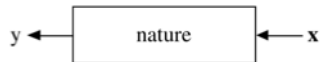
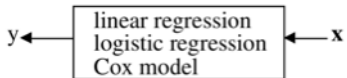


Class slides for Tuesday, September 22:
Armed conflict, part 1

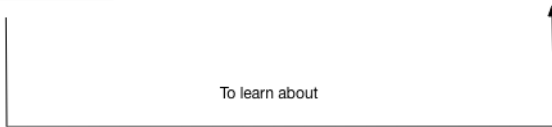
Matthew J. Salganik

COS 597E/SOC 555 Limits to prediction
Fall 2020, Princeton University

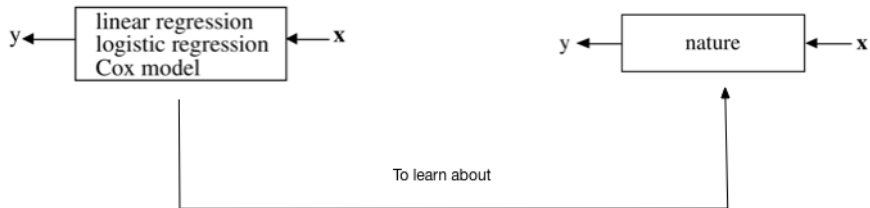
Social science (Data modeling)



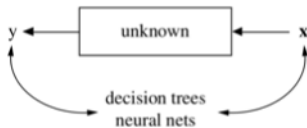
To learn about



Social science (Data modeling)

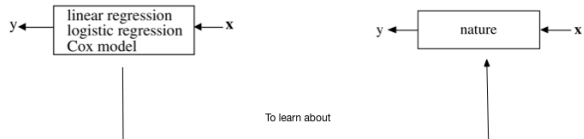


Computer science (Algorithmic modeling)



Compare \hat{y} and y to learn about algorithm

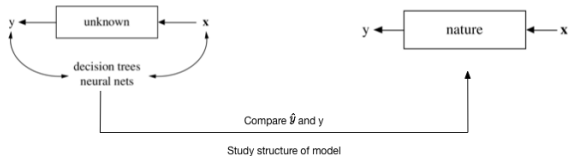
Social science (Data modeling)



Computer science (Algorithmic modeling)



Third way (Prediction for understanding)



Predicting armed conflict: Time to adjust our expectations?

Lars-Erik Cederman^{1*} and Nils B. Weidmann^{2*}

- ▶ Prediction can be at different timescales (daily, weekly, yearly) or different geographic scales (cities, states, countries). Some geographic and timescales may be more accurate than others. Some types of violence may also be more predictable than others (e.g., violence between nation-states vs violence by “terrorist” groups)

- ▶ Prediction can be at different timescales (daily, weekly, yearly) or different geographic scales (cities, states, countries). Some geographic and timescales may be more accurate than others. Some types of violence may also be more predictable than others (e.g., violence between nation-states vs violence by “terrorist” groups)
- ▶ Prediction of armed conflict will be difficult because:
 - ▶ complexity of the data generating process (this is asserted but not explained)
 - ▶ historical disjunctures (e.g., end of Soviet Union, countries merging and splitting)
 - ▶ data quality (both outcomes and predictors)
 - ▶ rare events

- ▶ Prediction can be at different timescales (daily, weekly, yearly) or different geographic scales (cities, states, countries). Some geographic and timescales may be more accurate than others. Some types of violence may also be more predictable than others (e.g., violence between nation-states vs violence by “terrorist” groups)
- ▶ Prediction of armed conflict will be difficult because:
 - ▶ complexity of the data generating process (this is asserted but not explained)
 - ▶ historical disjunctures (e.g., end of Soviet Union, countries merging and splitting)
 - ▶ data quality (both outcomes and predictors)
 - ▶ rare events
- ▶ As a thought experiment compare predicting armed conflict to predicting US Presidential election.

- ▶ Prediction can be at different timescales (daily, weekly, yearly) or different geographic scales (cities, states, countries). Some geographic and timescales may be more accurate than others. Some types of violence may also be more predictable than others (e.g., violence between nation-states vs violence by “terrorist” groups)
- ▶ Prediction of armed conflict will be difficult because:
 - ▶ complexity of the data generating process (this is asserted but not explained)
 - ▶ historical disjunctures (e.g., end of Soviet Union, countries merging and splitting)
 - ▶ data quality (both outcomes and predictors)
 - ▶ rare events
- ▶ As a thought experiment compare predicting armed conflict to predicting US Presidential election.
- ▶ Even if we had a perfectly accurate predictions, this might not be very useful for policy or theory.

Impossibility result in Gartzke & Coletta and Gartzke

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$
- ▶ Demand from actor A depends on the unseen (and unobservable) costs of for actor B.

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$
- ▶ Demand from actor A depends on the unseen (and unobservable) costs of for actor B.
- ▶ Sketch: If motivated participants can't do it, then researchers can't do it.

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$
- ▶ Demand from actor A depends on the unseen (and unobservable) costs of for actor B.
- ▶ Sketch: If motivated participants can't do it, then researchers can't do it.
- ▶ Limitations:
 - ▶ hard to put a number on the unpredictability

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$
- ▶ Demand from actor A depends on the unseen (and unobservable) costs of for actor B.
- ▶ Sketch: If motivated participants can't do it, then researchers can't do it.
- ▶ Limitations:
 - ▶ hard to put a number on the unpredictability
 - ▶ beware of tweaks to the set-up (e.g, repeated games, changes to the utility function)

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$
- ▶ Demand from actor A depends on the unseen (and unobservable) costs of for actor B.
- ▶ Sketch: If motivated participants can't do it, then researchers can't do it.
- ▶ Limitations:
 - ▶ hard to put a number on the unpredictability
 - ▶ beware of tweaks to the set-up (e.g, repeated games, changes to the utility function)
 - ▶ What about divorce?

- ▶ These kinds of limits to prediction arguments are very rare in the social sciences.
- ▶ $U_A(d) =$
 $Pr(\text{war})[\text{expected value from war}] + [1 - Pr(\text{war})][\text{value from accepted demand}]$
- ▶ Demand from actor A depends on the unseen (and unobservable) costs of for actor B.
- ▶ Sketch: If motivated participants can't do it, then researchers can't do it.
- ▶ Limitations:
 - ▶ hard to put a number on the unpredictability
 - ▶ beware of tweaks to the set-up (e.g, repeated games, changes to the utility function)
 - ▶ What about divorce? What about chess?

The perils of policy by p-value: Predicting civil conflicts

Michael D Ward

Department of Political Science, Duke University

Brian D Greenhill

Department of Political Science, University of Washington

Kristin M Bakke

Department of Political Science, University College London

Journal of Peace Research

47(4) 363–375

© The Author(s) 2010

Reprints and permission:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0022343309356491

jpr.sagepub.com



- Bring \hat{y} into $\hat{\beta}$ world

- ▶ Bring \hat{y} into $\hat{\beta}$ world
- ▶ social scientists often says that a variable X predicts Y if X is statistically significant in a regression model to prediction Y.

- ▶ Bring \hat{y} into $\hat{\beta}$ world
- ▶ social scientists often says that a variable X predicts Y if X is statistically significant in a regression model to prediction Y .
- ▶ assess model not by whether it produces new predictors that are statistically significant, but by whether it adds to predictive power.

- ▶ Bring \hat{y} into $\hat{\beta}$ world
- ▶ social scientists often says that a variable X predicts Y if X is statistically significant in a regression model to prediction Y .
- ▶ assess model not by whether it produces new predictors that are statistically significant, but by whether it adds to predictive power.
- ▶ They don't call for turning civil war prediction into a Kaggle contest.

Provocations:

- ▶ Maybe we can isolate unpredictability? peace \rightarrow crisis \rightarrow war. Gartzke argues we can predict peace \rightarrow crisis, but not crisis \rightarrow war. Maybe we could test that empirically?

Provocations:

- ▶ Maybe we can isolate unpredictability? peace \rightarrow crisis \rightarrow war. Gartzke argues we can predict peace \rightarrow crisis, but not crisis \rightarrow war. Maybe we could test that empirically?
- ▶ How can we limit ourselves to “useful” improvements in predictive performance?

Provocations:

- ▶ Maybe we can isolate unpredictability? peace \rightarrow crisis \rightarrow war. Gartzke argues we can predict peace \rightarrow crisis, but not crisis \rightarrow war. Maybe we could test that empirically?
- ▶ How can we limit ourselves to “useful” improvements in predictive performance?
- ▶ Can a model (e.g., Fearon & Laitin and Collier & Hoeffler) aid “understanding” even if it does not predict accurately?

Provocations:

- ▶ Maybe we can isolate unpredictability? peace \rightarrow crisis \rightarrow war. Gartzke argues we can predict peace \rightarrow crisis, but not crisis \rightarrow war. Maybe we could test that empirically?
- ▶ How can we limit ourselves to “useful” improvements in predictive performance?
- ▶ Can a model (e.g., Fearon & Laitin and Collier & Hoeffler) aid “understanding” even if it does not predict accurately?
- ▶ What should we do when we can’t just get more cases?

Class slides for Tuesday, September 22:
Armed conflict, part 1

Matthew J. Salganik

COS 597E/SOC 555 Limits to prediction
Fall 2020, Princeton University