Labour Economics 35 (2015) 1–15

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Labor market effects of sports and exercise: Evidence from Canadian panel data☆

Michael Lechner [a,c,e,f,g,h,*], Nazmi Sari [b,d,i,**]

[a] Swiss Institute for Empirical Economic Research, University of St. Gallen, Switzerland
[b] Department of Economics, University of Saskatchewan, Saskatoon, Canada
[c] CEPR, London, United Kingdom
[d] Saskatchewan Population Health Evaluation Research Unit, Saskatoon, Canada
[e] PSI, London, United Kingdom
[f] CESIfo, Munich, Germany
[g] IAB, Nuremberg, Germany
[h] IZA, Bonn, Germany
[i] Health Quality Council, Saskatoon, Canada

HIGHLIGHTS

- We analyse the effects of individual sports & exercise on labour market outcomes.
- We find substantial earnings effects.
- These effects occur only if the level of sports and exercise is high enough.

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 1 April 2015
Accepted 3 April 2015
Available online 10 April 2015

JEL classification:
Z28
Z29
I12
I18
J24
L83
C21

Keywords:
Physical activity
Canadian National Population Health Survey
Individual sport participation
Human capital
Labor market
Matching estimation

ABSTRACT

Based on the Canadian National Population Health Survey we estimate the effects of individual sports and exercise on individual labor market outcomes. The data covers the period from 1994 to 2008. It is longitudinal and rich in life-style, health, and physical activity information. Exploiting these features of the data allows for a credible identification of the effects as well as for estimating dose–response relationships. Generally, we find positive long-run income effects. However, an activity level above the current recommendation of the WHO for minimum physical activity is required to reap the long-run benefits.

☆ All computations have been performed in the Statistics Canada Research Data Center at the University of Saskatchewan in Saskatoon. We thank the Data Center and Statistics Canada for their permission to use the data for this research under project # 10-SSH-SKY-2603. All computations were implemented with our own programs using Gauss 11 installed in the Data Center. A previous version of the paper was presented at IZA, Bonn, ZEW, Mannheim, the Alpine Labour Seminar, Laax, the annual meeting of the German Economics Association in Hamburg, the annual meeting of the European Sports Economics Association in Antwerp, the EALE meeting in Ljubljana, as well as in seminars of the economic departments of the Universities of Aarhus, Copenhagen, and Southern Denmark at Odense. We thank participants, two referees of this journal as well as Helena Skyt Nielson for helpful comments and suggestions. The usual disclaimer applies.

* Correspondence to: M. Lechner, Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen, Varnbüelstrasse 14, CH-9000 St. Gallen, Switzerland.
** Correspondence to: N. Sari, Department of Economics, University of Saskatchewan, Arts 815, 9 Campus Drive, Saskatoon, SK S7N 5A5, Canada.
E-mail addresses: Michael.Lechner@unisg.ch (M. Lechner), Nazmi.Sari@usask.ca (N. Sari).
URL's: http://www.michael-lechner.eu (M. Lechner), http://homepage.usask.ca/sari/ (N. Sari).

http://dx.doi.org/10.1016/j.labeco.2015.04.001
0927-5371/

1. Introduction

Positive health effects of physical activity are well established in the literature.[1] In this paper we analyze whether physical activity also leads into higher productivity of the labor force and thus higher earnings? Although potentially highly relevant for public policy, the answer to this question may be less obvious than it appears. Even though the medical literature agrees that more activity is always better for health,[2] it is not obvious that such health effects translate one-to-one into earnings gains. For example, on the one hand increasing physical activity is likely to take-up additional time which leads to a reduction of other leisure or working time, with an uncertain effect on earnings. On the other hand, physical activity may also build other skills, like social capital (e.g. Seippel, 2006), team skills, and self-discipline that are expected to improve productivity, and thus earnings, per se. Therefore, in this paper we investigate the effects of various levels of sports and exercise on labor market outcomes directly.

The literature on the effects of sports and exercise on labor market outcomes (for working age adults) is limited. The main reason is the lack of data that are sufficiently (large and) rich to allow identification of the respective causal effects and that contain reasonably detailed information on both labor market outcomes and sports and exercise. Recently, Kavetsos (2011) analyzes this relation using cross-section data from 25 European countries. Based on a parametric IV approach with the regional prevalence of sports participation serving as instrument, he finds positive employment effects. Rooth (2011) uses an experimental setting to show that people signaling leisure time sports participation in their job application are more likely to be invited to a job interview. His experimental analysis is supplemented by an observational study based on Norwegian registry data, which suggests long-run earnings effects of physical fitness in the range of 2 to 5%. Finally, based on a large cross-sectional database for England, using semiparametric matching methods Lechner and Downward (2013) find a positive association between different types of sport activities and earnings.

There are two related studies based on German household panel data (German Socioeconomic Panel). Cornelißen and Pfeifer (2008) use random effects regression models and find positive earnings associations for men. Lechner (2009a) exploits the same panel data differently in an attempt to identify causal effects. The goal of his semiparametric approach is to use the panel data to deal with the issue of non-random individual selection into activity levels without having to resort to the rather restrictive parametric panel econometric models that implicitly impose assumptions on the effect heterogeneity and the relation of labor market outcome, sports participation, and (time constant) unobserved and observed confounders. The main idea of his research design, which we follow closely in this paper, is to divide the time periods into three different types: a base period ('−1'), a 'treatment' period ('0'), and subsequent post-treatment periods. The base period is used to measure the confounding variables, e.g. health or education among many others, and, to condition on pre-treatment outcomes and activity levels. Exogeneity of these conditioning variables is ensured by splitting the sample in strata defined by the physical activity status in this base period. By definition, within such a stratum these variables cannot be affected by unobserved factors related to sports and exercise participation. If lagged outcome variables are affected by unobserved variables (e.g. time preference) in a similar fashion over time (which is a typical assumption underlying fixed effects estimators), together with the semiparametric estimation approach, the lagged outcomes in the strata essentially control for unobserved 'fixed effects'. In his design, semiparametric matching estimation is used to minimize the dependence on arbitrary parametric econometric models that uncovered long lasting earning gains in the range of about 10%. Overall, such design used with informative variables to control for additional confounding has the potential to lead to credible and robust causal inference and avoids some of the issues that were present in the other papers mentioned above.[3] In Lechner's study, however, samples were rather small and some important confounders that could be time varying, like detailed health information, were missing. Furthermore, the most important activity measure was not detailed at all.

This paper uses the basic design of Lechner (2009a) and implements a similar estimation strategy, however, using substantially more informative data from Canada. To be more precise, the empirical strategy consists of the following steps: The analysis is based on a population that is of age 20 to 44 in 1994 (the base year) and followed until 2008. We estimate the effects of three levels of activity ('treatment') defined in 1996 (the panel survey is biannually conducted). The data are stratified according to activity level in 1994 and according to sex, since effects and participation in activities are known to be heterogeneous w.r.t. to activity level and sex. Covariates and pre-treatment outcomes that are used in matching are measured in 1994. Matching estimation is performed within each stratum to estimate the effects of various outcome variables measured from 1998 to 2008. Subsequently, the strata specific results are aggregated to compute overall effects.

We attempt to contribute to the literature in several dimensions: Firstly, we improve the credibility of the actual identification of causal effects (compared to Lechner, 2009a) by using data with more informative health information which allows controlling for health conditions (in 1994) in a more detailed way. Secondly, the more detailed physical activity information allows to some extent to uncover dose–response relationships, i.e. to investigate how the effects depend on the intensity of the activity (which was also absent from Lechner, 2009a). Furthermore, since the data cover a period from 1994 to 2008, effect dynamics as well as medium to long-run impacts are estimated.

We find generally positive earnings effects of 10% to 20% (after 8 to 12 years), but no systematic and statistically significant effects on other labor market outcomes, like employment status or hours worked. Interestingly, an important dose–response relationship appears: To get the full benefits of sports and exercise participation, it is necessary to be active at the highest of the three activity levels considered (which is above the current recommendations for minimum physical activity, e.g. World Health Organization, 2010). Thus, these results suggest that the current activity levels of large parts of the Canadian population, which are not so different compared to the activity levels observed in many other developed countries, are still far below the point for which a further increase would lead to negative returns in terms of earnings. As we measure the effects over time, and as it is plausible that for various reasons labor market effects need time to become effective, it is reassuring that in the very short-run the effect appears to be close to zero, because a larger initial non-zero effect might be suggestive of remaining selection problems.

The structure of the paper is as follows: In the next section, we introduce the data. Section 3 discusses some general features of sports and exercise in Canada. Section 4 outlines the research design and the estimation strategy used. Section 5 contains the results and some sensitivity checks. Section 6 concludes. There are several appendices that

[1] See for example the literature review by Warburton et al. (2006). More recent exhaustive literature reviews are provided by U.S. Department of Health and Human Services (2008), the Annex II of EU (2013) and Reiner et al. (2013), among several others.

[2] While Warburton et al. (2006) state that "There appears to be a linear relation between physical activity and health status, such that a further increase in physical activity and fitness will lead to additional improvements in health status" (p. 801), a recent study for Canada by Humphreys et al. (2014) finds positive, but decreasing health effects ("Increasing the intensity above the moderate level and frequency of participation in physical activity appears to have a diminishing marginal impact on adverse health outcomes", p. 1).

[3] Examples of such issues are restrictions on effect heterogeneity by using parametric models, possible reverse causality when measuring confounders, or the lack of credibility and inference problems of some suggested IV estimators.