Reproducibility and Applied Econometrics - The Effect of Studying on Grades

Nadav Tadelis

April 2018

Abstract

In this paper we establish a framework for reproducible empirical research. We use a non-standard 2SLS model to estimate the marginal effects of studying on grades. The paper is split into two distinct sections. The first part is the econometric analysis on the causal impact of studying. The second part details the steps taken to ensure reproducibility and suggests how to easily integrate these methods into a researcher's future projects.

1 Introduction

This paper explores what responsible and reproducible research practices look like in applied econometrics. We present an instrumental variables approach to estimating the causal effect of studying on grades and develop custom python scripts to implement an unusual 2SLS set up that allows for nonlinearity in our endogenous predictor. The latter part of this work discusses the current state of reproducibility in econometric research, and explains in detail the techniques we implement in the analysis.

In recent years there has been a strong push to increase the reproducibility and replicability of scientific research. Unfortunately this movement seems to have been centered on the hard sciences and has not yet become standard practice in the social sciences. It is possible that this is partly due to a lack of reproducibly researched papers in these fields. This paper attempts to show an example of what reproducibility looks like for empirical work in the social sciences, especially in applied econometrics.

This paper was written as my honors thesis for undergraduate studies in statistics at UC Berkeley. Due to deadlines for submission I was unable to spend an appropriate amount of time on the econometric analysis, and will point out some of the weak points in my models (specifically the instruments). Any comments would be greatly appreciated.

citation test [3, 2, 1]

References

[1] J. D. Angrist, G. W. Imbens, and D. B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91:444–455, 1996.

[2] M. Goossens, F. Mittelbach, and A. Samarin. *The LaTeX Companion*. Addison-Wesley, Reading, Massachusetts, 1993.

[3] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.