# Causal Inference with Observational Data

## 1. Introduction

Marc F. Bellemare

May 2018

# Introduction

Dave Giles wrote:

> *I am definitely not a fan of "cookbook econometrics." It's pointless, and frankly dangerous, to simply tell students what to do without telling them why ... [T]he real benefit in being led through the proof of a standard result in econometrics is that it enables you to see exactly where (and how) any underlying assumptions are actually used.*

I agree.

# Introduction

So why am I here to teach a “cookbook econometrics” course?

While it is difficult to argue with Prof. Giles’ claim that “being led through the proof ... enables you to see exactly where (and how) any underlying assumptions are actually used,” my goal with this class is both more humble and more advanced than what he has in mind: I want to teach you *how* to use the standard empirical methods used to make causal inference with observational data, and how to use them like seasoned applied econometricians do.

# Introduction

Your taking this class presupposes that you have already learned those standard methods in previous classes—preferably in a year-long, first-year doctoral econometrics sequence.

The problem is that in the typical first-year econometrics sequence, there is so much ground to cover that instructors typically have little to no time to discuss how the material is applied.

The end result is something like the following dialogue from the movie *Aliens* (1986).

# Introduction

*Ripley: How many drops for you is this, lieutenant?*
*Gorman: Thirty-eight. Simulated.*
*Vasquez: How many combat drops?*
*Gorman: Uh, two. Including this one.*
*Drake: Sh*t.*

# Introduction

This is especially problematic given the often wide gap between the theory and practice of econometrics.

Here is a simple example to motivate our work: At some point or other, everyone in this room was almost surely taught that one can use a Durbin-Wu-Hausman test to test for exogeneity.

This is true in theory. But is it true in practice?

# Introduction

Suppose, for instance, that a third-year PhD student wants to estimate a Mincerian equation, such that

$$\ln w_i = \alpha + \beta s_i + \gamma x_i + \epsilon_i, \quad (1)$$

where $w$ denotes the wage of individual $i$, $s$ denotes her years of schooling, and $x$ denotes her education.

Our young economist knows that a person $i$'s education is likely to be endogenous to that person's wage, but she is lucky to have a valid instrument for education, so she estimates two versions of the equation: an OLS version, and a 2SLS version.

# Introduction

She then runs a Durbin-Wu-Hausman test pitting $(\widehat{\alpha}_{OLS}, \widehat{\beta}_{OLS}, \widehat{\gamma}_{OLS})$ against $(\widehat{\alpha}_{2SLS}, \widehat{\beta}_{2SLS}, \widehat{\gamma}_{2SLS})$, whose null hypothesis that the two parameter vectors are identical implies that education is exogenous to wage, and she fails to reject the null.

She then writes up her results, explaining that her OLS results identify the causal effect of education on wage (maybe presenting the 2SLS results in an appendix), she submits his article for publication at a good journal, and she gets rejected.

Why?

# Introduction

Or, to stick with Hausman tests, suppose another young economist has longitudinal data, and wants to estimate a different version of a Mincerian equation, such that

$$\ln w_{it} = \alpha + \beta s_{it} + \gamma x_{it} + \delta_i + \epsilon_{it}. \tag{2}$$

Our young economist isn't sure whether $\delta$ should be a coefficient or a component of the error term, so he estimates two versions of the equation, one with fixed effects (FE), one with random effects (RE).

So he runs a Durbin-Wu-Hausman test pitting $(\widehat{\alpha}_{RE}, \widehat{\beta}_{RE}, \widehat{\gamma}_{RE})$ against $(\widehat{\alpha}_{FE}, \widehat{\beta}_{FE}, \widehat{\gamma}_{FE})$, whose null is such that the two parameter vectors are identical, and so that random effects is the right specification since in that case it is efficient. He fails to reject the null.

So he writes up his results, explaining that his RE results identify the causal effect of education on wage, then submits his article for publication at a top general journal. He, too, gets rejected.

Why?

# Introduction

Since the Credibility Revolution (Angrist and Pischke, 2010), several of us know the answer to the questions just posed, but this isn't exactly the kind of thing you learn in your core econometrics courses.

A course like this one is especially important given the tacit knowledge and social norms involved in the practice of applied econometrics: what to do, what not to do, what to show, what can be safely made "available upon request from the authors," what to emphasize, what to put in an appendix, and so on.

# Introduction

Another thing that compounds the problem is that field courses (e.g., ag, development, health, labor) tend to present only the cream of the crop of papers, so it is often difficult for students to know what differentiates those excellent papers from papers that are merely average, or from very bad papers. As with everything else, if you want to learn what works here, you need both a treatment and a control group.

But unless you do a lot of reviewing for journals, you probably don't get to read many bad papers in order to learn what works and what doesn't. And why should you? Life is too short!

# Introduction

Worse, the papers that are typically found on the syllabus of field courses are often heavily edited—they are versions of a paper where the editor has said: "Good enough. I'm convinced. You don't really need to include all those robustness checks or those alternative specifications."

Well, yes—the editor is convinced because those robustness checks and alternative specifications were included to begin with!

Case in point: The second-to-last version of my most recent article (Bellemare and Nguyen, 2018) had 10 tables. After the editor asked us to shorten the manuscript as a condition for acceptance, the published version ended up having three tables!

# Introduction

Another issue is that although it is easy to see what the authors of papers in top journals do when it comes to applied econometrics, there are a number of what Sherlock Holmes would have referred to as "dogs that didn't bark."

In other words, it is not easy to keep track of the things you learn in your classes but which successful authors do not use in their own work.

For example, a search for all economics articles on JSTOR published during the period 2012-2017 whose abstracts include the words "Heckman" and "selection" yielded four articles, none in well-known journals. Yet, the Heckman selection model is still frequently taught in PhD courses, and it is often given a weight equal to everything else.

# Introduction

In sum, this course is a complement to—not a substitute for—a first-year PhD econometrics sequence.

Even the best textbooks (e.g., Wooldridge's introductory textbook, or his more advanced textbook) only offer a snapshot of the social norms people follow when doing applied work at a single point in time.

But those norms evolve all the time.

# Introduction

What I bring to the table is almost 20 years of hands-on experience as an applied econometrician, of which 12 years as a faculty member at leading US research universities.

I also have the added bonus of having collected my own survey data and of having run my own experiments a number of times.

It is my sincere hope that this class can help fill the gap between theory and practice, and make your empirical results more credible and more convincing.