

The Promise and Limits of Machine Learning and Big Data in Macrofinance

Discussion by Christopher Palmer
University of California, Berkeley

MFM Winter Meeting

January 29, 2016

Outline

1. Machine learning and big data comments
 - a) Prediction versus causation
 - b) Virtues of agnostic black-box modeling?
 - c) Big Data
 - d) Lucas Critique
2. Brief comments on Giesecke, Sirignano, Sadhwani

Why Machine Learning?

- Kleinberg, Ludwig, Mullinaithan, Obermeyer (2015 AER P&P) answer:
- Prediction vs causation
 - Umbrella (FICO) vs. rain dance (monetary policy)
 - Risk management vs. regulation?
 - Variance vs. bias
 - Can prediction inform causation?
- As academics, need a lot of help to care about predictions instead of “how the world works”
 - Not content with a kitchen sink regression, we want to understand mechanisms

Value of Models

- Kay's answer on value of machine learning:
 1. To handle big data
 2. To handle “true” model's unknown nonlinearities
- “All models are wrong, some models are useful.” –George Box
- “The art of modeling is what you leave out.” –Bengt Holmstrom
- Promise of machine learning is that we won't need our woefully simplistic models.
- But w/o models, we also don't really gain any understanding.
- If we're lucky, we'll gain enhanced predictive power.
- This is no substitute for modeling—it is just for a different purpose than prediction.

Lucas Critique

- Key worry, inspired by “The Failure of Models that Predict Failure,” Rajan, Seru, Vig (2015 JFE)
- When we are interested in prediction instead of causation, we always have to be concerned with the stability of that model.
- Teaching to the test \Leftrightarrow Lending to the test
- Given the data requirements of machine learning, hard to assess model drift over time.

Big Data: No Substitute for Identification

- Helps with: power, weak instruments, local effects, etc.
- **Not** selection bias, measurement error, endogeneity
- Worry can be thought of as external validity:
- Yes, can get significant coefficients with $R^2 < .0001$ if sample large enough...
- But if pattern only really applies to a very small subset of a very large dataset, have we really learned anything general?
- Estimated treatment effects start to look like very very local treatment effects.

“Deep Learning for Mortgage Risk”

Giesecke, Sirignano, and Sadhwani

- Goal: Improve risk management by leveraging machine learning and big data.
- Real goal: Demonstrate proof-of-concept handling breadth and depth of loan-level data for perfect applications yet unspecified.
- Research Question: Can machine learning techniques outperform logit predictions of defaults?
 - *NOT* does X cause default, will macroprudential foster robust loans, is a crisis coming, etc.
- Method: Compare predictions out of sample for both methods.

Why Machine Learning Here?

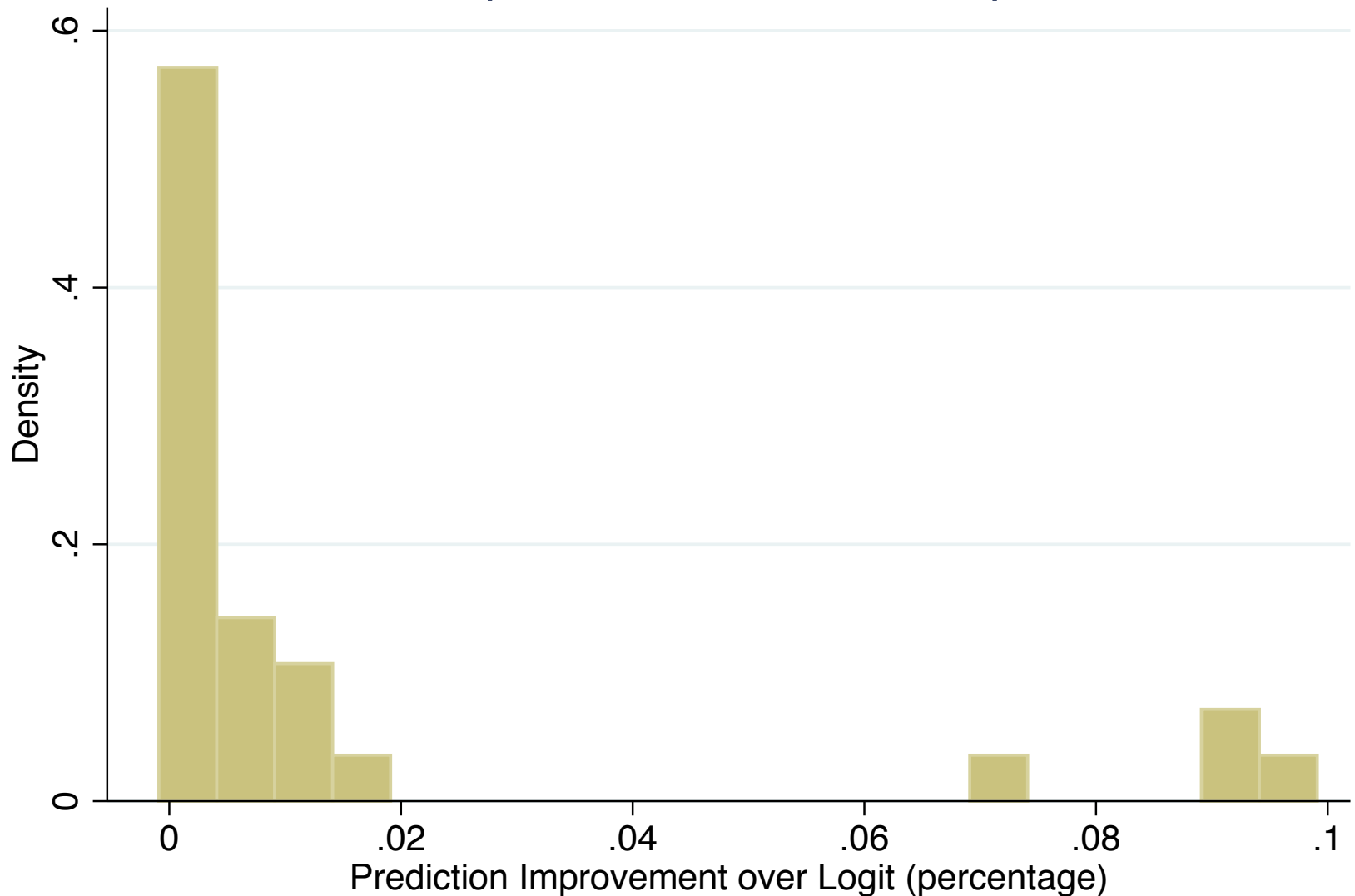
1. To handle big data. Why not a sample?
 - Once you look within zip codes, 120m nationwide becomes small.
 - Nonparametric methods are data hogs: “Curse of Dimensionality”
 - For rare events especially, need sufficient failures
2. To improve predictions via neural networks modeling

Prediction Improvements

- Even regularized multinomial logit can predict default state with high degree of accuracy (99%).
- Hardest: predict voluntary prepayment. (Logit 65%)
 - Fundamental problem for risk management, pricing of RMBS...
- Successfully predict voluntary prepayment 9 percentage points better (74% of time) than multinomial logit
 - All other performance states: 0-1 p.p. improvement
- Not done yet. Can do one-month predictions. What horizon is most important for risk management? Stress testing?
- Improvement in false positives or false negatives?

Value Added Predicting Prepayment

Out of Sample State Prediction Improvement



Conclusion

- We've seen mostly proof-of-concept papers in machine learning + macrofinance.
- Not their fault; should be eye opening. Nontrivial next step is to demonstrate the golden application that teaches us something new about the macroeconomy.
- To facilitate adoption of these methods, hold our hands to help us see what research questions can uniquely be answered with these methods, as opposed to their use in predictive analytics.
 - See Varian (2014) for some of this hand holding.