

Capability Ratios Predict Nothing: Introducing A Better Measure of Expected Dispute Outcomes

Robert J. Carroll
University of Notre Dame

Brenton Kenkel
Vanderbilt University

Motivation

In bargaining models, states' expectations about who would win a potential dispute play a major role in international negotiations.

Existing empirical work on conflict typically employs ratios of the Composite Index of National Capabilities to proxy for these expectations.

Is this measure any good at actually predicting outcomes? If not, how can we do better?

Data

Sample: Militarized international disputes, $N = 1,732$

Predictors: Capability index components, with multiple imputation for missing values

- ▶ Iron and steel production
- ▶ Military expenditures
- ▶ Military personnel
- ▶ Primary energy consumption
- ▶ Total population
- ▶ Urban population

Response: Outcome — ordered

- ▶ Side A Wins/Side B Yields
- ▶ Stalemate
- ▶ Side B Wins/Side A Yields

The Traditional Measure

$$\text{CINC}_i = \frac{1}{6} \sum_{c=1}^6 \frac{i\text{'s level of component } c}{\text{total world level of component } c}$$

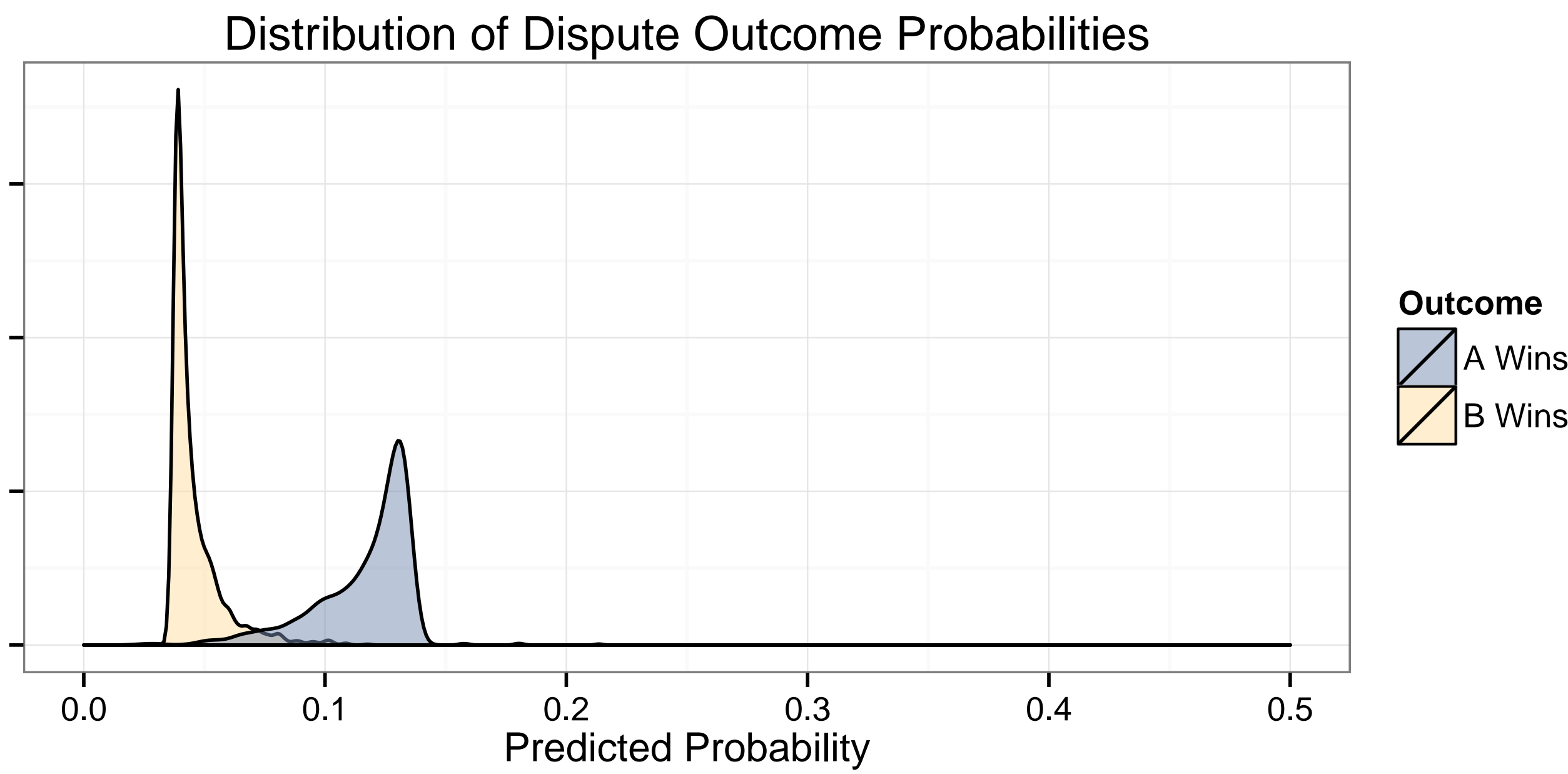
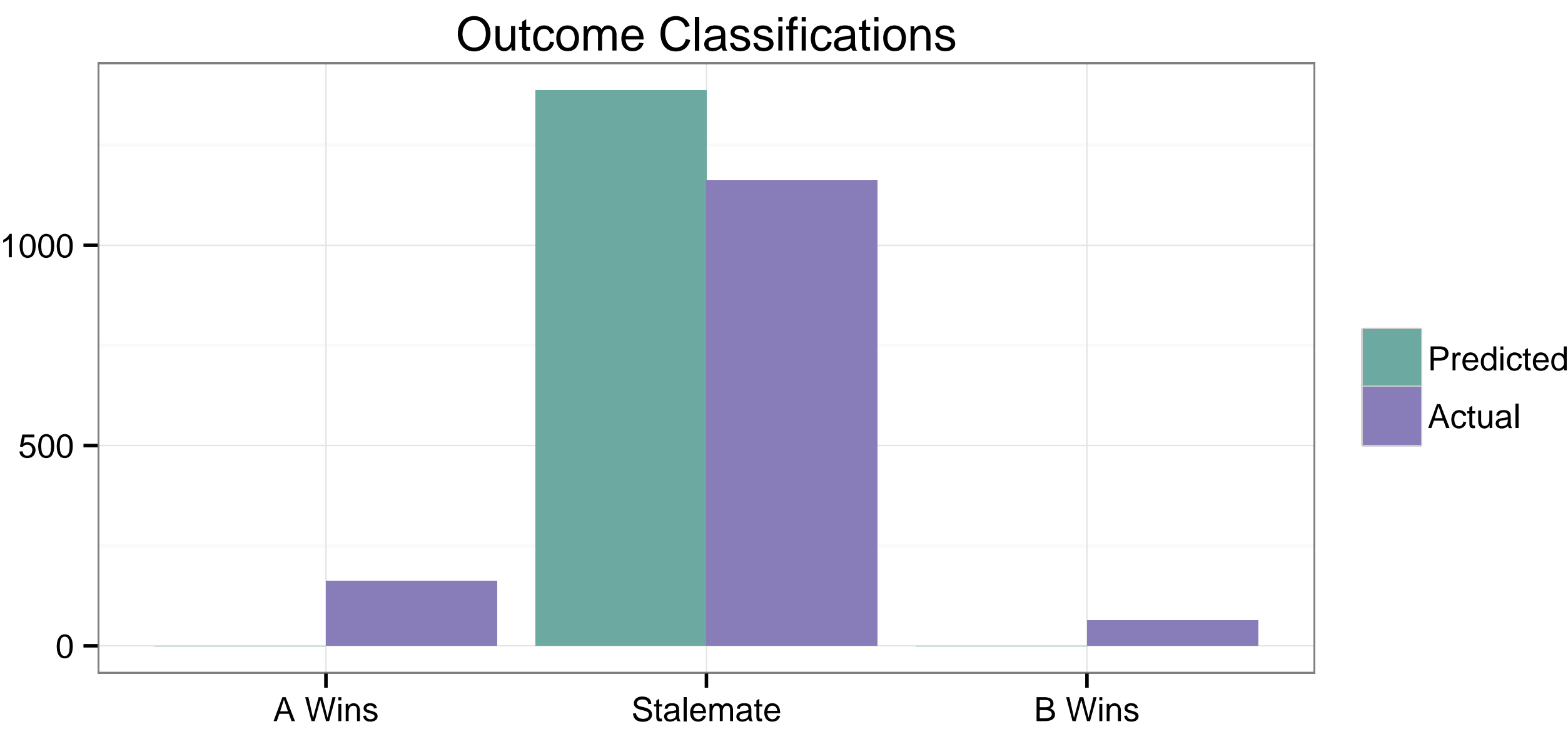
$$\text{Capability Ratio}_{i,j} = \frac{\text{CINC}_i}{\text{CINC}_i + \text{CINC}_j}$$

We want to control, as best as possible, for expected dispute outcomes. Does the Capability Ratio do that?

- ▶ Originally designed to measure power concentration at the system level (Singer et al 1972)
- ▶ Components not weighted
- ▶ Inflexible over time

Reasons for Skepticism about Capability Ratio

Results from ordered logit of Outcome on $\log(\text{Capability Ratio})$:

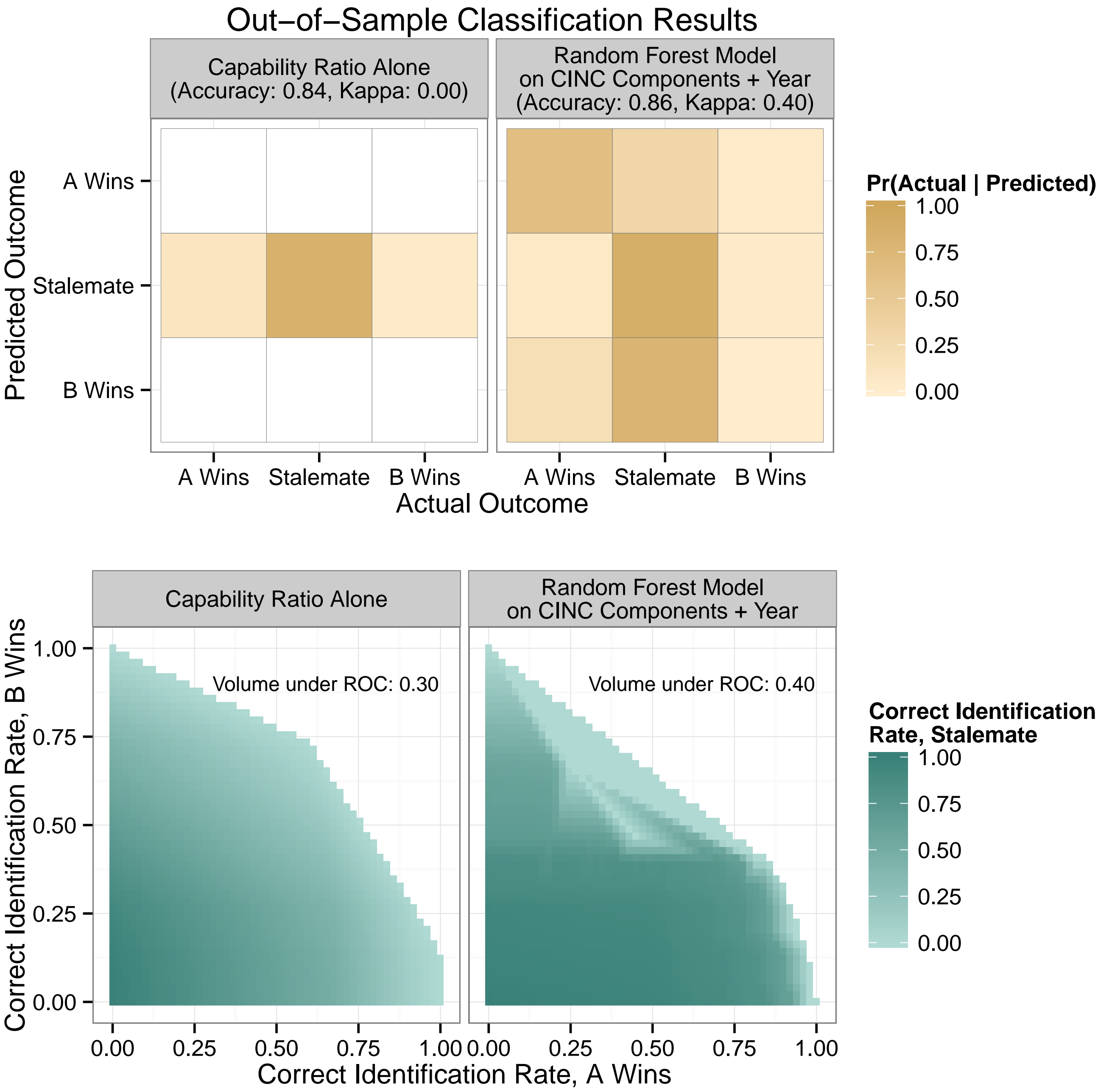


Finding an Alternative

Can we predict dispute outcomes better without collecting new data?

- Split data on CINC components and dispute outcomes into training set (80%) and test set (20%)
- Within the training set:
 - ▶ Use k -fold cross-validation to tune various models and estimate their ability to predict dispute outcomes out of sample
 - ▶ Random forests (Breiman 2001)
 - ▶ Support vector machines (Cortes and Vapnik 1995)
 - ▶ Always include all six components for each side, with or without year of dispute start
 - ▶ Choose the model with the best cross-validation performance, and estimate its parameters by fitting to the full training set
- Apply the fitted model to the test set to obtain an unbiased estimate of its out-of-sample error

Results



Conclusions

The Capability Ratio was not designed to predict dispute outcomes well, and indeed it does not. We construct a better measure using the same component data.

We argue that proxy variables should be well-tuned for the use at hand and validated using out-of-sample predictive performance — a lesson that applies elsewhere in IR and political science.

References

- Leo Breiman. 2001. "Random Forests." *Machine Learning* 45:5–32.
- Corinna Cortes and Vladimir Vapnik. 1995. "Support-Vector Networks." *Machine Learning* 20:273–297.
- Douglas Mossman. 1999. "Three-Way ROCs." *Medical Decision Making* 19:78–89.
- J. David Singer, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820–1965." In Bruce Russett (ed.), *Peace, War, and Numbers*, Beverly Hills: Sage, 19–48.
- R Packages: `Amelia`, `caret`, `ggplot2`, `kernlab`, `MASS`, `randomForest`.