Econometrics, Predictive Modeling, and Causal Estimation

Francis X. Diebold
University of Pennsylvania

April 28, 2017

The "Full Ideal Conditions" (FIC)

The data-generating process is:

$$y = X\beta + \varepsilon$$

$$\varepsilon \sim N(\underline{0}, \sigma^2 I)$$

$$E(X'\varepsilon) = 0,$$

and the fitted model matches it exactly.

The Predictive Modeling Problem

A major goal in econometrics is predicting y. In the language of estimation, the question is "If a new person arrives with covariates $x = (1, x_2, ..., x_K), x \in X$, what is my minimum-MSE estimate of her y? So we are estimating a conditional mean $E(y|x)$.

Predictive modeling (assuming linearity):

$$\widehat{E(y|x)} = \hat{\beta}_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_K x_K$$

The Causal Estimation Problem

A major goal in econometrics is predicting the effects of exogenous "treatments" or "interventions" or "policies". In the language of estimation, the question is "If I intervene and give someone a certain treatment $\partial x_j, x_j \in X$, what is my minimum-MSE estimate of her ∂y?" So we are estimating the partial derivative $\partial y / \partial x_j$.

Causal estimation (assuming linearity):

$$\widehat{\partial y / \partial x_j} = \hat{\beta}_j$$

In general, estimating a partial derivative $\partial y / \partial x_j$ is very different from estimating a conditional mean $E(y|x)$.

So, Two Different Estimation Problems. Alternatively, Two Different Prediction Problems...

Predict y given $x, x \in X$

vs.

Predict ∂y given $\partial x_j, x_j \in X$

We will use:

"Predictive Modeling"

vs.

"Causal Estimation"

Under conditions, both problems are solved simultaneously by LS.

What conditions?

The FIC!

The data-generating process is:

$$y = X\beta + \varepsilon$$

$$\varepsilon \sim N(\underline{0}, \sigma^2 I)$$

$$E(X'\varepsilon) = 0,$$

and the fitted model matches it exactly.

What Happens to Predictive Modeling When the FIC Fail?

Nothing.

LS regression always consistent for $E(y|x)$.

(Follows from the LS estimation criterion.)

What Happens to Causal Estimation When the FIC Fail?

Potential doom.

LS regression *inconsistent* for $\partial y/\partial x_j$

LS s.e.'s also inconsistent

The Old-School Causal-Estimation Response

Fix key FIC violations as follows:

(1) "Find" strong and exogenous "instrumental variables"
for consistent treatment effect estimation.

– A passive strategy based on non-experimental data.

(2) GLS modeling of heteroskedasticty and/or autocorrelation
for consistent s.e.'s

– A more active strategy.

"Complete modeling"

Penn
University of Pennsylvania

The New-School Causal-Estimation Response

Fix key FIC violations as follows:

(1) Perform/approximate a randomized experiment
for consistent causal estimation

– An active strategy based on quasi-experimental data.

"Design-based econometrics"

Important advance in thinking.

(2) HAC methods for consistent estimation of s.e.'s

– A more passive strategy.

"Incomplete modeling":
"Linear regression of y on x_j
with a few controls and HAC s.e.'s"

Penn
University of Pennsylvania

MASSIVE PROBLEM:
The New-School Causal-Estimation Response Doesn't Work for Predictive Modeling

Crucially Important Information is Discarded

– Systematic approximation of the DGP
is crucially important
(i.e., selection of x from X)

– Heteroskedasticity and autocorrelation information
is crucially important
(autocorrelation for point prediction,
heteroskedasticity for interval and density prediction)

– Nonlinearity
may be is crucially important
(e.g. we may want to predict y far from the mean x)

Penn
University of Pennsylvania

What To Do?

Declare that econometrics is about causal estimation,
but not about predictive modeling,
so that predictive modeling is someone else's problem.

Unfortunately, the sentence above is not a gag.

It's the agenda of an influential part of the
new school causal estimation community.

See, e.g., Angrist and Pischke, "Undergraduate Econometrics Instruction: Through Our Classes, Darkly," NBER w.p. 23144.

Penn
University of Pennsylvania

Re-Writing the Textbooks

– Two modern streams: predictive and new-school causal

– Both beneficially de-emphasize the traditional FIC approach, the former using new PM tools and the latter using new CE tools

– The textbooks need re-writing to discuss both streams and their tools, NOT to erase predictive modeling.

Penn
University of Pennsylvania