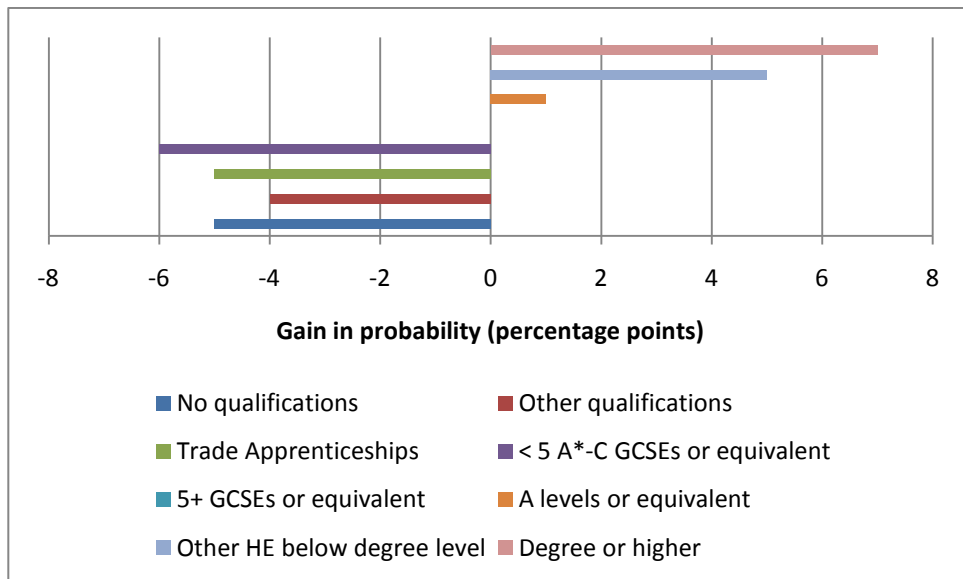


Chart 2.5 Gain in probability – sex

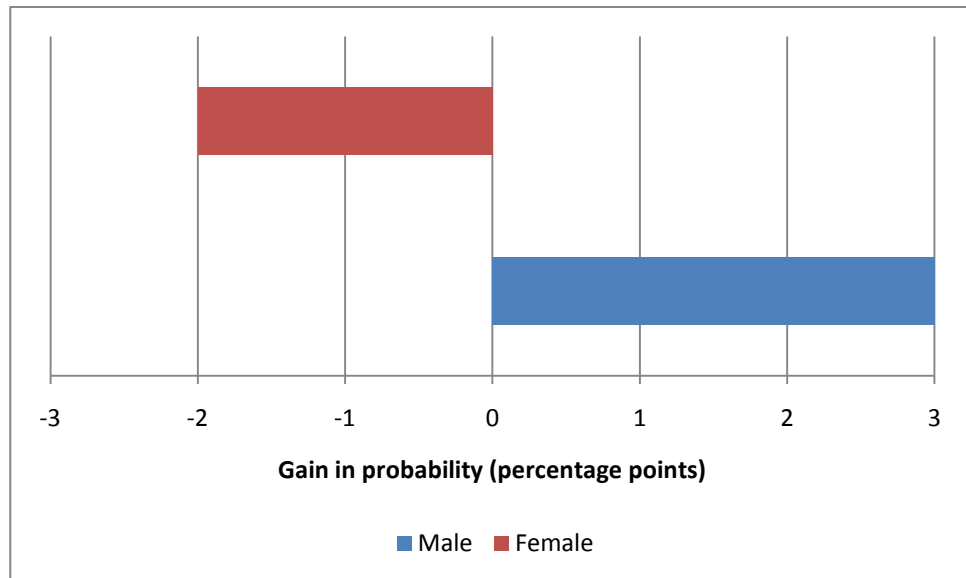
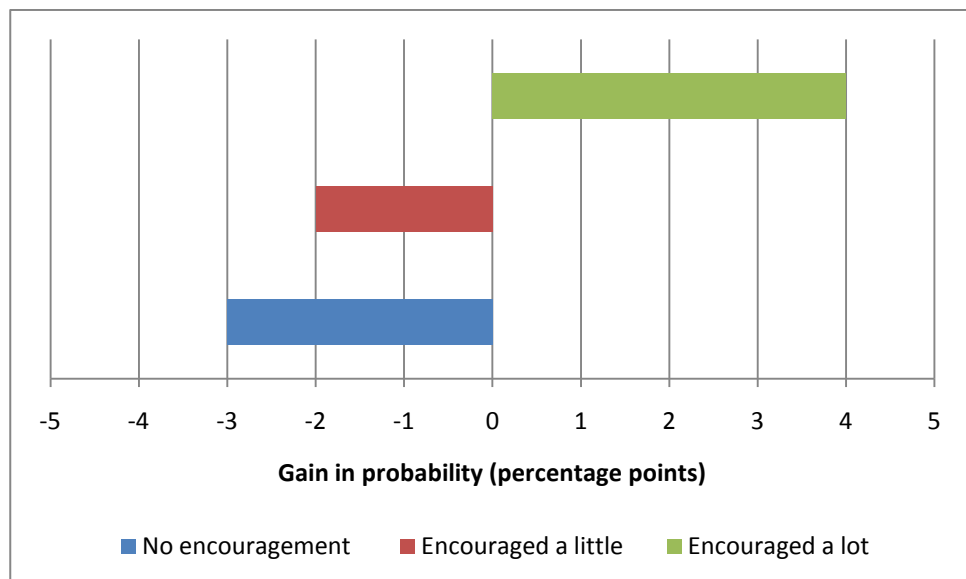


Chart 2.6 Gain in probability – parental encouragement to do sport



Current activities

The dataset was split into three 'base likelihood' groups based on the probability value generated by the logistic regression version of this model:

- Relatively *high* probability of participating in 30 minutes of moderate intensity level sport and active recreation on three or more days in the past week = >29%
- *Mid* level probability of meeting participating at this level = 15%-29%
- Relatively *low* probability of participating at this level = <15%

Despite the weak predictive power of the model based only on background characteristics, the addition of current activity variables added very little. The new models explained only an additional 2-5% of the observed variety in activity levels.

Only four current activity variables proved significantly predictive:

- Volume of TV viewing
- Newspaper readership [mid-level group only]
- Holiday activity
- Viewing sport websites

The last of these four has an obvious connection with recreational physical activity and sport and the two may have an interdependent relationship. Nevertheless, watching sport programmes on TV was *not* a significant predictor of sporting activity, illustrating how 'obvious' connections are not always observed.

It should be noted here that the master variable set is the same for each of the three base likelihood groups. Only the relative influence of each variable differs.

The positive relationships were with low volume TV viewing, reading broadsheet newspapers, foreign or frequent holidays, and viewing sport content on the web.

Although these relationships hold for all three base likelihood groups, the relative influence varies. A simple way of summarising this is to compute the simple standard deviation of the MCA-computed penalties and rewards for each variable.

This showed that penalties and rewards were strongest for the high base likelihood group on three of the four variables. The exception was newspaper readership where penalties and rewards were strongest for the middle base likelihood group.

We can do a similar calculation to find the specific characteristics that best distinguish the three base likelihood groups. The five most distinguishing features are:

Not viewing sport websites

(**penalty** = zero for low base likelihood group, -2PP for middle base likelihood group and -8PP for the high base likelihood group)

Reading a non-mainstream daily newspaper

(**penalty** = -4PP for low base likelihood group, -6PP for middle base likelihood group and -10PP for the high base likelihood group)

Watching four hours or more TV per day

(**penalty** = -2PP for low base likelihood group, -4PP for middle base likelihood group and -6PP for the high base likelihood group)

Watching less than one hour of TV per day

(**reward** = +3PP for low base likelihood group, +3PP for middle base likelihood group and +6PP for the high base likelihood group)

Reading *The Times*

(**reward** = +5PP for low base likelihood group, +5PP for middle base likelihood group and +2PP for the high base likelihood group)

The scores in brackets show the expected increase or decrease in the proportion of people participating in 30 minutes of moderate intensity level sport and active recreation on three or more days in the past week if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty' or 'reward'.

Tables 2.1 to 2.4 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Table 2.1 Percentage point gain in probability – accessing sport websites

	LOW	MIDDLE	HIGH
Accesses sport websites	+7	+6	+9
Does not access sport websites	0	-2	-8

Table 2.2 Percentage point gain in probability – newspaper readership

	LOW	MIDDLE	HIGH
Financial Times	*	*	+4
The Times	+5	+5	+2
The Guardian	+2	+4	+1
Daily Mail	+1	+4	+1
Metro	+1	+3	+1
Daily Express	0	+3	+1
Daily Mirror	0	+1	0
Local daily newspaper	0	-1	0
The Sun	-1	-2	-1
Daily Telegraph	-1	-2	-3
The Independent	-1	-4	-3
Daily Star	-4	-4	-3
Other daily newspaper	-4	-6	-10
None	-1	-2	0

* = numbers too small to report

Table 2.3 Percentage point gain in probability – daily hours of television viewing

	LOW	MIDDLE	HIGH
About 5 hours or more	-2	-4	-6
About 4 hours	-1	-2	-5
About 3 hours	+1	-2	-3
About 2 hours	+2	+3	-1
About 1 hour	+4	+3	+5
Less than 1 hour	+2	+3	+6
Never watch TV	+5	+3	+6

Table 2.4 Percentage point gain in probability – holiday activity

	LOW	MIDDLE	HIGH
Holidays in UK and abroad	+3	+3	+5
Holidays abroad only	+3	0	+1
Holidays UK only	0	-3	-4
No holidays	-2	-1	-3

Chapter 3: Modelling arts activity

Background characteristics

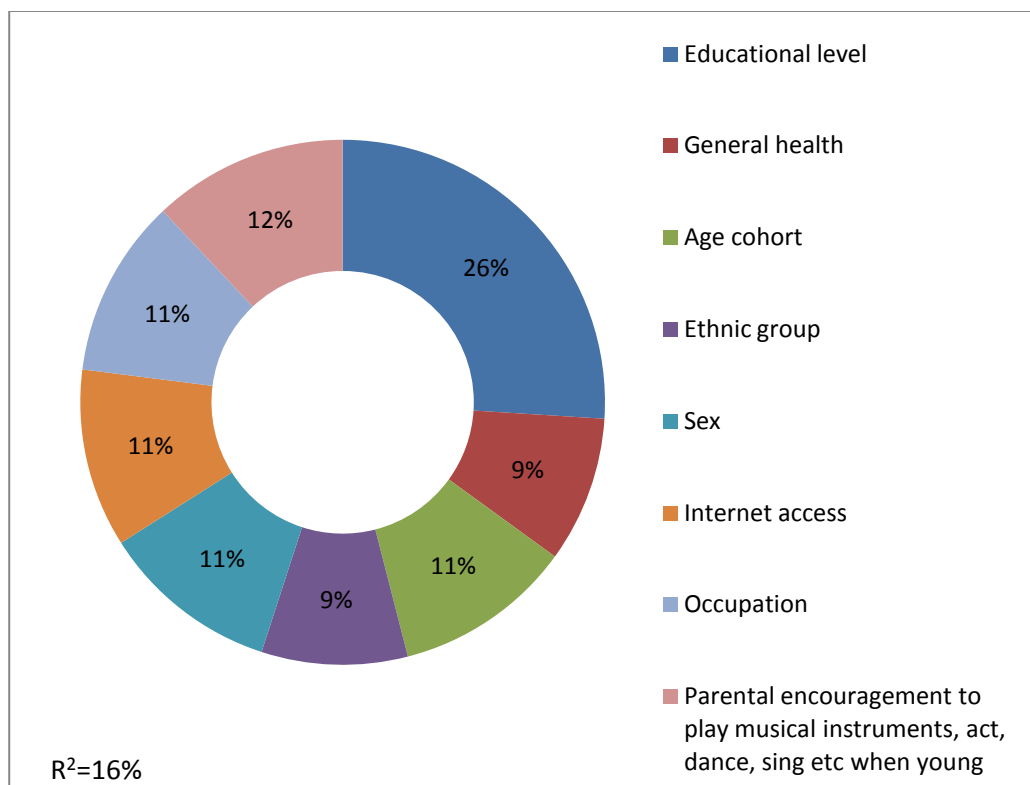
The model for arts activity was moderately successful. It is estimated that the stage 1 'background characteristics' model explains only 16% of the observed variety in activity levels^{ix}

Eight background characteristics proved significantly predictive:

- Educational level
- Level of parental encouragement when young
- Age cohort
- Sex
- Internet access
- Occupation
- General health
- Ethnic group

Chart 3.1 shows the relative weight of each of these variables in terms of independent predictive power. The chart sums to 100% where 100% equals the predictive power of the whole model.

Chart 3.1 Predictors of arts activity – background characteristics



As shown in Chart 3.1, the predictors of arts participation were: a high educational level, strong parental encouragement to participate in the arts when young, late middle age (55-74) and youth (16-19), internet access, a professional/managerial occupation, good general health, and a White or Mixed Race ethnicity. Women were also more likely than men to participate in the arts, even when holding the other characteristics constant.

The three most *positive* characteristics are:

- Educated to degree level (+16PP)
- Parents encouraged individual a lot when young (+11PP)
- Aged 55-74 (+5PP if aged 55-74, +7PP if aged 65-74) or 16-19 (+6PP)

The scores in brackets show the expected increase in the proportion participating in the arts if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'reward'.

The three most *negative* characteristics are:

- No qualifications (-14PP)
- Bad health or worse (-9PP if 'bad', -20PP if 'very bad')
- Asian ethnic group (-13PP)

The scores in brackets show the expected decrease in the proportion participating in the arts if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty'.

Charts 3.2 to 3.6 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Chart 3.2 Gain in probability – educational level

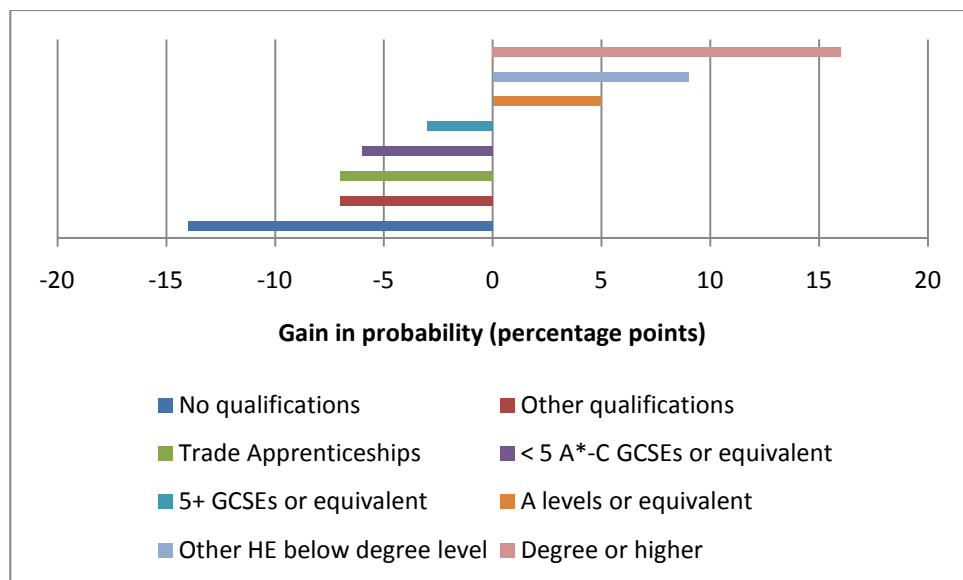


Chart 3.3 Gain in probability – parental encouragement towards performing arts when young

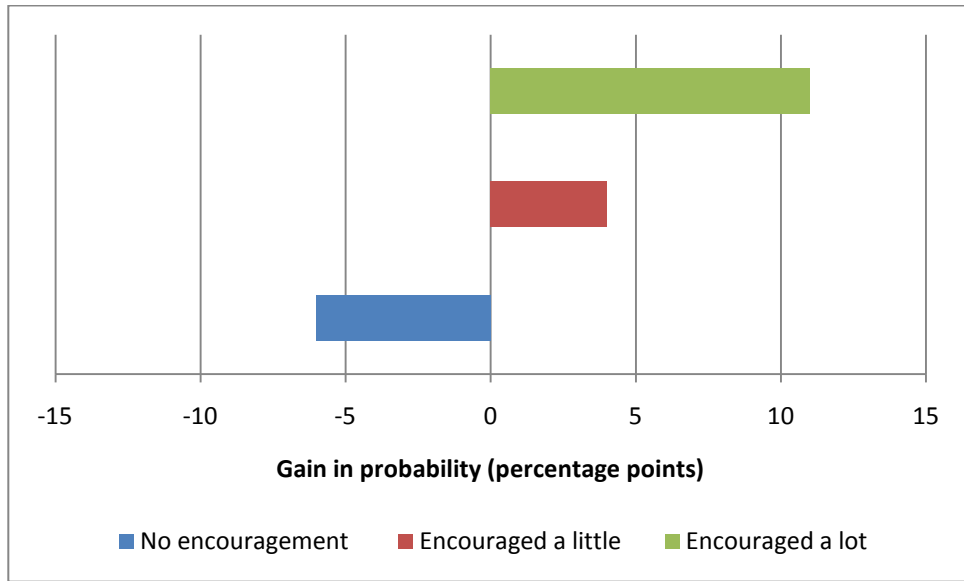


Chart 3.4 Gain in probability – age cohort

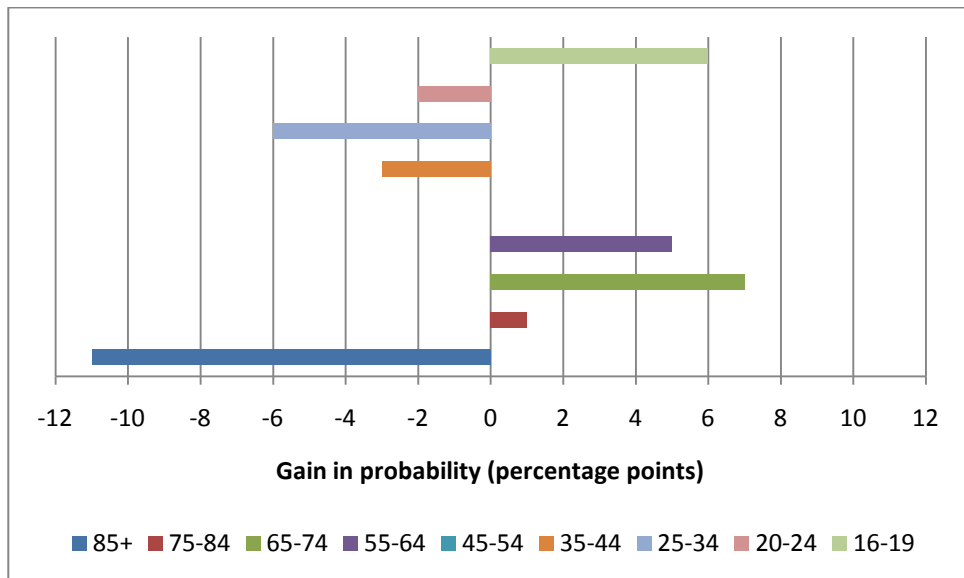


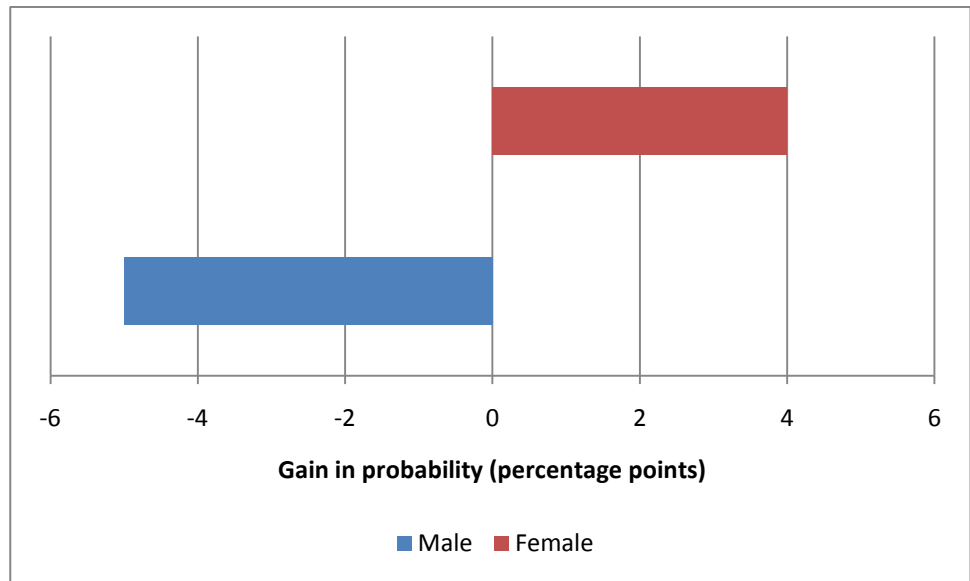
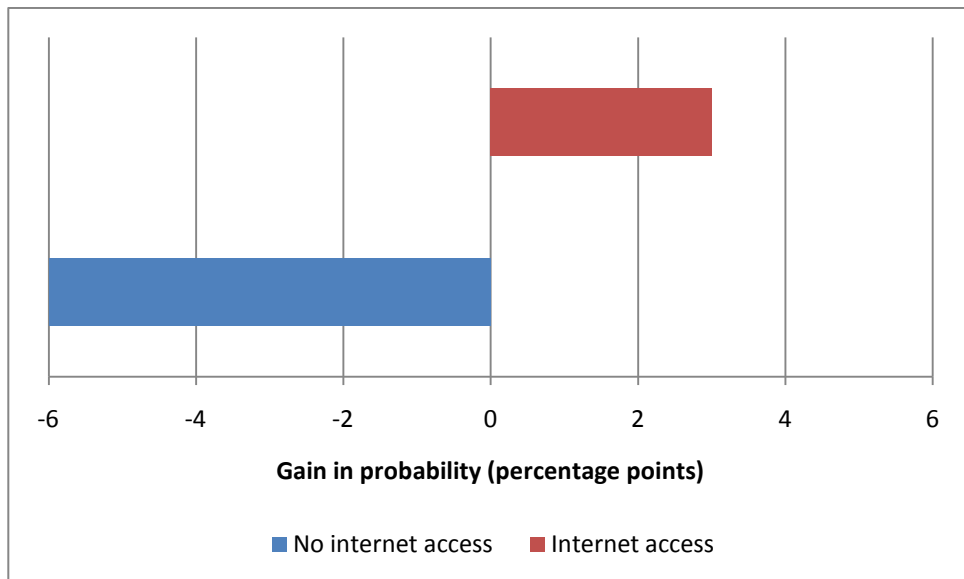
Chart 3.5 Gain in probability – sex**Chart 3.6 Gain in probability – internet access**

Chart 3.7 Gain in probability – occupation

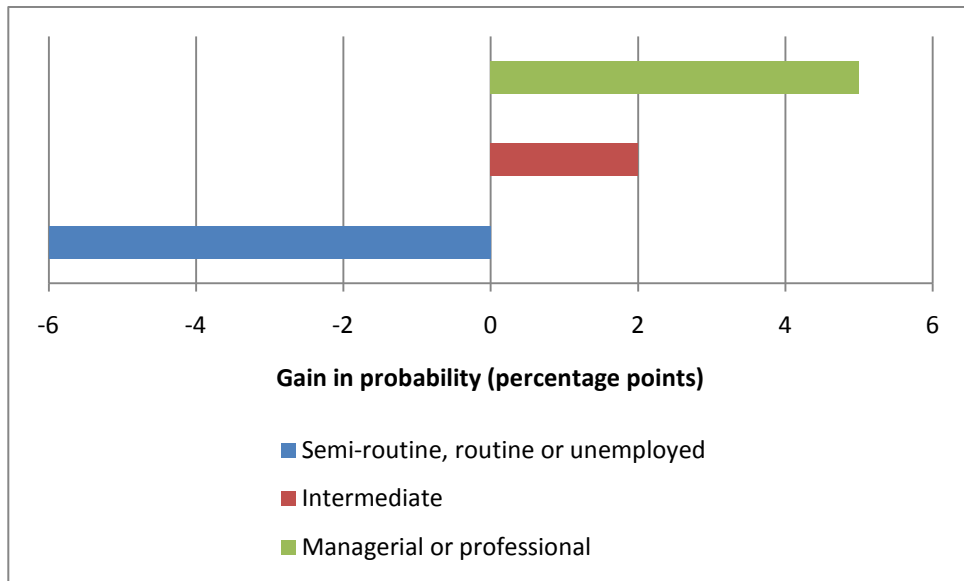


Chart 3.8 Gain in probability – general health

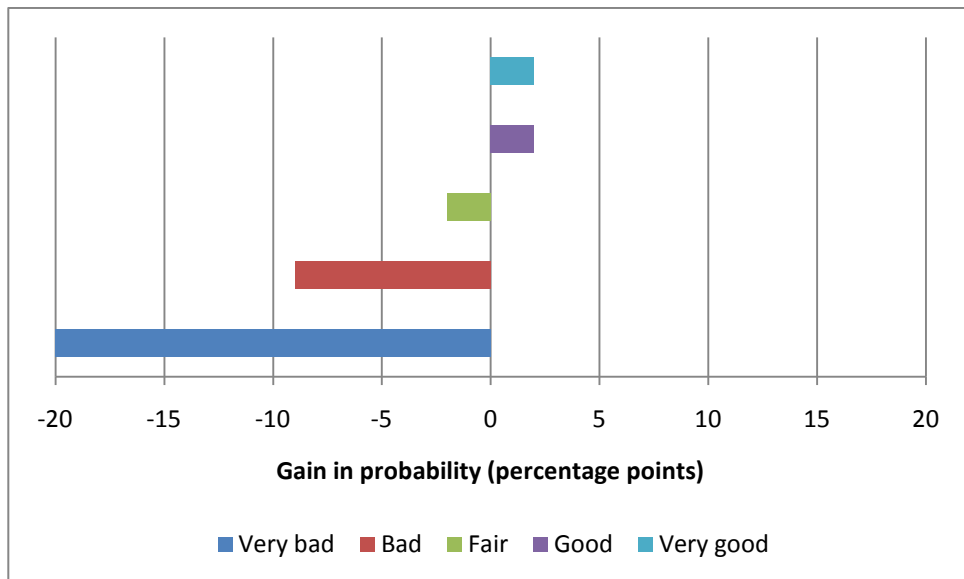
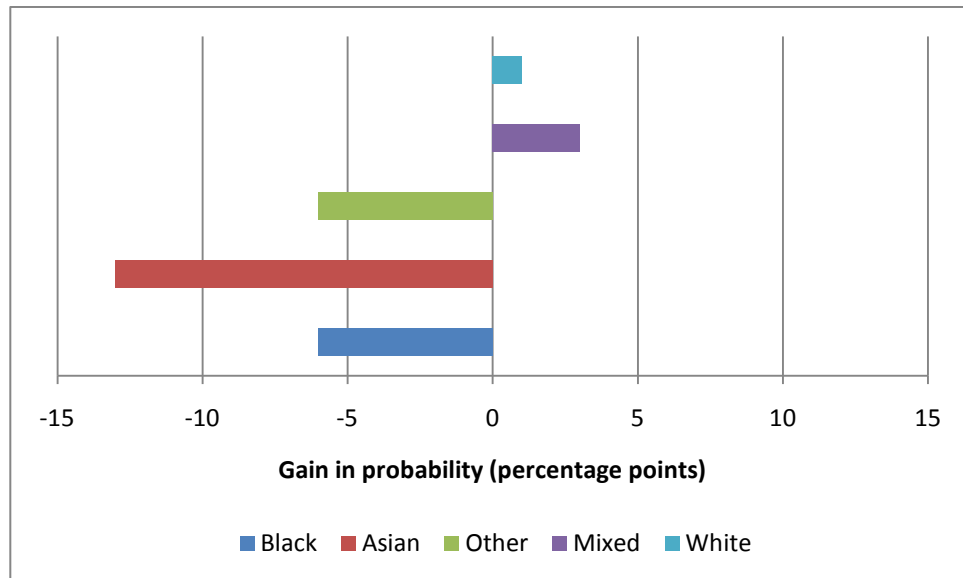


Chart 3.9 Gain in probability – ethnicity



Current activities

The dataset was split into three 'base likelihood' groups based on the probability value generated by the logistic regression version of this model:

- Relatively *high* probability of participating in the arts at least three times in the last 12 months = >72%
- *Mid* level probability of participating at this level = 51%-72%
- Relatively *low* probability of participating at this level = <51%

Despite the moderate predictive power of the model based only on background characteristics, the addition of current activity variables did not add a great deal. The new models explained an additional 10-13% of the observed variety in activity levels.

Nine current activity variables proved significantly predictive:

- Newspaper readership
- Volume of TV viewing
- Watching arts programmes
- Watching history programmes
- Watching drama programmes
- Accessing theatre/concert websites
- Accessing museum/gallery websites
- Holiday activity
- Voluntary work

Some of these variables (e.g. watching drama programmes and accessing theatre/concert websites) have an obvious connection with arts activity and may have an interdependent relationship with it.

It should be noted here that the master variable set is the same for each of the three base likelihood groups. Only the relative influence of each variable differs.

The positive relationships were with reading broadsheet newspapers, low volume TV viewing but also watching arts, history or drama programmes, foreign or frequent holidays, doing voluntary work and accessing relevant websites (museums, galleries, theatre venues etc.).

Although these relationships held for all three base likelihood groups, the relative influence varied. A simple way of summarising this is to compute the simple standard deviation of the MCA-computed penalties and rewards for each variable.

This showed that penalties and rewards were generally strongest for the low base likelihood group (outright for six of the nine variables and joint top for another two). The exception was volume of TV viewing, where the relationship was strongest with the middle base likelihood group.

We can do a similar calculation to find the specific characteristics that best distinguish the three base likelihood groups. The five most distinguishing features are:

Accessing theatre/concert websites

(**reward** = +22PP low base likelihood group, +14PP for middle base likelihood group and +6PP for the high base likelihood group)

Reading the *Daily Mirror* / non-mainstream newspaper

(**penalty** = zero for low base likelihood group, -2PP/-5PP for middle base likelihood group and -12PP/-7PP for the high base likelihood group)

Holidays at home and abroad

(**reward** = +16PP for low base likelihood group, +7PP for middle base likelihood group and +4PP for the high base likelihood group)

Watching arts programmes

(**reward** = +17PP for low base likelihood group, +12PP for middle base likelihood group and +6PP for the high base likelihood group)

Doing voluntary work

(**reward** = +15PP for low base likelihood group, +11PP for middle base likelihood group and +6PP for the high base likelihood group)

The scores in brackets show the expected increase or decrease in the proportion of people participating in the arts at least three times in the last 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty' or 'reward'.

Tables 3.1 to 3.9 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Table 3.1 Percentage point gain in probability – newspaper readership

	LOW	MIDDLE	HIGH
The Independent	*	+4	+5
Financial Times	*	*	-1
The Guardian	+12	+9	+6
Metro	+9	+5	+1
Daily Telegraph	+4	+4	+2
The Times	+3	+7	+2
Local daily newspaper	+3	+1	0
Daily Mail	+1	+2	+2
Daily Express	+1	+2	-3

Daily Mirror	0	-2	-12
The Sun	-3	-7	-13
Daily Star	-8	-2	*
Other daily newspaper	0	-5	7
None	-1	-1	0

* = numbers too small to report

Table 3.2 Percentage point gain in probability – daily hours of television viewing

	LOW	MIDDLE	HIGH
About 5 hours or more	-5	-6	-6
About 4 hours	0	-1	-1
About 3 hours	0	-1	-1
About 2 hours	+6	+1	0
About 1 hour	+2	+2	0
Less than 1 hour	+2	+5	+2
Never watch TV	+7	+15	+9

Table 3.3 Percentage point gain in probability – watching arts programmes

	LOW	MIDDLE	HIGH
Watches arts programmes	+17	+12	+6
Does not watch arts programmes	-1	-1	-1

Table 3.4 Percentage point gain in probability – watching history programmes

	LOW	MIDDLE	HIGH
Watches history programmes	+5	+5	+2
Does not watch history programmes	-3	-3	-2

Table 3.5 Percentage point gain in probability – watching drama programmes

	LOW	MIDDLE	HIGH
Watches drama programmes	+5	+5	+3
Does not watch drama programmes	-2	-2	-2

Table 3.6 Percentage point gain in probability – accessing theatre/concert websites

	LOW	MIDDLE	HIGH
Accesses theatre/concert websites	+22	+14	+6
Does not access theatre/concert websites	-2	-6	-8

Table 3.7 Percentage point gain in probability – accessing museum/gallery websites

	LOW	MIDDLE	HIGH
Accesses museum/gallery websites	+14	+8	+4
Does not access museum/gallery websites	0	-1	-2

Table 3.8 Percentage point gain in probability – holiday activity

	LOW	MIDDLE	HIGH
Holidays in UK and abroad	+16	+7	+4
Holidays abroad only	+3	0	-1
Holidays UK only	+4	+1	+1
No holidays	-4	-3	-4

Table 3.9 Percentage point gain in probability – voluntary work

	LOW	MIDDLE	HIGH
Does voluntary work	+15	+11	+6
Does no voluntary work	-2	-3	-3

Chapter 4: Modelling museum and gallery visits

Background characteristics

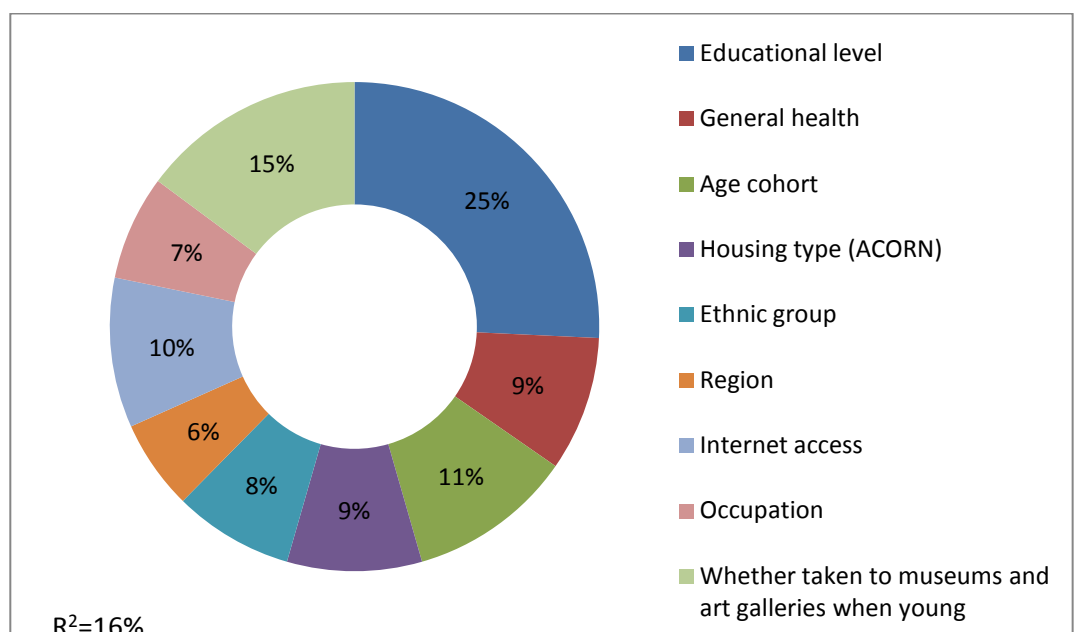
The predictive model for museum and gallery visits was only moderately successful. It is estimated that the stage 1 'background characteristics' model explains 16% of the observed variety in activity levels^x

Nine background characteristics proved significantly predictive:

- Educational level
- Frequency of visiting museums/galleries when young
- Age cohort
- Internet access
- Housing type (ACORN classification)
- General health
- Ethnic group
- Occupation
- Region

Chart 4.1 shows the relative weight of each of these variables in terms of independent predictive power. The chart sums to 100% where 100% equals the predictive power of the whole model.

Chart 4.1 Predictors of museum and gallery visits – background characteristics



As shown in Chart 4.1, the predictors of museum and gallery visits were: high educational level, frequent visits to museums or galleries when young, late middle age (55-74), internet access, the 'urban prosperity' ACORN segment, good general health, non-Black/Asian ethnic group, a professional/managerial occupation, and a location in the North or in London.

The three most *positive* characteristics are:

- Taken to museums/galleries at least 3-4 times a year when young (+19PP if at least once a month, +16PP otherwise)
- Educated to degree level (+16PP)
- 'Urban prosperity' ACORN segment (+9PP)

The scores in brackets show the expected increase in the proportion attending a museum or gallery at least once in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'reward'.

The three most *negative* characteristics are:

- No qualifications (-13PP)
- Black or Asian ethnic group (-14PP if Black, -8PP if Asian)
- Bad health or worse (-9PP if 'bad', -15PP if 'very bad')

The scores in brackets show the expected decrease in the proportion attending a museum or gallery at least once in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty'.

Charts 4.2 to 4.10 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Chart 4.2 Gain in probability – educational level

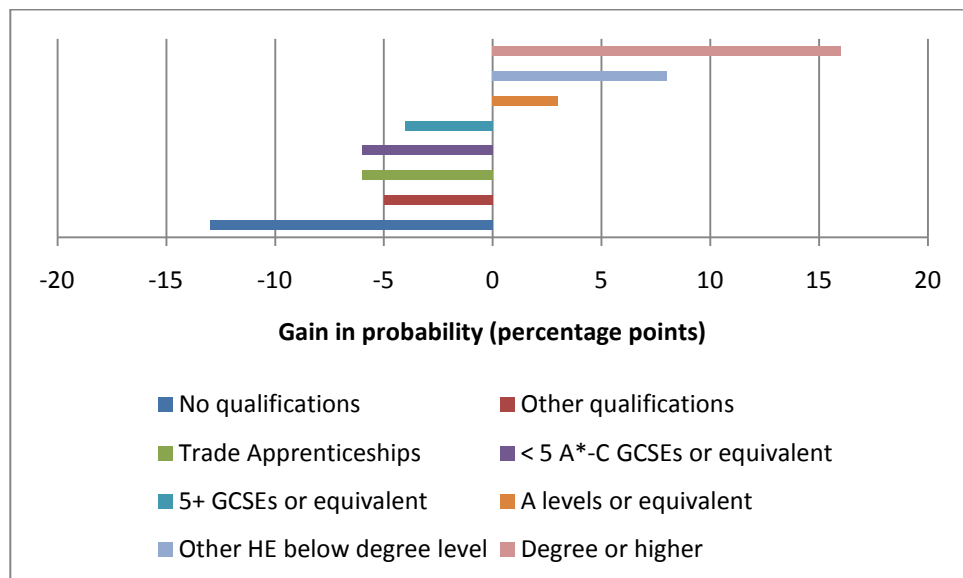


Chart 4.3 Gain in probability – whether taken to museums or art galleries when young

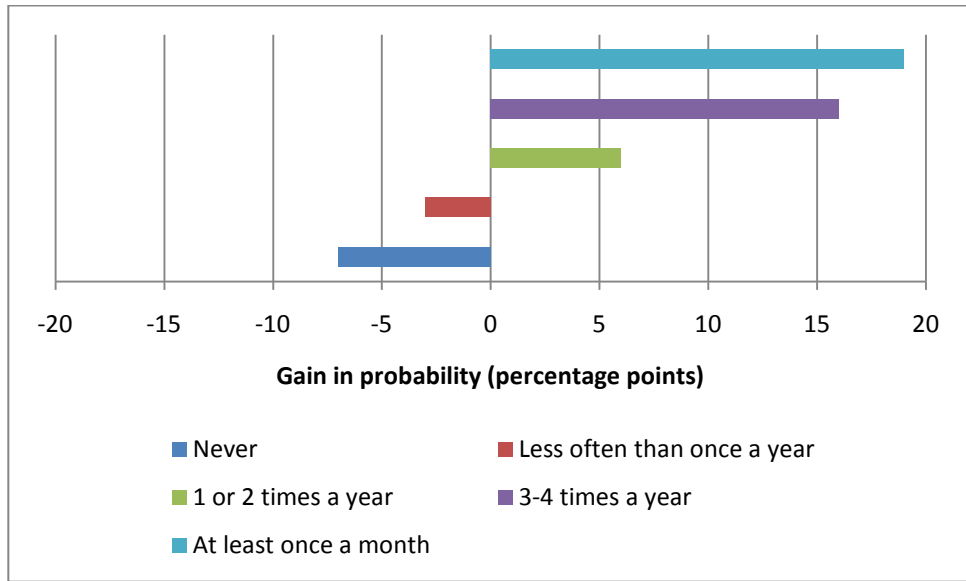


Chart 4.4 Gain in probability – age cohort

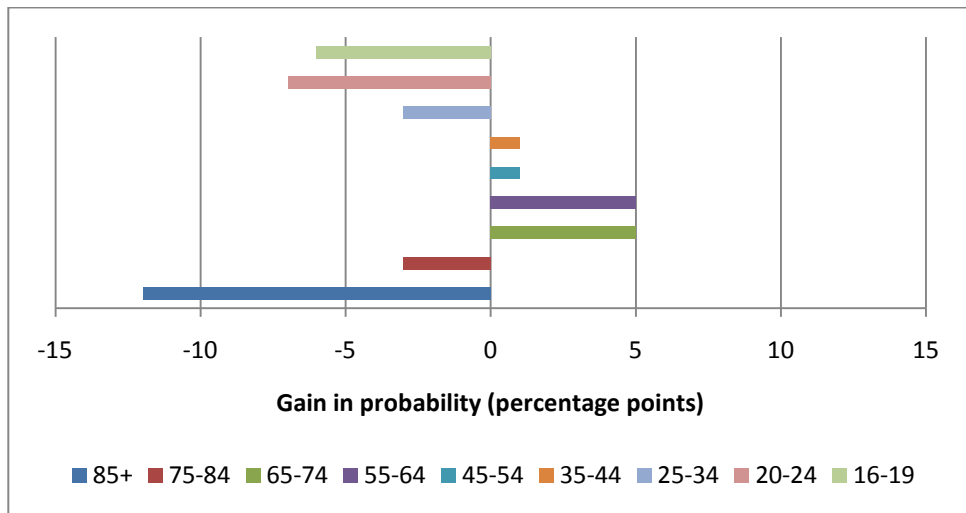


Chart 4.5 Gain in probability – internet access

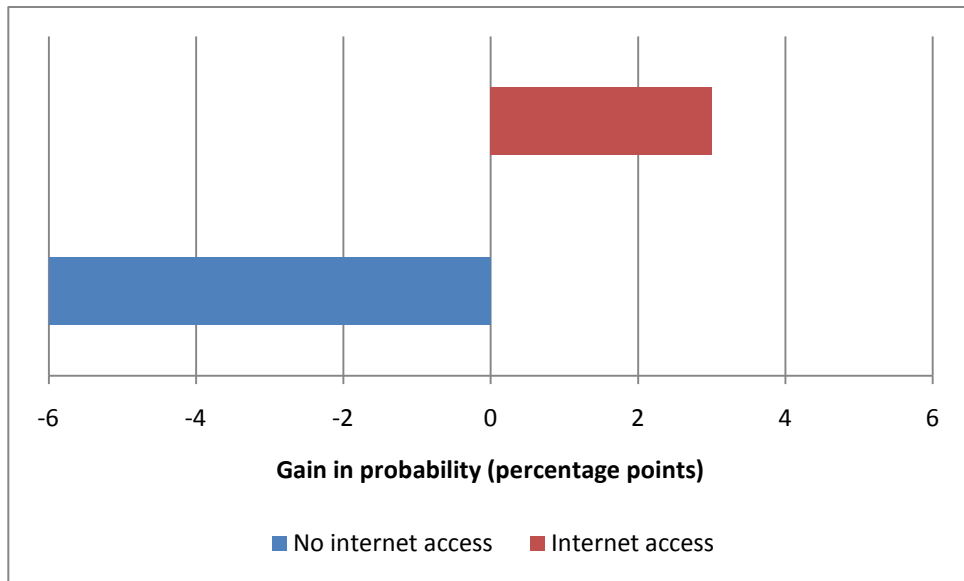


Chart 4.6 Gain in probability – ACORN classification

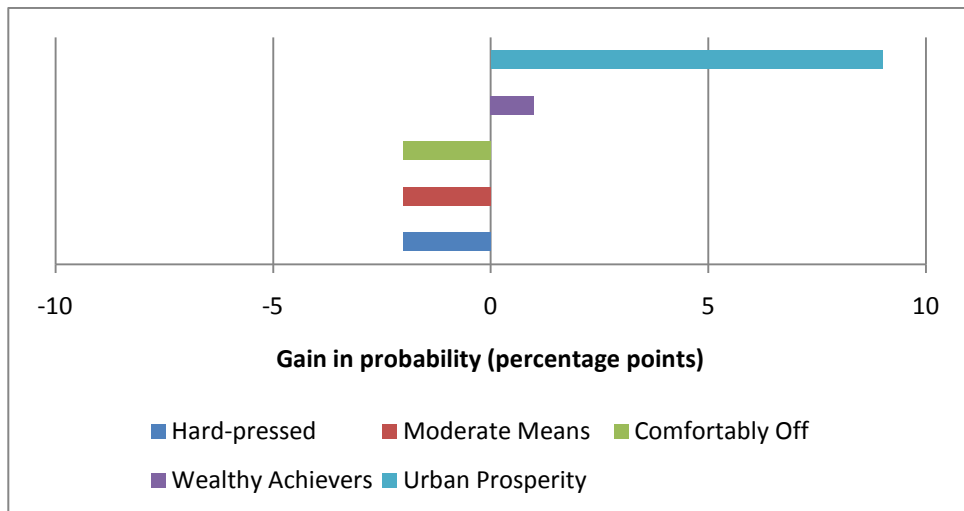


Chart 4.7 Gain in probability – general health

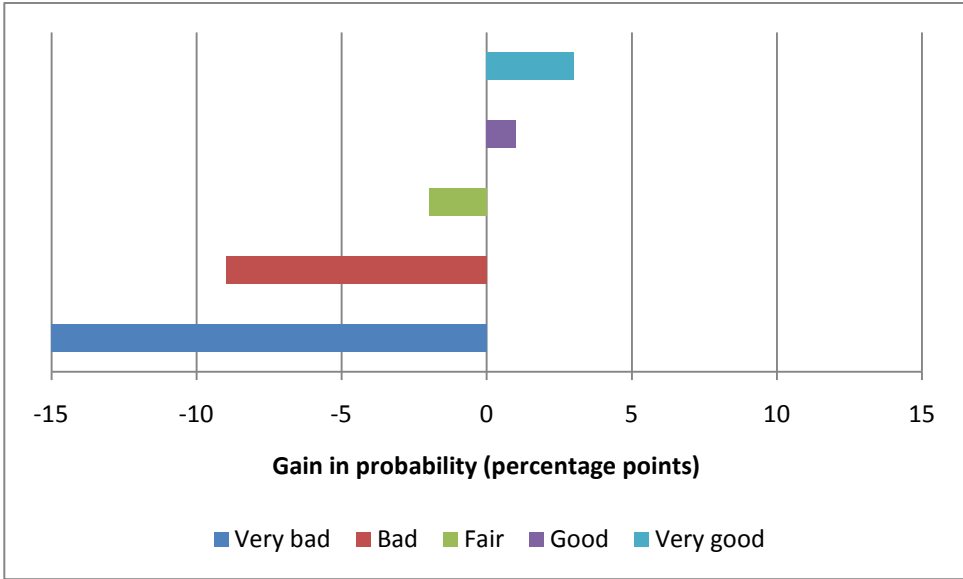
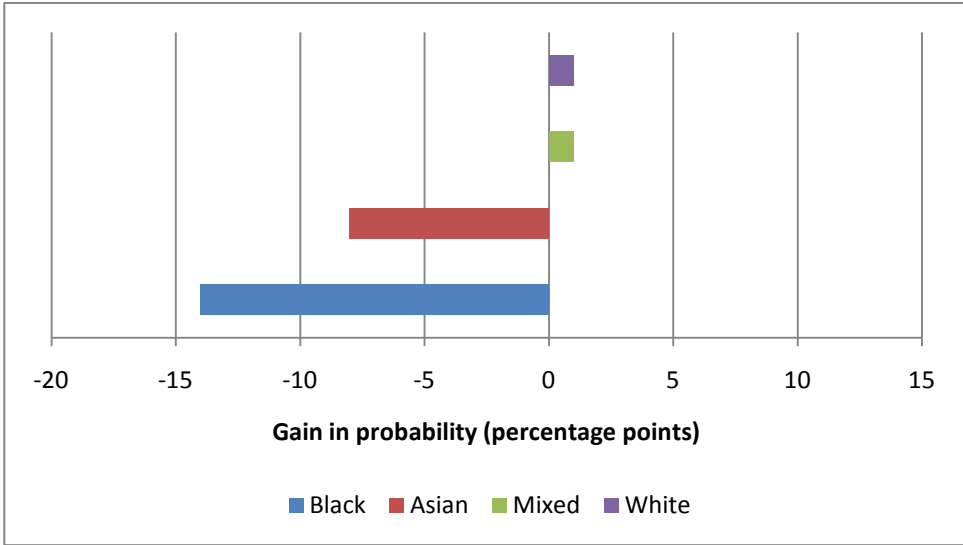
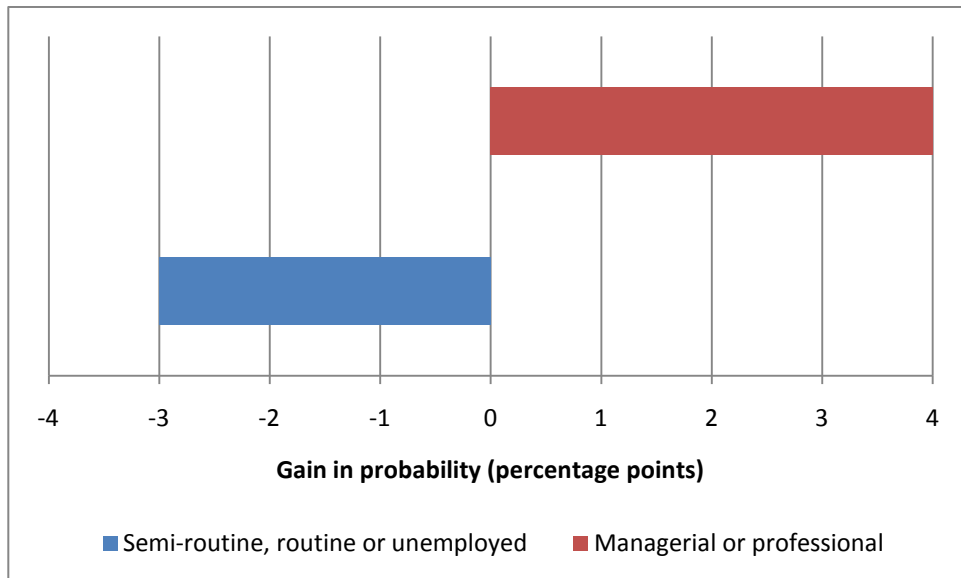


Chart 4.8 Gain in probability – ethnicity



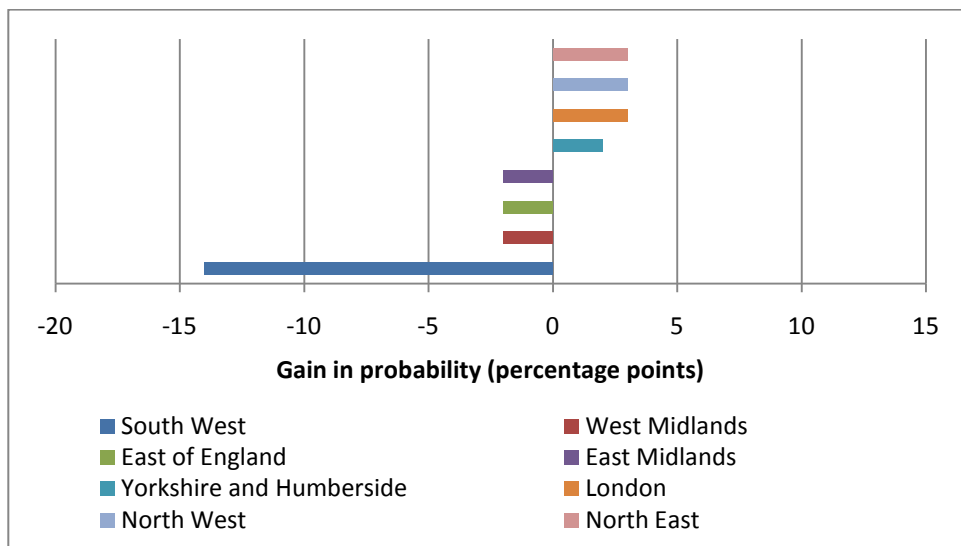
Note: no gain for 'other' ethnicity, so excluded from chart

Chart 4.9 Gain in probability – occupation



Note: no gain for immediate occupation, so excluded from chart

Chart 4.10 Gain in probability – English region



Note: no gain for the South East, so excluded from chart

Current activities

The dataset was split into three ‘base likelihood’ groups based on the probability value generated by the logistic regression version of this model:

- Relatively *high* probability of meeting PSA criteria = >55%
- *Mid* level probability of meeting PSA criteria = 33%-55%
- Relatively *low* probability of meeting PSA criteria = <33%

The addition of current activity variables improved the model considerably although the magnitude of the improvement varied between base likelihood groups. It was most effective for the high base likelihood group (the new model explained an additional 23% of the observed variety in activity levels) and least effective for the low base likelihood group (+13%).

Nine current activity variables proved significantly predictive:

- Accessing museum or gallery websites
- Holiday activity
- Arts activity
- Newspaper readership
- Watching history programmes
- Watching arts programmes
- Buying books
- Volume of TV viewing
- Voluntary work

Some of these variables (e.g. watching history programmes and accessing museum or gallery websites) have an obvious connection with museum and gallery visits and may have an interdependent relationship with it.

It should be noted here that the master variable set is the same for each of the three base likelihood groups. Only the relative influence of each variable differs.

The positive relationships were: accessing relevant websites, foreign or frequent holidays, a moderate to high level of arts activity, reading broadsheet newspapers, low volume TV viewing but also watching history or arts programmes, frequent book purchases, and doing voluntary work.

Although these relationships hold for all three base likelihood groups, the relative influence varies. A simple way of summarising this is to compute the simple standard deviation of the MCA-computed penalties and rewards for each variable.

This showed that the strongest penalties and rewards were shared out between the three base likelihood groups. For some variables, the strongest relationship was shown with the high base likelihood group; for others the strongest relationship was with the middle or low base likelihood groups.

We can do a similar calculation to find the specific characteristics that best distinguish the three base likelihood groups. The five most distinguishing features are:

Reading *The Times*

(**reward** = +22PP for low base likelihood group, +7PP for middle base likelihood group and +4PP for the high base likelihood group)

Accessing museum or gallery websites

(**reward** = +35PP low base likelihood group, +29PP for middle base likelihood group and +17PP for the high base likelihood group)

Not doing arts activity

(**penalty** = -6PP for low base likelihood group, -10PP for middle base likelihood group and -15PP for the high base likelihood group)

Reading *The Sun*

(**penalty** = -2PP for low base likelihood group, -6PP for middle base likelihood group and -9PP for the high base likelihood group)

UK Holidays only / a mix of UK/foreign holidays

(**reward** = +5PP/+10PP for low base likelihood group, +4PP/+9PP for middle base likelihood group and a mixture for the high base likelihood group: -2PP if UK holidays only, +6PP if both types)

The scores in brackets show the expected increase or decrease in the proportion attending a museum or gallery at least once in the past 12 months if everybody had

the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty' or 'reward'.

Tables 4.1 to 4.9 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Table 4.1 Percentage point gain in probability – accessing museum/gallery websites

	LOW	MIDDLE	HIGH
Accesses museum/gallery websites	+35	+29	+17
Does not access museum/gallery websites	-1	-5	-9

Table 4.2 Percentage point gain in probability – holiday activity

	LOW	MIDDLE	HIGH
Holidays in UK and abroad	+10	+9	+6
Holidays abroad only	+2	-1	-1
Holidays UK only	+5	+4	-2
No holidays	-4	-5	-6

Table 4.3 Percentage point gain in probability – engaging with the arts

	LOW	MIDDLE	HIGH
Yes	+8	+6	+4
No	-6	-10	-15

Table 4.4 Percentage point gain in probability – newspaper readership

	LOW	MIDDLE	HIGH
The Times	+22	+7	+4
Financial Times	*	*	-1
The Independent	*	+8	+4
Metro	+9	+6	+4
The Guardian	+8	+12	+8
Daily Star	+6	-3	*
Daily Telegraph	+4	+6	+3
Local daily newspaper	+1	+3	0
Daily Mail	+1	-2	-3

Daily Mirror	-1	-2	0
The Sun	-2	-6	-9
Daily Express	-3	-4	-8
Other daily newspaper	0	+3	+5
None	-2	0	-4

* = numbers too small to report

Table 4.5 Percentage point gain in probability – watching history programmes

	LOW	MIDDLE	HIGH
Watches history programmes	+4	+6	+3
Does not watch history programmes	-2	-4	-2

Table 4.6 Percentage point gain in probability – watching arts programmes

	LOW	MIDDLE	HIGH
Watches arts programmes	+8	+7	+8
Does not watch arts programmes	-1	-2	-1

Table 4.7 Percentage point gain in probability – buying books

	LOW	MIDDLE	HIGH
Buys books	+5	+3	+2
Does not buy books	-2	-3	-3

Table 4.8 Percentage point gain in probability – daily hours of television viewing

	LOW	MIDDLE	HIGH
About 5 hours or more	-2	-6	-5
About 4 hours	+1	-2	-4
About 3 hours	0	+1	-2
About 2 hours	+2	0	0
About 1 hour	+1	+4	+1
Less than 1 hour	+1	+4	+4

Never watch TV	+8	+10	+9
----------------	----	-----	----

Table 4.9 Percentage point gain in probability – voluntary work

	LOW	MIDDLE	HIGH
Does voluntary work	+6	+4	+4
Does no voluntary work	-1	-1	-2

Chapter 5: Modelling heritage site visits

Background characteristics

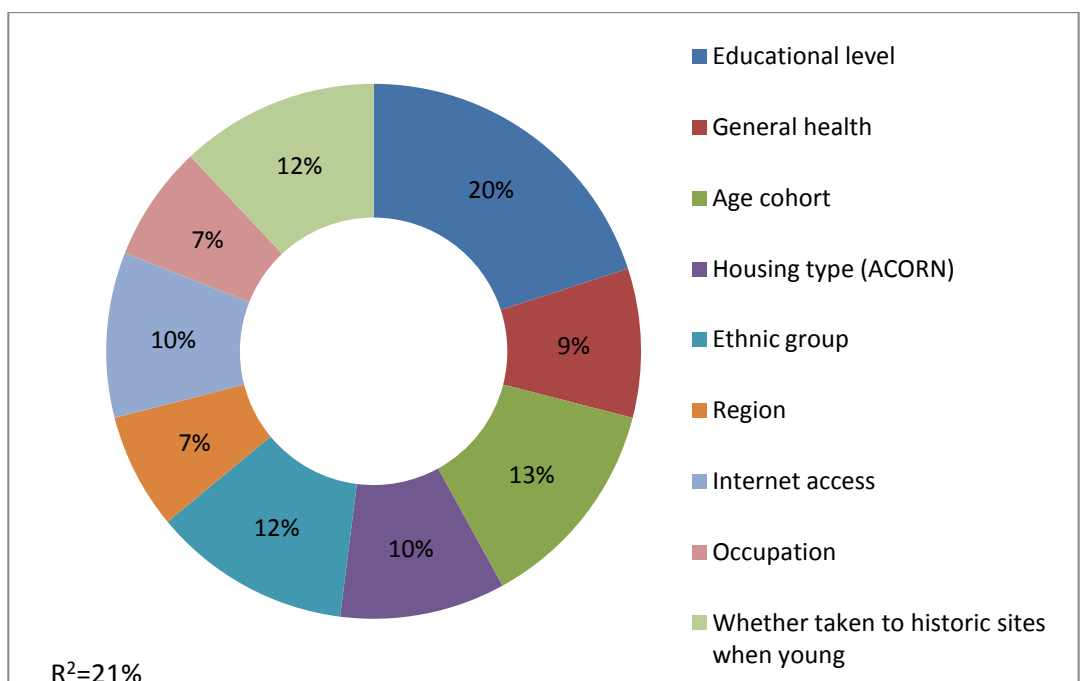
The predictive model for heritage site visits was the most successful of the five, although still not strong by conventional standards. It is estimated that the stage 1 'background characteristics' model explains 21% of the observed variety in activity levels^{xi}.

Nine background characteristics proved significantly predictive:

- Educational level
- Age cohort
- Ethnic group
- Frequency of visiting historic sites when young
- Housing type (ACORN classification)
- Internet access
- General health
- Region
- Occupation

Chart 5.1 shows the relative weight of each of these variables in terms of independent predictive power. The chart sums to 100% where 100% equals the predictive power of the whole model.

Chart 5.1 Predictors of heritage site visits – background characteristics



As shown in Chart 5.1, the predictors of visiting a heritage site were: a high educational level, late middle age (55-74), White ethnicity, frequent visits to historic sites when young, the 'urban prosperity' or 'wealthy achievers' ACORN segment, internet access, good general health, not living in London or the West Midlands, and a professional or managerial occupation.

The three most *positive* characteristics are:

- Taken to historic sites at least 3-4 times a year when young (+17PP if at least once a month, +10PP otherwise)
- Educated to degree level (+13PP)
- Aged 55-74 (+7PP if aged 55-64, +8PP if aged 65-74)

The scores in brackets show the expected increase in the proportion visiting at least two historic environment sites in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'reward'.

The three most *negative* characteristics are:

- Black, Asian or other non-white ethnic group (-21PP if Black, -15PP if Asian, -14PP if other non-white ethnic group)
- Bad health or worse (-11PP if 'bad', -18PP if 'very bad')
- No qualifications (-13PP)

The scores in brackets show the expected decrease in the proportion visiting at least two historic environment sites in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty'.

Charts 5.2 to 5.10 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Chart 5.2 Gain in probability – educational level

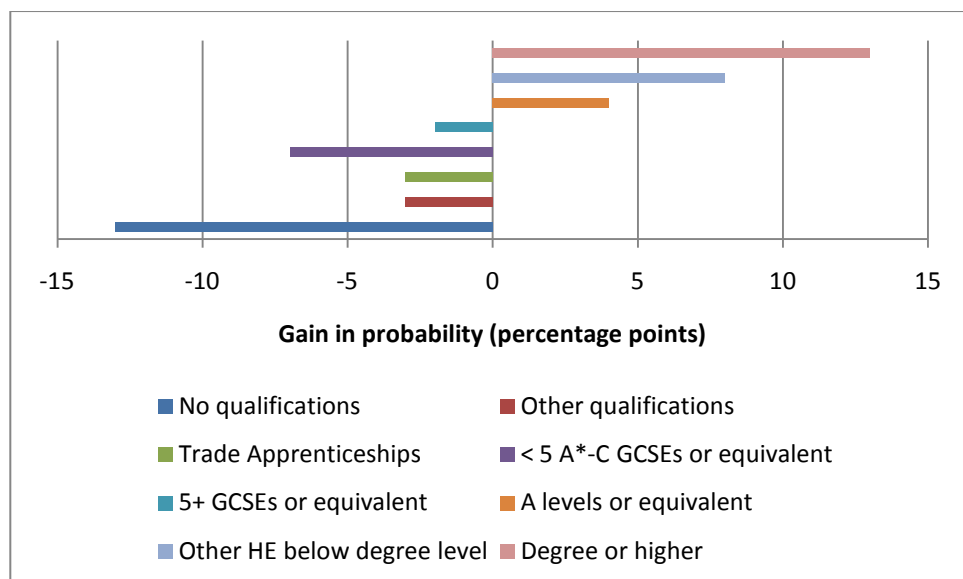
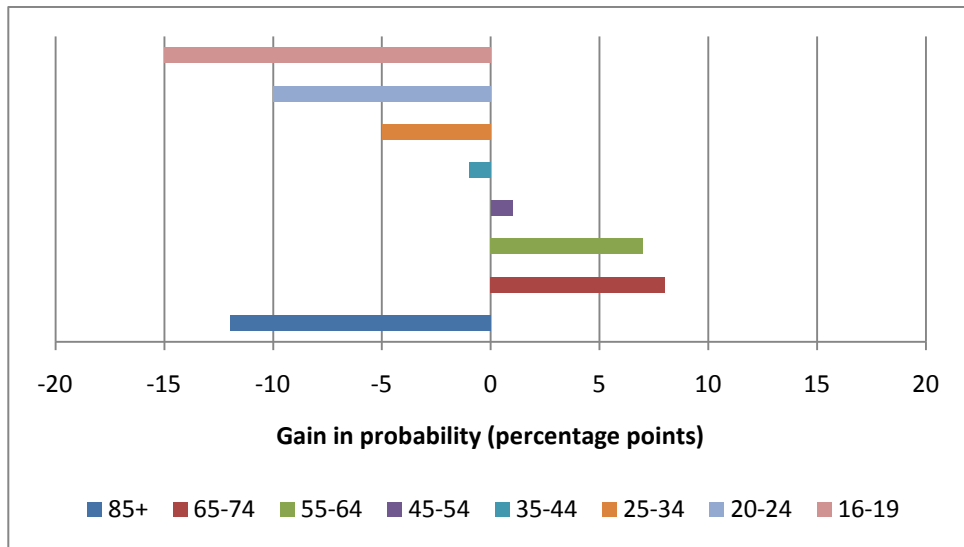


Chart 5.3 Gain in probability – age cohort



Note: no gain 75-84 year olds, so excluded from chart

Chart 5.4 Gain in probability – ethnicity

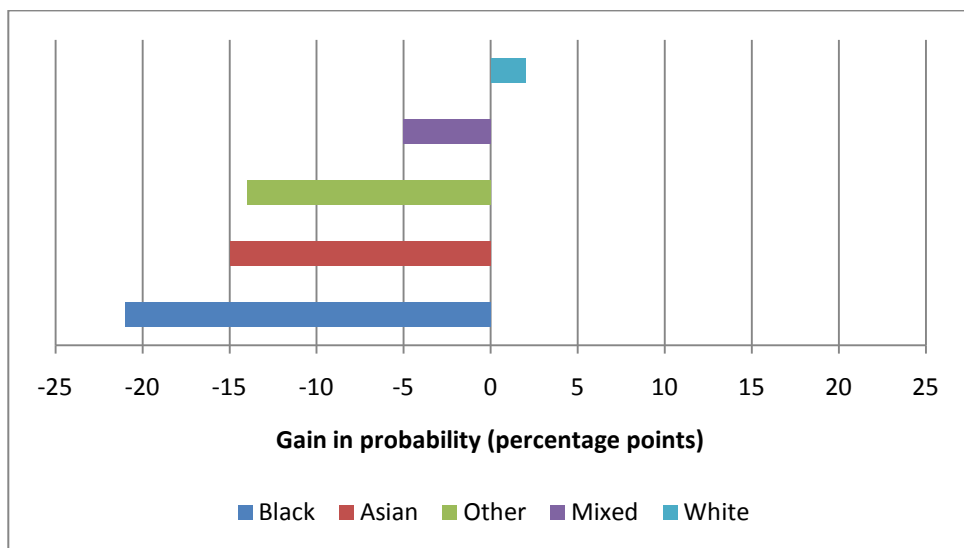


Chart 5.5 Gain in probability – whether taken to historic sites when young

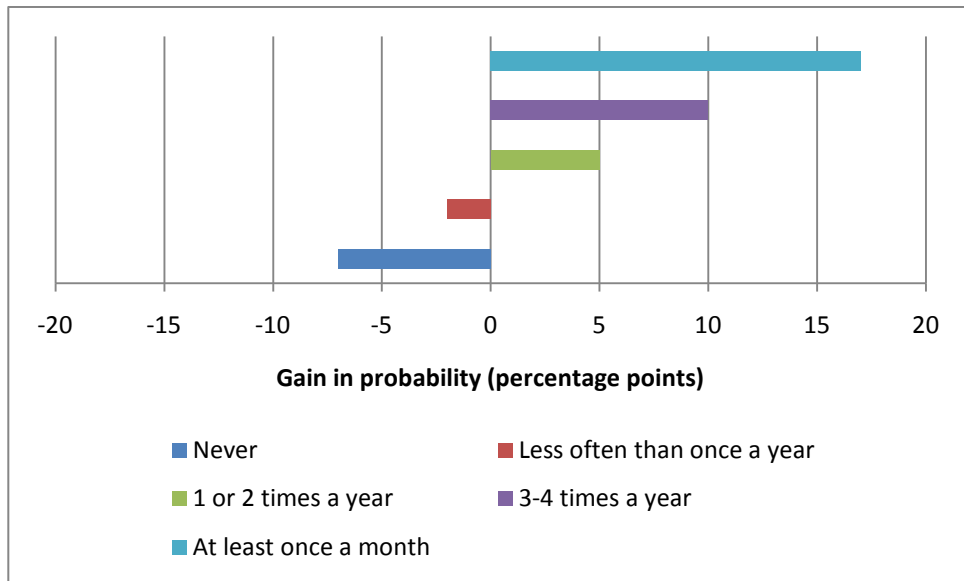


Chart 5.6 Gain in probability – ACORN classification

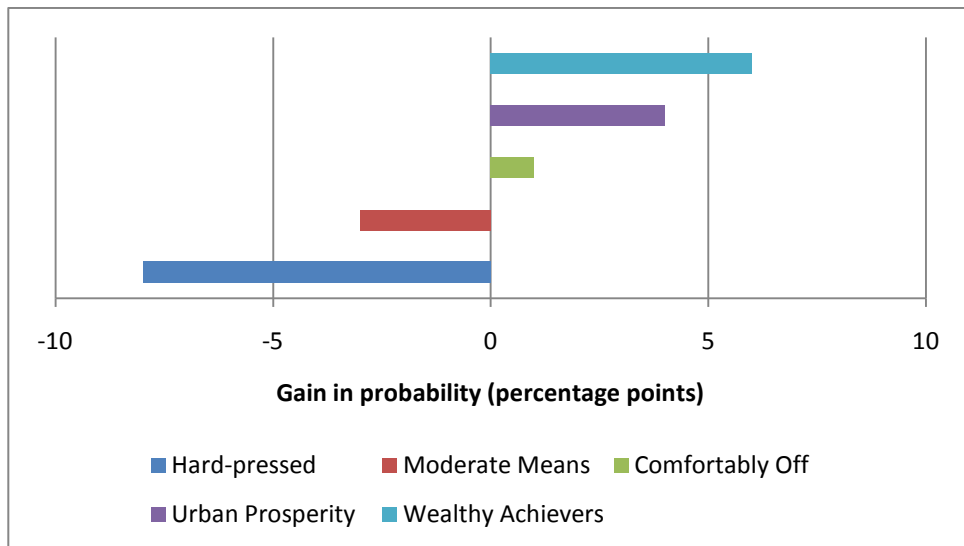


Chart 5.7 Gain in probability – internet access

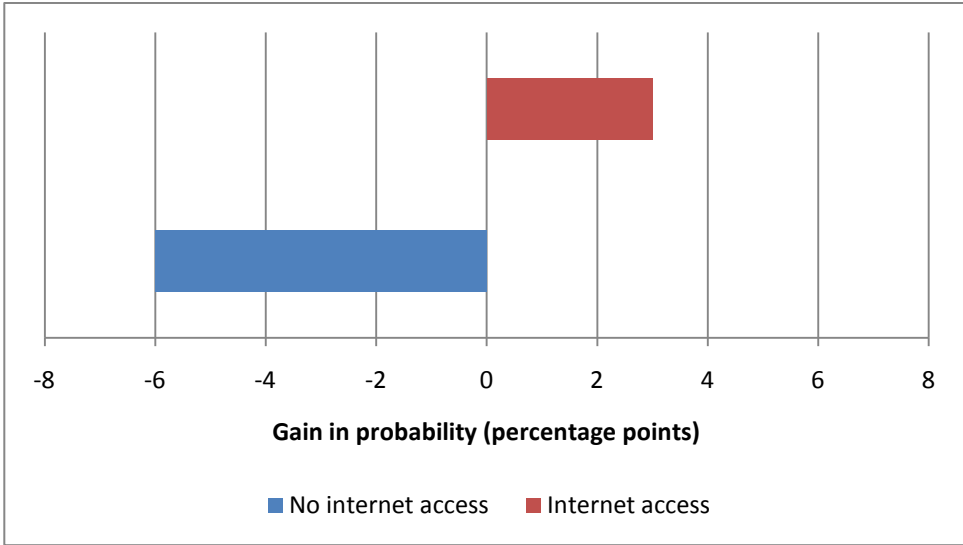


Chart 5.8 Gain in probability – general health

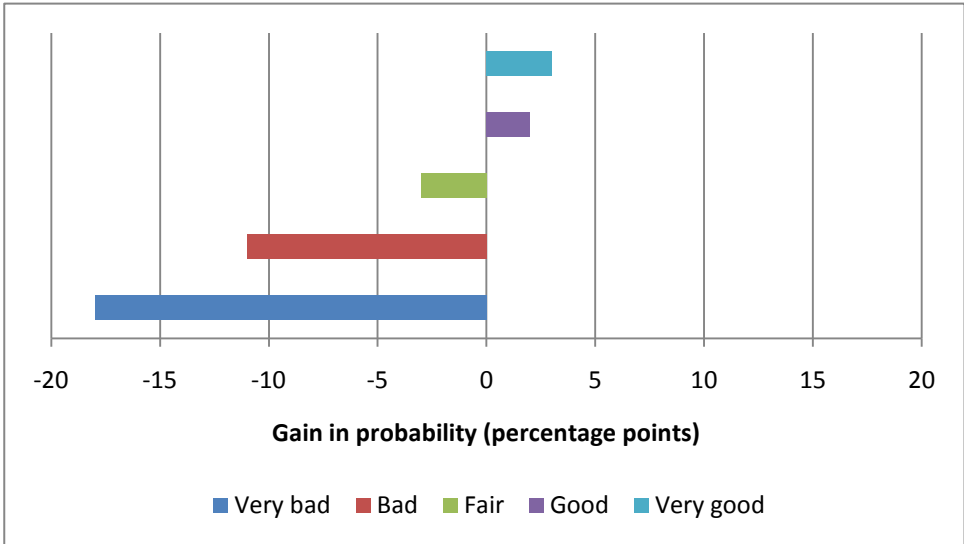
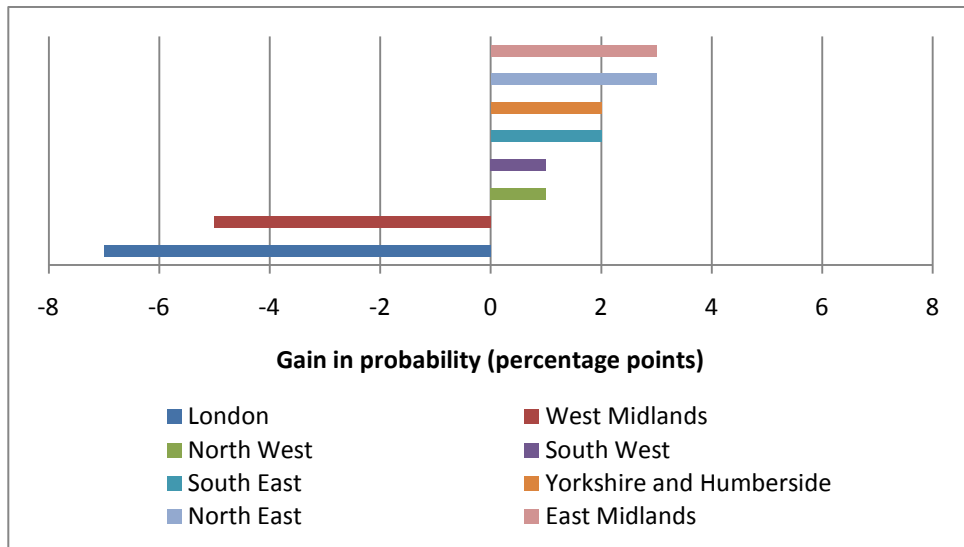
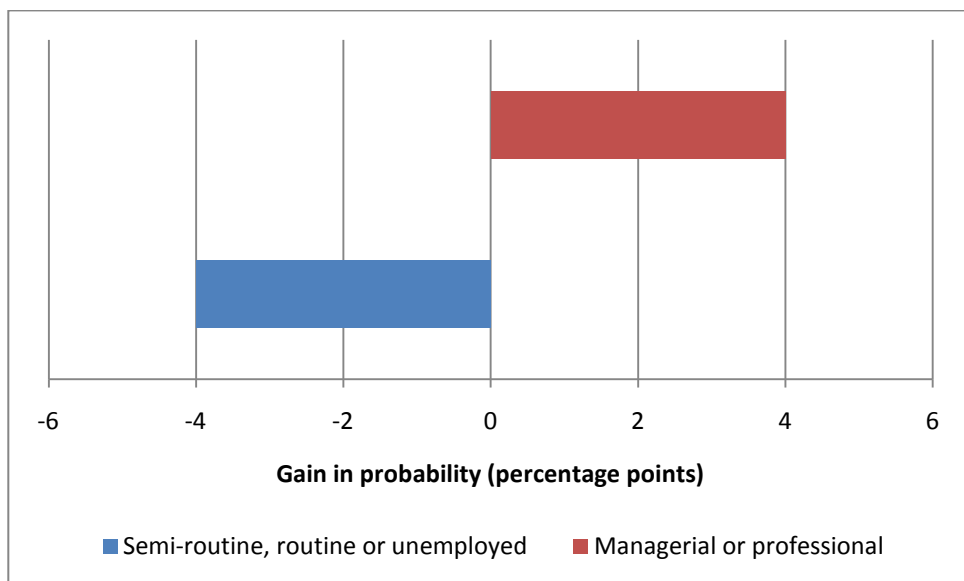


Chart 5.9 Gain in probability – English region



Note: no gain for the East of England, so excluded from chart

Chart 5.10 Gain in probability – occupation



Note: no gain for immediate occupation, so excluded from chart

Current activities

The dataset was split into three 'base likelihood' groups based on the probability value generated by the logistic regression version of this model:

- Relatively *high* probability of meeting PSA criteria = >70%
- *Mid* level probability of meeting PSA criteria = 48%-70%
- Relatively *low* probability of meeting PSA criteria = <48%

Despite the moderate predictive power of the model based only on background characteristics, the addition of current activity variables did not add a great deal. The new models explained an additional 12-14% of the observed variety in activity levels.

Eight current activity variables proved significantly predictive:

- Holiday activity
- Watching history programmes
- Accessing historical websites
- Accessing theatre or concert websites
- Voluntary work
- Newspaper readership
- Volume of TV viewing
- Watching current affairs programmes

Some of these variables (e.g. watching history programmes and accessing historical websites) have an obvious connection with visiting historic sites and may have an interdependent relationship with it.

It should be noted here that the master variable set is the same for each of the three base likelihood groups. Only the relative influence of each variable differs.

The positive relationships were with UK or a mix of UK and foreign holidays, low volume TV viewing but also watching history or current affairs programmes, accessing historical websites but also arts websites, doing voluntary work, and reading broadsheet newspapers.

Although these relationships held for all three base likelihood groups, the relative influence varied. A simple way of summarising this is to compute the simple standard deviation of the MCA-computed penalties and rewards for each variable.

This showed that penalties and rewards were generally strongest for the low base likelihood group (outright for five of the eight variables and joint top for another two). The exception was volume of TV viewing, where the relationship was strongest with the high base likelihood group.

We can do a similar calculation to find the specific characteristics that best distinguish the three base likelihood groups. The five most distinguishing features are:

UK only holidays / a mix of UK and foreign holidays

(**reward** = +10PP/+21PP for low base likelihood group, +9PP/+10PP for middle base likelihood group and +2PP/+7PP for the high base likelihood group)

Accessing historical websites

(**reward** = +23PP low base likelihood group, +14PP for middle base likelihood group and +8PP for the high base likelihood group)

Accessing theatre or concert websites

(**reward** = +13PP low base likelihood group, +8PP for middle base likelihood group and +4PP for the high base likelihood group)

Reading *The Times*

(**reward** = +11PP for low base likelihood group, +7PP for middle base likelihood group and +3PP for the high base likelihood group)

Reading *The Sun*

(**penalty** = -2PP for low base likelihood group, -6PP for middle base likelihood group and -9PP for the high base likelihood group)

The scores in brackets show the expected increase or decrease in the proportion meeting PSA criteria if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty' or 'reward'.

Tables 5.1 to 5.8 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Table 5.1 Percentage point gain in probability – holiday activity

	LOW	MIDDLE	HIGH
Holidays in UK and abroad	21	10	7
Holidays abroad only	3	0	-1
Holidays UK only	10	9	2
No holidays	-7	-10	-9

Table 5.2 Percentage point gain in probability – watching history programmes

	LOW	MIDDLE	HIGH
Watches history programmes	+9	+9	+5
Does not watch history programmes	-4	-4	-4

Table 5.3 Percentage point gain in probability – accessing historical websites

	LOW	MIDDLE	HIGH
Accesses historical websites	+23	+14	+8
Does not access historical websites	-1	-3	-4

Table 5.4 Percentage point gain in probability – accessing theatre/concert websites

	LOW	MIDDLE	HIGH
Accesses theatre/concert websites	+13	+8	+4
Does not access theatre/concert websites	-1	-3	-4

Table 5.5 Percentage point gain in probability – voluntary work

	LOW	MIDDLE	HIGH
Voluntary work	+11	+8	+4
Does no voluntary work	-2	-2	-2

Table 5.6 Percentage point gain in probability – newspaper readership

	LOW	MIDDLE	HIGH
The Times	+11	+7	+3
Financial Times	*	*	+5
Daily Telegraph	+7	+7	+3
Daily Mail	+6	+2	+1
Local daily newspaper	+2	+2	0
The Independent	*	+8	+4
The Guardian	0	+6	+4
Daily Express	0	0	-1
Daily Star	-1	-6	*
Metro	-2	-3	-3
The Sun	-2	-6	-9
Daily Mirror	-2	-1	-5
Other daily newspaper	0	+1	-3
None	-1	0	-2

* = numbers too small to report

Table 5.7 Percentage point gain in probability – daily hours of television viewing

	LOW	MIDDLE	HIGH
About 5 hours or more	-4	-5	-9
About 4 hours	+1	-2	-1
About 3 hours	+3	+3	-2
About 2 hours	+3	+1	0
About 1 hour	-1	0	+3
Less than 1 hour	-1	+1	+4

Table 5.8 Percentage point gain in probability – watching current affairs programmes

	LOW	MIDDLE	HIGH
Watches current affairs programmes	+6	+3	+2
Does not watch current affairs programmes	-2	-1	-2

Chapter 6: Modelling library visits

Background characteristics

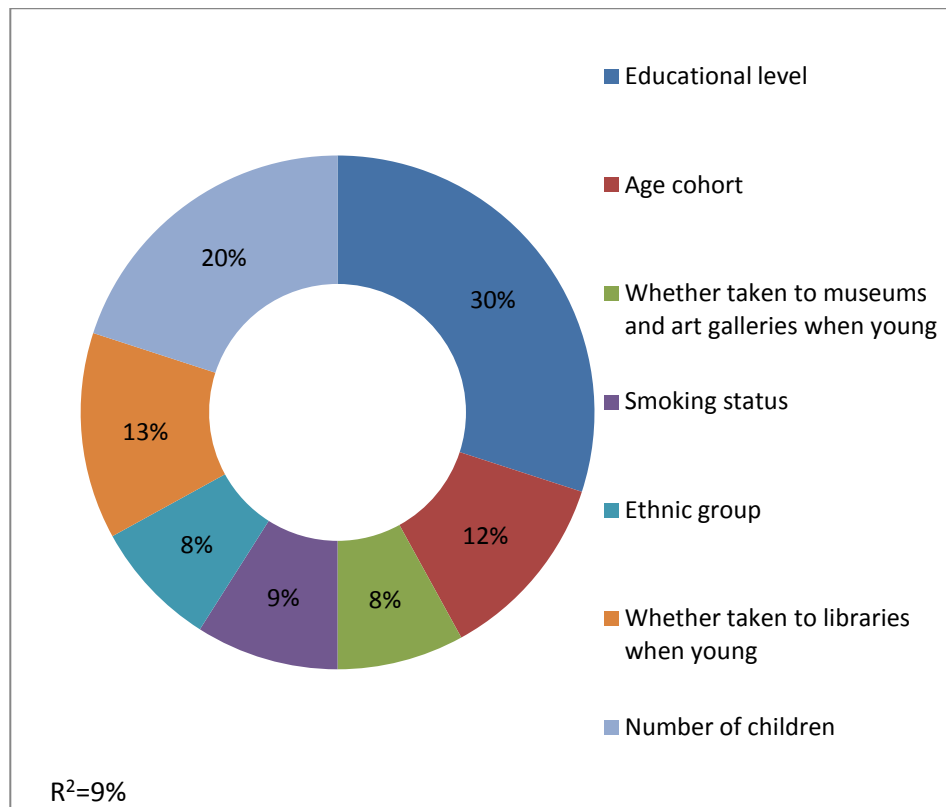
The predictive model for library visits was the weakest of the five. It is estimated that the stage 1 'background characteristics' model explains only 9% of the observed variety in activity levels^{xii}.

Seven background characteristics proved significantly predictive:

- Educational level
- Number of children in household
- Frequency of visiting libraries when young
- Age cohort
- Smoking status
- Frequency of visiting museums/galleries when young
- Ethnic group

Chart 6.1 shows the relative weight of each of these variables in terms of independent predictive power. The chart sums to 100% where 100% equals the predictive power of the whole model.

Chart 6.1 Predictors of library visits – background characteristics



As shown in Chart 6.1, the predictors of library visits were: high educational level, two or more children in the household, frequent visits to libraries, museums or galleries when young, early/middle old age (65-84), non-smoker status, and a non-White ethnic group.

The three most *positive* characteristics are:

- Two or more children in the household (+13PP)
- Educated to degree level (+13PP)
- Non-White ethnic group (+7PP if Asian, +6PP if Black, +10PP if other non-white or mixed ethnic group)

The scores in brackets show the expected increase in the proportion using a public library service at least once in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'reward'.

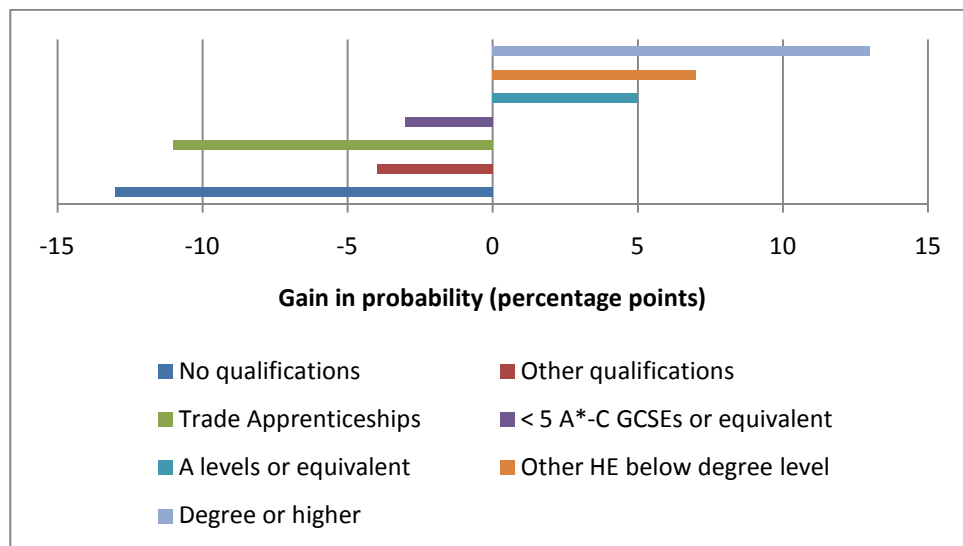
The three most *negative* characteristics are:

- No qualifications or trade apprenticeship as highest (-13PP)
- Never taken to libraries when young or taken less often than once a year (-6PP)
- No children (-4PP)

The scores in brackets show the expected decrease in the proportion using a public library service at least once in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA 'penalty'.

Charts 6.2 to 6.8 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Chart 6.2 Gain in probability – educational level



Note: no gain for 5+ GCSEs or equivalent, so excluded from chart

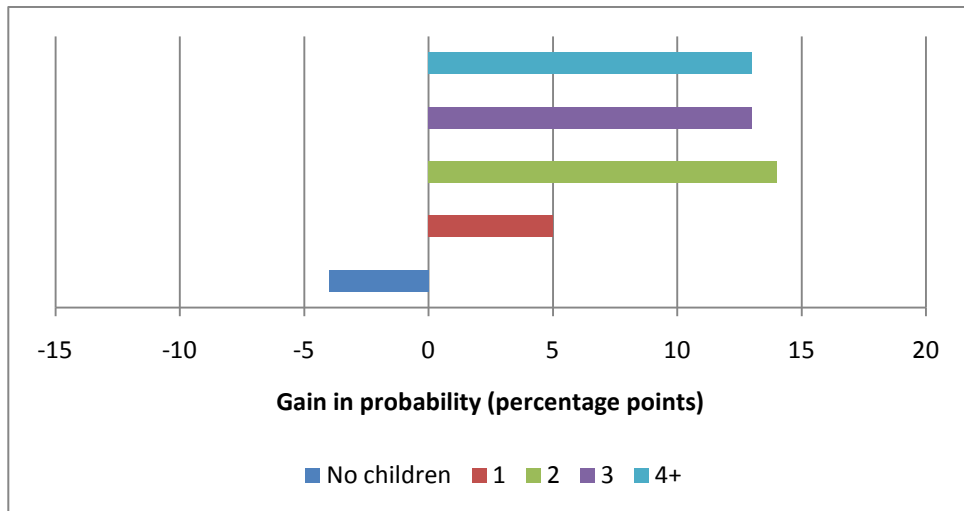
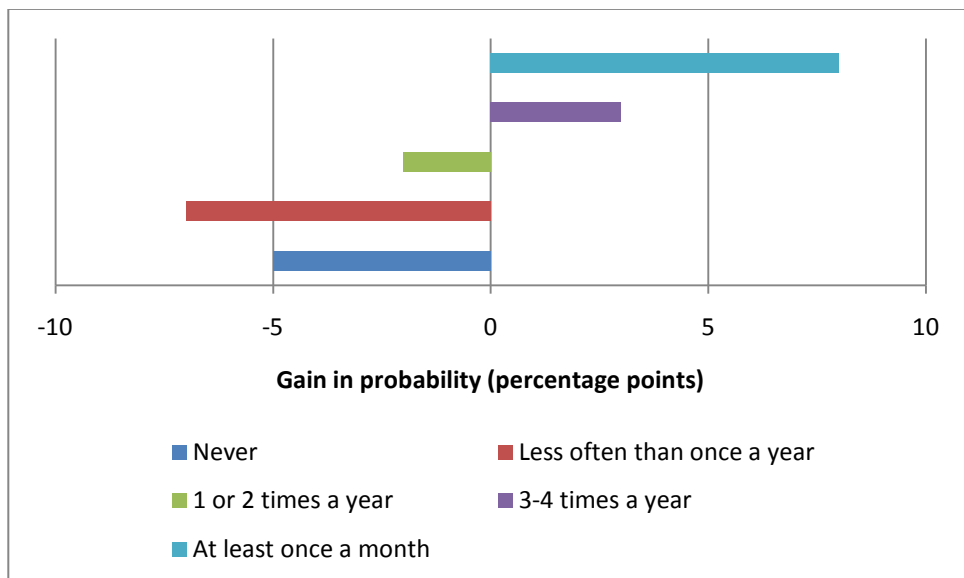
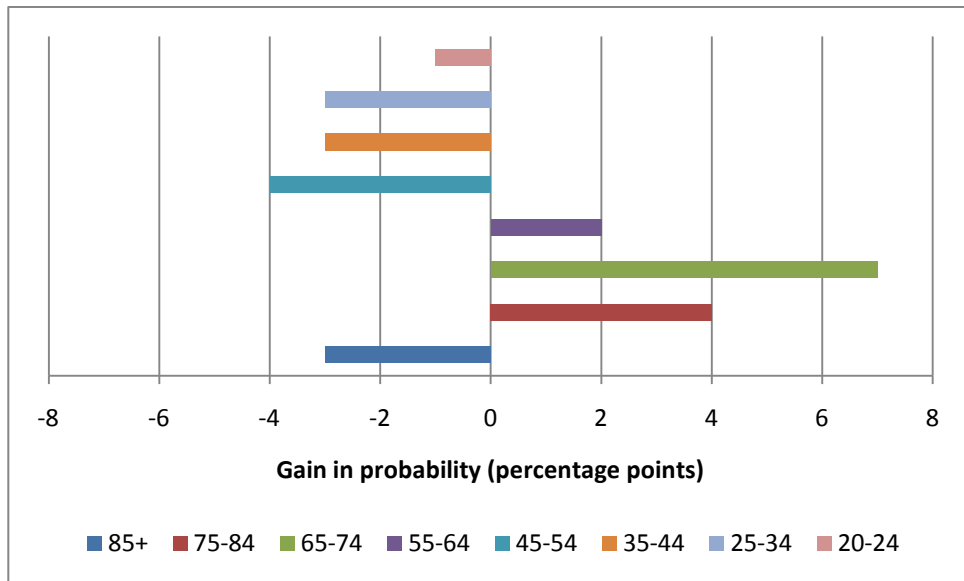
Chart 6.3 Gain in probability – number of children in household**Chart 6.4 Gain in probability – whether taken to libraries when young**

Chart 6.5 Gain in probability – age cohort



Note: no gain for 16-19 year olds, so excluded from chart

Chart 6.6 Gain in probability – smoking status

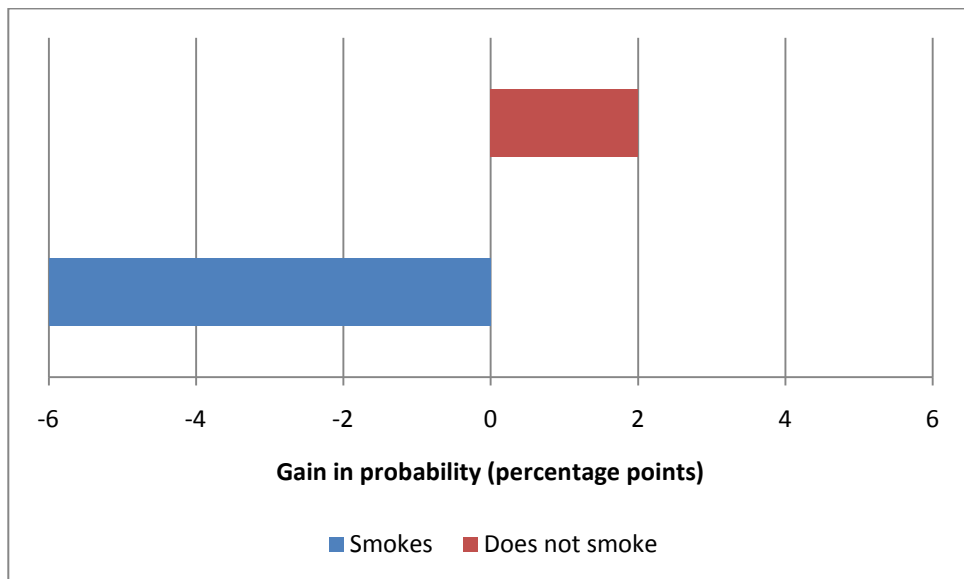
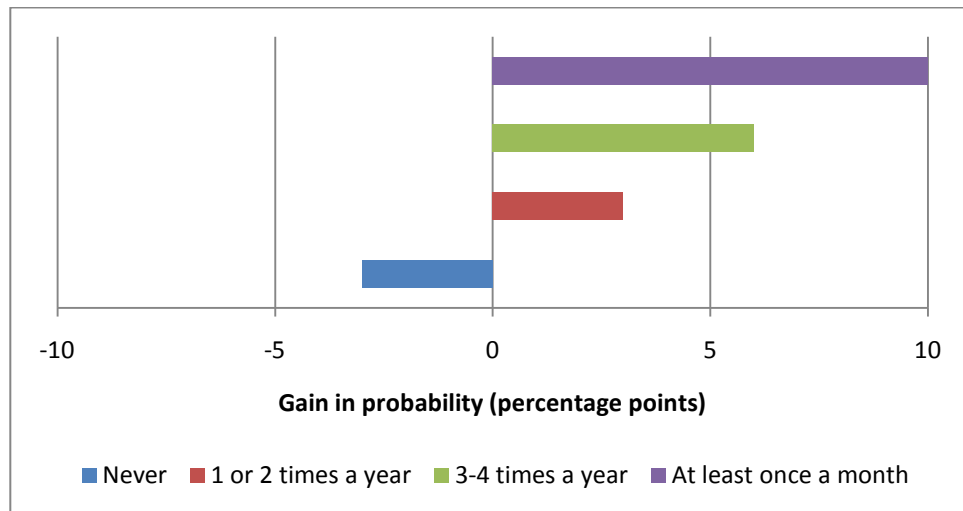
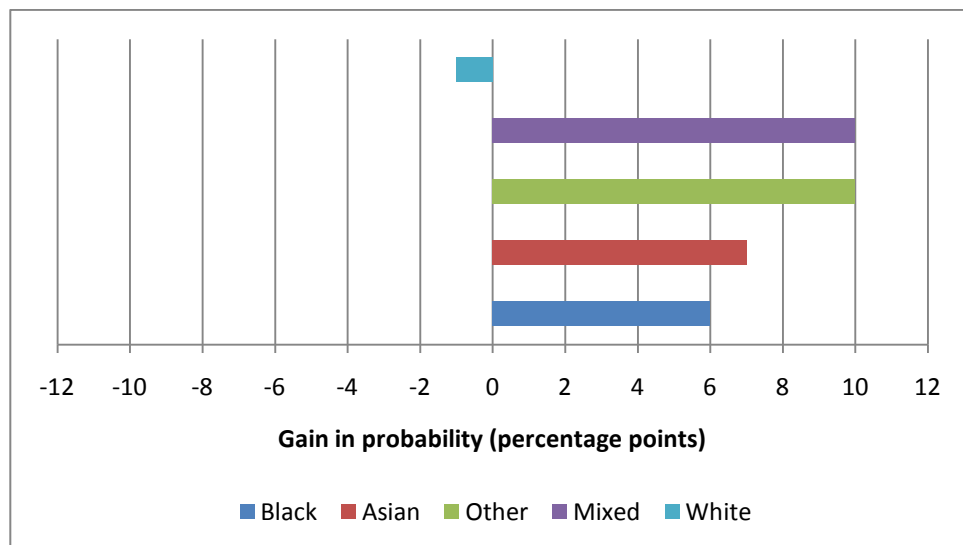


Chart 6.7 Gain in probability – whether taken to museums/art galleries when young



Note: no gain for 'less often than once a year', so excluded from chart

Chart 6.8 Gain in probability – ethnicity



Current activities

The dataset was split into three 'base likelihood' groups based on the probability value generated by the logistic regression version of this model:

- Relatively *high* probability of meeting PSA criteria = >54%
- *Mid* level probability of meeting PSA criteria = 40%-54%
- Relatively *low* probability of meeting PSA criteria = <40%

Despite the moderate predictive power of the model based only on background characteristics, the addition of current activity variables did not add a great deal. The new models explained an additional 13-16% of the observed variety in activity levels.

Only five current activity variables proved significantly predictive:

- Accessing library websites

- Visiting museums or galleries
- Reading for pleasure
- Voluntary work
- Newspaper readership

Some of these variables (e.g. accessing library websites and reading for pleasure) have an obvious connection with visiting a library and may have an interdependent relationship with it.

It should be noted here that the master variable set is the same for each of the three base likelihood groups. Only the relative influence of each variable differs.

The positive relationships were with accessing library websites, visiting museums or galleries, reading for pleasure, doing voluntary work, and reading broadsheet newspapers.

Although these relationships held for all three base likelihood groups, the relative influence varied. A simple way of summarising this is to compute the simple standard deviation of the MCA-computed penalties and rewards for each variable.

This showed that penalties and rewards were generally strongest for the low base likelihood group (outright for three of the five variables and joint top for another one). The exception was voluntary work, where the relationship was strongest with the middle base likelihood group.

We can do a similar calculation to find the specific characteristics that best distinguish the three base likelihood groups. The five most distinguishing features are:

Accessing library websites

(**reward** = +56PP low base likelihood group, +48PP for middle base likelihood group and +36PP for the high base likelihood group)

Reading the *Daily Star*, *Daily Telegraph* or non-mainstream newspaper:

(*Daily Star*: **penalty** = -4PP low base likelihood group, “**reward**” = +1PP for middle base likelihood group and **penalty** = -11PP for the high base likelihood group)

(*Daily Telegraph*: **reward** = +10PP low base likelihood group, +5PP for middle base likelihood group and zero for the high base likelihood group)

(Non-mainstream newspaper: “**reward**” = +1PP for the low and middle base likelihood groups, **penalty** = -8PP for the high base likelihood group)

Visiting museums or galleries

(**reward** = +14PP for low base likelihood group, +8PP for middle base likelihood group and +4PP for the high base likelihood group)

The scores in brackets show the expected increase or decrease in the proportion using a public library service at least once in the past 12 months if everybody had the characteristic but all other relevant characteristics are as observed. This is called the MCA ‘penalty’ or ‘reward’.

Tables 6.1 to 6.5 show the MCA-computed penalties and rewards associated with each level of each variable in the model.

Table 6.1 Percentage point gain in probability – accessing library websites

	LOW	MIDDLE	HIGH
Accesses library websites	+56	+48	+36

Does not access library websites	-1	-5	-8
----------------------------------	----	----	----

Table 6.2 Percentage point gain in probability – visiting museums/galleries

	LOW	MIDDLE	HIGH
Visits museums/galleries	+14	+8	+4
Does not visit museums/galleries	-5	-7	-6

Table 6.3 Percentage point gain in probability – reading for pleasure

	LOW	MIDDLE	HIGH
Reads for pleasure	+6	+3	+2
Does not read for pleasure	-8	-6	-5

Table 6.4 Percentage point gain in probability – voluntary work

	LOW	MIDDLE	HIGH
Does voluntary work	+4	+6	+4
Does no voluntary work	-1	-2	-2

Table 6.5 Percentage point gain in probability – newspaper readership

	LOW	MIDDLE	HIGH
Financial Times	*	*	-9
The Independent	*	+2	+4
Daily Telegraph	+10	+5	0
The Times	+7	+2	-2
The Guardian	+6	-1	+3
Daily Mail	+4	+1	+1
Daily Express	+3	-1	-1
Metro	0	-1	+2
Local daily newspaper	-1	+2	0
Daily Mirror	-2	-1	-1
The Sun	-3	-3	+1
Daily Star	-4	+1	-11
Other daily newspaper	+1	+1	-8
None	-1	-1	-1

* = numbers too small to report

Endnotes

ⁱ PSA21: to build more cohesive, empowered and active communities was agreed as part of the 2008-11 spending review period. While Communities and Local Government are the lead department for PSA 21, DCMS have responsibility for Indicator 6: to increase the percentage of adults (aged 16 and over) who participate in culture and sport by 2010/11.

ⁱⁱ Logistic regression produces odds ratios. Odds are derived from probabilities. If a variable is distributed 30% 'yes' and 70% 'no', we say that an individual's base probability of 'yes' is 0.3. The odds of 'yes' equal the probability of 'yes' divided by the probability of 'no'; i.e. $0.3/0.7 = 0.43$. If something is more likely than not, the odds will be greater than 1. If something is less likely than not, the odds will be less than 1 but always above 0. The odds ratio equals the odds for an individual with characteristic x divided by the odds for an individual without characteristic x, holding other relevant characteristics constant. However, while this is fairly simple to grasp when predictor variables are binary, it is more difficult when predictor variables are categorical with three or more possible values. The selection of the 'reference category' (equivalent to the absence of a characteristic) becomes critical. If the reference category is not a typical value, all of the other categories may have statistically significant odds ratios when compared to the reference category but not when compared with each other. In contrast, the MCA method directly shows how the proportion giving the answer 'yes' would change if a predictor variable distribution is switched from the observed distribution to a distribution in which 100% of the population has characteristic x. These are usually called 'penalties' and 'rewards'. One advantage logistic regression odds ratios has over MCA penalties and rewards is that they can be combined to produce an overall odds ratio for a sub-group defined on several parameters. MCA penalties/rewards cannot be summed in the same way.

ⁱⁱⁱ Note that the real distribution is estimated from the sample - as are the relationships described in the model.

^{iv} It should be borne in mind that these 'proportion adjustment' values assume that the distributions of all other relevant variables in the model are held at the observed level. That would be unlikely because there are small correlations between the various predictor variables in the model. However, these models do not contain any variables that are highly correlated with each other because this phenomenon (called 'collinearity') spoils the effect estimates. This is one of the reasons why educational level and occupation class are rarely included in the same model (unless converted into less powerful but also less correlated variables).

^v There is no single agreed way of dealing with apparently substantial effects based on small samples. The estimated effect sizes remain the best estimates even if they are not significantly different from zero according to statistical tests. To ignore a suspected major effect for a small minority may be a larger error than to comment upon the occasional spurious effect. The approach here is a compromise, recognising that estimates based on very small samples are inherently unstable but not ignoring substantial estimated effects from slightly larger (but still small) samples.

^{vi} This is the MCA-calculated R² value which tends to produce a result close to the Cox and Snell pseudo-R² in logistic regression. The more frequently quoted Nagelkerke R² is 0.14.

^{vii} This goes beyond formal statistical significance which is relatively easy to achieve with a very large sample like this. At minimum the predictor variable should shift the distribution of the target variable (in this case the binary PSA for sport/recreational physical activity) by at least a few percentage points when the value of the predictor variable changes.

^{viii} There are one or two instances where an interaction term has been included but the impact of inclusion is very marginal.

^{ix} This is the MCA-calculated R2 value. The more frequently quoted Nagelkerke R2 is 0.19.

^x This is the MCA-calculated R2 value. The more frequently quoted Nagelkerke R2 is 0.18.

^{xi} This is the MCA-calculated R2 value. The more frequently quoted Nagelkerke R2 is 0.24.

^{xii} This is the MCA-calculated R2 value. It equals the more frequently quoted Nagelkerke R2 value.