Economics Has a Math Problem - Bloomberg

A lot of people complain about the math in economics. Economists tend to quietly dismiss such complaints as the sour-grapes protests of literary types who lack the talent or training to hack their way through systems of equations. But it isn't just the mathematically illiterate who grouse. New York University economist Paul Romer -- hardly a lightweight when it comes to equations -- recently complained (http://paulromer.net/mathiness/) about how economists use math as a tool of rhetoric instead of a tool to understand the world.

Personally, I think that what's odd about econ isn't that it uses lots of math -- it's the *way* it uses math. In most applied math disciplines -- computational biology, fluid dynamics, quantitative finance -- mathematical theories are always tied to the evidence. If a theory hasn't been tested, it's treated as pure conjecture.

Not so in econ. Traditionally, economists have put the facts in a subordinate role and theory in the driver's seat. Plausible-sounding theories are believed to be true unless proven false, while empirical facts are often dismissed if they don't make sense in the context of leading theories. This isn't a problem with math -- it was just as true back when economics theories were written out in long literary volumes. Econ developed as a form of philosophy and then added math later, becoming basically a form of mathematical philosophy.

In other words, econ is now a rogue branch of applied math. Developed without access to good data, it evolved different scientific values and conventions. But this is changing fast, as information technology and the computer revolution have furnished economists with mountains of data. As a result, empirical analysis is coming to dominate econ.

One sign of this is the sudden burst of interest in machine learning in the economics field. Machine learning is a broad term for a collection of statistical data analysis techniques that identify key features of the data without committing to a theory. To use an old adage, machine learning "lets the data speak." In the age of Big Data (http://web.stanford.edu/~jdlevin/Papers/BigData.pdf), machine learning is a hot field in the technology business, and is a key tool of the rapidly expanding field of data science. Now, econ is catching the bug.

Two economists who have been pushing for the adoption of machine learning techniques in economics are Susan Athey and Guido Imbens of Stanford University. The two economists explained (https://twitter.com/JustinWolfers/status/622460229377245184) machine learning techniques to an interested crowd at a recent meeting of the National Bureau of Economic Research. Their overview stated that machine learning techniques emphasized causality less than traditional economic statistical techniques, or what's usually known as econometrics. In other words, machine learning is more about forecasting than about understanding the effects of policy.

That would make the techniques less interesting to many economists, who are usually more concerned about giving policy recommendations than in making forecasts. But Athey and Imbens have also studied (https://faculty-gsb.stanford.edu/athey/documents/HeterogeneousEffects.pdf) how machine learning techniques can be used to isolate causal effects, which would allow economists to draw policy implications.

Basically, Athey and Imbens look at the problem of how to identify treatment effects. A treatment effect is the difference between what would happen if you administer some "treatment" -- say, raising the minimum wage -- and what would happen without the treatment. This can be very complicated, because there are lots of other factors that affect the outcome, besides just the treatment. It is also complicated by the fact that the treatment may work differently on different people at different times and places. A final problem is that the data economists have to answer the question is usually very limited -- a big impediment for traditional econometrics, which generally assumes that the amount of data is comfortably large. Athey and Imbens deal with these issues by importing a method from data science, called a regression tree. Statistically literate readers can peruse their slides here (http://www.nasonline.org/programs/sackler-colloquia/documents/athey.pdf).

Another economist who has looked at the potential of machine learning is Hal Varian, a highly successful former professor who now serves as the chief economist at Google. In a 2013 paper, Varian released a paper discussing how new machine learning techniques developed by data scientists can help economists improve their understanding of reality. For example, he discusses how machine learning can help choose between different models (something economists often ignore), cope with uncertainty about which model is correct and avoid overfitting (overly complex explanations that can't predict anything). In a set of slides (http://web.stanford.edu/class/ee380/Abstracts/140129-slides-Machine-Learning-and-Econometrics.pdf) released in early 2014, Varian tied machine learning techniques to the recent rise of quasi-experimental methods in economics. This represents a fusion between traditional econometrics and new data science techniques.

Varian, Athey and Imbens are not the only examples of this mini-trend. Data science blogger Kenneth Sanford has a few more (http://blogs.sas.com/content/subconsciousmusings/2015/06/05/is-machine-learning-trending-with-economists/).

Clear thinking from leading voices in business, economics, politics, foreign affairs, culture, and more.

Share the View

So is economics going to become another branch of applied math? Will econometrics and data science merge? Berkeley economist Brad DeLong thinks so. "The work [of economics] will be done," he writes, "by data scientists, computer modelers, and historians of various stripes." That is almost certainly too extreme a prediction. But the interest in machine learning is just one more sign that economics may be starting to shed its peculiar fixation on theory and join its cousins in the data-driven future.