The sports and exercise life-course: a survival analysis of recall data from Ireland

Abstract

Recall data from a representative sample of 3080 adults in Ireland is used to investigate transitions into and out of regular participation in sports and exercise – an important contributor to overall physical activity. The method produces a continuous picture of participation across the life-course, allowing key transition periods in the life-course to be identified and the determinants of transitions to be analysed with multivariate models. Late adolescence emerges as an important period, when many people drop out from team sports, especially females. Participation in adulthood mostly involves taking up individual sports and exercise activities. The likelihood of making this transition is strongly associated with socio-economic status. Transitions in activity during adulthood do not display significant sex differences, suggesting that the gender gap for involvement in sports and exercise has its roots in childhood. The method also allows age and cohort effects to be distinguished, revealing higher participation among more recent cohorts. The findings must be interpreted carefully, since they are reliant on the accuracy of personal recall. Yet they have implications for how physical activity policy applies over the life-course, suggesting possible returns to targeting lower socio-economic groups in early adulthood, to offering a broader range of activities to young females, and to researching and promoting those activities most likely to be of interest to current young adults as they age.

The health benefits of physical activity are well established. Higher levels of activity are associated with reduced risks of coronary heart disease (Batty, 2002), stroke (Wendel-Vos, Schuit, Feskens, Boshuizen, Verschuren, Saris et al., 2004), diabetes (Jeon, Lokken, Hu & van Dam, 2007), various cancers (e.g. Tardon, Lee, Delgado-Rodriguez, Dosemeci, Albanes, Hoover et al., 2005; Monninkhof, Elias, Vlems, van der Tweel, Schuit, Voskuil et al., 2007) and improved skeletal health (Branca, 1999). Evidence for these and other health benefits has been extensively reviewed in the United States by the Physical Activity Guidelines Advisory Committee (2008), which concluded that there was very strong evidence linking physical activity to better health-related fitness, lower risk of developing disabling medical conditions and lower rates of various chronic diseases.

Increasing physical activity is a medical recommendation and a public health policy objective (e.g. World Health Organization, 2002). Physical activity policies across the developed world look to participation in sports and exercise activities as an important part of leisure-time physical activity. Public health policy can therefore gain from improved understanding of why some people participate in sports and exercise and others do not.

Determinants of participation in sports and exercise

Recognition of the link between physical activity and health has inspired a considerable research effort, part of which aims to understand the determinants of participation in sports and exercise. Participation surveys have been used to uncover the socio-economic and socio-demographic characteristics associated with participation. The standard statistical approach is cross-sectional regression, which shows that women, older people and those of lower socio-economic status are less

likely to participate in sports and exercise (e.g. Farrell & Shields, 2002; Stamm & Lamprecht, 2005). Other factors, including parental sporting involvement, health status, ethnicity, transport access and marital status also emerge as significant in a proportion of such studies. There is evidence that social and environmental factors mediate these relationships, especially the influence of socio-economic status (Cerin & Leslie, 2008). For policymakers, the implications are that interventions to increase physical activity, especially among lower socioeconomic groups, may need to be multilevel (Giles-Corti & Donovan, 2002), supplementing education on the benefits of physical activity with social and environmental supports.

A number of analyses have employed longitudinal data to assess the influence of childhood activity on later participation in sports and exercise. Reviewing longitudinal work, Malina (2001) states that the study of Cardiovascular Risk in Young Finns is comprehensive. Telama, Xiaolin, Jorma, Ilkka, Olli & Olli (2005) analyse this random sample of more than 1500 children, surveyed in 1980 and 2001. Scores on a physical activity index, based mostly on sporting activity, displayed a low to moderate correlation between the dates (0.33 – 0.44 for males, 0.14 – 0.26 for females). Correlations between childhood and adulthood sporting activity of similar magnitude have been also recorded in the Leuven Growth Study for Flemish males (Lefevre, Philippaerts, Delvaux, Thomis, Vanreusel, Vanden Eynde et al., 2000) and females (Scheerder, Thomis, Vanreusel, Lefevre, Renson, Vanden Eynde et al., 2006) and in the Amsterdam Growth and Health Study (van Mechelen, Twisk, Post, Snel and Kemper, 1999). These findings suggests that while habits formed when young do influence participation in sports and exercise in later life, transitions in participation between childhood and adulthood are likely to be important. Transitions during young adulthood appear to limit the tracking of physical activity during adulthood to

moderate levels also (Anderssen, Jacobs, Sidney, Bild, Sternfeld, Slattery et al., 1996; Barnett, Gauvin, Craig & Katzmarzyk, 2008). Although the strength of the relationship between physical activity and obesity remains contentious (e.g. Chaput, Leblanc, Pérusse, Després Bouchard & Tremblay, 2009), some longitudinal research records a relatively low correlation also between adult obesity and childhood weight, after controlling for build (Wright, Parker, Lamont & Craft, 2001).

There is less consistency regarding reported associations between changes in activity and various background characteristics. For example, Barnett et al. (2008) comment that different studies have found low socio-economic status to be associated with both increasing and decreasing activity. A possible explanation is offered here.

This paper adopts a different approach. Recall data from a representative sample of 3080 adults in Ireland is used to reconstruct individual sports and exercise histories. As well as permitting a large representative sample and avoiding attrition, the method records participation over the life-course continuously rather than as a series of snapshots. This has advantages for statistical analysis, because it invites multivariate survival techniques. The primary aim of the paper is to use such techniques to identify specific transitions in the sports and exercise life-course and to examine the determinants of these transitions. However, the results obviously rely on the validity of recall, which is therefore discussed in some detail in the next section.

Recall data

The Survey of Sport and Physical Exercise was carried out by the Survey Unit of Ireland's *Economic and Social Research Institute (ESRI)* in 2003. Fieldwork followed a pilot survey and was carried out by professional interviewers experienced in

conducting social surveys. The questionnaire obtained detailed information about sports and exercise activity, plus a standard set of background characteristics. A random sample of adults aged 18 and over was drawn from Ireland's electoral register (population 2,888,000). Interviews were conducted face-to-face at home. The response rate was 67%, giving a final sample of 3080 (1575 female, 1505 male). The sample is broadly representative (Fahey, Layte & Gannon, 2004): sex, age, geographic and socio-economic profiles accord closely with the Irish Census of 2002. Where participation rates are reported below, the sample is reweighted by sex, age and region to give exact conformity with the Census. The Irish adult (18+) population is relatively young, with a median age of 41 in 2003. It also has, by international comparison, a strong negative association between age and educational attainment.

The definition of sports and exercise activities employed is broad. Non-competitive personal exercise (e.g. swimming, going to the gym, running) is included, along with hillwalking and competitive walking (i.e. athletics), but not recreational walking (i.e. going for a walk). Informal activity with friends and family is considered equivalent to organized activity with a sports or health club.

At the beginning of the survey, respondents were informed that this definition would be used throughout and were shown a list of more than 60 sports and exercise activities that complied with it (supplied in the Appendix). A few sports on the list may be deemed to involve little physical activity (e.g. billiards/snooker, fishing). All the results below are robust to their exclusion. Respondents were asked, separately for each activity, whether they had participated during the previous twelve months. There followed an open question about any activity undertaken that was not on the list but complied with the definition. Compliance was determined during data analysis, with a

few additional activities being included, such as fencing and American football. For each activity undertaken, further information was gathered, including the frequency of participation and age at which the activity was taken up. A subsequent section of the survey presented the same list of activities again and asked, separately for each, whether the respondent had ever previously participated "on a regular basis". The term "regular" was not precisely defined in this question, but previous sections had employed a consistent definition of at least once a month. For these activities, the age at which the respondent stopped participating was obtained.

Thus, the survey responses allow the construction of individual sporting histories, recording for every year of age whether the individual was participating in sports and exercise regularly and, if so, which specific activities.

Recall failure

Complete sports and exercise histories can be constructed for 2896 (94.0%) of the sample. The remaining 6.0%, who responded "don't know" to at least one historical question relating to age of take-up or drop-out from an activity, are excluded from further analysis. This introduces a modest downward bias in participation rates, since those unable to recall are more likely to have been among the individuals (78%) who had participated at some point in their lives. Further analysis of the excluded group reveals that recall failure is not significantly related to background characteristics, save that those aged under 30 have a lower rate of failure (3.8%). Hence, the degree of downward bias in participation differs between this group and the rest of the sample. Relative to the differences explored in the following sections this differential bias is small, however, and the regression models are robust to the exclusion of those aged under 30.

Recall error

Recall data has previously been used to address various research questions in economics and sociology, including factors influencing unemployment, educational attainment and health service usage. The potential for measurement error and other possible biases has itself been examined. Dex (1991) extensively reviews earlier studies; Bound, Brown & Mathiowetz (2001) provide a meta-analysis. The conclusions from both studies are consistent: recall data is more accurate when the recall period is shorter, the activity being measured is salient, and the behaviour is habitual over a long period. Thus, on the positive side, periods of participation in sports and exercise tend to extend over years and to involve salient events. On the negative, recall is across decades. Furthermore, given the loose definition of "regular participation" in the survey, a degree of measurement error will arise in cases where the regularity of participation was marginal.

Dex (1991) also suggests that face-to-face surveys and aided recall (lists of prompts) improve accuracy. The present survey was carried out face-to-face and respondents were heavily prompted. On encountering the section on past participation, they had already been exposed to the full list of activities twice (as well as the current participation questions, a section had asked about social participation). Respondents were asked separately about each activity listed. The relevant section also obtained information about the context in which the activity usually took place (location, other participants, formal versus informal setting etc.). These questions produced fewer "don't knows" than the questions relating to starting and stopping ages, suggesting that respondents found it easier to recall participating in activities than to provide ages of taking them up and dropping out. It therefore seems unlikely that entire

participation periods are missing from the data, i.e. that some respondents stated that they had never regularly participated in an activity when, in fact, they had. The greater concern is the accuracy of the ages stated for taking up and dropping out, perhaps especially the latter, which accounts for a greater proportion of missing values. Although the distribution of ages given for taking up activities is fairly smooth, the ages supplied for dropping out after 30 years tend to clump around ages divisible by five, indicating measurement error caused by approximate recall. Preliminary analyses indicated no association between this tendency and the explanatory variables.

Sample consistency

A representative cohort of the adult population in 2003 will be a biased sample of the equivalent cohort in 1974. Some people will have migrated and others died. In addition, the 2003 sample may contain immigrants who lived elsewhere in 1974. If migration and death are correlated with participation, biases could be introduced.

Immigrants were not identified in the survey, but large-scale immigration to Ireland is a very recent phenomenon and the proportion of immigrants on the electoral register by 2003 would have been very low. The greater issue is emigration. Population outflows from Ireland peaked in the 1950s and again in the 1980s. The former wave is the more problematic, because the large majority of emigrants who left in the latter wave had returned by 2003 (Fahey, Fitz Gerald and Maître, 1998). Thus, repeating the analysis only for respondents under 50 provides a control for biases due to migration, while also controlling for biases due to survival. In the models reported below, estimated coefficients do not change significantly when the sample is limited to those aged under 50.

Background characteristics

The aim is to relate sports and exercise histories to socio-demographic and socio-economic background characteristics. Variables such as sex and year of birth can be considered fixed, but socio-economic status can change and the available explanatory variables are recorded only at the time of the survey, so the variables as measured in 2003 must act as proxies for previous values. Income, social class and (to a lesser extent) educational attainment, could differ between 2003 and previous years, although each tends to be highly correlated over time and between subsequent generations of the same family. The likely impact of this measurement error is to lessen estimated associations between the socio-economic variables and participation in sports and exercise.

Summary of data concerns

On the upside, the use of aided recall and the low prevalence of missing values offers reassurance that participation episodes are unlikely to missing from the data. Meanwhile, limiting the sample to particular age ranges permits tests for sample biases arising from differential recall failure and changes in coverage. The models are robust to these tests, suggesting that such biases are not overly strong. On the downside, recall data is likely to introduce measurement error for ages at transitions. Using variables relating to socio-economic circumstances in 2003 as proxies for previous values also implies measurement error in these variables. The likely effect is to weaken the estimated associations. Significant effects are, therefore, likely to be more solid in terms of direction than estimated magnitude.

Graphical analysis of the sports and exercise life-course

The data provide information on participation in regular sports or exercise for all individuals up to age 18, beyond which the sample begins to be reduced as a proportion are yet to reach the age in question. That is, as age increases there is a steady increase in the number of censored observations. To deal with this censoring, techniques of survival analysis are employed.

The 'sports hill'

Figure 1 provides a depiction of participation in sports and exercise across the life-course, henceforth referred to as the 'sports hill'. It is, in effect, an adapted form of the Kaplan-Meier analysis (Kaplan and Meier, 1958). Beyond age 18, as age increases, the proportion participating is calculated only from the remaining risk set, i.e. those who have reached the age in question. Thus, the similarity between the sports hill and Kaplan-Meier analysis is the method used for handling right censored observations. The difference is that individuals can change status in either direction (take-up or drop-out). For instance, more than half of all activities undertaken at age 40 were taken up after age 20. The sports hill estimates the probability of participation at each age, conditional on having reached that age, rather than conditional on continuous participation up to that age.

Figure 1 provides the sports hill for all activities, team sports and individual activities. It has a distinctive shape, peaking at age 15, when over 60% participated regularly. There is an identifiable kink at age 11, which coincides with the transition from primary to secondary school in the Irish education system. Participation exhibits a

sharp fall-off in the late teens, followed by decline throughout adulthood which is steeper in earlier adulthood and shallower in middle age.

This pattern can be better understood by considering team and individual activities separately. Although almost all sports and exercise activities can be adapted into team games, the distinction is between what are essentially team games (football, basketball etc.) and the rest, including activities based on races (cycling, running etc.), one-on-one competition (racket sports, golf etc.) or personal exercise (going to the gym). Team sports account for the peak and steep decline after age 15. Participation in individual activities does not peak until age 20 and declines only gradually. The flattening of the sports hill in middle age reflects the balance between team and individual activities, because the more rapid drop-out from team sports ceases to be a factor beyond age 40.

The sports hill by sex and educational attainment

Figure 2 provides strikingly different sports hills for males and females. Sex differences are apparent at a young age, especially for team sports. Females took up team sports later, near the beginning of secondary school, but rapidly dropped out again. This effect appears to be the primary cause of the sharp decline in overall participation in late adolescence. For individual activities the gender gap is much narrower. Males surpassed female participation only once they had begun dropping out from team sports. The gender gap for individual activities also varies across the life-course, widening in young adulthood, narrowing in the 30s, then widening again. It is possible that notable life-course events such as parenthood, menopause or retirement play a role. Of these, the data allow a test only for an effect of parenthood, since information on the age of children is available, while information relating to

menopause and retirement is not. The data do not suggest significant changes in participation associated with becoming a parent.

Figure 3 charts sports hills for three levels of educational attainment. Those with higher attainment have much higher participation rates. Separate curves for team and individual activities (not shown) reveal that the association with educational attainment is present for both types, but larger for individual activities.

This univariate analysis is instructive regarding transition points in the life-course and suggestive regarding influences, but it is also limited. The gaps in participation rates in Figure 3, for example, may be influenced by other factors correlated with age and educational attainment, such as cohort. There are also correlations between sex and socio-economic variables that could affect the participation gap in Figure 2. Nevertheless, the sports hill is very helpful for deciding what types of multivariate model might best disentangle potential determinants. Specifically, because participation as a child and as a young adult appears to be strongly related to transitions into and out of educational institutions, it makes sense to model participation either side of these transition points. The next section therefore explores the determinants of participation during youth at ages 15 and 20, i.e. at the peak of participation during secondary schooling and again once individuals have moved on. Above age 20, the shape of the sports hill implies that transitions into and out of participation occur more steadily. Given this, a survival model is appropriate to examine the determinants of taking up new activities and of dropping out.

Youth participation

Binary logistic regression is used to model the determinants of having participated regularly at ages 15 and 20. Three dependent variables are employed: all activities, team sports and individual activities. The dependent variable takes the value '1' if the respondent participated regularly at age 15 (20), '0' if not.

Independent variables

The choice of independent variables is informed by previous work, which shows that sex, age and socio-economic status are the characteristics most strongly associated with participation (see above). The models are robust to the inclusion of additional available variables (residential location, household size and composition, marital status, urban-rural area type), which are omitted for reasons of parsimony.

In previous studies, age and cohort are unavoidably confounded. That is, older people may participate less because participation declines with age or because their cohort participated less throughout life, although in the literature the effect is usually put down to age rather than cohort (e.g. Farrell & Shields, 2002). The method used here allows these effects to be distinguished. The independent variable 'cohort' corresponds to the mean age of the sample minus the age of the respondent, divided by ten. Hence, an increase of one unit equates to being born one decade later. If the impact of cohort on participation were non-linear it would be more appropriate to make this continuous variable categorical, but preliminary tests supported the assumption of linearity. Attitudes to females doing sports have changed in recent decades, so an interaction term between sex and cohort is included. Interactions

involving sex, cohort and each of the other independent variables were also tested, but were non-significant or only marginally significant and are omitted.

The survey asked respondents to indicate whether, during the years they were in secondary school, their parents participated regularly in sports and exercise. This variable turns out to be powerful and is included in all specifications.

The main variable used to indicate socio-economic status is educational attainment. A categorical social class variable is also available, but with the exception of the coefficient for the highest class, 'professional' (which enters as a dummy variable in some models), the variable is always non-significant once educational attainment and income are included. As is typical in household surveys, some respondents did not supply income information, so the inclusion of income reduces the sample. For each dependent variable, two models are reported, one that does not include income and one that does. (If, instead, a dummy variable is added for those who did not supply income information, it is not significant.) The raw household income figure is equivalized using the modified OECD income scale (1.0 for respondent, +0.5 for additional adults, +0.3 for additional children), logged to counter the skew of the distribution and divided by the inter-quartile range, such that one unit relates to the difference between the 25^{th} and 75^{th} percentile of the income distribution; more intuitively, between being moderately well off and moderately badly off.

Models for having participated at age 15

Table 1 presents estimated odds ratios for having participated regularly in sports and exercise at age 15, derived from six logistic regression models (three dependent variables with and without income specified). Considering first the impact of sex and

cohort, males are roughly three times more likely to have participated at age 15, although the effect is confined to team sports. More recent cohorts are significantly more likely to have participated at age 15. This cohort effect applies to both team and individual activities, but for team sports the odds ratio on the 'Male*Cohort' interaction term implies that it is the result of females having been more likely to participate in team sports in recent cohorts.

There is a consistent and strong association between parent's and children's activity. The effect is largest for individual activities where both parents were participants. Given the age profile of parents of schoolchildren, it is more likely that the activity undertaken by the parents (especially the mother) was also an individual one.

Educational attainment is highly significant, especially for individual activities. Those who went on to obtain postgraduate qualifications were some three times more likely to have participated than those who obtained only lower second-level qualifications. Introducing the income variable, which is itself significant, moderates the influence of educational attainment somewhat, but its association with participation remains strong for individual activities. Given that in most cases these socio-economic indicators relate to an individual's status several decades after the behaviour in question, the extent of this socio-economic influence on participation is striking.

Models for having participated at age 20

Table 2 provides similar analysis at age 20. Males were over four times more likely to participate at age 20; ten times in relation to team sports, although the interaction term suggests this large gender gap narrowed slightly in later cohorts. These odds ratios are consistent with and help to quantify the univariate analysis of Figure 2. Many

individuals of both sexes gave up sports and exercise between ages 15 and 20, but the rate of dropout was more severe for females, resulting in a much widened gender gap by age 20.

The influence of cohort is weaker at age 20. This suggests that participation during the secondary school years rose in recent decades, but that the increase did not entirely feed through to participation after leaving school.

There remains a significant influence of coming from a more active family, although the impact of physically active mothers is diminished relative to age 15, except where both parents were active. One potential hypothesis might relate this sex-specific change in parental influence to the widening gender gap in participation just described, but interactions between the sex of the respondent and the pattern of parent's activity are non-significant (not shown).

The socio-economic variables have a stronger association with participation than at age 15. This finding should be treated cautiously, however, because socio-economic indicators recorded at the time of the survey are likely to reflect status at age 20 more accurately than status at age 15. Belonging to a professional occupation (a variable omitted from the regressions for age 15, where it is non-significant) is also significant for participation at age 20, although its interaction with sex implies that the effect applies only to females.

Take-up and drop-out as adults

The likelihood of participation in sports and exercise activities is highest among teenagers and young adults. Following this peak in activity, two flows determine the level of participation during adulthood: the rate of taking up activities and the rate of

dropping out. A proportional hazards method, Cox regression (Cox, 1972), is used to examine these transitions.

The advantage of Cox's proportional hazards method is that it does not require the probability distribution of status changes over time to be known. This is helpful for transitions not addressed by extensive previous research (Hosmer & Lemeshow, 1999), as with participation in sports and exercise, but requires that the proportional hazards assumption be satisfied. For all models reported, this condition is tested for each covariate using the method of Grambsch & Therneau (1994) and Hosmer & Lemeshow (1999), whereby a time-dependent interaction term is tested for significance. (An illustration is described below.)

Who took up sports and exercise activities?

Table 3 presents relative rates of risk (the exponents of the coefficients in Cox regression) for having taken up an activity after age 20. The sample includes both those already participating at age 20 and those not, with initial status itself entering as a covariate categorized by type of activity. In addition to the explanatory variables introduced previously, car ownership is significant and is included in the specification. Some 15% of Irish adults do not have access to a car and they were less likely to take up an activity.

Sex is non-significant. A larger sample might find a gender difference, i.e. the finding may be a type II error, but the estimated influence of sex is anyway far less than that of cohort or socio-economic factors. People born one decade later are estimated to have taken up sports and exercise activities at a 30% faster rate. Across models (1) and (2), high educational attainment, income, professional status and car ownership

are strong and significant determinants of taking up an activity. Since all these variables are positively correlated, a typical individual with a high socio-economic position was several times more likely to have taken up an activity over a given period.

Initial participation status at age 20 is also a significant. Those already participating in one type of activity took up new activities at a faster rate, whether the existing activity was a team or individual one. The relative risk is not as high for those who were already participating in both types, which probably reflects a level of saturation in participation. Once initial status is controlled for, having had parents who participated is non-significant.

Model (3) shows a test of the proportional hazards assumption, where the covariate of interest is educational attainment. This example illustrates the diagnostic test, but also makes the telling point that the impact of educational attainment on participation extends beyond years spent at college. The proportional hazards assumption requires that the influence of each covariate on the transition in question remains proportionately the same over time. One might hypothesize that individuals who went to college would have been especially likely to take up an activity in their early twenties, but that this advantage would have diminished relative to other factors in later adulthood. This hypothesis implies that the relative risk on the interaction term 'Third-level*Ln(year)' should be significantly less than one and, if confirmed, would render the proportional hazards model invalid. In fact, the interaction terms for educational attainment are all non-significant. Furthermore, this finding applies to all other covariates in the model. The impact of educational attainment is lasting and the proportional hazards model seems appropriate.

Who dropped out?

The technique is next applied to having dropped out from sports and exercise after age 20, with two changes. First, preliminary work revealed that the sex variable violates the proportional hazards assumption ($p<0.00l$). This result is unsurprising, because the sports hills for males and females (Figure 2) not only display contrasting levels but contrasting shapes. Consequently, separate models of drop-out were estimated for males and females, for which the diagnostic test of the proportional hazards assumption holds. This separation exacerbates the second difference with respect to the take-up models, namely that because half of the sample was not participating at age 20, the drop-out models apply to smaller samples, reducing the likelihood of establishing statistical significance.

Table 4 provides relative rates of risk for having dropped out after age 20. This time, there are no significant effects of cohort. Contrastingly, there is a continuing impact of parent's participation. Individuals with two active parents, especially females, dropped out significantly more slowly. Socio-economic effects are important but less apparent than for taking up activities, although income is significant for males. But by far the biggest factor in whether an individual dropped out is the activity they were undertaking at age 20. Participants in only team sports dropped out at three to four times the rate for males and eight times the rate for females, compared with people who also participated in individual activities (or, in the case of males, only participated in individual activities). This is interesting given the findings for taking up activities above. Relative to someone doing no activities at age 20, an individual doing team sports was somewhat more likely to take up another activity, almost invariably an individual one. But relative to someone who already participated in an

individual activity, they were no more likely to take up another activity and very much more likely to drop out.

Discussion

The analysis identifies key transitions in the sports and exercise life-course and provides models for the determinants of these transitions. Among the various significant associations, perhaps the most notable are: the importance of the period around end of secondary schooling, when many activities are dropped; the change in the relationship between sex and participation that occurs around this time; and the existence of a strong cohort effect. Each has potential policy implications.

Many individuals who continue regular physical activity progress from team sports to individual sports and exercise activities in late adolescence. The likelihood of making this transition and hence of continuing to participate in later adulthood is strongly linked to background characteristics, especially socio-economic status. This may partly explain the relatively low correlation between physical activity as a child and as an adult (e.g. Telama et al., 2005). The different sporting life-courses by social group may also explain the contrasting results across longitudinal studies regarding the effects of various background characteristics on the likelihood of increased or decreased activity (e.g. Barnett et al., 2008). The precise ages, especially around adolescence, at which activity is recorded will have large consequences for the relationships estimated when comparing repeated snapshots of participation.

While income matters, educational attainment seems to be a particularly strong and enduring determinant of participation, especially of taking up individual activities. A number of separate processes could be behind this relationship. Those who obtain

higher levels of educational attainment may be more likely to have parents with resources, more likely to have parents who themselves participated, more likely to attend better schools, more likely to have social and environmental supports for physical activities and, by remaining in full-time education for longer, more likely to have convenient and subsidized access to sports and exercise opportunities in early adulthood. These are indirect effects. Direct effects are also possible: better educated individuals may be more likely to understand the benefits of sports and exercise.

Policymakers focus considerable effort on trying to get children from all social groups involved in sports and exercise, in the hope that this will instil good physical activity habits across the socio-economic spectrum. The analysis provided here suggests that while there may be a pay-off to such policies, the pay-off is likely to be reduced by transitions that occur in young adulthood. There is therefore a good case for seeking effective interventions at this later point in the life-course, especially interventions targeted at lower socio-economic groups. For example, given the strong influence of continuing education, policies targeting participation opportunities at recent school-leavers merit consideration.

The findings also suggest that the relationship between participation and sex changes between childhood and adulthood. The gender gap in childhood is not consistent across activities, but mostly concerns team sports. Sex differences in the rate of take up as an adult appear to be small or non-existent, while males drop out from activities faster than females during early adulthood. More simply, participation in sports and exercise activities moves closer to gender equality once individuals enter an adult environment, where presumably they have greater choice and autonomy with respect to leisure-time activities. The corollary of this conclusion is that females face a

physical activity disadvantage as children, albeit one that has lessened somewhat in recent years. Their fleeting involvement with team sports suggests a relatively poor return to efforts to involve females in such activities. Opportunities for involvement in a broader range of activities might help to make physical activity among young women more popular and enduring.

The findings reveal that participation in sports and exercise, both as children and adults, is higher in more recent cohorts, at least in Ireland. Thus, the declining participation by age usually found in cross-sectional analysis may to a significant degree reflect a *cohort* rather than *age* effect. This might surprise those who highlight rising levels of obesity, including in Ireland (National Taskforce on Obesity, 2005). Yet increased participation in sports and exercise does not necessarily imply increased physical activity overall. Other contributing factors include modes of transport, workplace activity, labour saving devices and so on. The suggestion is only that recent generations of young adults have higher participation in sports and exercise and, therefore, may continue to participate more in later in life also. For policymakers, maintenance of this increased level of participation into middle age and beyond may have a significant public health return. Analysis and promotion of those activities most likely to appeal to current young adults as they age could contribute successfully to a long-term physical activity policy.

The strength of the recall data and method employed here is that it provides a continuous picture of the sports and exercise life-course, which cross-sectional and even longitudinal datasets generally cannot. Nevertheless, the potential weaknesses of the method must not be forgotten. The analysis covers only a subset of physical activity; the most important omission being recreational walking. It is possible that

the findings are unique to Ireland, although perhaps unlikely, since the cross-sectional patterns are very similar to those recorded elsewhere (e.g. Farrell & Shields, 2002; Cerin & Leslie, 2008). Most importantly, the method relies entirely on the accuracy of individual recollections of sports and exercise activity. Participation rates and estimated associations are hence likely to contain measurement error and must be considered approximate. Future research could perhaps establish the validity of recall data in this area, by incorporating recall information in later waves of longitudinal studies.

References

Anderssen, N., Jacobs, D.R.J., Sidney, S., Bild, D.E., Sternfeld, B., Slattery, M.L., & Hannan, P. (1996). Change in secular trends in physical activity patterns in young adults: a seven-year longitudinal follow-up in the Coronary Artery Risk Development in Young Adults Study (CARDIA). *American Journal of Epidemiology, 143*(4), 351-362.

Barnett, T.A., Gauvin, L., Craig, C.L., & Katzmarzyk, P.T. (2008). Distinct trajectories of leisure time physical activity and predictors of trajectory class membership: a 22 year cohort study. *International Journal of Behavioural Nutrition and Physical Activity, 5:57*.

Batty, G. (2002). Physical activity and coronary heart disease in older adults. *European Journal of Public Health, 12*(3), 171-176.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In J.J. Heckman, & E. Leamer (Eds.), *Handbook of Econometrics, Volume 5* (pp.3705-3843). Elsevier.

Branca, F. (1999). Physical activity, diet and skeletal health. *Public Health Nutrition, 2*, 391-396.

Cerin, E., & Leslie, E. (2008). How socio-economic status contributes to participation in leisure-time physical activity. *Social Science & Medicine, 66*(12), 2596-2609.

Chaput, J.-P., Leblanc, C., Pérusse, L., Després, J.-P., Bouchard, C., & Tremblay, A. (2009). Risk Factors for Adult Overweight and Obesity in the Quebec Family Study: Have We Been Barking Up the Wrong Tree? *Obesity*. Advance online publication: http://www.nature.com/oby/journal/vaop/ncurrent/abs/oby2009116a.html

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society: Series B, 34*, 187-220.

Dex, S. (1991). *The Reliability of Recall Data: A Literature Review*. Working papers of the ESRC Research Centre on Micro-social Change, Paper 11, Colchester: University of Essex.

Fahey, T., Fitz Gerald, J., & Maître, B. (1998). The Economic and Social Implications of Demographic Change. *Journal of the Statistical and Social Inquiry Society of Ireland, XXVII*(V), 185-222.

Fahey, T., Layte, R., & Gannon, B. (2004). *Sports Participation and Health Among Adults in Ireland.* Dublin: Economic and Social Research Institute.

Farrell, L., & Shields, M. A. (2002). Investigating the economic and demographic determinants of sporting participation in England. *Journal of the Royal Statistical Society: Series A, 165*(2), 335-348.

Giles-Corti, B., & Donovan, R. (2002). The relative influence of individual, social and physical environment determinants of physical activity. *Social Science & Medicine, 54*, 1793-1812.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, *81*, 515-526.

Hosmer, D. W., & Lemeshow, S. (1999). *Applied Survival Analysis: Regression Modeling of Time to Event Data.* New York: John Wiley & Sons.

Jeon, C., Lokken, R., Hu, F., & van Dam, R. (2007). Physical activity of moderate intensity and risk of type 2 diabetes: a systematic review. *Diabetes Care, 30*(3), 744-752.

Kaplan, E., & Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*, 457-481.

Lefevre, J., Philippaerts, R.M., Delvaux, K., Thomis, M., Vanreusel, B., Vanden Eynde, B., Claessens, A.L., Lysens, R., Renson, R., & Beunen, G.P. (2000). Daily Physical Activity and Physical Fitness From Adolescence to Adulthood: A Longitudinal Study. *American Journal of Human Biology*, *12* 487-497.

Malina, R.M. (2001). Tracking of physical activity across the lifespan. Washington, DC: President's Council on Physical Fitness and Sports.

Monninkhof, E.M., Elias, S.G., Vlems, F.A., van der Tweel, I., Schuit, A.J., Voskuil, D.W., & van Leeuwen, F.E. (2007). Physical activity and breast cancer: a systematic review. *Epidemiology*, *18*(1), 137-157.

National Taskforce on Obesity (2005). *Obesity: The Policy Challenges*. Dublin: Department of Health and Children.

Physical Activity Guidelines Advisory Committee (2008). *Physical Activity Guidelines Advisory Committee Report, 2008*. Washington, DC: U.S. Department of Health and Human Services.

Scheerder, J., Thomis, M., Vanreusel, B., Lefevre, J., Renson, R., Vanden, E.B., & Beunen, G.P. (2006). Sports Participation Among Females From Adolescence to Adulthood. *International Review for the Sociology of Sport*, *41*(3-4), 413-430.

Stamm, H., & Lamprecht, M. (2005). Structural and cultural factors influencing physical activity in Switzerland. *Journal of Public Health*, *13*(4), 203-211.

Tardon, A., Lee, W., Delgado-Rodriguez, M., Dosemeci, M., Albanes, D., Hoover, R., & Blair, A. (2005). Leisure-time physical activity and lung cancer: a meta-analysis. *Cancer Causes Control, 16*, 389-397.

Telama, R., Xiaolin, Y., Jorma, V., Ilkka, V.Ã., Olli, W., & Olli, R. (2005). Physical activity from childhood to adulthood: A 21-year tracking study. *American Journal of Preventive Medicine, 28*(3), 267-273.

van Mechelen, W., Twisk, J.W.R., Post, G.B., Snel, J., & Kemper, H.C.G. (1999). Physical activity of young people: the Amsterdam Longitudinal Growth and Health Study. *Medicine & Science in Sport & Exercise, 32*(9), 1610-1616.

Wendel-Vos, G.C., Schuit, A.J., Feskens, E.J., Boshuizen, H.C., Verschuren, W.M., Saris, W.H., & Kromhout, D. (2004). Physical activity and stroke. A meta-analysis of observational data. *International Journal of Epidemiology, 33*(4), 787-798.

World Health Organization (2002). *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*. Geneva: World Health Organization.

Wright, C.M., Parker, L., Lamont, D., & Craft, A.W. (2001). Implications of childhood obesity for adult health: findings from thousand families cohort study. *British Medical Journal, 323* 1280-1284.

Appendix

The following is the list of sports and exercise activities with the headings and in the order they were presented to survey respondents.

Indoor Games: Badminton; Basketball; Billiards; Snooker; GAA Handball; Olympic Handball; Racquetball; Squash; Table Tennis; Tenpin Bowling; Volleyball; Netball.
Field Games: Baseball; Softball; Cricket; Rugby; Hockey; Gaelic Football; Hurling; Ladies Football; Camogie; Soccer; Soccer – five-a-side.
Water Sports: Canoeing; Kayaking; Rowing; Sailing/Yachting; Sub Aqua; Surfing; Swimming; Water Skiing; Wind Surfing.
Gym-based Activities: Aerobics/Keep Fit; Boxing; Martial Arts/Judo/Karate; Weight Lifting; Wrestling; Gymnastics.
Outdoor Pursuits: Archery; Athletics; Bowling – lawn; Bowling – road; Cycling – for leisure; Cycling – for sport; Cycling – for transport; Equestrian; Fishing; Golf; Hang Gliding; Paragliding; Cross Country; Jogging; Motor Racing; Mountaineering/Rock Climbing; Orienteering; Pitch & Putt; Roller Blading/Skating; Shooting – target; Shooting – clay pigeon; Tennis; Triathlon; Tug-of-War; Hillwalking; Winter Sports

Figure captions

Figure 1: The “sports hill” – proportion engaged in regular sports and/or exercise activity at each age (of those who had reached that age)

Figure 2: Sports hills by sex

Figure 3: The sports hill by educational attainment

Table 1: Odds ratios for participation in sports and exercise at age 15, estimated by logistic regression. The dependent variable in columns (1) and (2) is participation in all types of activity, with income excluded and included in the model respectively. Columns (3) and (4) repeat the analysis for team sports; columns (5) and (6) for individual activities.

	Exp(β)					
	All activities		*Team sports*		*Individual activities*	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Male	2.86***	3.02***	3.65***	3.67***	0.91	0.89
Cohort	1.31***	1.31***	1.36***	1.36***	1.13***	1.11**
Male*Cohort	0.82***	0.84***	0.72***	0.74***	1.00	1.06
Parent's participation (Ref: Neither)						
Father only	2.11***	2.05***	1.48***	1.54***	1.56***	1.42***
Mother only	1.89**	1.70**	1.65**	1.83**	1.50*	1.30
Both participated	2.32***	2.58***	1.26*	1.27	2.56***	2.53***
Educational attainment (Ref = Lower 2nd level)						
No qualifications	0.58***	0.64***	0.54***	0.60***	0.61***	0.68**
Higher 2nd level	1.79***	1.77***	1.31**	1.26*	1.77***	1.88***
Third-level	2.02***	1.76***	1.33**	1.18	2.55***	2.49***
Postgraduate	2.53***	2.02***	1.21	0.95	3.24***	2.97***
Income		1.30***		1.19**		1.30***
Constant	0.60***	0.11**	0.36***	0.13***	0.27***	0.05**
N	2860	2189	2860	2189	2860	2189
-2LL	3258.6	2451.9	3497.8	2659.8	3271.0	2459.2
Nagelkerke R^2	0.24	0.26	0.19	0.20	0.17	0.19
Hosmer-Lemeshow (p-value)	0.60	0.78	0.49	0.30	0.70	0.98

(* p < 0.10, ** p < 0.05, *** p < 0.01)

Table 2: Odds ratios for participation in sports and exercise at age 20, estimated by logistic regression. The dependent variable in columns (1) and (2) is participation in all types of activity, with income excluded and included in the model respectively. Columns (3) and (4) repeat the analysis for team sports; columns (5) and (6) for individual activities.

	Exp(β)					
	All activities		*Team sports*		*Individual activities*	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Male	4.61***	4.33***	11.02***	10.95***	1.74***	1.60***
Cohort	1.12***	1.13***	1.27***	1.23***	1.11**	1.12**
Male*Cohort	0.92	0.91	0.80***	0.81***	0.98	1.00
Parent's participation (Ref: Neither)						
Father only	1.85***	1.77***	1.57***	1.55***	1.53***	1.40***
Mother only	1.32	1.02	1.29	1.07	1.40	1.07
Both participated	2.49***	2.47***	1.34*	1.34*	2.31***	2.13***
Educational attainment (Ref = Lower 2nd level)						
No qualifications	0.75**	0.88	0.78	0.81	0.77	0.89
Higher 2nd level	1.80***	1.80***	1.23	1.11	2.04***	2.12***
Third-level	2.85***	2.61***	1.32*	1.11	3.38***	3.36***
Postgraduate	3.96***	3.37***	1.46*	1.28	4.29***	3.93***
Income		1.20**		1.05		1.27***
Professional	1.46**	1.26	2.01***	2.04***	1.31*	1.10
Male*Professional	0.61**	0.81	0.53***	0.54**	0.58***	0.77
Constant	0.26***	0.08***	0.07***	0.13***	0.19***	0.04***
N	2725	2118	2725	2118	2725	2118
-2LL	3192.0	2476.6	2666.8	2084.2	3138.1	2429.5
Nagelkerke R^2	0.26	0.26	0.28	0.29	0.18	0.19
Hosmer-Lemeshow (p-value)	0.25	0.13	0.72	0.66	0.51	0.63

(* p < 0.10, ** p < 0.05, *** p < 0.01)

Table 3: Relative risk of taking up a sport or exercise activity after age 20, estimated by Cox regression. Column (1) presents main effects; column (2) additionally controls for income; column (3) adds an interaction between educational attainment and time.

	Exp(β)		
	(1)	*(2)*	*(3)*
Male	1.10	1.10	1.10
Cohort	1.31***	1.30***	1.31***
Parent's participation (Ref: Neither)			
Father only	1.06	1.00	1.06
Mother only	0.81	0.80	0.82
Both participated	1.13	1.07	1.13
Educational attainment (Ref = Lower 2nd level)			
No qualifications	0.83	0.90	0.75
Higher 2nd level	1.17	1.07	0.97
Third level	1.63***	1.46***	1.68**
Postgraduate	2.00***	1.53**	2.94***
No car	0.55***	0.55***	0.55***
Professional	1.30***	1.28**	1.30***
Participation history (Ref = Never participant)			
Dropout	0.90	1.29*	1.11
Team only	1.57***	1.87***	1.73***
Individual only	1.47***	1.92***	1.63***
Both	1.24	1.47**	1.37
Income		1.43***	
No qualifications*Ln(year)			1.05
Higher 2nd level*Ln(year)			1.10
Third-level*Ln(year)			0.99
Postgraduate*Ln(year)			0.79
N	2,625	2,049	2,625
Event (took up sport)	666	538	666

(* p < 0.10, ** p < 0.05, *** p < 0.01)

Table 4: Relative risk of dropping out from sports and exercise activities after age 20, estimated by Cox regression. In columns (1) and (2), the model applies to males only, with income excluded and included respectively; in columns (3) and (4), females only.

	Exp(β)			
	Male		*Female*	
	(1)	*(2)*	*(3)*	*(4)*
Cohort	0.93	0.96	1.05	1.07
Parent's participation (Ref: Neither)				
Father only	0.81	0.85	1.06	1.26
Mother only	1.13	0.87	0.45	0.35
Both participated	0.71*	0.62**	0.40***	0.40***
Educational attainment (Ref = Lower 2nd level)				
No qualifications	0.98	0.90	0.76	0.82
Higher 2nd level	0.97	1.03	0.77	0.73
Third-level	0.72*	0.92	0.88	0.85
Postgraduate	0.43***	0.49**	0.52*	0.37**
Type of sport at age 20 (Ref = both)				
Team only	3.33***	3.89***	8.00***	7.76***
Individual only	1.01	0.97	1.67	1.58
Income		0.68***		1.12
N	860	685	492	381
Event (dropped out)	343	288	187	143

(* p < 0.10, ** p < 0.05, *** p < 0.01)

Figure 1

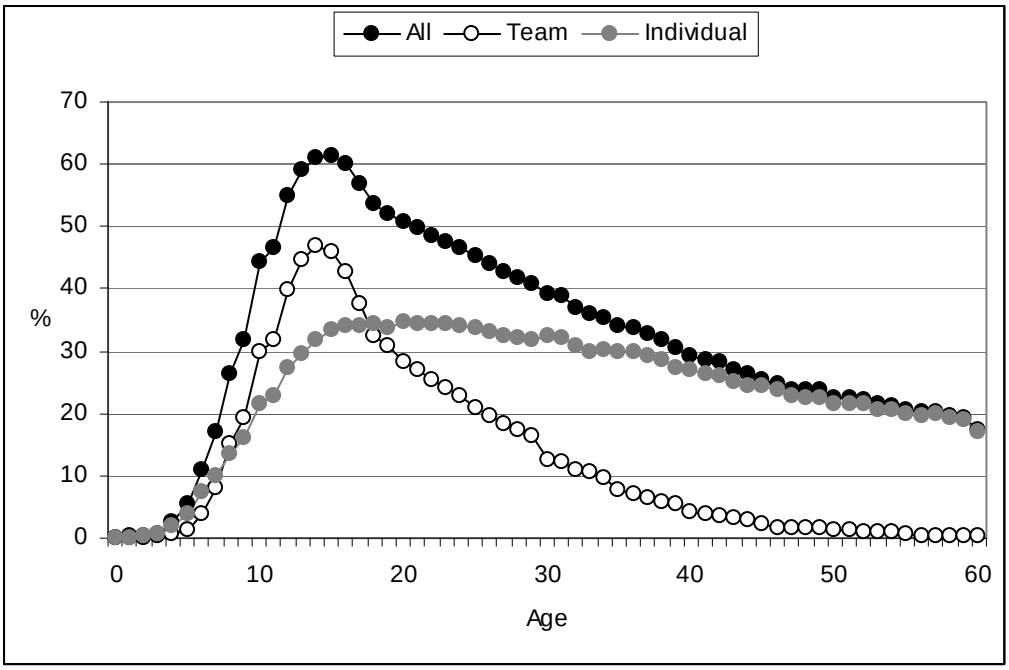


Figure 2

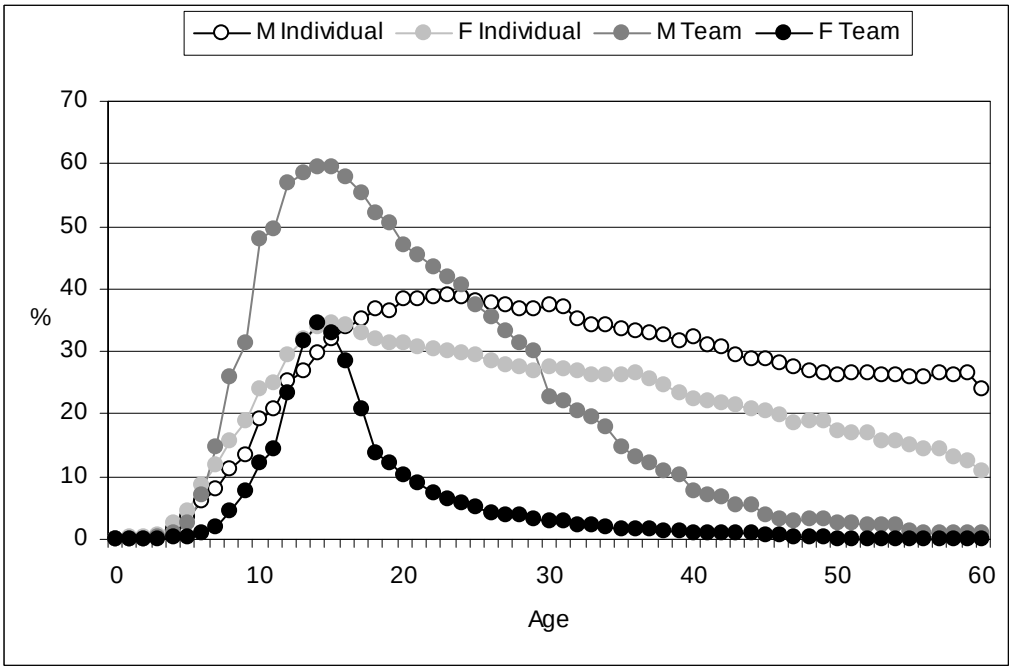
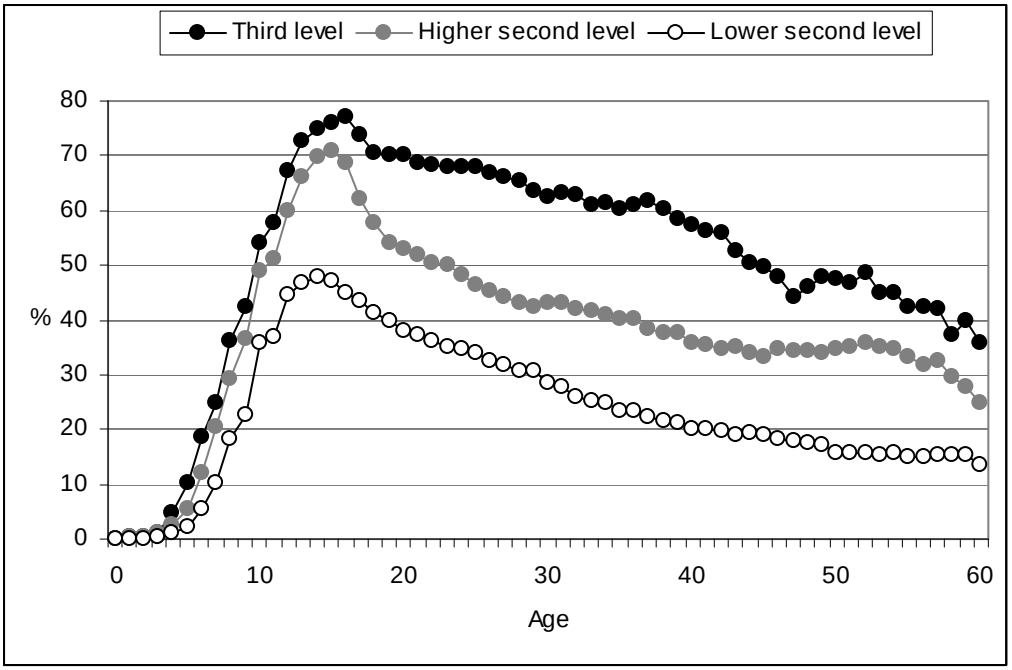


Figure 3


Third level
Higher second level
Lower second level
80
70
60
50
% 40
30
20
10
0
0
10
20
30
40
50
60
Age