

**Preliminary analysis of the UK 2000 Time Use Survey
for the Department for Culture, Media and Sport**

Examining participation in sporting and cultural activities

**Dr Patrick Sturgis, University of Surrey &
Dr Jonathan Jackson, London School of Economics and Political Science**

July 2003

What underlies the apparent negative relationship between education and sports participation?

In previous analyses it has been shown that possessing higher educational qualifications is associated with lower levels of participation in sport. This is certainly somewhat counter-intuitive as we would generally expect those with higher qualifications to have higher rates of participation on a number of different grounds: greater awareness of the benefits of exercise; continuity of sports participation at college or university; and being more likely to move into occupations that provide sports facilities, to name but three. The regression model showing a negative relationship between qualifications and sport is reproduced in Table 1 below.

Table 1 OLS Regression Model Predicting Participation in Sport

<i>Variables</i>	<i>B</i>	<i>Std. Error</i>	<i>95% P Value</i>
(Constant)	1.77	0.20	0.000
sex (male = 1)	0.52	0.04	0.000
age (years)	-0.05	0.01	0.000
age (years) squared	0.00	0.00	0.000
ns-sec# higher managerial and professional	-0.05	0.08	0.566
ns-sec lower managerial and professional	0.13	0.07	0.054
ns-sec small employers and own account workers	0.01	0.09	0.945
ns-sec lower supervisory and technical	-0.15	0.08	0.058
ns-sec semi-routine	-0.03	0.07	0.675
ns-sec routine	-0.12	0.07	0.092
ns-sec never worked and long-term unemployed	-0.01	0.10	0.893
ns-sec students and unclassifiable	0.02	0.08	0.812
white ethnic group	0.35	0.09	0.000
marital status+ single	0.26	0.06	0.000
marital status divorced	0.32	0.07	0.000
highest qualification* degree	-0.16	0.07	0.022
highest qualification A level	-0.18	0.10	0.082
highest qualification GCSE	0.23	0.04	0.000
general health	-0.12	0.02	0.000
# cultural events in past 4 weeks	0.69	0.04	0.000
# cultural events in past 4 weeks squared	-0.07	0.01	0.000
household income (banded)	0.06	0.01	0.000
whether household owns a car or not (owns = 1)	0.10	0.05	0.056
whether child (<16) in household or not (yes = 1)	-0.10	0.04	0.025
housing tenure (owns home = 1)	0.09	0.05	0.049
Region\$ south	-0.03	0.05	0.578
region London	-0.15	0.08	0.055
region north	0.00	0.05	0.962
region Wales/Scotland	0.02	0.06	0.802
N. Ireland	-0.29	0.12	0.012

N= 7942; R²=0.26; F=94, df=29, p=0.000; Source = UKTUS 2000

Omitted categories: # = intermediate occupations; + = married; * = no qualifications; \$ = Midlands

From Table 1, we can see that - controlling for all other variables in the model - possessing a degree or higher qualification reduces the estimated number of

sporting activities reported by 0.16. To illustrate what this means in real terms, consider that an 'average person' would report approximately 1.8 sports in the previous 4 weeks (this is the constant term in Table 1). If this 'average person' had a degree we would expect them to have reported 0.16 *fewer* activities than an equivalent person who does not hold a degree. Clearly then, this is not an effect of great magnitude and, as is sometimes the case with weak effects, the direction and significance of the estimate changes, depending on the other variables included in the model.

Table 2 below shows the estimates from a model that includes only the three indicators of educational qualification (the coefficients are, therefore, interpretable relative to the omitted category, those with no qualifications).

Table 2 OLS Regression Model Predicting Participation in Sport II

<i>Variables</i>	<i>B</i>	<i>Std. Error</i>	<i>95% P Value</i>
(Constant)	.839	.027	.000
highest qualification* degree	.498	.064	.000
highest qualification A level	.621	.095	.000
highest qualification GCSE	.951	.038	.000

Without any control variables in the model, all qualification levels show an increase in the reporting of sports relative to those with no qualifications. Those with a degree report approximately an additional half of a sporting activity (.498 in the B column) during the period in question. Relative to the other qualification levels, though, those with a degree still report somewhat lower numbers of sports.

Table 3 OLS Regression Model Predicting Participation in Sport

<i>Variables</i>	<i>B</i>	<i>Std. Error</i>	<i>95% P Value</i>
(Constant)	1.77	0.20	0.000
age (years)	-0.05	0.01	0.000
age (years) squared	0.00	0.00	0.000
ns-sec# higher managerial and professional	-0.05	0.08	0.566
ns-sec lower managerial and professional	0.13	0.07	0.054
ns-sec small employers and own account workers	0.01	0.09	0.945
ns-sec lower supervisory and technical	-0.15	0.08	0.058
ns-sec semi-routine	-0.03	0.07	0.675
ns-sec routine	-0.12	0.07	0.092
ns-sec never worked and long-term unemployed	-0.01	0.10	0.893
ns-sec students and unclassifiable	0.02	0.08	0.812
highest qualification* degree	-0.16	0.07	0.022
highest qualification A level	-0.18	0.10	0.082
highest qualification GCSE	0.23	0.04	0.000
general health	-0.12	0.02	0.000
# cultural events in past 4 weeks	0.69	0.04	0.000
# cultural events in past 4 weeks squared	-0.07	0.01	0.000

However, once we add in a few control variables (age, social class, health status and number of cultural events attended in past 4 weeks) we see that the direction of the coefficient reverses to become negative, making it an example of 'Simpson's paradox'. Substantively, what this means is that it is not qualifications *per se* that are associated with lower levels of this definition of sports participation but the associated characteristics of individuals with this level of qualification. People with degrees tend also to be from higher socio-economic groups, younger, in good health, and participators in the arts. So, it is these latter characteristics that are positively correlated with participation and, once they are controlled for, we see a reversal in sign of the relationship.

Consider the Measurement

Before we too hastily conclude that obtaining a degree makes us less likely to participate in sport, however, it is also worth reflecting on how 'sports participation' is actually measured in this analysis. The dependent variable in Table 1 is a simple count of the total number of different sporting activities participated in during the previous 4 weeks reported in the individual questionnaire. On this variable, someone who went jogging six days a week but did no other sport during the period in question would have a score of 1. Someone who had played darts, crazy golf and pool on a single day during the same period, however, would have a score of 3. Most people would not consider the sports participation of the latter individual to be greater than that of the former but this is how it would appear on this variable. So what this variable really measures is, perhaps, more the *breadth* than the *depth* of sporting participation.

If we turn to a slightly different measure of the same basic concept (sports participation), we can see that the relationship between qualifications and participation reverses. Table 4 shows the results of a logistic regression model predicting a dependent variable derived from the diary data but using the same set of predictors as in Table 1. This variable is a binary indicator of whether the individual recorded doing any of the activities falling under the one-digit code '6' in the diary (sports and outdoor activities). This model, then, is predicting the (log of the) probability that an individual will participate in any of these activities *on the average day*.

The first column in table 4 (B) is the logit and is not readily interpretable in a direct sense. Of more interest for substantive interpretation are the p value (indicating whether the relationship is statistically significant) and the odds ratio (indicating how much the odds of participating in sport go up or down for each unit change in the predictor in question).

For instance, the row in Table 4 containing the estimates for sex of respondent, has a p value of 0.000 which indicates that this relationship is statistically significant (technically, the probability that the true population value of this coefficient is zero is less than one in one thousand). The odds

ratio for sex is 1.32, indicating that the odds of participating in sport increase by approximately 32% for men relative to women.

Table 4 Logistic Regression Model Predicting Participation in Sport

<i>Variables</i>	<i>B</i>	<i>P value</i>	<i>Odds Ratio</i>
(Constant)	-1.88	0.000	0.15
sex (male = 1)	0.28	0.000	1.32
age (years)	-0.01	0.124	0.99
age (years) squared	0.00	0.025	1.00
ns-sec# higher managerial and professional	-0.15	0.273	0.86
ns-sec lower managerial and professional	0.01	0.928	1.01
ns-sec small employers and own account workers	-0.18	0.110	0.83
ns-sec lower supervisory and technical	-0.18	0.070	0.83
ns-sec semi-routine	-0.15	0.073	0.86
ns-sec routine	-0.10	0.259	0.90
ns-sec never worked and long-term unemployed	0.15	0.160	1.17
ns-sec students and unclassifiable	-0.03	0.755	0.97
white ethnic group	-0.14	0.299	0.87
marital status+ single	0.36	0.000	1.43
marital status divorced	0.16	0.110	1.17
highest qualification* degree	0.35	0.000	1.43
highest qualification A level	0.06	0.435	1.06
highest qualification GCSE	0.04	0.612	1.04
general health	-0.21	0.000	0.81
# cultural events in past 4 weeks	0.19	0.001	1.21
# cultural events in past 4 weeks squared	-0.02	0.108	0.98
household income (banded)	0.00	0.751	1.00
whether household owns a car or not (owns = 1)	0.00	0.997	1.00
whether child (<16) in household or not (yes = 1)	0.22	0.001	1.25
housing tenure (owns home = 1)	0.15	0.023	1.16
Region\$ south	0.02	0.792	1.02
region London	-0.33	0.000	0.72
region north	-0.08	0.285	0.93
region Wales/Scotland	0.06	0.470	1.06
N. Ireland	-0.04	0.771	0.96

Source = UKTUS 2000; Omitted categories: # = intermediate occupations; + = married; * = no qualifications; \$ = Midlands

The corresponding figures for highest qualification (degree) are 0.000 and 1.43, indicating that having a degree makes people nearly 50% more likely to participate in one of the activities in question on the average day and that this coefficient is highly unlikely to be due to sampling variation.

In sum, then, the results of this more in-depth analysis of how educational qualifications are related to sports participation are that how our measures of participation are operationalised is of key importance. When considering what we might term the *breadth* of an individual's participation (the heterogeneity of activities) over a fairly long time frame, there is evidence supporting a weak negative association. However, when we look at the

probability of participating in any of a range of different sports on the average day, there is evidence of a stronger relationship in the opposite direction.

What is the impact of having children in the household on the cultural and sporting activities of parents/carers?

Parental status is likely to be an important variable in understanding life course changes in cultural and sports participation of individuals and groups. With the arrival of young children, many parents find they have limited opportunity to maintain their previous levels of participation in both areas of activity. However, as children grow and begin, themselves, to participate in sport and cultural events both at school and through other institutional or less formal networks, it is likely that this may impact on the participation levels of their parents and, possibly, other household members.

In order to investigate the moderating effect of children on the cultural and sports participation of other household members, a series of regression models was fitted to data from the household and individual questionnaires and the diaries. Table 5 shows the results of a logistic regression model predicting whether an individual reported participating in any of the following sporting activities in their diary entries for a particular day:

Codes Used to Construct Sports Participation Variable from Diary Data

- 610 Unspecified physical exercise
- 611 Walking and hiking
 - 6111 Taking a walk or hike that lasts at least 2 miles or 1 hour
 - 6119 Other walk or hike
- 612 Jogging and running
- 613 Biking, skiing and skating
 - 6131 Biking
 - 6132 Skiing or skating
- 614 Ball games
 - 6140 Unspecified ball games
 - 6141 Indoor pairs or doubles games
 - 6142 Indoor team games
 - 6143 Outdoor pairs or doubles games
 - 6144 Outdoor team games
 - 6149 Other specified ball games
- 615 Gymnastics
- 616 Fitness
- 617 Water sports
 - 6170 Unspecified water sports
 - 6171 Swimming
 - 6179 Other specified water sports
- 619 Other specified physical exercise
- 620 Unspecified productive exercise
- 621 Hunting and fishing
- 629 Other specified productive exercise
- 631 Sports related activities

- 6310 Unspecified sports related activities
- 6311 Activities related to sports
- 6312 Activities related to productive exercise

The predictors in this first model are limited to dummy variables indicating the banded age of the respondent. Estimates are provided for the ‘average day’, week days and weekend days as it is likely that participation rates will vary across these periods, possibly interacting with household type.

Table 5 Logistic Regression predicting Sports Participation

Predictor	Average day			Weekdays			Weekends		
	B	Sig.	O.R.	B	Sig.	O.R.	B	Sig.	O.R.
25TO34*	-0.332	0.000	0.717	-0.332	0.000	0.718	-0.333	0.011	0.716
35TO44	-0.405	0.000	0.667	-0.369	0.000	0.691	-0.477	0.000	0.620
45TO54	-0.396	0.000	0.673	-0.432	0.000	0.649	-0.327	0.014	0.721
55TO64	-0.117	0.134	0.890	-0.014	0.882	0.986	-0.350	0.013	0.705
65TO74	-0.233	0.006	0.792	-0.151	0.141	0.860	-0.414	0.007	0.661
75PLUS	-0.367	0.000	0.693	-0.268	0.016	0.765	-0.590	0.001	0.555
Constant	-1.468	0.000	0.230	-1.563	0.000	0.210	-1.256	0.000	0.285

* Omitted category = 16-24; Source UKTUS 2000; base = all respondents 16+; **bold** = significant at 95% confidence level.

Table 5 shows that the probability of participating in one of these sporting activities is highest for the 16-24 age group. These are the omitted category in the model and so all other coefficients in table 5 should be interpreted as being relative to the probability for this baseline group. For the average day, all age groups apart from those aged 55-64 are significantly less likely to participate in these activities. Those aged 25 to 34, for instance, are 28% less likely to have taken part in one of these activities, while the figure for those aged 35-44 is 33%. There is also some variation in sports participation by age over the course of the average week, with the pattern emerging most clearly and strongly on the weekend.

Next we turn to look at the impact of young children in the household on sports participation and how it varies by age group. Table 6 presents the same basic model as shown in table 5 but, this time, a variable indicating whether a child under the age of 16 is resident in the household.

The inclusion of the ‘child in household’ variable has little impact on the direction and significance of the estimates for the age group dummies, although there are a few minor changes from those shown in Table 5. Of more importance are the coefficients for the main effect of the ‘child in household’ variable itself. On both the ‘average day’ and on weekdays, the presence of a child under the age of 16 in the household significantly *increases* the probability of sports participation amongst the other household members. On the average day, the presence of a child in the household increases the probability of sports participation by 30%, while on weekdays the figure rises

to 43%. Interestingly, the effect is much weaker and non-significant on weekends.

Table 6 Logistic Regression predicting Sports Participation II

Predictor	Average day			Weekdays			Weekends		
	B	Sig.	O.R.	B	Sig.	O.R.	B	Sig.	O.R.
25TO34*	-0.294	0.000	0.745	-0.280	0.002	0.756	-0.322	0.015	0.725
35TO44	-0.325	0.000	0.723	-0.262	0.005	0.769	-0.453	0.001	0.635
45TO54	-0.414	0.000	0.661	-0.455	0.000	0.634	-0.333	0.013	0.717
55TO64	-0.190	0.017	0.827	-0.110	0.249	0.896	-0.373	0.010	0.688
65TO74	-0.313	0.000	0.731	-0.257	0.014	0.773	-0.440	0.005	0.644
75PLUS	-0.451	0.000	0.637	-0.378	0.001	0.685	-0.617	0.000	0.540
CHILD	0.265	0.000	1.303	0.357	0.000	1.429	0.080	0.407	1.083
Constant	-1.913	0.000	0.148	-2.166	0.000	0.115	-1.389	0.000	0.249

* Omitted category = 16-24; Source UKTUS 2000; base = all respondents 16+; **bold** = significant at 95% confidence level.

While the models presented in Table 6 demonstrate that having young children in the household increases the probability of sports participation, it gives no indication of the extent to which this effect is manifested differentially across age bands. We might reasonably expect, for instance, that the impact of children on activity patterns will vary across the 25-34 and the 35-44 year age bands. Children in the latter households are likely, on average, to be older themselves and, therefore, the number and range of sporting activities participated in by the family (or other household unit) may be somewhat different from households containing younger children. In order to investigate this possibility, some interaction terms were added to the model presented in Table 6.

To interpret a regression model containing interaction (multiplicative) terms it is necessary to understand that the coefficients of the elements of the product term are not average effects which remain constant at all levels of the variable in question – as is the case in the ‘main effects’ models presented in Tables 5 and 6 – but rather represent *conditional* effects whose magnitudes vary across the values of the other variable in the interaction term. Thus, the coefficients for the constituent parts of an interaction term tell us how much and in which direction we should expect the dependent variable to change given a one unit change in the variable in question *when the other variable in the interaction term assumes the value zero*. The coefficient of the cross-product term tells us how much we should expect this value to change with each unit increase in the other variable in the interaction. Thus if an interaction term comprises variables x1 and variable x2, then a significant interaction would indicate that the amount of change in the dependent variable given a one unit change in x1 will depend on the level assumed on variable x2.

Table 7 presents the same model as shown in table 6 but with the addition of four interaction terms for age bands 25-34 to 55-64 by the variable ‘child in

household'. None of these interaction terms reaches statistical significance at the 95% level of confidence, indicating that the impact of having children in a household on sports participation is not moderated by the ages of the other household members.

Table 7 Logistic Regression predicting Sports Participation III

	Average day			Weekdays			Weekends		
	B	Sig.	O.R.	B	Sig.	O.R.	B	Sig.	O.R.
25TO34*	-0.600	0.025	0.549	-0.510	0.124	0.601	-0.774	0.090	0.461
35TO44	-0.554	0.032	0.575	-0.548	0.086	0.578	-0.530	0.233	0.589
45TO54	0.075	0.801	1.078	-0.040	0.915	0.961	0.268	0.588	1.307
55TO64	0.268	0.632	1.307	0.213	0.758	1.237	0.445	0.645	1.561
65TO74	-0.304	0.001	0.738	-0.245	0.026	0.783	-0.439	0.008	0.645
75PLUS	-0.441	0.000	0.643	-0.366	0.002	0.694	-0.615	0.001	0.540
CHILD	0.233	0.043	1.262	0.313	0.028	1.368	0.076	0.701	1.079
25-34*CHILD	0.192	0.218	1.212	0.141	0.462	1.152	0.291	0.279	1.338
35-44*CHILD	0.156	0.321	1.168	0.191	0.318	1.211	0.055	0.843	1.056
45-54*CHILD	-0.280	0.097	0.756	-0.235	0.266	0.791	-0.350	0.217	0.705
55-64*CHILD	-0.230	0.427	0.795	-0.159	0.655	0.853	-0.419	0.404	0.658
Constant	-1.859	0.000	0.156	-2.091	0.000	0.124	-1.383	0.000	0.251

* Omitted category = 16-24; Source UKTUS 2000; base = all respondents 16+; **bold** = significant at 95% confidence level.

Turning now to how young children in the household impact on the probability of other household members participating in cultural activity, Table 8 shows the results of a logistic regression model predicting whether an individual reported participating in any of the following cultural activities in their diary entries for a particular day:

Codes Used to Construct Cultural Participation Variable from Diary Data

- 520 Unspecified entertainment and culture
- 521 Cinema
- 522 Theatre and concerts
- 5220 Unspecified theatre or concerts
- 5221 Plays, musicals or pantomimes
- 5222 Opera, operetta or light opera
- 5223 Concerts or other performances of classical music
- 5224 Live music other than classical concerts, opera and musicals
- 5225 Dance performances
- 5229 Other specified theatre or concerts
- 523 Art exhibitions and museums
- 529 Other specified entertainment and culture
- 5291 Visiting a historical site
- 5292 Visiting a wildlife site
- 5293 Visiting a botanical site
- 5294 Visiting a leisure park
- 5295 Visiting an urban park, playground or designated play area
- 5296 Other specified entertainment or culture

Table 8 Logistic Regression predicting Cultural Participation

	Average day			Weekdays			Weekends		
	B	Sig.	O.R.	B	Sig.	O.R.	B	Sig.	O.R.
25TO34*	0.140	0.294	1.15	-0.169	0.329	0.84	0.598	0.006	1.82
35TO44	-0.032	0.817	0.97	-0.296	0.094	0.74	0.377	0.092	1.46
45TO54	-0.334	0.026	0.72	-0.740	0.000	0.48	0.205	0.381	1.23
55TO64	-0.370	0.022	0.69	-0.791	0.000	0.45	0.170	0.492	1.19
65TO74	-0.376	0.030	0.69	-0.305	0.141	0.74	-0.538	0.091	0.58
75PLUS	-0.643	0.001	0.53	-0.838	0.001	0.43	-0.327	0.303	0.72
Constant	-3.098	0.000	0.05	-3.162	0.000	0.04	-2.955	0.000	0.05

* Omitted category = 16-24; Source UKTUS 2000; base = all respondents 16+; **bold** = significant at 95% confidence level.

Table 8 shows that, on the average day, participation in cultural activities declines with age. Although, there is no significant difference amongst those aged 16-44, those in the higher age bands are all significantly less likely to participate in these cultural events than their younger counterparts. For instance, those aged 45-54 are 25% less likely to participate than those aged 16-24, while the corresponding figure for those aged 75 or over is nearly 50%, as we might expect. There is considerable variation in this general trend, depending on whether we focus on week days or the weekend. On the weekend, the 25-34 age group are 80% more likely to participate in these cultural activities than those aged 16-24 while no further significant differences emerge amongst the other age bands.

As with the models fitted for sports participation, we now turn to a consideration of the impact of young children on the cultural consumption of other household members. Table 9 adds in effects for the 'child in household' variable to the models presented in Table 8.

Table 9 Logistic Regression predicting Cultural Participation II

	Average day			Weekdays			Weekends		
	B	Sig.	O.R.	B	Sig.	O.R.	B	Sig.	O.R.
25TO34*	0.119	0.375	1.13	-0.205	0.239	0.81	0.594	0.007	1.81
35TO44	-0.074	0.599	0.93	-0.368	0.042	0.69	0.369	0.106	1.45
45TO54	-0.325	0.030	0.72	-0.724	0.000	0.48	0.207	0.377	1.23
55TO64	-0.329	0.045	0.72	-0.718	0.002	0.49	0.178	0.478	1.19
65TO74	-0.330	0.062	0.72	-0.224	0.293	0.80	-0.529	0.100	0.59
75PLUS	-0.595	0.004	0.55	-0.754	0.005	0.47	-0.318	0.322	0.73
CHILD	-0.138	0.154	0.87	-0.235	0.074	0.79	-0.027	0.850	0.97
Constant	-2.871	0.000	0.06	-2.777	0.000	0.06	-2.910	0.000	0.05

* Omitted category = 16-24; Source UKTUS 2000; base = all respondents 16+; **bold** = significant at 95% confidence level.

None of the coefficients for the 'child in household' variable are significant at the 95% level of confidence, although the parameter for the weekdays model comes quite close (0.074). Ignoring, for the moment, that these relationships fall without conventional significance, the coefficients in all three models

indicate that having young children in the household *reduces* the probability of participation in cultural activities, though this effect is apparent almost entirely during the week.

As described earlier on page 7, the effects in Table 9 are ‘main effects’, that is the effect of each variable is assumed to be constant at all levels of the other variables in the model. This is not always a realistic assumption and there is reason to believe in this instance, that the presence of children in a household may impact differentially on parents and carers of different ages. Table 10 shows the same basic model as in Table 9 but with additional terms included for the interaction of ‘children in household’ with the first four age bands.

Table 10 Logistic Regression predicting Cultural Participation III

	Average day			Weekdays			Weekends		
	B	Sig.	O.R.	B	Sig.	O.R.	B	Sig.	O.R.
V25TO34	-0.925	0.035	0.40	-1.215	0.037	0.30	-0.650	0.349	0.52
V35TO44	-0.077	0.859	0.93	0.120	0.833	1.13	-0.302	0.663	0.74
V45TO54	-0.440	0.400	0.64	0.249	0.710	1.28	-1.239	0.151	0.29
V55TO64	-3.433	0.095	0.03	-7.807	0.353	0.00	-2.017	0.351	0.13
V65TO74	-0.234	0.221	0.79	-0.205	0.370	0.82	-0.307	0.387	0.74
V75PLUS	-0.496	0.023	0.61	-0.734	0.009	0.48	-0.086	0.810	0.92
CHILD	-0.393	0.055	0.67	-0.287	0.256	0.75	-0.607	0.086	0.55
INT1	0.670	0.012	1.95	0.659	0.058	1.93	0.788	0.066	2.20
INT2	-0.058	0.839	0.94	-0.399	0.303	0.67	0.389	0.390	1.48
INT3	0.080	0.795	1.08	-0.605	0.138	0.55	0.878	0.079	2.41
INT4	1.623	0.118	5.07	3.586	0.394	36.08	1.226	0.270	3.41
Constant	-2.461	0.000	0.09	-2.694	0.000	0.07	-1.986	0.001	0.14

* Omitted category = 16-24; Source UKTUS 2000; base = all respondents 16+; **bold** = significant at 95% confidence level.

Looking first at the model for the ‘average day’, we can see that, relative to those aged 16-24, those aged 25-34 are 60% *less* likely to have taken part in any of the activities in question. However, the interaction term for this age band (INT1) is positive and significant, indicating that amongst those in this age band, having children under 16 in the household *increases* the probability of cultural participation by nearly 100%. To put this more simply, people aged 25-34 are less likely to participate in cultural activities than those aged 16-24 but only if they don’t have children. This effect is only significant at the 95% confidence level for the average day estimates, although it is only just outside this conventional cut-off point for the weekdays and weekend models. It is not apparent for any of the other age bands.

Examining group differences in cultural and sporting participation using CHAID modelling

This section approaches the question of sporting and cultural participation in a slightly different way to standard regression modelling. Dividing diary respondents into categories that differ according to their probability of participation in a sporting or cultural activity, we can see what 'natural' groupings exist and how these are defined by important socio-economic and demographic variables. Using simple classification rules that identify these natural we are able to identify the most important factors that predict sporting and cultural participation.

To illustrate the general idea, consider that it is possible to divide people into groups according to their likelihood of participating in sport. These groups might be defined initially by respondents' age, with older people less likely to participate. *Within* this age variable, we might next find that gender discriminates between individuals of the same age, and within gender socio-economic status may be important and so on. The resulting picture is of groups ordered in ascending likelihood of participating in sport. We might find, for instance, that those least likely to participate in sport are married men, over 40, with no formal qualifications working full-time in a manual occupation.

Such an analysis can be achieved using a statistical technique known as Chi-Squared Automatic Interaction Detector (CHAID). Using this method, a tree-like diagram is produced, based upon a series of discriminant analyses that segment the population according to variables that interact in predicting a specified outcome. The first step identifies the strongest statistically significant predictor of participation (or whatever else is the focus of analysis). As categorical variables are necessary for using this technique (or at least, when one uses continuous variables, a process of automatic categorisation occurs), the sample is split into the relevant number of categories, where participation differs across groups. Then, each of these groups is split into yet smaller subgroups based on additional predictors that are the most statistically significant within previously defined groups. The minimum group size is specified *a priori* by the analyst. There is no magic number in this regard but we have used 250 in all the analyses that follow because, below this size, solutions are unlikely to be particularly robust or stable.

The following analyses are of dichotomised variables from the diaries, merged with individual and household questionnaire data. We examine separately predictors of whether respondents took part in one or more sporting and cultural activities (see pages 6 and 9 for definitions of the dependent variables) on a weekday or the weekend. The following independent variables were used to form the participation likelihood groups.

It is important to keep this list in mind when examining the classifications produced, as some of the more interesting findings often involve the absence of variables that one might have intuitively anticipated would be important.

Independent variables included in the CHAID analyses

Age
Sex
Ethnicity
Socio-economic classification
Marital status
Job type
Educational qualifications
Living arrangements
Health
Household income
Car availability
Number of adults
Number of children
Tenure
Geographical region
Unemployment rate
Population density
Voluntary work
Looking after people

Beginning with participation in cultural activity during the week, Figure 1 shows that the single most important factor is economic activity. Starting from the very top of the tree, there are 8,586 individuals and of these, 6.9% took part in at least one cultural activity. All respondents are initially, therefore, sub-divided into 4 categories: 10.8% of full time students took part in one such activity; 6.9% of those unemployed, looking after the family and home, and economically inactive for other reasons; 4.0% of those working part-time, retired, or long-term sick and disabled; and 3.0% of the full-time employed (and the small amount of 'missing'). So, it seems that the more time people have on their hands (at least with relation to the amount of the day they spend working), the more they appear take part in cultural activities during the week.

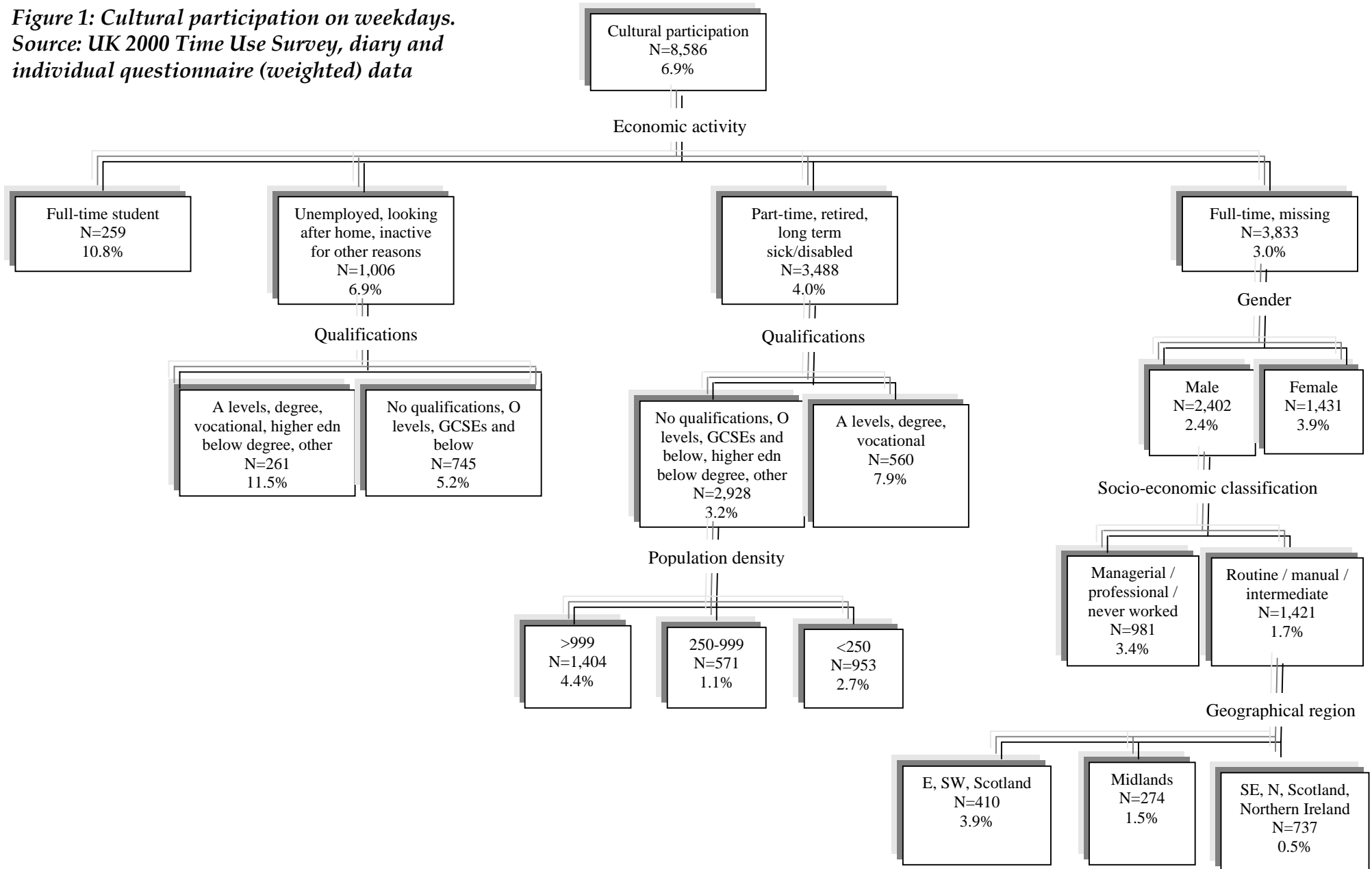
But as the tree shows, there are additional sub-groupings within three out of the four segments. While full-time students are not further divided, the second and third segments are broken up by level of qualification. Namely, having A levels or higher increases the incidence of participation. Indeed, the highest incidence group includes those who are unemployed/looking after the home, and have A levels or above (11.5% compared to the average rate of 6.9%). And there is a further branch of the tree from the lower category of educational attainment, among those who are part-time/retired/long-term

sick and disabled – looking at this set of three ‘nodes’ shows that participation has a curvilinear relationship with population density – the highest rates of participation come at the lowest and highest population densities.

Finally, the fourth set of groups. Those who are full time employed and missing¹ (3.0%), are split up by gender. Females are marginally more likely than males to participate in cultural activities within this employment status (3.9% versus 2.4%). There are two further variables significant within this male sub-group: socio-economic status (managerial/professional & never worked, 3.4%, versus ‘all other’ 1.7%); and then within ‘all other’ socio-economic status, region was a significant factor (Eastern, South West, Scotland, 3.9%, Midlands, 1.5%, London & SE, Wales, North, Northern Ireland, 0.5%). It should be noted that his last category has the lowest proportion of cultural participation – full time/male/not professional/living in South East England, the North, Wales and Northern Ireland.

¹ Missing means that no answer was provided to this question. CHAID treats missing values as an independent category rather than dropping them from analysis.

Figure 1: Cultural participation on weekdays.
 Source: UK 2000 Time Use Survey, diary and individual questionnaire (weighted) data

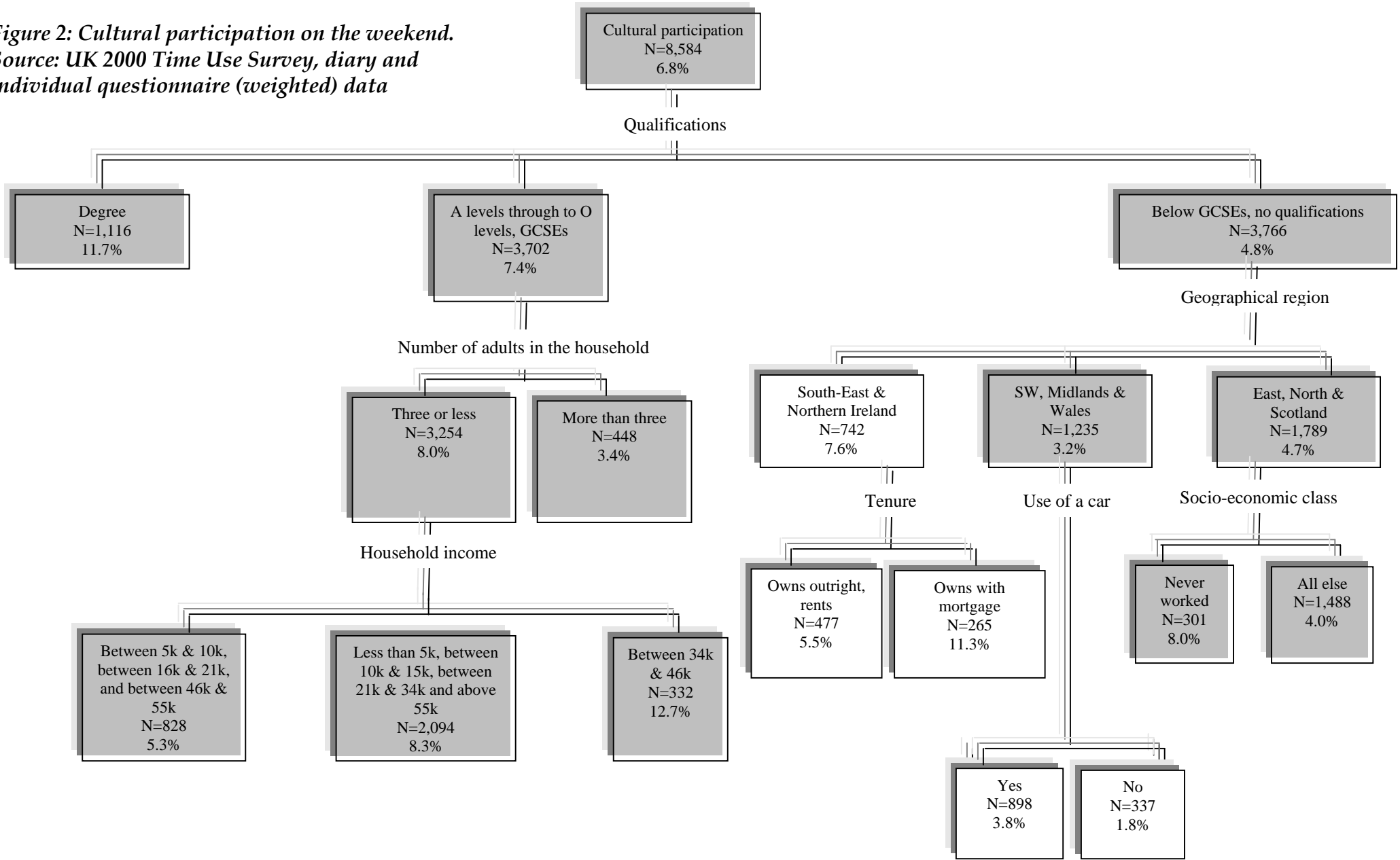


We now turn to cultural participation during the weekend, examining whether patterns differ to that found during the week. The most significant factor for weekend cultural activity is level of educational qualification. Having a degree is associated with the most activity (11.7% versus 6.8% for all respondents) – indeed, this is the second highest incidence category in the entire sample. There are no further statistically significant cleavages within this group.

The next category of qualification is GCSEs through to A levels (7.4%). This then divides into those with more than three adults in the household (3.4%) and those with three or less (8.0%). Next, within those with three or fewer adults in the household, household income is important: households earning between 34k & 46k have an incidence of 12.7% – the highest of all categories; those earning less than 5k, between 10k & 15k, between 21k & 34k and above 55k have an incidence of 8.3%; and those earning between 5k & 10k, between 16k & 21k, and between 46k & 55k have an incidence 5.3%.

Finally, of those with no qualifications or ones below GCSEs or O levels (4.8%), there are differences across region of residence. Living in South-East England and Northern Ireland means a higher level of participation (7.6%); living in East and Northern England & Scotland slightly less (4.7%); and living in South West England, the Midlands, and Wales even lower levels are apparent (3.2%). Each of these regional variables are further sub-divided by housing tenure: those who own with a mortgage (11.3%); those who own outright or rent (5.5%); socio-economic status (never worked 8.0%, all else 4.0%); and, use of a car (Yes, 3.8%, no 1.8%) respectively.

Figure 2: Cultural participation on the weekend.
 Source: UK 2000 Time Use Survey, diary and individual questionnaire (weighted) data

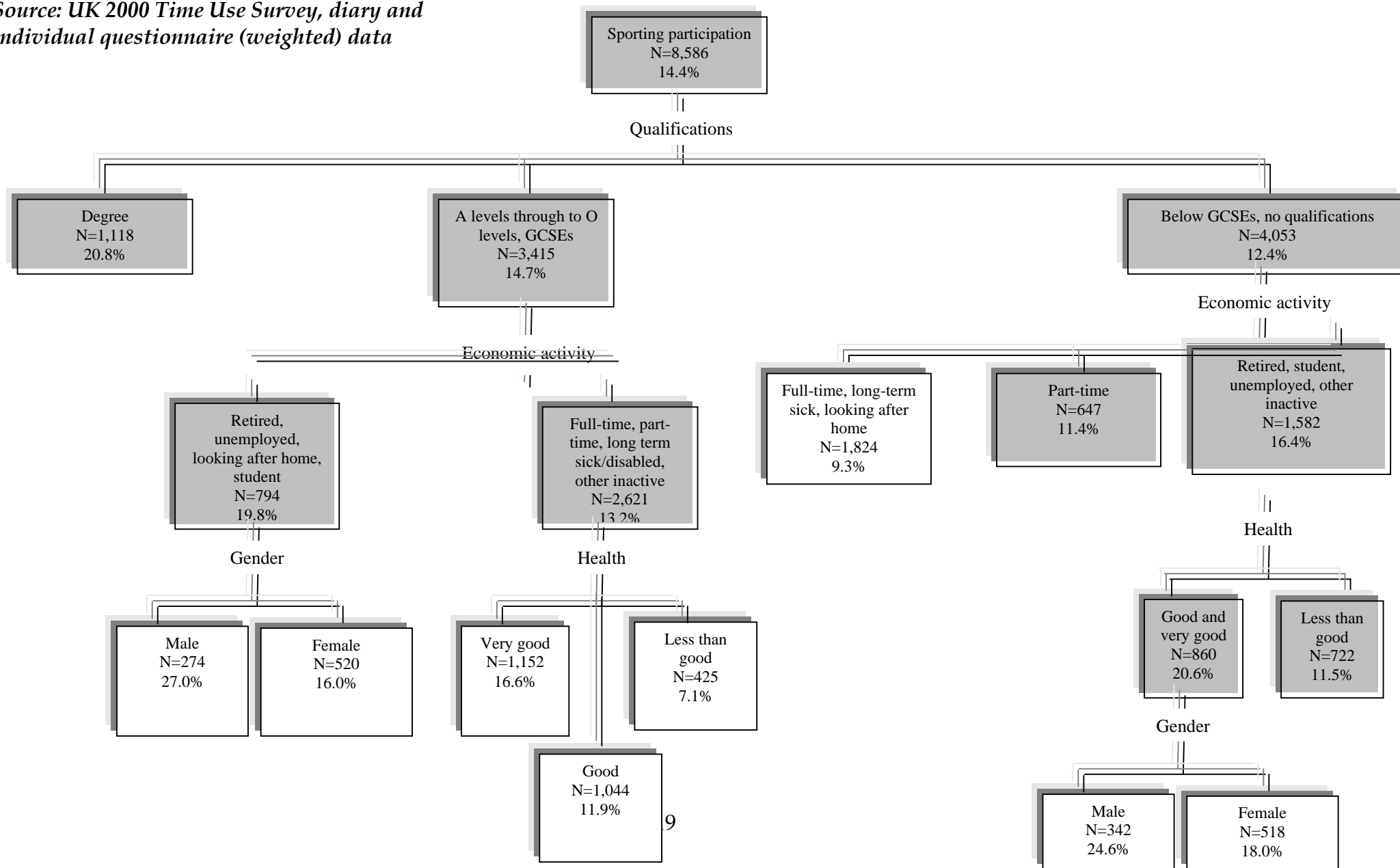


Looking now at the trees produced for sporting participation, Figure 3 shows that once again, educational qualification is an important discriminatory factor, at least for weekday sporting activity. Having a degree increases participation (20.8%) relative to attaining a level between GCSEs and A levels (14.7%), and having no qualifications or qualifications below GCSE (12.4%).

While having a degree does not sub-divide further, the other two qualification groupings do. Those who are students, retired, unemployed, or looking after the home have a higher level of participation (19.8%) than those working full-time and part-time, or being long-term sick or disabled (13.2%) for those with qualifications between GCSE and A level. Of those who are retired, unemployed, or looking after the home, men are more likely (27.0% - the single biggest group in the sample) than women (16.0%) to participate in sport during the week. The other category of economic activity also has a further sub-grouping, this time arranged related to self-reported health status: those in poorer health are less likely to participate in sport.

For respondents who have no educational qualifications, or qualifications below GCSE level, being retired, a student, or unemployed increases levels of sporting participation (16.4%), at least compared to those working part-time (11.4%) and long-term sick, looking after the home and working full-time (9.3%). Of the first group, having good health is important (20.6% versus 11.5%), and healthy males are more likely to do sport at least once during the week than healthy females (24.6% versus 18.0%).

Figure 3: Sporting participation on weekdays.
 Source: UK 2000 Time Use Survey, diary and individual questionnaire (weighted) data



Finally, we consider sporting activity over the weekend (Figure 4). Again, education is the first and most important segmenting factor to come out of the CHAID analysis, although there is a sharper contrast between 'no qualifications' and those up to A level, compared to weekday participation. But there are differences in the second level of branching - gender divides those in the middle education category, and for males, living in an area with a higher population density is associated with the highest level of sporting activity (27.8%). For females, health is important.

Next, age emerges as an important sub-grouping for those with no qualifications. Being 24 years and below has a participation rate of 24.3%, compared to 12.6%. Of those who are older than 24, males are more likely to play sport during the weekend (15.1% against 10.7%), but being between 25 and 44 does increase this likelihood. For females above 25 years, the rate of unemployment is a final sub-grouping.

Figure 4: Sporting participation on the weekend.
 Source: UK 2000 Time Use Survey, diary and individual questionnaire (weighted) data

