SOCIAL INTERACTIONS AND THE DEMAND FOR SPORT: AN ECONOMIC ANALYSIS

PAUL DOWNWARD and JOSEPH RIORDAN*

This paper explores the decision to participate in sports activities in the United Kingdom and the subsequent frequency of participation. The paper draws links between economic and other theories of social interaction to motivate the discussion and links these theories to assessing policy initiatives in the United Kingdom. Cluster analysis is combined with a Heckman analysis to examine the empirical evidence provided by the General Household Survey in 2002. The results suggest that social and personal capital are of paramount importance in determining sports participation, and consequently, it is these features that policy should focus upon. However, the legitimacy of policy activism requires philosophical justification. (*JEL* B41, C2, D11, D12)

I. INTRODUCTION

The promotion of mass participation in sport, as a form of physical activity, is now firmly on the public policy agenda in the United Kingdom and elsewhere.[1] The health and well-being of citizens form part of popular discourse, evidenced by repeated references to "obesity" epidemics in the media and indicated by the establishment of new policies, policy agents, or a refocusing of previous efforts to address this issue. For example, in the United Kingdom, a new central government Minister for Public Health has been established to work in partnership with the Department for Culture, Media and Sport, the Department for Communities and Local Government, the Department for Transport, the Department for Education and Skills, and sports delivery bodies to raise participation. This is indicative of a more general pattern in most economies though tensions in policy priorities exist (see, e.g., Downward et al., forthcoming; Green and Houlihan, 2005; Houlihan, 1997).

The purpose of this paper is to assess such policy initiatives by an explicit focus upon economic theory and subsequent empirical investigation. In the next section, a brief resume of the policy context in the United Kingdom is provided. Section III then reviews the main elements of the theoretical and empirical literature on sports participation. It is shown, using an elementary model based on Becker (1974), how investment in personal consumption capital and social capital, through social interactions, can conceptually account for lifestyle and complementary consumption in sports.

Section IV then presents details of the data and variables used in the analysis. Sections V

*An earlier version of this paper was presented at the International Association of Sports Economics Conference at Bochum in May 2006. We are grateful for comments received at this presentation. We are also grateful for the comments of two anonymous referees and the editor of this journal for their helpful and constructive comments, which have improved the exposition of this paper. Data for this study were made available through the Economic and Social Research Council UK Data Archive, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, United Kingdom. The results and interpretations of the data are solely the responsibility of the authors.

Downward: Senior Lecturer, Institute of Sport and Leisure Policy, School of Sport and Exercise Sciences, Loughborough University, Leicestershire LE11 3TU, United Kingdom. Phone +44 (0) 1509 226365, Fax +44 (0) 1509 223935, E-mail p.downward@lboro.ac.uk

Riordan: Senior Lecturer, AFE Department, University of Hertfordshire, De Havilland Campus, Hatfield, AL10 9AB, United Kingdom. Phone 0044 (0)1707 285539, Fax 0044 (0)1707 285455, E-mail j.g.j.riordan@herts.ac.uk

1. There are a variety of definitions of sport that emanate from particular literatures, that is, embracing sociological, psychological, philosophical, and economic concepts. Some discussion of these definitions is provided in Downward et al. (forthcoming). For current purposes, sports are viewed as activities involving physical activity and undertaken for recreational, that is, nonobligated, purposes in a formal or an informal setting and accepted as such in policy and public discourse. In this regard, the sports monitored in official data meet this definition.

ABBREVIATION

GHS: General Household Survey

Contemporary Economic Policy (ISSN 1074-3529)
Vol. 25, No. 4, October 2007, 518–537
doi:10.1111/j.1465-7287.2007.00071.x

and VI then discuss the twofold quantitative approach used in the paper and present relevant results. Section VII discusses the policy implications of the paper. Conclusions then follow.

II. UK SPORTS PARTICIPATION POLICY

In the case of the United Kingdom, lying below central government agencies, UK Sport and four sports councils, "Sport England," "sportScotland," "The Sports Council for Wales," and "The Sports Council for Northern Ireland," deliver public policy initiatives and allocate resources. As Carter (2005) indicated, there is a blurred distribution of functions and funding between the agencies. However, broadly speaking, UK Sport focuses on elite sport and the other councils' mass participation in devolved regions. Sport England (2004a) has provided a conceptual framework to explain changes in participation in sport, which is summarized in Table 1, and identifies "drivers" and "settings" of change.

The intuition behind these drivers and settings is simple. For example, as one ages or works longer hours or volunteers less to provide support for sport, then participation will fall either through individual choice or, in the latter case, because of supply constraints. Not surprisingly, moreover, such changes will be mediated through decision-making social environments such as the home, workplace, or educational environment.

This raises two important and related points for this paper. The first is that, conceptually speaking, Sport England (2004b) recognized that the preferences of individuals are, at least partially, endogenous. Moreover, they recognized that the context of decisions, or equivalently the social structures surrounding agents, affects the choices of agents in return. The second important point is that economic argument does not play an explicitly prominent role and the evidence base is also primarily descriptive. Consequently, this paper seeks to examine mass participation in sport in the United Kingdom but to make the economic theoretical issues at stake more transparent and, importantly, subject them to statistical testing by a large-scale data set.

TABLE 1
Sport England Analysis of Determinants of Participation

Seven Drivers of Change	Five Settings for Change
Aging population	The home
Time pressures	The community
Well-being and obesity	The workplace
Levels of investment	Higher and further education
Utilizing education	Primary and secondary school
Variations in access	
Volunteers and professionals	

Source: Sport England (2004b, p.10).

III. THEORETICAL AND EMPIRICAL LITERATURE REVIEW

A. Theories of Sports Participation

The economic theory that has been employed to explain sports participation can be understood as drawing upon two traditions: an orthodox, neoclassical perspective and a heterodox or wider social science perspective (Downward et al., forthcoming). In both of these contexts, however, the analysis of sport has tended to be indirect and even reflects different general theoretical concerns about how to model decision making.[2] For example, in the latter case, contributions from Scitovsky (1976) and Earl (1983, 1986) are drawn upon by Gratton and Tice (1991) to explore the psychological foundations of consumer choice in sport and, in particular, learning by doing. Post-Keynesian consumer analysis (see, e.g., Lavoie, 1994) also draws upon this concept and also insights from the studies of leisure by Veblen (1925) and Galbraith (1958) and by implication (Bourdieu, 1984, 1988, 1991) that individual preferences are shaped by social values. The predictions of such theories would be that prior experience in sports activities is likely to raise participation in any specific activity and that social interactions, or lifestyles, will also affect participation.

In contrast, the main contribution to understanding sports participation from a neoclassical economics perspective has been to apply the income-leisure trade-off model

2. Space precludes a full discussion of these distinctions; consequently, attention is focused upon the neoclassical theories. It is worth noting, however, that the differences between these and the heterodox theories are quite subtle. The theories share the same predictions, but the latter emphasizes different methodological and philosophical underpinnings and consequently normative or policy implications.

of labor supply (e.g., see Gratton and Taylor, 2000). In this context, leisure, that is, sport, is defined as the dual of work, the latter of which provides income for consumption.

A more comprehensive neoclassical theoretical foundation for the analysis of sports participation can be constructed, however, with reference to Becker (1965, 1974). The latter paper is directly concerned with the accumulation of personal consumption capital and social interactions in consumption. In this way, the key *predictions* associated with sports participation noted above can be formalized.

To begin with, Becker (1965) integrated the allocation of time explicitly into the labor supply and consumption decisions of individuals. It is recognized that individuals make these decisions as part of a household. The distinction between consumption and production is removed as the analysis emphasizes that "time" and "market goods" combine as resources in household production to generate the basic commodities that yield utility from consumption. In this regard, the traditional income-leisure trade-off model becomes a special case of Becker's approach in which the (opportunity) cost of leisure consists entirely of foregone earnings, and the (opportunity) cost of commodities is foregone consumption of other commodities. Becker (1974) extended the allocation of time model so that the characteristics of other agents directly enter the utility function of a given agent. Consequently, agents can invest resources to accumulate what they consider to be desirable characteristics or to reduce what they consider to be undesirable characteristics. There is no explicit focus upon leisure. Rather, Becker (1974) provided an analysis of both the accumulation of personal capital and social capital through the potential for social interactions.

This implies that the usual proxy variables for tastes, in econometric work, such as socioeconomic characteristics, should be viewed at least partially as the results of agents' decisions. It follows that variations in their impact should not necessarily be seen as evidence of unstable preferences but potentially resource adjustment over time in line with decision making based upon stable preferences. These features can be illustrated by presenting a simple static exposition of Becker (1974).

Assume that there is one commodity that yields utility, "U," and this is participation in a sports activity, "P."

$$U = U(P) \tag{1}$$

Following Becker (1965), participation needs to be "produced" by the agent for consumption by allocating market goods, "C_1," such as sports equipment, and time, "t," to the activity. Likewise, following Becker (1974), certain characteristics of the agent, or other agents, "C_2," can affect this production process. In the former case, this could be the skill levels required to produce the activity, such as hand-eye coordination, et cetera. In the latter case, this could be the image or reputation or status that is associated with participating in the activity. For example, playing golf might be associated with social prestige, rugby with masculinity, et cetera. This suggests a household production function for the sports activity:

$$P = P(C_1, C_2, t). \tag{2}$$

For simplicity, ignoring the input of time, this implies, on the assumption of a Cobb-Douglas utility function:

$$U = C_1^{\alpha} C_2^{\beta}. \tag{3}$$

Consider that characteristic C_2 is contingent upon an initial endowment of C_1, given the superscript "0" and the exercise of effort E. This might be because owning some market goods to participate in the activity, or previous experience of the activity indicates an initial level of skills or reputation being possessed, but that effort can be expended to enhance these skills or reputational characteristics:

$$C_2 = C_1^0 + E. \tag{4}$$

The money budget constraint facing the agent can be written as:

$$p_1 C_1 + p_2 E = M. \tag{5}$$

Out of current money income, greater consumption of C_2 is facilitated by the expenditure of resources on E at their shadow price. However, Equations (4) and (5) imply a "wealth" constraint:

$$p_1 C_1 + p_2 C_2 = M + p_2 C_1^0, \tag{6}$$

whereby the left-hand side indicates the expenditure on the market goods or characteristics and the right-hand side the current money

income available to the agent plus the value of the previous endowment of goods, skills, or characteristics already accumulated. Solving for the demand function for C_2 leaves

$$(7) \qquad C_2 = \left(\frac{\beta}{\alpha+\beta}\right)\frac{M}{p_2} + \left(\frac{\beta}{\alpha+\beta}\right)C_1^0,$$

which is the standard Marshallian demand function augmented by the last term on the right. The interpretation of this term clearly suggests that the demand or consumption of C_2, for example, the skills required to participate in a sport, is higher than it would have been because of the introduction of the Equation (4).

Consequently, if we interpret the Equation (4) as representing investment in the consumption skills needed to undertake a sport, then this suggests that previous consumption in the same activity can increase current consumption of the same activity because skill acquisition, through the production function (Equation 2), is linked to participation in the activity. An implication of this finding is that if the skills are common across participant activities, then consumption of one sports activity can also increase the consumption of another sports activity. There is an obvious rationale, therefore, for expecting the demand for any one sports activity to be positively related to the demand for other sports in the context of personal consumption capital investment.

Significantly, too, if we interpret the Equation (4) as originally proposed by Becker (1974), with C_2 being thought of as a series of desirable characteristics of other agents that the agent wishes to replicate, then, in this case, the first term on the right-hand side of Equation (4) becomes the initial endowment of these characteristics held by the agent the investment of effort in accruing these characteristics because they yield utility. This suggests that, for example, an increase in money income would yield a positive investment in the characteristics C_2, adding to the stock of existing characteristics because

$$(8) \qquad \frac{\partial C_2}{\partial M} = \left(\frac{\beta}{\alpha+\beta}\right)p_2^{-1}.$$

In contrast, if the characteristics were considered undesirable, that is, $\beta < 0$ in Equation (3), then the derivative in Equation (8) could turn out to be negative, implying that a rise in income leads to a reduction in the characteristics accrued by the agent from previous stocks, for example, the original social environment in which the agent was based. Rises in income may facilitate a shift away from these characteristics in favor of investment in others. This suggests the possibility that sets of characteristics are likely to be associated with distinct groupings of people over time as variables change and dependent on different sets of preferences. This is a description of social capital accumulation. Significantly, as implied by the production function (Equation 2), participation in activities could depend upon such social capital accumulation. In other words, ''lifestyles'' will emerge and these will affect sports participation. Significantly, it can be argued that these possibilities are implied in the existing empirical literature.

B. Empirical Literature

The economic literature providing large-scale empirical testing of sports participation is relatively sparse. Early U.S. studies covering a variety of outdoor recreational activities suggest that there is a consistent rise in participation for younger, male, white, and more educated respondents as well as those with a higher income. There is also some evidence that having children in the household reduces participation (Adams, Davidson, and Seneca, 1966, 1968; Cicchetti, Davidson, and Seneca, 1969; Davidson, 1967; Davidson, Tower, and Waldman, 1969). Caution should be attached to the results; however, as typically, given the historic context, ordinary least squares regressions were employed on binary data measuring participation or not in various activities.

In the United Kingdom, more recent studies are by Gratton and Tice (1991), Farrell and Shields (2002), Downward (2004), and Downward (forthcoming). Consistent with the more recent data, appropriate binary variable estimators are employed. However, with the exception of Downward (forthcoming), none of the others used weighted data.[3] Broadly

3. The potential importance of this is indicated by noting that Farrell and Shields (2002) do not obtain any regional effects on participation. In contrast, Downward (forthcoming) did. The official weights for the data are partially based on regional information from census data. In Downward (forthcoming), the North and Wales are significantly associated with participation in Rugby, Scotland with Golf, and the East with cycling. All these results are highly plausible in the United Kingdom.

speaking, these studies all provide support for the greater impact of socioeconomic characteristics such as the form of employment and level of education upon sports participation as opposed to work hours and household income levels which might be indicative of traditional substitution and income effects. Farrell and Shields (2002) and Downward (2004) also particularly indicated the importance of gender and household factors such as the presence of children having effects on the participation rates of particular sports. For example, they found that males tend to participate more than females in sports and declines in sports participation are associated with increasing age, being married, and the presence of children in households. The latter is particularly the case for females. In addition, "lifestyle" factors such as drinking and self-reported better health tend to raise participation, while smoking reduces it. Finally, Gratton and Tice (1991) and Downward (forthcoming) indicated that consumption in other sports is strongly associated with participation in any particular sport.

In the light of the theoretical discussions above, a number of points are worth noting from these empirical studies. The first is that the direct impact of income and the opportunity costs of work, that is, work hours, upon decisions to participate in sports is relatively small. In a time allocation framework, one might expect this as the shadow prices of resource allocation are understated by purely market-based indicators. The second is that social structures appear to be potentially important determinants of sports participation because groups of similar personal and socioeconomic variables appear to be significant. This is potential evidence of the impact of social capital or lifestyles on participation. Finally, there is evidence that previous participation in sports, that is, personal consumption capital, promotes participation in sports. The remainder of this paper provides a more direct empirical analysis of the impact of personal consumption capital and social capital on the decision to participate in sports and its frequency.

IV. DATA, VARIABLES, AND METHOD

Following Downward (forthcoming), data from the 2002 General Household Survey (GHS) are employed in this research. The GHS is a continuous survey, which began in 1971, and is conducted by the Office for National Statistics. It collects data on a range of topics, by face-to-face interview, from private households in Great Britain. As well as core topics such as household and family characteristics, education, health, income, and demographics, it also investigates other topics, such as sport and leisure, periodically. The previous occasion in which sport and leisure was investigated is 1996.

In the Sport and Leisure module for 2002, data on participation are collected for activities undertaken in the 4 wk before the interview took place, coupled with their frequency, and participation for the 12 mo before the interview took place. In this paper, the focus is upon the decision to participate or not, and their frequency, in the last 4 wk before the interview for 14,819 adults aged 16 yr or over. In addition, a broad range of personal and socioeconomic characteristics are identified as factors that may influence these decisions, either collectively or individually. Table 2 provides details of their name and measurement according to the relevant broad set of characteristics. The previous theoretical and empirical literature reviewed above provided guidance over the selection of the variables.

As one of the main aims of the paper is to examine the social interactions that affect sports participation, as evidenced by the accumulation of social capital, a twofold research design was employed. A cluster analysis was first undertaken to identify subsamples of individuals with common lifestyles as measured by correspondent personal, socioeconomic, and sporting characteristics as indicative of the outcomes of social interactions.

Subsequently, membership of these clusters, measured as a dummy variable, is employed in a Heckman model alongside the original variables, including proxies for personal consumption capital accumulation, to examine the choice to participate in sport and the frequency of this participation. Each of these stages is now examined in more detail.

V. CLUSTER ANALYSIS

Cluster analysis is appropriate to identifying lifestyles as the outcomes of social interactions because it comprises a set of multivariate statistical techniques with the aim of identifying and classifying objects, that is, the cases and not the variables, into similar types and

TABLE 2
Social Economic and Sports Variables

Conceptual Category	Variable Name	Variable Measurement
Individual/social characteristics		
Age	age	Years—continuous
Sex	sex	Male or female; 1 = male, 0 = female
Ethnicity[a]	whbrit	White British; 1 = yes, 0 = no
Educational attainment	degree, schoola, schoolo	First degree (equivalent) or more, A levels, O levels, or equivalent; 1 = yes, 0 = no (other or no qualifications is the omitted category)
Marital status	married[a]	Married or not; 1 = yes, 0 = no
Household composition	nadmales	Number of adult males in the household
	nadfems	Number of adult females in the household
	n0to4	Number of preschool children in the household
	n5to15	Number of school-age children in the household
Health	genhlth	General view of health; 1 = not good, 2 = fairly good, 3 = good
	illness	Presence of longstanding illness; 1 = yes, 0 = no
	limitact	If longstanding illness affects activity; 1 = yes, 0 = no
Smoking	cignow	Smoke cigarettes nowadays; 1 = yes, 0 = no
	cigarreg	Smoke cigar at least once a month; 1 = yes, 0 = no
	pipe	Smoke a pipe at least once a month; 1 = yes, 0 = no
Drinking	drinknow	Drink alcohol nowadays; 1 = yes, 0 = no
	drinkamt	Amount of alcohol drunk; 1 = hardly to 5 = heavily
Economic characteristics		
Employment status[a]	empman, prof, nonman, personal, skillman, semiskill, unskill	Employer or manager, professional, nonmanual, personal services, skilled manual, partially skilled manual or technical worker, unskilled; 1 = yes, 0 = no (other employment status is the omitted category)
	working, retired, keephous[a]	In work, retired from work, or keeping house; 1 = yes, 0 = no (not working, unable to work, the omitted category)
Region[a]	north, mids, south, wales, scotland	Northern England and Yorkshire, East and West Midlands and East Anglia, South West and South East England, Wales, and Scotland; 1 = yes, 0 = no (London is the omitted category)
Access to motor vehicle	usevcl1	Own or use a motor vehicle; 1 = yes, 0 = no
Income	weekinc	Gross household income in £000s
Hours worked	tothrs	Total usual hours of work per week
Unpaid hours	unpaidhr	Total weekly unpaid hours work
Sports characteristics		
Sports participation	sp4walk; sp401–sp440	Participation in walking of at least two miles, sports activities 1–40 in the last 4 wk; 1 = yes, 0 = no
Sports club participation	sp401sc–sp440sc	Sports participation in a sports club; 1 = yes, 0 = no

continued

TABLE 2
continued

Conceptual Category	Variable Name	Variable Measurement
Other club participation	sp401oc–sp440oc	Sports participation in another club; 1 = yes, 0 = no
Sports frequency	sptim1–sptim40	Number of times undertaking the activity in the last 4 wk
Sports volunteering	voltime	Hours spent on sports volunteering; 1 = less than 1 h/wk to 5 = 5 h or more per week
The number of sports activities	numsportw	The number of sports participated in during the last 4 wk
Leisure volunteering	voltime2	Hours spent on arts and other volunteering; 1 = less than 1 h/wk to 5 = 5 h or more per week
The number of arts and leisure activities	numcultw	The number of arts and other leisure activities participated in during the last 4 wk. These activities such as watching TV, listening to the radio, reading, painting, dancing, and the arts

Notes: Updated from Downward (forthcoming).
[a]In these variables, a much wider set of characteristics were investigated initially. However, problems of small sample sizes associated with specific sports required some aggregation of categories.

has been used in the medical and biological sciences, as well as social science research (Byrne, 2003; Romesburg, 2004). Cluster analysis groups the cases or individuals according to similarities in the values of the variables that are used to describe the behavior of cases.

There are a wide variety of methods of cluster analysis, but they are traditionally either hierarchical or relocational. In the former, individual cases are formed into successively larger clusters by allocating cases to clusters or combining clusters sequentially until one single cluster is constructed. In other methods, cases are iteratively reallocated to best fit a predetermined number of clusters. In either case, a distance measure is required to calibrate the similarity or dissimilarity of cases. With ratio, interval, and ordinal data, Euclidean distances or varieties of coefficients can be calculated that lie between various ranges. With nominal data, "matching" coefficients based on proportions of shared characteristics can be calculated (Romesburg, 2004).

In this paper, because the data set contains variables measured on a variety of scales, and it is also very large, "two-step" cluster is employed making use of SPSS (SPSS, 2001; Zhang, Ramakrishnon, and Livny, 1996). This method combines the maximum likelihood distance measures developed by Banfield and Raftery (1993) for continuous variables and by Melia and Heckerman (1998) for categorical variables to allow for combinations of these. The maximum likelihood procedure can be used to best fit cases to a predetermined number of clusters or to identify the number of clusters that best fit the data. To reduce computation problems, a relocational approach is used first to initially estimate the clusters based on information criteria. Second, cluster allocations are refined by maximizing the distance between the closest clusters in a hierarchical approach. The overall number of cases remains the same over the two stages.[4] In this paper, as the emphasis is upon exploration of the possibility of lifestyles, the number of clusters was identified from the data.

This is not to suggest that there were no theoretical expectations about the results.

4. Although maximum likelihood methods are employed, it should be noted that there is considerable flexibility in the use of cluster analysis; consequently, a rationale should be provided for each application. In this regard, the cluster analysis should be regarded as exploratory.

Rodgers (1978) developed classifications of leisure, recreation, and sport, which have become accepted categories and are still implied in the 1993 Council of Europe, European Sports Charter. Rodgers (1978) argued that sports have four essential elements present. Sports comprise physical activity, for a recreational purpose, taking place within frameworks of both competition and institutions. Recreational activity would not include institutional competition, while leisure may not include physical activity but reflect the use of nonobligated time. One might expect the sports and leisure activities in the GHS to be grouped according to such characteristics.[5]

The cluster analysis yielded three distinct and interpretable clusters for 9,738 cases. The first cluster contained only 281 cases, the second cluster contained 2,012 cases, and the final, largest cluster contained 7,445 cases. This suggests that of the total sample of 14,819 cases, 9,738 cases had distinct profiles. A further 5,081 cases produced an indistinct pattern of behavior. In the subsequent regression analysis, these indistinct cases are treated as the omitted base category for the cluster membership variable. Space precludes a profile of all the variables, so Table 3 focuses upon the sports participation characteristics of the cases and Table 4 some of their main socioeconomic characteristics as discussed in other empirical literature. Because the cluster analysis groups cases according to distributions of values of variables across a multivariate setting, one can always describe the clusters according to any particular variables of interest.

In Table 3, the sports and leisure activities listed from the GHS are cross tabulated against cluster membership. The activities are also grouped according to the modal frequency of cases engaging in a particular activity. The three columns under each cluster heading indicate the percentage of cases undertaking that activity, out of a total given by the value for N in the last column of that row. Consistent with expectations, the first group of activities captures activities classified by the GHS as leisure activities. It is in the second and third clusters that sports, as classified by the GHS, are located. However, the second group of activities is predominantly recreational activities, while the latter group is predominantly team or specialist activities. In the light of earlier discussions, therefore, the clusters are labeled as Leisure, Recreation, and Sport, respectively.

Table 4 indicates some of the characteristics of cases as reported in previous empirical work. The table cross tabulates these characteristics against cluster membership. The upper part of the table reports the percentage frequencies of particular attributes across the clusters. It should be noted here that the percentages are calculated against the total number of cases of each *set* of characteristics rather than each row total. This helps to show how the characteristics are distributed across the clusters for any given variable. The last column does, however, report the number of cases according to each specific characteristic of variables. For noncategorical variables, the lower part of the table reports the mean and standard deviations of variables to indicate the characteristics of cases.

The results across clusters suggest that cases in the Leisure cluster are more likely to be older, female, and with a lower income and a more diffuse educational profile. They are also less likely to participate in sports activities. For example, the average number of sports undertaken is 1 in the Leisure cluster as opposed to 3 in the Recreational cluster and 5 in the Sport cluster.

In contrast, there is evidence that younger males are more likely to belong to the Sport or Recreation clusters. The educational achievement is also likely to be higher for cases in the Sport and Recreation clusters as opposed to the Leisure cluster. The patterns for health, however, are broadly similar but, as implied above, both average incomes and their dispersion are greatest for the Sport cluster cases and least for Leisure cluster cases, and there is an absence of children for cases in the Sport cluster.

Intuitively, examining these variables profiles indicates lifestyle transitions associated with age. Notably, participation in the number of leisure activities does not differ across clusters, which indicates that sports participation is the form of leisure that does adjust to lifestyle stage. Broader leisure pursuits that may not involve physical activity do not.

5. One might intuitively expect that sports may also cluster around technical characteristics such as racquet sports, team sports, et cetera. One should caution against this expectation, however. On the one hand, the cluster analysis is undertaken over the whole range of variables, so sports comprise only one aspect of lifestyles. On the other hand, there may exist social and cultural barriers between apparently similar sports.

TABLE 3
Cluster Sports Participation Profile

Sport and Leisure Activity Undertaken in Past 4 Wk	Cluster 1 "Sport"	Cluster 2 "Recreation"	Cluster 3 "Leisure"	*N*
Walk of two or more miles	3.77	27.64	68.59	3,770
Snooker	7.86	39.65	52.49	865
Watched TV	2.89	20.66	76.45	9,643
Listened to radio	3.01	21.77	75.23	8,812
Listened to records/tapes	3.10	22.10	74.58	8,475
Read books	3.00	23.73	73.27	6,342
Sung/played an instrument	4.11	30.88	65.01	1,046
Performed in a play	5.36	36.90	57.74	168
Painting	3.55	28.95	67.51	874
Dancing	3.84	30.15	66.01	1,068
Enrolled on a course	3.96	27.97	68.08	733
Attending leisure class	5.63	43.38	50.99	657
Written stories/poetry	5.21	33.44	61.35	326
Running an arts event	2.84	18.93	78.23	9,478
Swimming indoors	5.41	48.33	46.26	1,349
Swimming outdoors	9.77	65.80	24.43	307
Cycling	7.15	47.56	45.28	965
Indoor bowls	0.00	70.67	29.33	75
Outdoor bowls	10.00	67.50	22.50	40
Tenpin bowling	7.76	46.84	45.40	348
Keep fit/aerobics	4.37	52.80	42.82	1,303
Martial arts	42.39	42.39	15.22	92
Weight training	12.38	73.94	13.68	614
Weight lifting	13.85	74.62	11.54	130
Gymnastics	0.00	90.48	9.52	21
Football indoors	22.87	65.43	11.70	188
Football outdoors	18.21	54.05	27.75	346
Cricket	3.45	82.76	13.79	58
Tennis	10.98	76.88	12.14	173
Badminton	10.33	78.80	10.87	184
Squash	20.14	68.35	11.51	139
Table tennis	16.36	67.27	16.36	110
Jogging/running	13.77	68.64	17.59	523
Angling	10.58	53.44	35.98	189
Ice-skating	4.88	95.12	0.00	41
Golf	5.37	57.22	37.41	540
Skiing	18.18	75.00	6.82	44
Horse riding	7.62	67.62	24.76	105
Climbing	32.39	63.38	4.23	71
Motor sports	18.03	40.98	40.98	61
Shooting	27.62	44.76	27.62	105
Rugby	80.00	11.43	8.57	35
American football	100.00	0.00	0.00	3
Gaelic sports	100.00	0.00	0.00	2
Hockey	89.47	0.00	10.53	19
Netball	76.47	11.76	11.76	17
Basketball	53.85	35.90	10.26	39
Athletics	70.59	17.65	11.76	17
Sailing	52.24	29.85	17.91	67
Canoeing	74.07	25.93	0.00	27

continued

TABLE 3
continued

Sport and Leisure Activity Undertaken in Past 4 Wk	Cluster 1 "Sport"	Cluster 2 "Recreation"	Cluster 3 "Leisure"	*N*
Windsurfing	82.35	17.65	0.00	17
Curling	100.00	0.00	0.00	3
Volleyball	90.48	9.52	0.00	21

TABLE 4
Summary Cluster Profile[a]

Variable	Description	Sport	Recreation	Leisure	*N*
Sex	Male	2	12	34	4,727
	Female	1	9	42	5,011
Ethnicity (whbrit)	White British	3	19	70	8,894
	Non-White British	0	2	7	844
Marital status	Single, never married	1	7	17	2,413
	Married	1	12	47	5,818
	Married and separated	0	1	2	275
	Divorced	0	2	8	952
	Widowed	0	0	3	280
n0to4	0	2	18	66	8,401
	1	0	2	8	1,055
	2	0	1	2	262
	3	0	0	0	20
n5to15	0	2	15	56	7,088
	1	0	3	11	1,432
	2	0	2	7	939
	3 or more	0	1	2	279
nadfems	0	0	2	6	827
	1	2	16	60	7,632
	2	0	2	9	1,086
	3 or more	0	0	1	193
nadmales	0	0	2	11	1,284
	1	2	15	54	6,976
	2	1	3	9	1,222
	3 or more	0	1	2	256
Health	Not good	0	1	11	1,162
	Fairly good	1	5	22	2,706
	Good	2	15	43	5,870
First degree or more		1	10	21	3,085
A levels		0	3	10	1,360
O levels		1	5	21	2,589
Other		0	2	25	2,704
Household income	Mean	1,690	888	612	9,738
£ week	SD	3,465	880	538	
Number of sports (numsportw)	Mean	5	3	1	9,738
	SD	3	2	1	
Number of leisure activities (numcultw)	Mean	4	4	4	9,738
	SD	1	1	1	
Age	Mean	36	41	45	9,738
	SD	12	13	13	

Note: SD = standard deviation.
[a]Percentages rounded up to integers.

VI. REGRESSION ANALYSIS

To examine the decision to participate or not in sport and the frequency of participation, cluster membership can be employed as a variable in a Heckman model as well as the number of activities participated in and also including all the individual and socioeconomic variables. The former two variables measure the social interactions and personal consumption capital accumulation of the individuals, respectively, while the latter variables allow for the exploration of the effects of independent influences of these factors on decisions.

The choice of the Heckman model to undertake the regression analysis needs some justification, as a number of competing estimators are possible. The Heckman model can be considered as the appropriate method to examine the choice to participate in sport and then its frequency if it is conceivable that the sample of individuals undertaking sports of different frequencies is censored, that is, contingent upon the initial choice to participate in the sport. This implies that the sample of observed frequencies of participation could be a nonrandom sample. This view would be consistent with the time allocation framework of this paper. Taking the example of the GHS data, it could be understood that for the 4-wk period of data collection, agents first plan to allocate time to participate in an activity. However, over the 4-wk period, participation frequency may become contingent upon changes in circumstances. In contrast, it could be the case that the choice set comprises voluntary decisions to participate on any number of independent occasions, which could include not at all. It remains, therefore, that this assumption needs to be tested, and the Heckman model allows for this possibility.[6]

$$(9) \qquad F_i = x_i\beta + \eta_i \;\; F_i > 0 \text{ only if } p_i = 1.$$

$$(10) \qquad P_i = z_i\gamma + \varepsilon_i \;\; P_i = 1 \text{ and } 0 \text{ otherwise}$$

where η is $N(0, \sigma)$, ε is $N(0, 1)$, and Corr $(\eta, \varepsilon) = \rho$.

Equation (9) indicates that the frequency of participation, F, for any case "i" is a linear function of a set of variables x plus a random error assumed to follow the normal distribution. However, the frequency of participation can only be observed if the individual "i" participates in the sport or not, as described by Equation (10) for P_i. Participation, P_i, also depends on a set of variables z plus a random error that in this case is assumed to be bivariate normal. If Equation (9) is estimated directly without account being taken of Equation (10), and the correlation between the random errors (ρ) is non zero, then the estimates from Equation (9) will be biased. The Heckman model, in contrast, estimates Equation (9) accounting for Equation (10) using either a two-step or a maximum likelihood method.

In the former case, a control for sample selection can be obtained by including Mills Lambda, calculated as $\lambda = \sigma\rho$, in Equation (9). If the coefficient on this term is significant, according to a standard t-test (or large-sample equivalent), then sample selection bias was evident and purged from the regression, which now produces consistent estimates. In the latter case, a direct Wald test of $\rho = 0$ can be undertaken by comparing the joint likelihood of Equations (9) and (10) being independent equations against the likelihood of their being nonindependent equations. This test follows a chi-squared distribution.

The choice between the Heckman estimation methods to an extent depends upon practical considerations. For example, the maximum likelihood approach can be used on weighted data, which is desirable as discussed earlier, whereas the two-step method cannot. However, the maximum likelihood approach can be unstable and fail to converge (Statacorp, 2003). In this research, therefore, which used StataSE8 to provide the estimates, the maximum likelihood method was tried first on weighted data, and if this failed to converge, the two-step method was employed to test for the presence on nonindependent equations.[7] In all cases, cluster sampling was employed to account for the nonindependence of cases as they were sampled as part of households,

6. The Tobit estimator is another contender. However, this estimator lacks the robust standard error corrections, cluster sampling, and weighting options possessed by the Heckman model that are viewed as important in capturing the properties of the data and implied biases from interdependent household sampling (Statacorp, 2003, p. 249). These options are available for the ordinary least squares and logistic regression estimators which are used to model the frequency to participate if sample selection bias is rejected and to model the decision to participate or not, respectively. With independent equations, moreover, it would seem reasonable to accord the value of zero no special status. It turns out that this is the case for all but participation in any sport and netball.

7. This was the case in swimming only. See Table 6.

and Huber-White robust standard errors were used to control for nonspherical disturbances. If sample selection bias was rejected on the basis of the Heckman estimates, separate ordinary least squares and logistic regressions were estimated for the frequency and choice to participate, respectively, allowing for cluster sampling, robust standard errors, and weighting.

A final issue that is important in using the Heckman model is to consider the identification of Equation (9). In this paper, the alternative functional forms, implied in the error terms, identify the equations. It can be argued that this is a weaker approach to adopt than restricting Equation (9) in terms of Equation (10). However, this reflects the exploratory nature of the research, which has been stressed throughout the paper because, as implied above, there is no strong reason identified in the theories of participation reviewed to assume that the factors that affect the decision to participate in a sport are necessarily different to those that determine how frequently participation takes place. The imposition of specific identifying restrictions would thus appear arbitrary.[8]

Table 5 presents the Heckman regression results for participation in any sport (anysport) and the total number of times that activities were participated in over the 4 wk before the interview as a measure of aggregate sports participation.[9] The maximum likelihood estimation was successful, the regressions significant overall and the test of independent equations rejected, suggesting that the choice to participate in sports activities and the frequency of participation are not independent decisions in the aggregate. These results are indicated by the Wald test statistics at the bottom of the table. In the first column, significant variables are noted. The second and third columns present the estimated coefficients and the large sample "z" statistics, respectively.[10]

The coefficient estimates suggest that participation in any sport and its frequency are likely to increase if there is participation in a number of sports (numsportw), that is, as a result of the accumulation of personal consumption capital. However, it is only the frequency of participation that will increase as the result of social interactions, as indicated by membership of the Sport (sport) or Recreation (recreation) clusters. The opposite is the case for cases belonging to the Leisure (leisure) cluster.

The likelihood of participation also increases with membership of sports or other clubs (spsc, spoc), having access to a vehicle (usevcl1), being a skilled manual worker (skillman), and someone who drinks alcohol (drinknow). In contrast, increasing age (age), living in the North (north) or Scotland (scotland), being the individual responsible for housekeeping (keephous), undertaking voluntary work (voltime), or being semiskilled (semiskill) reduce the likelihood of participating in any sport. The first result is intuitive. The second does point to some regional constraints on participation. The latter three are conceivably connected with gender, time, and income-time constraints, respectively. Significantly, the likelihood of participating in any sport actually reduces for cases in the Sport cluster (sport). The implication here is that this cluster is a relatively distinct set of individuals, as implied earlier by its relatively small size.

The frequency of participation also rises for those who have a perception of being more healthy (genhlth). It also rises for those who participate in a larger number of sports (numsportw) as well as those belonging to the Sport (sport) and Recreation clusters (recreation) but not the Leisure cluster (leisure). There is an element of aggregation likely in the former result, but on balance, these results reinforce the view that investment in personal consumption capital and lifestyles or social interactions that have a sports component reinforce the frequency of sports activity.

8. There is nothing inherent in the GHS data that suggest this, but rather this decision relies on a point of methodology and lack of specification of theoretical priors. As noted in the results below, different sets of variables do appear in each equation, but to use these subsequently would imply data mining. The specific pattern of these results is also somewhat diffuse, as might be expected when one examines specific activities across a large population. The main aim of the research should therefore be borne in mind at this point, which is to identify broadly if social interaction affects the decision to participate in activities and their frequency.

9. Only significant variables at the 5% level are reported and constant values suppressed for economy. Full results are available on request from the authors. In each table, significant variables are presented in the same order as the variables are indicated in Table 2.

10. In all regression results, a 5% significance level is adopted. Estimates are also presented for two decimal places.

TABLE 5
Regression Results: anysport

Participation: anysport			Total Frequency: anysport		
Independent Variables	**Coefficient**	z	**Independent Variables**	**Coefficient**	z
age	−0.01	−15.09	schoola	−1.20	−2.22
skillman	0.21	5.96	nadmales	0.66	2.20
semiskill	−0.06	−1.35	genhlth	1.03	4.49
drinknow	0.21	3.25	empman	−1.42	−2.33
keephous	−0.26	−4.61	nonman	−1.20	−2.30
scotland	−0.25	−4.55	north	1.28	2.22
north	−0.09	−2.64	usevcl1	−2.55	−4.54
usevcl1	0.14	3.58	weekinc	−0.53	−2.53
spsc	0.44	8.03	unpaidhr	0.19	2.34
spoc	0.45	9.87	numsportw	2.22	8.99
voltime	−0.05	−2.44	Sport	11.78	7.93
numsportw	0.14	12.90	Recreation	7.55	10.39
Sport	−0.28	−2.89	Leisure	−2.08	−4.14
			voltime2	−1.28	−5.00

Wald $\chi^2(45) = 1{,}506.82$, $P > \chi^2 = 0.00$; Wald $\chi^2(1)$ ($\rho = 0$) $= 90.17$, $P > \chi^2 = 0.00$; $N = 11{,}722$

Other factors that raise the frequency of participation are cases being in the North (north), cases with more adult males in the household (nadmales), and cases that work unpaid (unpaidhr). The results for the North suggest particularly strong commitments once constraints on participation are overcome. The latter two cases could plausibly be linked to males reinforcing patterns of behavior for one another through shared preferences, as males are more likely to be associated with the Sport and Recreation clusters. Working unpaid hours indicates an opportunity to be flexible in time allocation that is not possible in typical work relations. Significantly, in this regard, increased incomes (weekinc) reduce the frequency of participation, as does being a manual (empman) and nonmanual employee (nonman) or possessing qualification of at least A-level standard (schoola) and having access to a vehicle (usevcl1). These are characteristics of work-time constraints on participation. Finally, volunteering in leisure activities (voltime2) reduces the frequency of sports participation. This further suggests the distinction between leisure and sports lifestyles.

Broadly speaking, the results indicate that for sports participation, investment in personal consumption capital and social capital can increase the chance of cases participating in sport, as well as their frequency of participation. However, work-related income-time constraints can mitigate against more frequent participation.

To disaggregate the results, regression analyses for swimming, cycling, keep fit, and weight training are explored as recreational sports and rugby and netball for more specialized sports. Tables 6–8 present these pairs of sports, respectively. In the case of the recreational sports, it is notable that the equations were found to be independent. The same is also true of rugby. In these cases, separate weighted logistic and ordinary least squares robust regressions were estimated on the full sample, allowing for cluster sampling on households for the choice to participate and also the frequency of participation, respectively. In the tables, therefore, for the participation equations, the first column reports significant variables, the second column estimated coefficients, and the third column the large sample "z" statistic used to test the significance of the coefficients. In the frequency equations, the first two columns report the same information but the third column "t" statistics.

Table 6 supports the main results that participation in swimming and cycling are both more likely to occur with cases participating in other activities (numsportw) and for those in the Leisure cluster (leisure). However, they are less likely to participate in these activities if the cases belong to the Sport cluster (sport).

TABLE 6
Regression Results for Swimming and Cycling

Participation: Indoor Swimming		
Independent Variables	**Coefficient**	***z***
sex	−0.95	−9.88
married	0.25	2.53
nadmales	−0.28	−3.29
n0to4	0.56	7.14
n5to15	0.18	4.06
cignow	−0.31	−3.32
keephous	−0.42	−2.29
usevcl1	0.34	2.51
numsportw	0.67	20.65
Sport	−0.69	−2.47
Recreation	0.37	2.27
Leisure	0.35	2.45
Mills Lambda	1.36	0.18
Wald $\chi^2(45) = 1{,}194.10$, $P > \chi^2 = 0.000$; $N = 11{,}401$		

Frequency: Indoor Swimming		
Independent Variables	**Coefficient**	***t***
sex	−0.34	−7.14
n5to15	0.08	2.26
n0to4	0.21	3.17
cignow	−0.11	−2.79
keephous	−0.21	−2.23
sp401sc	4.25	5.51
sp401oc	5.14	14.42
voltime	−0.06	−2.21
numsportw	0.22	11.21
Recreation	0.78	6.10
voltime2	−0.10	−2.72
$F(47, 7068) = 19.38$, $P > F = 0.000$; $N = 11{,}726$		

Participation: Cycling		
Independent Variables	**Coefficient**	***z***
sex	0.41	4.19
nadfems	−0.20	−2.46
n5to15	0.11	2.16
drinkamt	−0.12	−2.55
cignow	−0.21	−2.07
usevcl1	−0.37	−2.72
numsportw	0.72	22.22
Sport	−0.84	−2.87
Leisure	0.44	2.74
numcultw	0.08	2.02
Wald $\chi^2(1)$ $(\rho = 0) = 0.82$, $P > \chi^2 = 0.365$; Wald $\chi^2(45) = 1{,}057.67$, $P > \chi^2 = 0.000$; $N = 11{,}689$		

Frequency: Cycling		
Independent Variables	**Coefficient**	***t***
sex	0.22	3.02
nadfems	−0.24	−3.48
cignow	−0.20	−2.80
empman	−0.39	−3.30
nonman	−0.21	−2.00
mids	0.37	2.86
usevc11	−0.65	−5.67
sp403sc	5.38	3.01
sp403oc	3.36	2.98
numsportw	0.50	13.57
$F(47, 7068) = 8.92$, $P > F = 0.000$; $N = 11{,}726$		

TABLE 7
Regression Results for Keep Fit and Weight Training

Participation: Keep Fit

Independent Variables	Coefficient	*z*
age	−0.02	−4.15
sex	−1.37	−12.92
schoola	0.29	2.05
genhlth	0.24	3.44
illness	0.24	2.23
cignow	−0.38	−3.97
Sport	0.68	2.56
Recreation	0.98	5.67
Leisure	−0.37	−2.44
numcultw	0.20	5.75

Wald $\chi^2(1)$ ($\rho = 0$) = 3.46, $P > \chi^2 = 0.063$; Wald $\chi^2(44) = 694.92$, $P > \chi^2 = 0.000$; $N = 11{,}114.00$

Frequency: Keep Fit

Independent Variables	Coefficient	*t*
age	−0.01	−2.00
sex	−0.36	−5.53
genhlth	0.13	3.02
cignow	−0.14	−2.34
north	0.26	2.33
unpaidhr	−0.02	−3.33
sp407sc	6.48	6.05
sp407oc	6.13	24.29
voltime	−0.12	−2.56
Sport	0.84	3.08
Recreation	1.38	8.30
Leisure	−0.28	−3.40
voltime2	−0.26	−5.23

$F(46, 7068) = 26.77$, $P > F = 0.000$; $N = 11{,}326$

Participation: Weight Training

Independent Variables	Coefficient	*z*
age	−0.05	−7.17
sex	0.68	3.9
nadfems	−0.35	−3.14
genhlth	0.25	2.09
drinkamt	0.16	2.34
retired	−1.06	−2.06
Sport	1.20	3.88
Recreation	0.98	3.81
Leisure	−1.30	−4.87

Wald $\chi^2(1)$ ($\rho = 0$) = 1.49, $P > \chi^2 = 0.222$; Wald $\chi^2(43) = 624.94$; $N = 11{,}331.00$

Frequency: Weight Training

Independent Variables	Coefficient	*t*
age	−0.01	−6.11
sex	0.22	4.96
degree	−0.17	−2.45
married	−0.11	−2.19
nadfems	−0.09	−2.26
genhlth	0.06	2.04
pipe	−0.29	−4.79
north	0.19	2.17
sp409sc	6.14	7.11
sp409oc	6.89	19.57
Sport	1.25	4.48
Recreation	0.83	7.30
Leisure	−0.22	−5.08
voltime2	−0.14	−4.26

$F(46, 7068) = 19$, $P > F = 0.000$; $N = 11{,}726$

TABLE 8
Regression Results for Rugby and Netball

Participation: Rugby			**Frequency: Rugby**		
Independent Variables	**Coefficient**	***z***	**Independent Variables**	**Coefficient**	***t***
numsportw	0.19	3.27	sp413sc	5.27	8.35
Leisure	−2.55	−3.26	prof	−0.04	−2.65
Sport	1.62	3.10	empman	−0.02	−2.30
sex	2.35	2.90	skillman	−0.01	−2.12
age	−0.08	−2.66	working	−0.03	−2.01
Recreation	−1.75	−2.52			
wales	2.75	2.40			
numcultw	0.30	2.07			
north	2.12	1.98			
Wald $\chi^2(1)$ ($\rho = 0$) = 1.92, $P > \chi^2 = 0.165$; Wald $\chi^2(39) = 227.55$, $P > \chi^2 = 0.000$; $N = 7{,}871$			$F(47, 7068) = 3.91$, $P > F = 0.000$; $N = 11{,}726$		

Participation: Netball			**Frequency: Netball**		
Independent Variables	**Coefficient**	***z***	**Independent Variables**	**Coefficient**	***z***
age	−0.04	−5.12	whbrit	3.32	2.34
sex	−1.31	−3.92	degree	3.25	2.73
voltime	0.29	5.49	empman	−5.79	−2.22
Sport	1.07	4.80	prof	−5.81	−2.00
Recreation	−0.83	−2.55	nonman	−5.75	−2.00
Leisure	−0.90	−3.79			
Wald $\chi^2(1)$ ($\rho = 0$) = 17.15, $P > \chi^2 = 0.000$; Wald $\chi^2(13)$ = undefined*; $N = 11{,}726$					

*The very small sample of netballers implies constraints on the data set such that the overall statistic cannot be calculated.

Swimming and cycling are also more likely to take place in the presence of children in the household. The main difference is that swimming is more likely for married (married) and female cases (sex) and where children of both preschool age (n0to4) and school age (n5to15) are present in the household. For cycling, this is more likely to be the case for males (sex) with school-age children in the household (n5to15).[11] Moreover, participation is less likely in swimming with the presence of more adult males (nadmales) in the household and less likely for cycling with more adult females (nadfems) in the household. Access to a car (usevcl1) has opposite effects on participation, increasing the likelihood of participation in swimming and, predictably, decreasing it for cycling.

The number of sports participated in (numsportw) raises the frequency of participation in both cases as does belonging to a sports (sp401sc, sp403sc) or other club (sp401oc, sp403oc). Membership of the Recreation cluster (recreation) also raises the frequency of participation in swimming. As with the decision to participate, being female or male (sex) and the presence of preschool- (n0to4) and school-age children (n5to15) or school-age children (n5to15) raises the frequency of participation in swimming and cycling, respectively. In contrast, the number of adult females (nadfems) in the household reduces the frequency of participation in cycling. Significantly, being responsible for keeping house (keephous) in the case of swimming and being employed in a manual (empman) or non-manual occupation (nonman) reduces the frequency of participation in cycling. These results clearly describe gender-oriented family activities of a leisure or recreational nature.

In the cases of the recreational activities keep fit and weight training, the effects of the family variables disappear. However, in both activities, increasing age (age) reduces the likelihood of participation. As participation in keep fit is more likely for females and for weight training more likely for males, this is suggestive of younger gender-oriented activities. There is no evidence of personal consumption capital effects, as measured by the number of sports (numsportw), but interestingly in the case of keep fit, the number of leisure activities (numcultw) is significant. As some of these may embrace elements of music, movement, or aesthetics, this might be indicative of different consumption skills being required. However, in both activities, there is evidence that social capital does affect the likelihood of participation. Sport (sport) and Recreation (recreation) cluster cases are more likely to participate in keep fit and weight training, but this is not so for Leisure (leisure) cases. Good health (genhlth), however, does increase the likelihood of participation in both activities. Similar results apply in the case of the frequency of participation, though this is also enhanced by membership of a sports (sp407sc, sp409sc) or other club (sp407oc, sp409oc). Other notable results are that voluntary (voltime, voltime2) and unpaid work (unpaidhr) reduce the frequency of participation.

In the case of the more specialized sports of rugby and netball, there is evidence of personal consumption and social capital effects. While the number of sports (numsportw) only increases the likelihood of participation in rugby, in both activities the likelihood of participation rises for those in the Sport cluster (sport) but not the Recreation (recreation) and Leisure (leisure) clusters. Notably too, the sports are associated with younger cases (age) and also gender (sex) as consistent with their traditions. Strong regional effects are also identified for Rugby and intuitively reflect its original professional location in the North (north) and it being the national game in Wales (wales). The frequency of participation rises for rugby for those in a sports club (sp413sc) but declines with an array of work-related characteristics. The latter is also true of netball.

VII. POLICY IMPLICATIONS

The policy implications of the above analysis can now be discussed in terms of assessing the target of sports policy and the potential policy levers that can be identified from the research. To begin with, the empirical analysis broadly identifies that one should not view participation in specific activities in isolation but that, if relevant, sports participation policy should target broad sets of activities. This is particularly because of the presence of personal consumption capital effects (numsportw)

11. The negative sign on the dichotomous variable "sex" indicates males less than females.

in the analysis of any sport, swimming, cycling, and rugby but is also implied with the general significance of the cluster membership variables, which measure social interactions. The analysis also shows that the frequency of participation in a specific activity is likely to fall as a result of various paid work-related characteristics and also voluntary and unpaid work, as indicative of an income-time constraint. In contrast, frequency of participation is likely to rise when cases are members of sports or other clubs. This suggests that once a threshold of investment is made in consumption activity, this reinforces commitment to the activity. In addition, the analysis implies that the frequency of participation may only be dependent upon the choice to participate in the aggregate. In terms of Section III, this implies a degree of separability in the utility function, as the decision to participate in any particular activity a number of times may only be conditional upon a prior allocation of resources to sports activities in general. It is once these constraints are set that worktime and other constraints influence the allocation of time to activities. This is clearly an issue that merits further research.

In terms of Sport England's model of drivers of change in participation, only aging and time emerge as distinct factors, in the sense of individual variables being statistically significant, with the latter connected to work and volunteering constraints. The other drivers, well-being, investment, education, and access, appear to be broadly subsumed within the personal consumption and social interaction effects identified above, in as much that they are part of the profile of cluster cases, despite specific occasions when, for example, significant regional effects might identify differences in investment in sport or significant gender and access to a vehicle indicate issues of access more generally defined. This suggests that policy levers should target age, gender, and broader lifestyles if general increases in participation are desired. To target, say, specific clubs may only enhance the frequency of participation of particular and minority cases. This raises issues associated with both the targeting and the domain of sports policy. A refocus away from competitive activity organized by traditional sports clubs and their governing bodies, as part of an implied hierarchy into elite-level competition, and a reorientation back toward mass participation may be required. These have been championed under various "Sport for All" campaigns in the past (Council of Europe, 1980). The movement toward a rationalization of the sports policy delivery bodies along elite or mass participation lines and the appointment of the new Minister for Public Health looking to link sports policy to wider policy development is, therefore, potentially a timely step in the right direction in the United Kingdom as the nature of sports funding and provision has increasingly been oriented toward elite sports development (Downward et al., forthcoming; Green and Houlihan, 2005).

In what sense, however, should policy agencies intervene in the delivery of sports? This is not a straightforward proposition and one that merits some further discussion. Consider again the methodological approach of Becker (1976, 1992). Section III identified this as embracing economic agents with stable preferences maximizing welfare, as perceived by them, subject to income, time, information, and other limiting resources with markets and social structures allocating resources according to their shadow prices. Under such circumstances, policy intervention is ruled out. This is because the model sketched in Section III can be shown to imply a version of the Coase theorem, derived as the "Rotten Kid" theorem, if the interdependency between consumption reflects a policymakers utility depending on the policy recipients utility (Becker, 1974). Consequently, any transfers of income between the policymaker and the recipient, for example, to facilitate sports participation, will not affect the consumption or welfare of either. This is even if the policymaker intends to enhance the welfare of the recipient and the latter does not reciprocate. Under such circumstances, there can be no strong rationale for active sports policy other than ensuring equal access to information about sports and physical activity to the population to allow agents to make choices.

In contrast, it could be argued that the efficacy of any policy intervention must, ultimately, be ascertained from an evaluation of a policy initiative. Further, as discussed in Section III, the predictions of the time allocation model are consistent with the theories derived from a broader social science tradition, which might, through their assumptions, provide a basis for policy activism. For example, if one argues that preferences can actually change

as a result of a new opportunity to participate in previously unfamiliar activity and that constraints to voluntary action exist, because agents do not possess optimizing capability or because particular characteristics and social circumstances act to exclude consumption opportunities, then an equation like Equation (7) might be viewed as showing that a lack of prior experience of particular activities or the possession of particular characteristics act as a barrier to participation. Likewise, income differentials, interpreted as reflecting variances in economic opportunity, will affect participation. Under such circumstances, active policy becomes desirable and should target both the constraints and the agent choice in seeking to promote greater participation.

Obvious examples would be to ensure that facilities are available to all, preventing exclusion on personal or social criteria through legislation, coupled with the flow of resources to support areas in which choices are desired but not attainable, for example, because of economic underdevelopment or cultural restraint.

VIII. CONCLUSIONS

This paper has addressed the determinants of both the choice to participate in sport and the frequency of that participation. Based on a model of social interactions, which synthesizes predictions from a broad literature, it is argued that investment in personal consumption and social capital will be integral to understanding sports participation. This paper has explored a large-scale data set in the United Kingdom and found results that support these predictions. Along with specific individual factors that affect participation, such as age and gender, it is argued that these results are broadly consistent with the current analysis of sports policy bodies in the United Kingdom. It is also argued that current changes in the organization of sports policy in the United Kingdom and its overseeing through a broader central governmental minister are appropriate innovations in policy in as much that policy activism should focus upon broad sets of activities and also shape choices as well as eliminate constraints, for example, through education. However, it is argued that this may raise policy tensions between calls for mass participation and elite sports development.

REFERENCES

Adams, F. G., P. Davidson, and J. J. Seneca. *The Social Value of Water Recreational Facilities Resulting from an Improvement in Water Quality in an Estuary*. reprinted in Davidson, L. (ed.), 1966.

———. *An Analysis of Recreational Use of the TVA Lakes*. reprinted in Davidson, L. (ed.), 1968.

Banfield, J. D., and A. E. Raftery. "Model-Based Gaussian and Non-Gaussian Clustering." *Biometrics*, 49, 1993, 803–21.

Becker, G. "A Theory of the Allocation of Time." *Economic Journal*, 75, 1965, 493–517.

———. "A Theory of Social Interactions." *Journal of Political Economy*, 82, 1974, 1063–91.

———. *The Economic Approach to Human Behaviour*. Chicago: The University of Chicago Press, 1976.

———. "The Economic Way of Looking at Life." *Journal of Political Economy*, 101, 1992, 385–409.

Bourdieu, P. *Distinction: A Social Critique of the Judgement of Taste*. London: Routledge, 1984.

———. "Program for a Sociology of Sport." *Sociology of Sport Journal*, 5, 1988, 153–61.

———. "Sport and Social Class," in *Rethinking Popular Culture, Contemporary Perspectives in Cultural Studies*, edited by C. Mukerji and M. Schudson. Berkeley, CA: University of California Press, 1991, 357–73.

Byrne, D. *Interpreting Quantitative Data*. London: Sage, 2003.

Carter, P. *Review of National Sport Effort and Resources*. London: DCMS, 2005.

Cicchetti, C. J., P. Davidson, and J. J. Seneca. *The Demand and Supply of Outdoor Recreation: An Econometric Analysis*. Washington, DC: Bureau of Outdoor Recreation, 1969.

Council of Europe. *European Sport for all Charter*. Strasbourg, France: Council of Europe, 1980.

Davidson, L. (ed.). *The Collected Writings of Paul Davidson, Vol. 2, Inflation, Open Economies and Resources*. London: Macmillan, 1991.

Davidson, P. *An Exploratory Study to Identify and Measure the Benefits Derived from the Scenic Enhancement of Federal-Aid Highways*. reprinted in Davidson, L. (ed.), 1967.

Davidson, P., J. Tower, and A. Waldman. *The Economic Benefits Accruing from the Scenic Enhancement of Highways*. reprinted in Davidson, L. (ed.), 1969.

Downward, P. M. "Assessing Neoclassical Microeconomic Theory via Leisure Demand: A Post Keynesian Perspective." *Journal of Post Keynesian Economics*, 26, 2004, 371–95.

———. "Exploring the Economic Choice to Participate in Sport: Results from the 2002 General Household Survey." *International Review of Applied Economics*, forthcoming.

Downward, P. M., A. Dawson, and T. Dejonghe. *The Economics of Sport: Theory, Policy and Evidence*. London: Elsevier, forthcoming.

Earl, P. E. *The Economic Imagination*. Brighton, UK: Wheatsheaf, 1983.

———. *Lifestyle Economics*. Brighton, UK: Wheatsheaf, 1986.

Farrell, L., and M. A. Shields. "Investigating the Economic and Demographic Determinants of Sporting

Participation in England." *Journal of the Royal Statistical Society (A)*, 165, 2002, 335–48, Part 2.

Galbraith, J. K. *The Affluent Society*. Harmondsworth, UK: Peguin, 1958.

Gratton, C., and P. Taylor. *The Economics of Sport and Recreation, E and F.N*. London: Spon, 2000.

Gratton, C., and A. Tice. "The Demand for Sport: A Two-Stage Econometric Model Estimated from the Health and Lifestyle Survey." Manchester Polytechnic Department of Economics Discussion Paper, 1991.

Green, M., and B. Houlihan. *Elite Sport Development: Policy Learning and Political Priorities*. London: Routledge, 2005.

Houlihan, B. *Sport, Policy and Politics: A Comparative Analysis*. London: Routledge, 1997.

Lavoie, M. "A Post Keynesian Approach to Consumer Choice." *Journal of Post Keynesian Economics*, 16, 1994, 539–62.

Melia, M., and D. Heckerman. *An Experimental Comparison of Several Clustering and Initialization Methods*. Microsoft Research Technical Report MSR-TR-98-06, 1998.

Rodgers, B. *Rationalising Sports Policies; Sport in its Social Context: Technical Supplement*. Strasbourg: Council of Europe, 1978.

Romesburg, H. C. *Cluster Analysis for Researchers*. NC: Lulu Press, 2004.

Scitovsky, T. *The Joyless Economy*. Oxford: Oxford University Press, 1976.

Sport England. *Driving up Participation*. London: The Challenge for Sport, 2004a.

———. *The Framework for Sport in England*. London: Sport England, 2004b.

SPSS. *Technical Report: The SPSS Two Step Clustering Component*. SPSSinc TSCWP-0101, 2001.

Statacorp. *Stata Statistical Software: Release 8.0*. College Station, TX: Stata Corporation, 2003.

Veblen, T. *The Theory of the Leisure Class*. London: George Allen and Unwin, 1925.

Zhang, T., R. Ramakrishnon, and M. Livny. "BIRCH: An Efficient Data Clustering Method for Very Large Databases." in *Proceedings of the ACM SIGMOD Conference on Management of Data, Montreal, Canada*, 1996, 103–14.