What to Do with Unclassifiable Outcomes in International Relations Data

Brenton Kenkel
Department of Political Science, University of Rochester

Motivation

IR datasets often have observations where the dependent variable cannot be coded easily:

- Continued disputes and joiners in COW
- Measures disagree (MID initiation: Side A or Revisionist?)
- Values just above/below arbitrary thresholds

From a statistical standpoint, these constitute nonrandom missing data or measurement error, so deleting them from the dataset will cause bias.

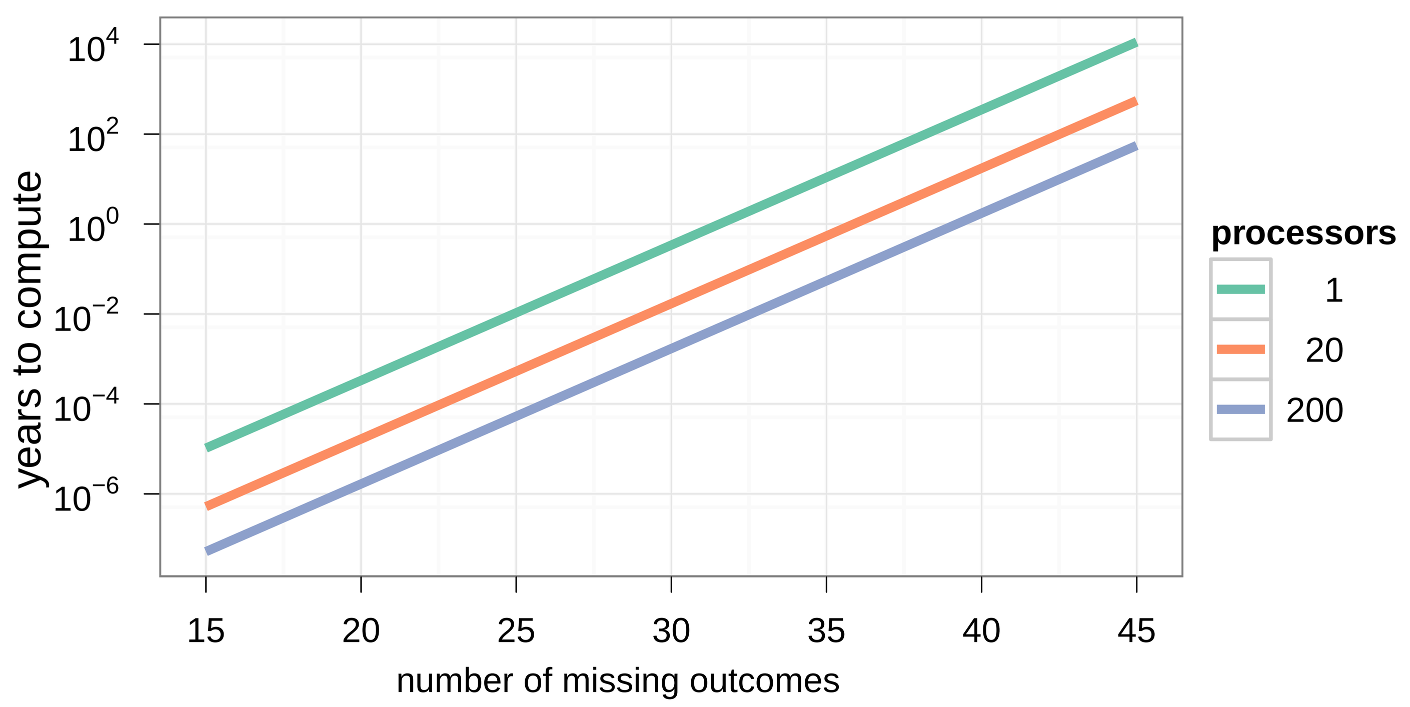
I develop a new method to check the robustness of logistic regression results when some outcomes are unclassifiable or unobserved, without introducing bias or making implausible assumptions.

Methodology

Ideal Procedure: Run logistic regression on all possibilities, check for no substantive difference in results.

observed	controls			possibilities				controls	
Y	X_1	X_2		$Y^{(1)}$	$Y^{(2)}$	$Y^{(3)}$	$Y^{(4)}$	X_1	X_2
?	8	5	fill in the missing values $\Longrightarrow$	*0*	*0*	*1*	*1*	8	5
?	2	4		*0*	*1*	*0*	*1*	2	4
0	7	2		0	0	0	0	7	2
1	1	10		1	1	1	1	1	10
0	3	5		0	0	0	0	3	5
1	2	3		1	1	1	1	2	3

The Problem: Computationally infeasible.



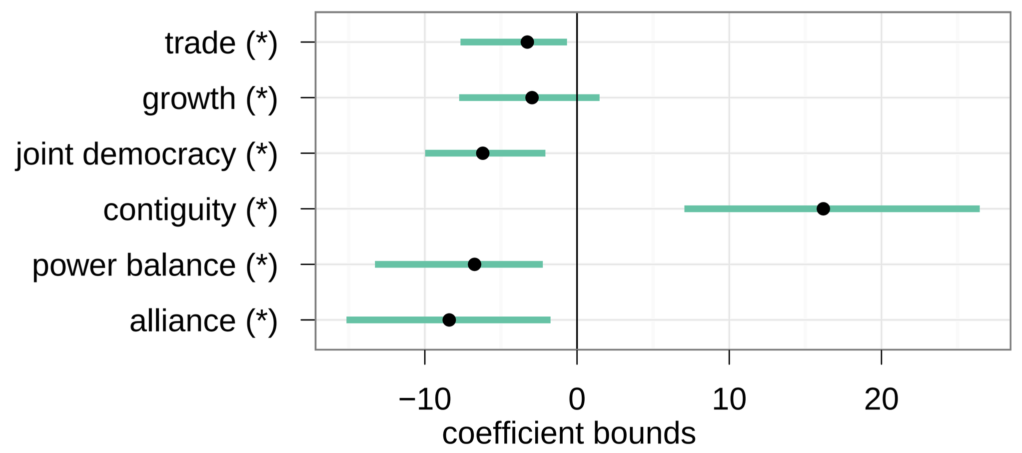
My Approach:

- Treat unclassifiable outcomes as nonignorable missing data
- Use moment conditions to approximate the results of the "ideal" procedure much more quickly (similar to Manski & Tamer 2002)

Application 1: Oneal and Russett (1997) on Liberal Peace

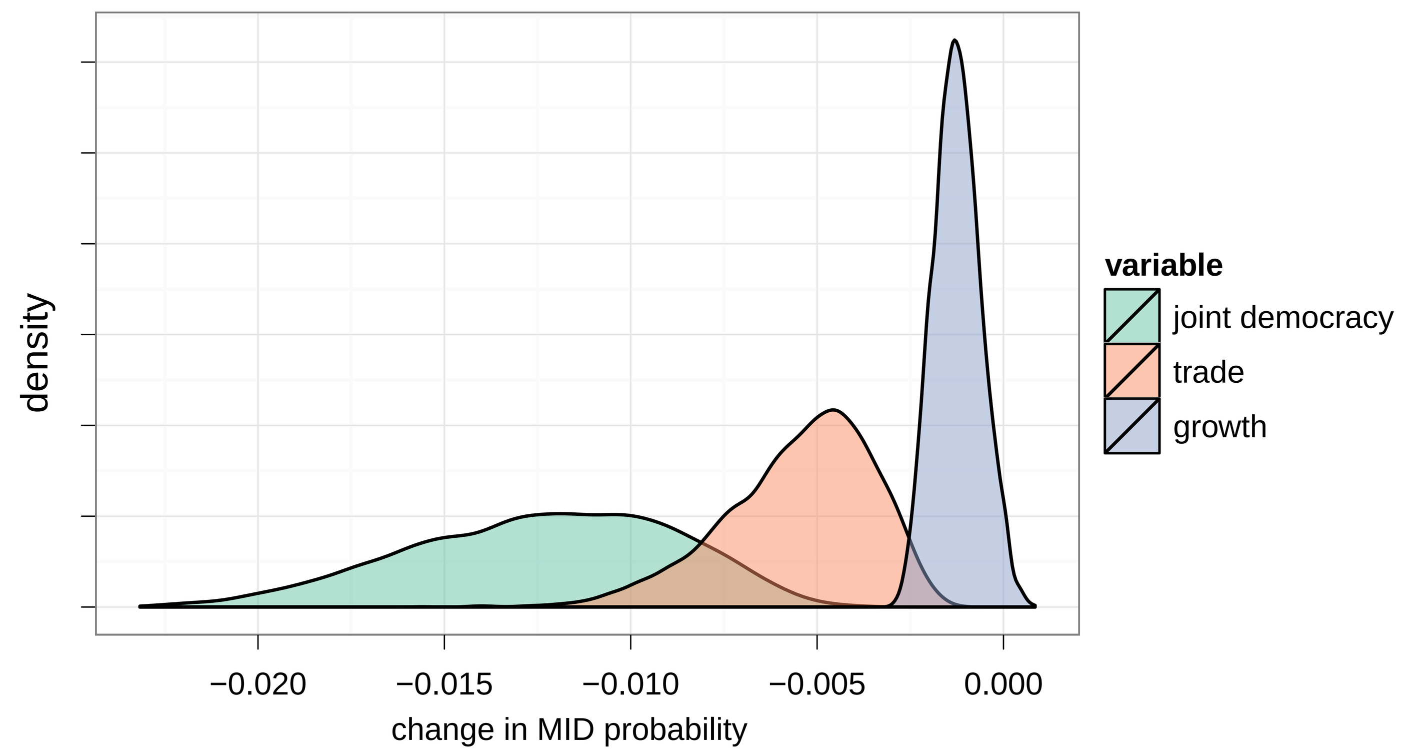
Sample: Politically relevant dyads ($N = 20{,}990$)
Outcome: Militarized dispute onset (405 cases)
Unclassifiable values: Continued disputes (542 cases)

Main Results



Lines: Estimated bounds
Points: Original coefficients
Starred: Statistically significant in original analysis

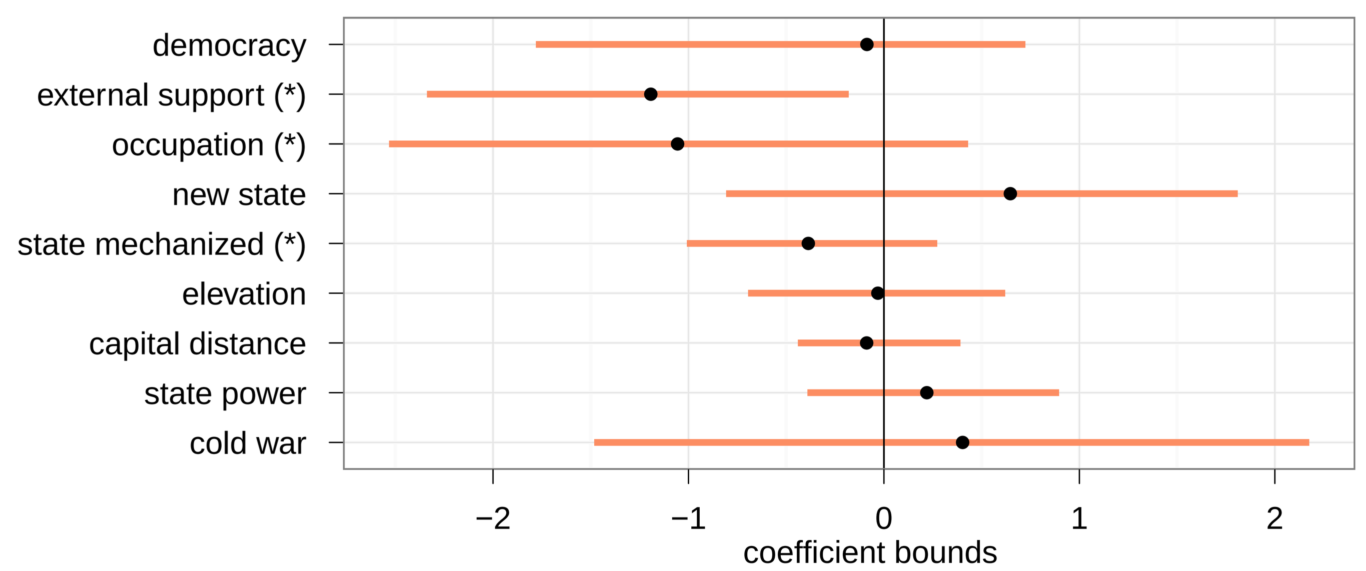
Substantive Effects



Application 2: Lyall (2010) on Counterinsurgencies

Sample: Counterinsurgency efforts ($N = 286$)
Outcome: Victory against insurgency (153 cases)
Unclassifiable values: "Difficult to code" draw (39 cases)

Main Results



Interpreting the Results

The bounds contain the set of coefficients that could be obtained as a *point estimate* under some assumption about the missing values—no population inference is implied.

If the bounds contain 0, the sample definitely does not provide evidence in favor of a directional hypothesis. However, failure to contain 0 only establishes that the sign of the sample estimate is robust to missingness/measurement error.

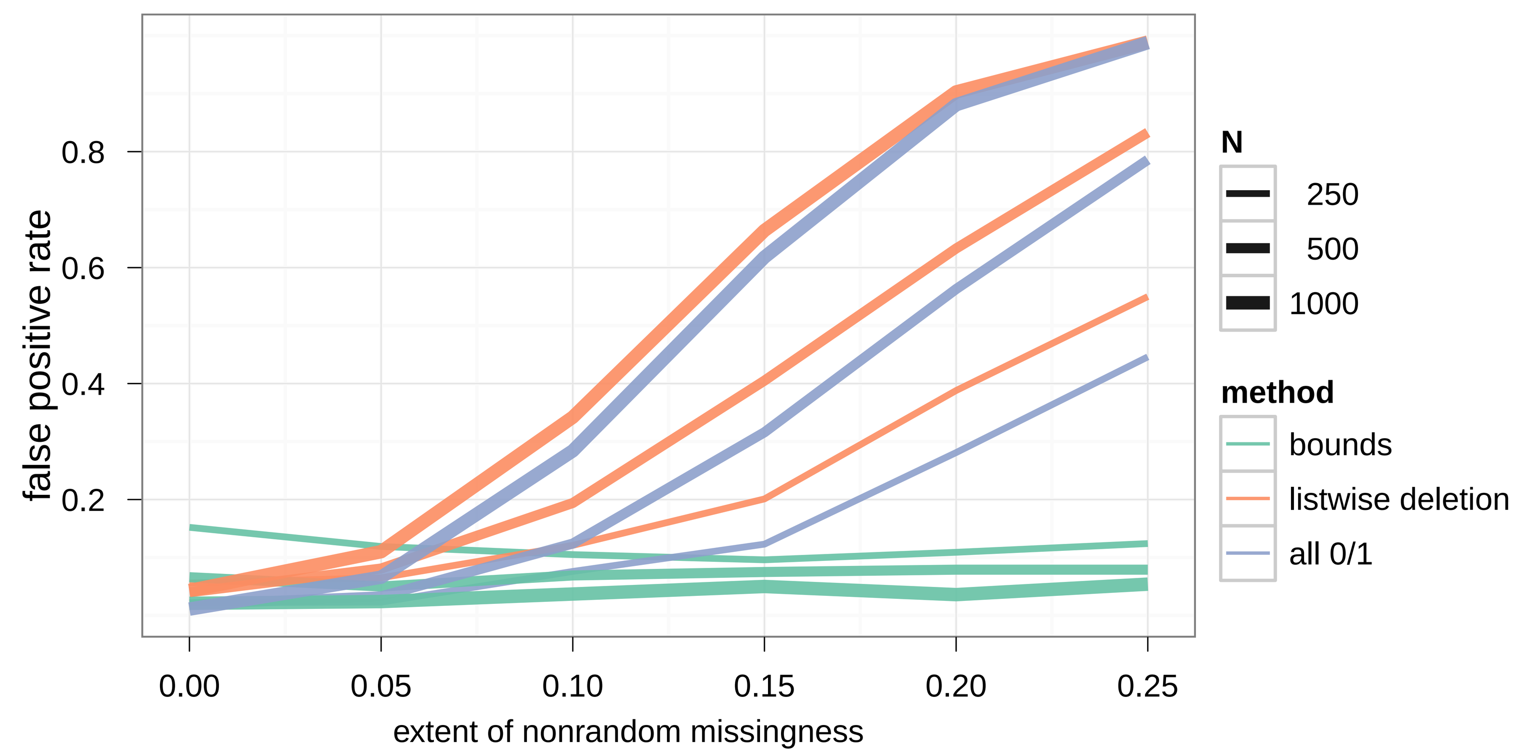
Monte Carlo Simulation

Comparison of the new method to common robustness checks in IR:

- Listwise deletion of unclassifiable outcomes
- Setting all to 0 or all to 1

Suppose the value of Y does not depend on X, but the probability of missingness depends on Y and X (hence is nonrandom). How well does each method perform at avoiding false positives?

Results



As N increases, the bounding method performs better (fewer false positives), while the other two get worse.

References

Jason Lyall. 2010. "Do Democracies Make Inferior Counterinsurgents? Reassessing Democracy's Impact on War Outcomes and Duration." *International Organization*, 64(1): 167–192.

Charles F. Manski and Elie Tamer. 2002. "Inference on Regressions with Interval Data on a Regressor or Outcome." *Econometrica* 70(2): 519–546.

John R. Oneal and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985." *International Studies Quarterly* 41(2): 267–293.