

Getting Started I: *Introduction to Econometrics*

1. What's the objective?

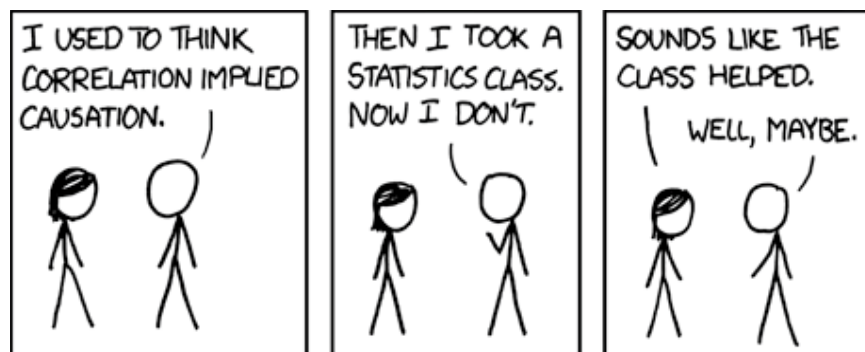
- a. Favorite coefficient models (the *variable in the spotlight*; you really only care about one relationship; *more is more*)
- b. Behavioral/descriptive models (understanding relationships; *parsimony preferred*)
- c. Forecasting models (predicting ahead, *out-of-sample*; *less is more*)

2. Estimating economic/statistical relationships

- a. Correlation drives econometrics
- b. To what extent are changes in X (alone) associated with changes in Y
 - i. *Ceteris Paribus* (everything else the same)
 1. Are you really controlling for all other factors? What did you miss?
 - ii. Where's the user's manual? *Oops, no manual!*
 1. What did you leave out of the model? ... and does it matter?
- c. Data:
 - i. Randomized controlled studies ... *if only*
 - ii. Real world (not so simple)... *observational studies* ... how representative is your sample?

3. Causality v. correlation: *Post hoc ergo propter hoc* (after this, therefore because of this)

- a. It's so tempting to make causal inferences, but never forget, correlation is not causation.



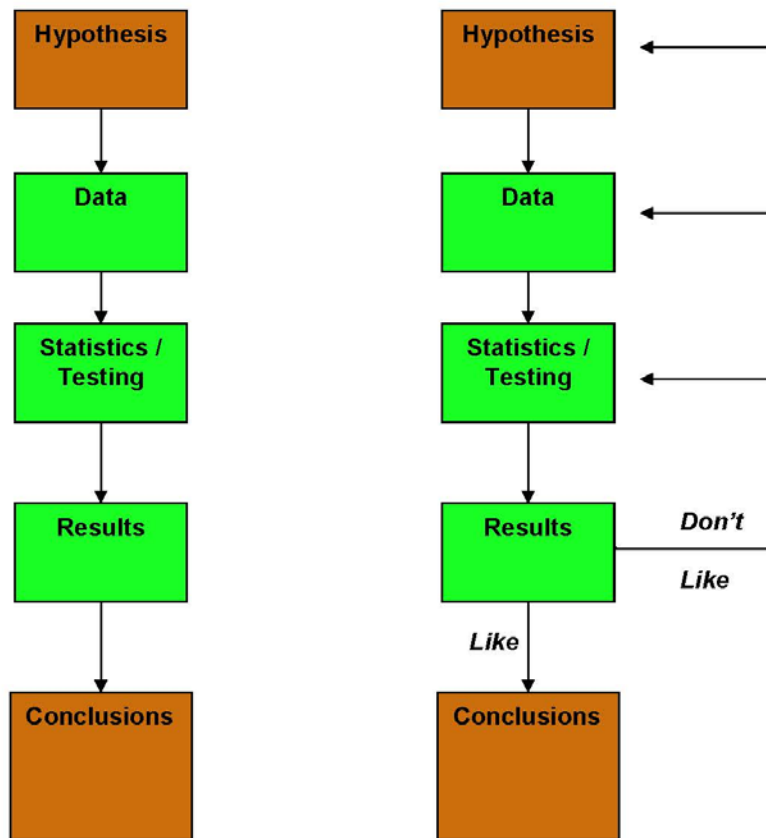
Introduction to Econometrics

- b. Classic example: # firefighters (or fire trucks) and damage from fires; great story but what do the data show? (CT data show no correlation)
- c. Spurious correlation:¹ <http://www.tylervigen.com/>
- d. So this not just a statistical exercise... Keep asking: Does this make sense? What have I learned? What mysteries remain? What do I do next?

4. Doing empirical economics

- a. What is the question?
- b. You have to start somewhere:
 - i. You have a brain... use it!
 - ii. (Economic) Theory weighs in
- c. Hone in on the specific hypothesis to be tested.

The Scientific Method: Fantasy v. Reality



¹ Some YouTube videos: <https://www.youtube.com/watch?v=g-g0ovHjQxs>, <https://www.youtube.com/watch?v=8B271L3NtAw> and <https://www.youtube.com/watch?v=lbODqslc4Tg>

- d. Bring on the data
- i. **Data integrity is #1 concern in doing econometrics! Beware data GIGO!**²
1. You can never spend enough time looking at your data ... bad data? biased sample? ... and if you spot issues:
 - a. Try to get better data/sample ... laziness is no excuse!
 - b. And if you can't get better data/sample, you can sometimes fix things using *oh so sophisticated* methods...
 - c. And even when you can't do that... maybe you can still say something about directions and magnitudes of effects/relationships?
 - d. So cheer up... it's not hopeless!
 - ii. Dependent variable (LHS variable)... typically Y ... what's to be explained
 - iii. Independent/explanatory/control variables/covariates (RHS variables)... typically X's
 - iv. **SLR v. MLR models**
 1. Simple Linear Regression (SLR) Models: You only have one explanatory (RHS) variable? I know these models are easier to work with... *but are you the laziest person ever?... or what?*
 2. Multiple Linear Regression (MLR) Models: Bring on the explanatory variables! ... the more the merrier! Everything is a bit more complicated... *but maybe life is complicated! get over it!*
 - v. You can easily spend 90%+ of your time building your datasets... don't skimp at this stage... your data drives your results!
 1. Why didn't you bring more data to the party? Laziness is no excuse!
 - vi. **Endogeneity (Omitted Variable Impact/Bias) is the #2 concern in doing econometrics!** ... You can't worry enough about this! Will your results stand up when Z is added to the model?

#1: Data Integrity
#2: Endogeneity
#3: ... not close!



1. *Damn those ^%#\$%#@#*&# confounding variables! (correlated with X's and Y's... and left out of (omitted from) the model*
2. Focus on what's in your model, but focus more on what you left out... gazillions of explanatory factors.. maybe one was important? How do you know without taking a look? Laziness is no excuse!

² GIGO: Garbage In; Garbage Out.

Introduction to Econometrics

- e. Estimate the model (specify a functional relationship? or maybe not!)
 - i. But not just one model: robust (sensitivity) analysis
 - ii. Econometrics as storytelling... every model answers some question... but is it the right question? ... or one you care about?
 - iii. In the end you'll have a preferred model... but the compelling analysis doesn't rely on a single model
 - iv. Be careful though... How many regressions did you run? (see *skepticism* below)
- f. Estimated effects: signs and significance
 - i. Look at those signs... Are they positive or negative? What is the direction of the estimated relationship? Does the sign of the estimate make sense? And if not, then what are you going to do about it?
 - ii. Signs relate to (partial) correlations in the data (controlling for everything else)
 - iii. Then look at the specific parameter estimates... Are you impressed? Are they statistically significant? (you can reject the Null Hypothesis of no/zero/nada/zilch effect... @ some credible significance level)
- g. *Economic significance (meaningfulness)* v. *statistical* significance
 - i. It might be a statistically significant effect... But does it really matter? Is the estimated impact meaningful? Should anyone care? When you brag about your results, will everyone just laugh at you?
 - ii. Just because the effect is precisely estimated doesn't mean that it's large in magnitude... or worthy of attention.
 - iii. Testing for statistical significance is relatively easy... determining whether an estimated effect is meaningful is not so easy, and very much judgmental and *ad hoc* (common sense; eyeball test; elasticities; *beta* regressions)
- h. Results/conclusions/hypotheses (not) rejected
 - i. How comfortable are you generalizing your conclusions? That was, after all, the driving force behind your analysis.
 - ii. And what about that smell test? ... will everyone just laugh at you?

5. Be a complete skeptic!

- a. Bring a healthy skepticism to all published research... and be pleasantly surprised!

6. Art v science

- a. [Data Matters](#) ... *and Art matters too!*
- b. Know your subject, this is not just a statistical exercise... and don't forget the smell test!

DATA MATTERS

