Sports Participation as an Investment in (Subjective) Health:
A Time Series Analysis of the Life Course

Professor Paul Downward
School of Sport Exercise and Health Sciences
Loughborough University
Leicestershire
LE11 3TU
UK
P.Downward@lboro.ac.uk
+44 (0)1509 226365

Dr Peter Dawson
School of Economics
University of East Anglia
Norwich
NR4 7TJ
UK
Peter.Dawson@uea.ac.uk
+44 (0)1603 597612

Professor Terence C. Mills
School of Business and Economics
Loughborough University
Leicestershire
LE11 3TU
UK
T.C.Mills@lboro.ac.uk
+44 (0)1509 222703

Key words: Sports, Health, Time Series Analysis, Trickle-down effect
ABSTRACT
Background: The causal relationship between sports participation, as physical activity, and subjective health is examined accounting for the London 2012 Olympic Games, which it was hoped would ‘inspire a generation’ by contributing to public health. Improvements to weaknesses in the literature are offered. First, stronger causal claims about the relationship between sports participation and health; second, the actual minutes and intensity of different measures of participation are used.

Methods: The rolling monthly survey design of the annually reported Taking Part Survey (TPS) is used to create time series data. This is analysed using a time-series modelling strategy.

Results: Increases in the level of subjective health requires accelerating sport participation, but no effect from the 2012 Olympics is revealed. Reductions in the level of health are brought about by increases in sports participation in early adulthood, although this gets reversed in middle age. However, a reduction in health re-emerges for older males compared to females.

Conclusions: For the population as a whole sport can contribute to health, with diminishing impact, but impacts vary across the life course and genders. Policy accounting for these variations is necessary. Policy aspirations that London 2012 would produce health benefits from increased sports participation are misplaced.
INTRODUCTION
There is international policy concern about the need to increase physical activity to improve the health of individuals and to use sport to achieve this.[1-3] In the UK, this policy aspiration lies in the strategic priorities for sport announced in 2002 of growing mass sports participation both directly, and indirectly through hosting and being successful at major sports events. [4, 5] This policy maintained that hosting the 2012 Olympics would ‘inspire a generation’ to engage in sport and improve the nation’s health and well-being. [6]

It is well known that physical activity can improve health through improved respiration and cardiovascular performance, increased muscle and bone strength, and reduction in the incidence of Type II diabetes. [7] There are also potential psychological improvements to mood and well-being. [8, 9] The causal impact of physical activity, over a longer time horizon and amongst the general population, on the overall health of individuals is less understood. Two relationships are possible. Grossman’s seminal work on health production argues that health can be viewed as a stock of capital yielding a flow of healthy time. [10] Consequently, health stocks will depreciate with age and infirmity but can be augmented by investment in health care. Increases in the flow of time spent on sports participation can thus be viewed as a direct investment in healthcare which then enhances the stock of health. Alternatively, sport could be a resultant flow of ‘healthy’ time.

Causal effects have been investigated in the literature using large-scale data representative of the wider population. [11-13] However, the measures of sport
participation have been binary indicators of participation. [14] This research measures the actual flow of minutes of participation at different intensities.

METHODS
Data are taken from the large-scale Taking Part Survey (TPS), commissioned by the Department for Culture, Media and Sport (DCMS) in the UK. Beginning in 2005, the TPS is a cross-sectional national survey of sporting, cultural, heritage, media and other activities for England of approximately 14,000 individuals. Different individuals are surveyed in each annual wave. Data capture is by Computer Assisted Personal Interviewing (CAPI). Whilst the data are typically published on an annual basis, primary sampling units are surveyed on a rolling monthly basis over the year, with monthly allocations of these sampling units following a string pattern based on a random starting point. [15] The data employed in this research covers the period from July 2005 until June 2013. To facilitate a longitudinal analysis, monthly averages of the values of relevant variables across each individual were taken over the waves of the survey, yielding a time series of 97 months over which each model is estimated.

The variables averaged across are; the overall (stock of) health, as measured by the variable ‘Genheal’. This is a subjectively scored general health variable in which respondents are asked to rate on a five point scale ‘How is your health in general?’ A score of ‘1’ implies ‘Very bad’, while a score of ‘5’ implies ‘Very good’. It has been shown that such generalised measures of health have good reliability and validity and have long been employed in studies examining overall health. [16, 17]
Second, three measures of sports participation are developed. The TPS asks three questions about 68 sports activities: whether the activity was undertaken in the last four weeks prior to the interview?; on how many days over the last four weeks did the activity take place?; and what was the typical duration, in minutes, of the activity? Combining these questions identifies the total minutes of sport undertaken in the last four weeks as 'totmin'. The survey also asks about the number of days in which the activity was undertaken with moderate intensity for at least 30 minutes. This allowed two further measures of sports participation to be calculated. The total minutes of sports participation on at least one day a week of 30 minutes moderate intensity, ‘1x30min’; and the total minutes of sports participation on at least three days a week of 30 minutes moderate intensity, ‘3x30min’. Each of these measures of total minutes thus has a lower bound of 30 minutes of moderate intensity exercise providing it takes place on at least the relevant number of days in the week. For example, if an individual undertook three activities of moderate intensity each week for 45 minutes each and one activity not of moderate intensity for 60 minutes in the week, ‘totmin’ would include all of the activities in a total of 195 minutes. ‘3x30min’ and ‘1x30min’ would both include the minutes for the first three activities as values of 135 minutes each. If, further, one of these first three activities was not of moderate intensity then the value of ‘3x30min’ would drop to zero and the other two activities would contribute to ‘1x30min’ as 90 minutes.’ These indicators are used by policy agencies in the UK to measure the inputs to health from sport. [18, 19]

In order to analyse the data an autoregressive distributed lag (ARDL) time-series statistical model was employed in which the average monthly influence of the minutes
of sports participation on average monthly general health was investigated. The general model is described in equation 1.

\[
Y_t = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i Y_{t-i} + \sum_{i=0}^{\infty} \beta_i (X_{t-i}) + u_t
\]  

(1)

Here the subscript ‘t’ represents the month. Treating ‘Y’ as general health and ‘X’ as sport participation, lags of ‘Y’ are included in the model to capture the history of the variable, which will measure the past influence of other factors that remain after the influence of the sports participation variable ‘X’ and its lags have been accounted for. The \(\beta\) coefficients on the latter variables measure the dynamic marginal effect of the sports participation variable on general health independently of other influences. The \(\alpha\) coefficients on the lagged values of ‘Y’ measure these other influences. The lag structure of the model is obtained by a general-to-specific testing approach that begins with longer lags and eliminates these successively using, for example, F-tests of exclusions, whilst ensuring that autocorrelation in the disturbances \((u)\) is eliminated so that no predictable effects through time remain in ‘Y’. Reversing the definitions of Y and X and re-estimating the model allows feedback from health to sports participation to be examined.

**RESULTS**

Figure 1 graphs the behaviour of the variable ‘Genheal’ over the sample period. There is some minor cyclical variation and some volatility in 2009, but the variation in the variable is densely packed around the mean at the upper end of the scale (mean 3.95, standard deviation 0.053), which is typical of such measurements. [16]
Figure 2 presents the mean values of the sports participation variables. The data suggest that differences between sports participation at different intensities and frequencies are of overall scale rather than underlying trend with means (standard deviations) of 295.35 (39.46), 277.59 (38.22) and 195.48 (34.13) for totmin, 1x30min and 3x30min, respectively. This suggests that some common underlying structure to the practice of participation. The extremely stable behaviour of sports participation in connection with socio-economic factors is well documented in the sport literature. [18, 20]

This literature suggests that males participate in sport more than females. [21, 22] Moreover, both the sociological and psychological literatures suggest that there are traditional competitive forms of participation for men, in which concepts of masculinity are created and identities communicated and reinforced through competitive and even aggressive actions, which might undermine health through reckless and risk-taking behaviour. This may be contrasted with the traditionally less competitive and more aesthetic context of female activity, though sport can be an arena in which this form of female identity is challenged. [23-25] For these reasons the statistical analysis is disaggregated by gender. The means (standard deviations) are 406.47(68.73), 381.61(66.23) and 272.64(59.04) for totmin, 1x30min and 3x30min respectively for males. For females, the corresponding statistics are 214.33(35.99), 201.05(35.37) and 137.57(31.65) for totmin, 1x30min and 3x30min respectively. These results suggest that males participate for longer in sport at any intensity.

INSERT FIGURE 2 HERE
Table 1 presents the results of the analysis for the whole sample in row 1, and for males and females in rows 2 and 3. The table only reports the dynamic marginal impact of sports participation on health because there was no evidence of any feedback from health to sports participation. For brevity the estimated parameters on the lags of health, which control for other influences, are omitted.
Table 1. ARDL results for the whole sample and for gender

<table>
<thead>
<tr>
<th>Cohort</th>
<th>X structure</th>
<th>1x30min</th>
<th>3x30min</th>
<th>Totmin</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>$X_t-X_{t-2}$</td>
<td>0.000351(3.63)</td>
<td>0.000360(3.22)</td>
<td>0.000319(3.53)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R^2=0.2842$</td>
<td>$R^2=0.2635$</td>
<td>$R^2=0.2788$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Delta R^2=0.1063$</td>
<td>$\Delta R^2=0.0856$</td>
<td>$\Delta R^2=0.1009$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$F_{\text{pred}} = 0.26[0.99]$</td>
<td>$F_{\text{pred}} = 0.26[0.99]$</td>
<td>$F_{\text{pred}} = 0.24[0.99]$</td>
</tr>
<tr>
<td>Male</td>
<td>$X_t-X_{t-4}$</td>
<td>0.000266(4.58)</td>
<td>0.000304(4.54)</td>
<td>0.000259(4.60)</td>
</tr>
<tr>
<td></td>
<td>$X_{t-1}-X_{t-5}$</td>
<td>0.000177(2.98)</td>
<td>0.000227(3.33)</td>
<td>0.000170(2.95)</td>
</tr>
<tr>
<td>Female</td>
<td>$X_t-X_{t-3}$</td>
<td>0.000169(1.75)</td>
<td>0.000237(2.21)</td>
<td>0.000177(1.83)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R^2=0.3360$</td>
<td>$R^2=0.3379$</td>
<td>$R^2=0.3336$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Delta R^2=0.1874$</td>
<td>$\Delta R^2=0.1893$</td>
<td>$\Delta R^2=0.1850$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$F_{\text{pred}} = 0.49[0.90]$</td>
<td>$F_{\text{pred}} = 0.49[0.92]$</td>
<td>$F_{\text{pred}} = 0.44[0.94]$</td>
</tr>
</tbody>
</table>

Notes: $F_{\text{pred}}$ is a predictive forecast test for a break at August 2012. 
t-statistics are in parentheses; marginal significance levels in brackets.
Columns 2 to 5 of Table 1 present the estimated preferred forms of the ARDL model. Column 2 indicates the structure of the lags on participation \( (X) \), as determined by appropriate hypothesis tests, while columns 3-5 report the estimated coefficients, along with the significant t-values which reject the null hypothesis that the coefficients are equal to zero. The \( R^2 \) statistic reporting the goodness of fit of the model along with \( \Delta R^2 \), indicating the change in \( R^2 \) as a result of augmenting a pure autoregressive form of the model, containing only lagged values of health, with sports participation variables, are also presented. The \( F_{pred} \) statistics included in each cell report the predictive forecast test for a break in the series at August 2012 and test for the presence of an Olympic effect. These tests reveal no shifts in the parameters of the fitted models and, as a consequence, there is no evidence that the 2012 Olympic Games shifted behaviour, thus casting doubt on the ‘trickle-down effect’ hypothesis that hosting and being successful at major events can galvanise sports participation, which supports previous research in the UK. [26-29]

The results indicate a positive causal relationship between sports participation and health that is consistent with the model of health production. This is indicated by the positive signs on all the coefficients in the first row of each of the last three columns for the ‘All’, ‘Male’ and ‘Female’ cohorts. The results also indicate that an increase in the change in sports participation, i.e., acceleration in sports participation, is linked to an increase in the level of health. For example, in the first row of the last column of Table 1, a change in the change (an acceleration) in sports participation measured over period ‘\( t-2 \)’ to ‘\( t \)’, as indicated by column 2, that is over the last two months, increases self-reported general health by 0.000319 of the unit scale running from ‘1’ (Very bad self-
reported health) to ‘5’ (Very good self-reported health). Increasingly more sports participation is required per period of time to produce increases in the level of health. This is consistent with there being a declining productivity of sport as a means of generating health gains. Naturally, the greatest gains from sports participation will occur as sports participation begins.

The results also indicate that the impact of these greater increases in sports participation on health are larger for males than for females, as indicated by the relative size of the coefficients, and that the dynamic response takes place over a longer time frame, as indicated by the lag structure of the model indicated in column 2. This latter result is consistent with the expectation in the literature, noted earlier, that males are typically more inclined to participate in sport than females.

In order to capture further insights into the impact of life transitions and of likely changes in the composition of sports participated in with ageing, four age groups were analysed. The age groups adopted were 16 to 25, to capture opportunities through education and early adulthood; 26 to 45, to capture adulthood; 46 to 65, to capture middle age; and, finally, 65 and older to capture older age, as indicated in the literature. [30] The results are presented in Table 2.

The results show that for younger males and females between 16-25 years of age, and unlike the aggregate results presented earlier, changes in the level, rather than acceleration in sports participation, is causally linked to changes in general health. This suggests that within this age group there is no declining health productivity from sport.
Stronger ‘impact’ effects are also noted for females, although here the dynamic response is delayed compared to males.
<table>
<thead>
<tr>
<th>Cohort</th>
<th>X structure</th>
<th>1x30min</th>
<th>3x30min</th>
<th>totmin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male 16-25</td>
<td>X_t</td>
<td>0.000155(4.68)</td>
<td>0.000169(4.91)</td>
<td>0.000157(4.78)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.2319</td>
<td>R²=0.2463</td>
<td>R²=0.2381</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.1828</td>
<td>ΔR²=0.1972</td>
<td>ΔR²=0.1890</td>
</tr>
<tr>
<td>Female 16-25</td>
<td>X_t-2</td>
<td>0.000357(2.94)</td>
<td>0.000373(2.87)</td>
<td>0.000352(2.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.2775</td>
<td>R²=0.2746</td>
<td>R²=0.2778</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0718</td>
<td>ΔR²=0.0689</td>
<td>ΔR²=0.0721</td>
</tr>
<tr>
<td>Male 26-45</td>
<td>X_t-2</td>
<td>-0.000130(1.74)</td>
<td>-0.000092(1.20)</td>
<td>-0.000118(1.63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.2585</td>
<td>R²=0.2251</td>
<td>R²=0.2554</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0258</td>
<td>ΔR²=0.0124</td>
<td>ΔR²=0.0227</td>
</tr>
<tr>
<td>Female 26-45</td>
<td>X_t-1-X_t-2</td>
<td>-0.000241(1.81)</td>
<td>-0.000214(1.68)</td>
<td>-0.000261(1.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.3288</td>
<td>R²=0.3254</td>
<td>R²=0.3327</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0251</td>
<td>ΔR²=0.0217</td>
<td>ΔR²=0.0290</td>
</tr>
<tr>
<td>Male 46-65</td>
<td>X_t-3-X_t-5</td>
<td>0.000213(2.15)</td>
<td>0.000123(1.07)</td>
<td>0.000200(2.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.1233</td>
<td>R²=0.0890</td>
<td>R²=0.1189</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0458</td>
<td>ΔR²=0.0115</td>
<td>ΔR²=0.0414</td>
</tr>
<tr>
<td>Female 46-65</td>
<td>X_t-X_t-3</td>
<td>0.000256(2.07)</td>
<td>0.000256(2.07)</td>
<td>0.000279(2.31)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.0452</td>
<td>R²=0.0452</td>
<td>R²=0.0553</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0452</td>
<td>ΔR²=0.0452</td>
<td>ΔR²=0.0553</td>
</tr>
<tr>
<td></td>
<td>X_t-3</td>
<td>-0.000347(1.57)</td>
<td>0.000046(2.00)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.0000577[2.46]</td>
<td>R²=0.0917</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X_t-4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X_t-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male 65+</td>
<td>X_t-3</td>
<td>-0.000294(1.79)</td>
<td>-0.000516(2.32)</td>
<td>-0.000288(1.77)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.1504</td>
<td>R²=0.1703</td>
<td>R²=0.1498</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0305</td>
<td>ΔR²=0.0504</td>
<td>ΔR²=0.0299</td>
</tr>
<tr>
<td>Female 65+</td>
<td>X_t-2</td>
<td></td>
<td></td>
<td>0.000718(2.69)</td>
</tr>
<tr>
<td></td>
<td>X_t-5</td>
<td>0.0000618(2.28)</td>
<td>0.000529(1.74)</td>
<td>0.000734(2.75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R²=0.0553</td>
<td>R²=0.0722</td>
<td>R²=0.1303</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔR²=0.0375</td>
<td>ΔR²=0.0313</td>
<td>ΔR²=0.1073</td>
</tr>
</tbody>
</table>

Note: t-statistics shown in parentheses.
The pattern of impact of sports participation on male health is different for the 26 to 45 age group. Here increases in the level of sports participation now ‘depreciates’ the level of male health, as indicated by the negative signs on the coefficients. This is indicative of current sports activity becoming less healthy and this could be because male sport is traditionally more competitive, as noted earlier, and that in these years team sport activity, for example, is at its peak. An alternative, complementary, explanation may be that a relatively over-competitive approach is adopted regardless of the absolute duration and intensity of the activity. Similar negative effects are noticed for females in this age group but here, in contrast, acceleration of sports participation over the months’t-1’ to ‘t-2’, reduce the level of health. This could be evidence that as females in this age group more rapidly expand their participation, this brings with it health problems. This could be suggestive of a possible perverse impact from encouraging female participation in what have traditionally been male sports, as noted earlier. For example, it is well documented that anterior cruciate ligament injury incidence is much higher for females than for males in sport generally and particularly in the rapidly growing sport of female football. [31, 32]

These results contrast with the 46 to 65 age groups. Although there is some variation in the female participation estimates for 3x30mins, the results suggest that for both males and females acceleration of sports participation increases the level of health. These results, which resonate with the aggregate sample, are indicative of middle-aged individuals adapting and increasing their physical activity through sport to a form of engagement that can enhance health but that they need to do this at increasing levels. This is clearly indicative of an important transitional phase of activity. Finally, the results for the 65 and older age groups suggest that a relationship between levels of
sports participation and health returns. Maintaining existing levels of activity, presumably as the ability to increase sports participation begins to fall, means that for males a depreciation of the health stock takes place, but this is not the case for females. This suggests that towards the end of the life cycle of sports participation, males adopt practices that fail to maintain their stocks of health, unlike females.

**DISCUSSION**

**Main finding of this study**

Overall, the results of this paper suggest that the main causal link in the population between sports participation and health runs from the former to the latter. Moreover, there is evidence that acceleration in sports participation is required to elicit increases in the level of health. These results are consistent with a health production model. The paper also indicates that there is no evidence that the 2012 Olympics harnessed this relationship. This challenges policy presumptions about the impact of hosting major sports events on galvanising participation to produce positive health outcomes for society.

**What is already known on this topic?**

It is widely documented in the literature that sport and physical activity can improve health and, potentially, psychological mood and well-being. [7-9]

**What this study adds**

This study has created a time-series data set measuring self-reported general health and for the first time the actual minutes of sports participation of various intensities by exploiting the rolling monthly sampling strategy of the Taking Part Survey between July
2005 and June 2013. As well as the insights noted above, which support the view that sport can contribute to health, disaggregated analysis cautions against simple generalised conclusions. There is variation over the life course and across genders. From a policy perspective, this suggests that sport is not just a neutral form of physical activity but that its practice can have detrimental effects on health. Policy should also focus on the form of practice as well.

Limitations of the Study

The lack of any longitudinal data for the same individual means that true cohort and lifestyle transition effects cannot be observed. The bivariate nature of the analysis, with large proportions of the variation in health outcomes left unexplained, suggests that the inclusion of other mediating and perhaps confounding influences could also help to inform the results. In the current study design, however, averaging over individuals by making use of the rolling monthly sample design to generate time-series data means that other socio-demographic characteristics would just measure proportions of the population. Nonetheless, the bivariate analysis yields robust statistical relationships between the direct marginal impact of sport, controlling for and conditional on the history of health as captured by the lag structure of the models that are estimated. With this in mind, what is observed between health and sports participation in this research should still be of direct concern to policy and worthy of further reflection upon.
ACKNOWLEDGEMENTS

We would like to thank the Department for Culture Media and Sport for pre-release access to the Taking Part data.

REFERENCES


Rasciute, S., & Downward, P.M. Health or happiness? What is the impact of physical activity on the individual. *Kyklos* 2010;, **63**(2): 256–270


23 Abbott, B.D. and Barber, B.L. Differences in functional and aesthetic body image between sedentary girls and girls involved in sports and physical activity: Does sport type make a difference?, *Psychology of Sport and Exercise* 2011; **12**(3): 333-342.


30 Lechner, M. and Downward, P. Heterogeneous sports participation and labour market outcomes in England, IZA DP No. 7690


Figure 1: Mean Value of General Health

Figure 2: Sports Participation