UNIVERSITY OF LIVERPOOL | Management School

Working Paper in Economics

 # 202017

June 2020

Does sport affect health and well-being or is it the other way around? A note on reverse-causality in empirical applications

Jing Guan
J.D. Tena

Does sport affect health and well-being or is it the other way around? A note on reverse-causality in empirical applications

by

Jing Guan, and J.D. Tena

Abstract

Simultaneity of endogenous variables is an issue in econometrics analysis as it affects the identification of structural parameters. Dealing with it implies the use of untestable assumptions such as, for example, choosing relevant instrumental variables, relying on the strong ignorability condition or imposing restrictions in the direction of causality among the endogenous variables. In order to overcome this problem, many recent empirical papers in the sport literature have proposed the use of seemingly aseptic approaches based on the estimation of simultaneous equations for the endogenous variables with correlated error terms. In this note, we discuss the required conditions to consider this approach as a valid alternative to instrumental variables in order to deal with the simultaneity problem.

Key words: *Simultaneity, SUR, structural models*

1. Introduction

The purpose of this note is to discuss the use of systems of simultaneous equations in the empirical literature to estimate the impact of health behaviour variables, such as sport and/or physical activity, on health and wellbeing.[1] Recent years have witnessed the availability of surveys that allow for the observation of these variables together with other individual socio-economic characteristics. Although this information has generated a burgeoning research literature aiming to estimate the main determinants of health and wellbeing[2], an important concern in this type of analysis regards the simultaneous observation of the different variables in the model. This issue becomes especially problematic as many of the databases are cross-sectional data which makes the identification of causal impacts a very difficult task due to the impossibility to identify whether life style variables affects health outcomes or it is the other way around.

Simultaneity is an old problem in many ambits of economics and, in fact, it has been an issue of research even before the creation of the Cowles research institute to the development of econometrics in 1932; see the canonical example in economics about the identification of demand and supply (Wright, 1928). In this context, it is generally accepted that identification can only be achieved by means of untestable assumptions. Some examples are the use of instrumental variables that are selected under the exclusion restriction and matching or propensity score regressions under the strong ignorability assumption; see Woolridge (2003) and Imbens and Woolridge (2009) and references therein for relevant examples of these methodologies. Even in the absence of specific instrumental variables, identification can still be achieved by other means, such as assuming a direction of causality among the endogenous variables (recursiveness assumption) or imposing the effect of some specific shocks to be negligible in the long term (long-run restrictions) among others; see Christiano et al. (1999) and Blanchard and Quah (1989) respectively.

[1] See, for example, Contoyannis and Jones (2004), Bali and Jones (2008) and Humphreys et al. (2014) among others.
[2] Just to cite two examples, see Forrest and McHale (2011) and Wicker and Frick (2015).

In the recent health and sport literature, it has become fashionable the use simultaneous equation models to deal with simultaneity in cross-sectional databases. These models typically do not use instruments to identify the direction of causality of simultaneous variables but, instead, this is imposed by the recursiveness assumption. We argue this is a sensible approach only when there are solid theoretical arguments to justify this restriction. However, this is not the case if there is double causality between health behaviour and health outcome. We also discuss the case where simultaneity affects more than one response variable. In this case, contrary to the claims of previous papers in the literature, the estimation of reduced form specification is generally not a valid alternative to deal with the simultaneity problem of response variables regardless of whether a seemingly unrelated regression (SUR henceforth) strategy is used or not to account for the fact that errors in the different equations are potentially correlated.

This brief note does not attempt neither to survey the extensive body of literature which deals with this type of estimation nor to discuss the main econometric properties of the estimators. Our main purpose is to discuss the theoretical implications of the use of simultaneous equation models to deal with endogeneity adopted by an important strand of the sport economic literature

2. General discussion

A common interest in the empirical literature discussed in this note is the causal impact estimation of health behaviour, typically sports or physical activity, on a set of response variables which can include health, well-being or academic performance measures. This implies to deal with two the following two types of identification problems: 1) simultaneity between exercise and response variables; and 2) simultaneity across the different response variables. We discuss these issues sequentially.

2.1. Impact of physical activity on a single response variable

If physical exercise and a single response variable are simultaneously observed, identification can still be achieved by finding instruments that only affect one of the two endogenous variables. If this is not possible, we require a restriction assumption on the direction of causality. For example, Contoyannis and Jones (2004) and Balia and Jones (2008) consider recursive multivariate probit models where reduced form specifications are considered for health behaviour activities while a latent health stock is contemporaneously affected by health behaviour. Similarly, Humphreys et al. (2014) also impose unidirectional causality from physical activity to health outcome. They justify this assumption on theoretical arguments as physical activity is an input in the regression of health. Therefore, in all these papers, the use of a simultaneous system of equations can be useful to deal with the endogeneity problem due to omitted variables that can affect both variables of interest. However, this approach does not solve the identification problem due to the simultaneous observation of physical exercise and health outcome which is achieved by the imposition of the recursiveness assumption.

Given this discussion, a sensible approach would be to compare the estimation results under the recursiveness assumption with those obtained from the use of instrumental variables. For example, Humphreys et al. (2014) consider a variable called sense of belonging to social community as instrument. They argue that this variable is only assumed to affect physical activity but not health. Their causal impact estimates are robust to the use of these two alternative methodologies.

2.2. Impact of physical activity on a multiple response variables

The simultaneous observation of different response variables creates an additional endogeneity problem as, in most cases, it is not possible to set a direction of causality across them based on economic theory. For clarity of exposition, let's abstract from the endogeneity problem concerning physical exercise, denoted by $p_{i,t}$ and consider the following system of structural equations for two response variables, $y_{1i,t}$ and $y_{2i,t}$,

$$y_{1i,t} = a + by_{2i,t} + \delta p_{i,t} + \beta' z_{i,t} + \vartheta' w_{1,i,t} + \varepsilon_{1,i,t} \quad (1)$$

$$y_{2i,t} = c + dy_{1i,t} + \phi p_{i,t} + \gamma' z_{i,t} + \phi' w_{2,i,t} + \varepsilon_{2,i,t} \quad (2)$$

where the vector of variables $w_{1,i,t}$ and $w_{2,i,t}$ are specific to each of the two equations; $z_{i,t}$ is a vector of common covariates for the two endogenous variables; and $\varepsilon_{1,i,t}$ and $\varepsilon_{2,i,t}$ are fundamental structural shocks, $E(\varepsilon_{1,i,t}, \varepsilon_{2,i,t}) = 0$. Variables $z_{i,t}$, $w_{1,i,t}$ and $w_{2,i,t}$ can include exogenous and predetermined variables that are uncorrelated with $\varepsilon_{1,i,t}$ and $\varepsilon_{2,i,t}$.

As in the previous section, the exclusion restriction is a well-known identification condition of the system above. It establishes that identification of our parameters of interest, , δ and ϕ, together with all the other parameters in equations (1)-(2) is only possible if variables $w_{1,i,t}$ or $w_{2,i,t}$ can be observed. However, if this is not the case, equations (1) and (2) are observationally equivalent. In empirical applications, it is not always possible to find instruments that only affect one particular response variable but not the other and this necessarily means that a direction of causality must be imposed, i.e. either b or d must be set equal to zero, to achieve identification.

The aforementioned simultaneity problem cannot be solved by a joint estimation of the reduced form version of equations 3 and 4. This issue has been considered in Rasciute and Downward (2010) and references therein to estimate the impact of physical exercise on health and happiness. They claimed to solve the simultaneity problem between health and well-being estimating the following reduced form version of the system described by equations (1) and (2) by means of a SUR model:

$$y_{1i,t} = e + \pi p_{i,t} + \tau' z_{i,t} + v_{1,i,t}, \quad (3)$$

$$y_{2i,t} = f + \rho p_{i,t} + \lambda' z_{i,t} + v_{2,i,t}. \quad (4)$$

However, this estimation does not solve the simultaneity problem between $y_{1i,t}$ and $y_{2i,t}$. First, error terms in expression (3)-(4) $v_{1,i,t} = \frac{\varepsilon_{1,i,t} + b\varepsilon_{2,i,t}}{1-db}$ and $v_{2,i,t} = \frac{d\varepsilon_{1,i,t} + \varepsilon_{2,i,t}}{1-db}$ are correlated, regardless of whether there is not any omitted variable in the model, just if either $b \neq 0$ or $d \neq 0$ .More importantly, even if this correlation is taking into account by means of a SUR model, $\pi = \frac{\delta + b\phi}{1-bd}$ and $\rho = \frac{d\delta + \phi}{1-bd}$ are unbiased estimates of the structural parameters of interest δ and ϕ only if there is not simultaneity between $y_{1i,t}$and $y_{2i,t}$, i.e. if both $b = 0$ and $d = 0$. Therefore, taking into account the correlation between $v_{1,i,t}$ and $v_{2,i,t}$ can improve estimation efficiency but it will result in a biased causal estimation unless the simultaneity between $y_{1i,t}$ and $y_{2i,t}$ is properly addressed.

Concluding remarks

Dealing with simultaneity in cross-sectional databases is a difficult task which requires the use of untestable identification assumptions either in the choice of instruments or the direction of causality. Therefore, when an exclusion restriction is not found, this problem can only be solved by a joint estimation of equations for each of the simultaneously observed variables if we have strong arguments to accept that health behaviour affects health outcome but not the other way around. The problem of simultaneity can also regards the estimation of the effect of health behaviour on several outcome variables. In this case, contrary to what is claimed by some papers in the literature, if response variables are simultaneously related, in the absence of exclusion restrictions, a SUR reduced form specification produces biased estimation of causal effects. A more sensible approach in this circumstance would be to study the robustness of the results to changes in the direction of causality imposed by the recursiveness assumption.

References

Balia S, Jones A. 2008. Mortality, lifestyle and socio-economic status. Journal of Health Economics 27: 1–26.

Blanchard, O. J. and D. Quah. 1989. The dynamic effects of aggregate demand and supply disturbances. American Economic Review 79: 655–673.

Christiano, L.J., Eichenbaum, M. and Evans, Ch. L. (1999) Monetary Policy Shocks: What have We Learned and to What End?, Handbook of Macroeconomics, J. B. Taylor & M. Woodford (ed.), Vol. 1, 65 – 148.

Contoyannis P, Jones A. 2004. Socio-economic status, health and lifestyle. Journal of Health Economics 23: 965–995.

Forrest, D., & McHale, I. (2011). Subjective well-being and engagement in sport: Evidence from England. The economics of sport, health and happiness: the promotion of well-being through sporting activities. Glos, UK: Edward Elgar Publishing, 184-199.

Humphreys, Brad R, McLeod, Logan and Ruseski, Jane E (2014). Physical Activity and Health Outcomes: Evidence from Canada. Health Economics, 23: 33-54.

Imbens, G.W. and J.M. Wooldridge (2009) Recent Development in the Econometrics of Program Evaluation, Journal of Economic Literature, 47(1), 5-86.

Rasciute, S., and Downward, P. (2010). Health or Happiness? What Is the Impact of Physical Activity on the Individual? Kyklos, 63(2), 256–270.

Wicker, P., & Frick, B. (2015). The relationship between intensity and duration of physical activity and subjective well-being. The European Journal of Public Health, 25(5), 868-872.

Woolridge, J.M. (2003) Further Results on Instrumental Variables Estimation of Average Treatment Effects in the Correlated Random Coefficient Model", Economic Letters, 79, 185-191.

Wright, Ph. G. (1928). The Tariff on Animal and Vegetable Oils. New York: MacMillan.