Trends in Social Science Methods Relevant to Intelligence Analysis

National Academy of Sciences
Decadal Survey of Social and Behavioral Sciences for Applications to National Security
Emerging Trends and Methods in International Security
October 11, 2017

Dr. Andrew Bennett
Professor of Government
Georgetown University

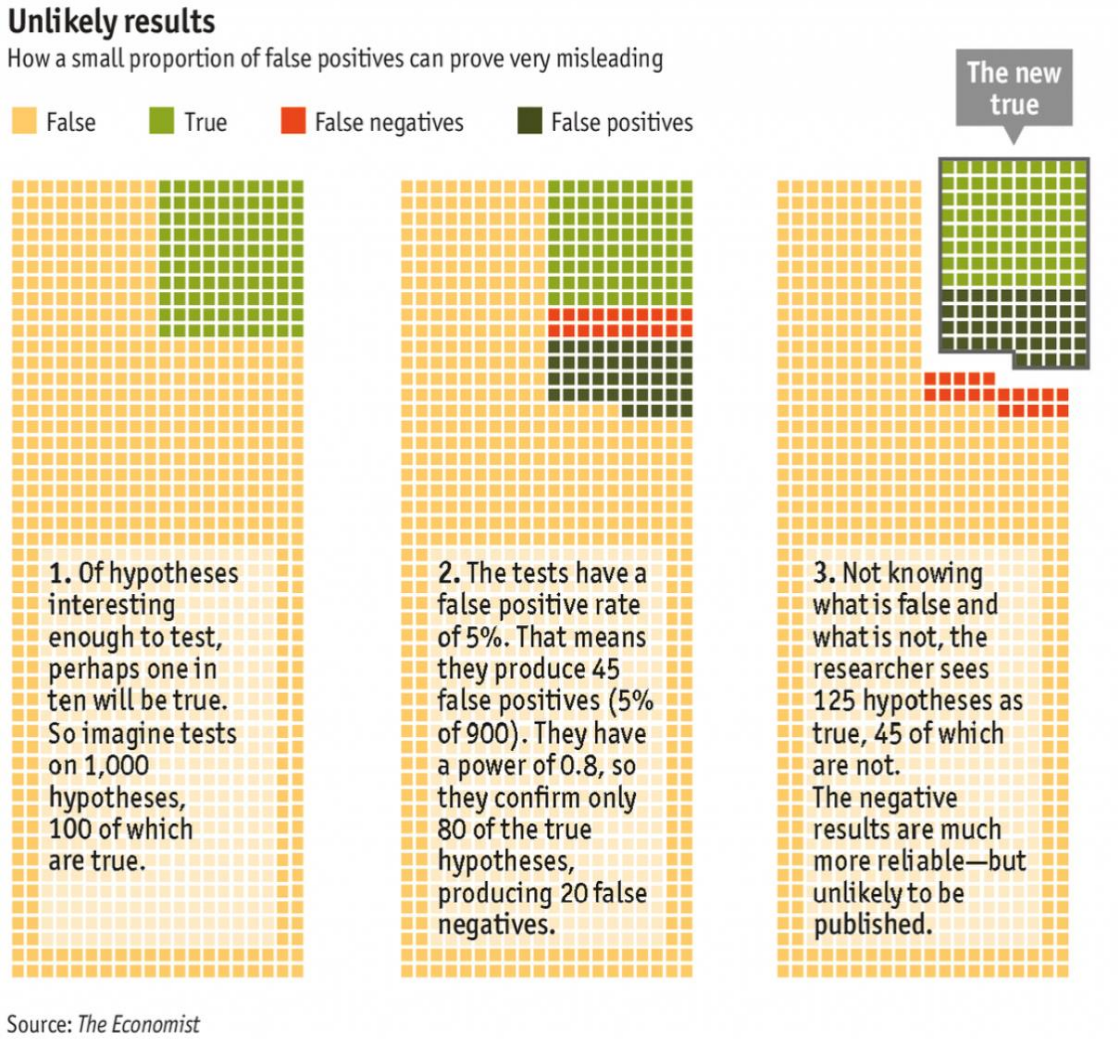
With thanks to Philip Schrodt, Jason Seawright, and Jay Ulfelder for suggestions.

GEORGETOWN UNIVERSITY

The Replication Crisis in the Sciences

- Many sciences are experiencing a “replication crisis,” due to the finding that studies in top journals often cannot be replicated.
- This is likely due to publication biases, p-hacking, and other poor practices.
- Journals are starting to use pre-registration and other steps to address these problems, but deeper changes in practices are needed and are underway.

Publication Bias and the Replication Crisis: An Illustration



GEORGETOWN UNIVERSITY

From Replication Crisis to Methodological Renaissance

- There is a countervailing trend to the replication crisis:

--big data is increasingly available in real time

--methods are rapidly improving in machine learning, computer-assisted content analysis, agent-based modeling, natural experiments, group-based "superforecasting," case studies, and multimethod research.

- Behind the headlines of the replication crisis a revolution in social science methods is taking place.

Example: Evolving Methods of the Political Instability Task Force (PITF)

- Some parts of the IC are already taking advantage of new methods, as is evident in the PITF.
- The PITF started out two decades ago as a very inductive forecasting exercise with a typical p-value hypothesis testing approach.
- A key predictor of state crises was infant mortality; this remains in many current models (as a predictor, not a causal variable).
- PITF is now using machine learning for prediction, plus out-of-sample testing, together with more real-time data. This uses a mix of causal and predictive variables.
- This encompasses three developments: big data availability, computer-assisted content analysis, and machine learning/artificial intelligence.

Big Data Availability

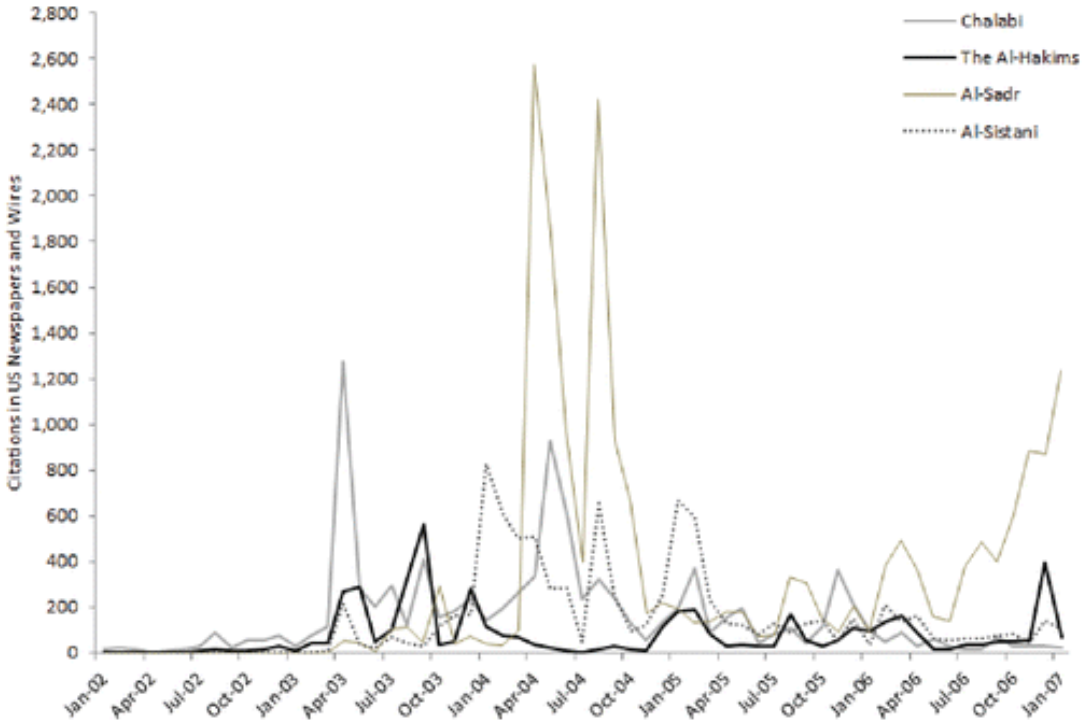
- Big data is increasingly available in real time on everything from social media, to events data, to GIS data.
- Examples include satellite photos of burning villages in Myanmar, detailed electoral data, and twitter trends.
- PITF's first efforts to use more real-time and geographically granular data have not yet greatly improved over earlier models, but first generation big data analyses are already as good as models refined over two decades.

Computer-Assisted Content Analysis (CCA)

- Texts are a key part of big data and CCA methods are becoming more sophisticated.
- Most classifiers are still basic word counters but efforts to incorporate syntax are improving.
- This includes fast, scalable natural language processing of audio and visual material as well as digital texts.
- Many improvements are in the private sector rather than academia; using them may require changing IC restrictions on using open source data and software.
- Even basic word counts can be revealing; trends over time indicate salience, similarities in texts reveal networks and coalitions, undirected searches can identify unexpected categories of texts or ideas.

CCA Example: Media (Mis)reads of Which Iraqi Leaders were Key

Figure 1. Leaders, January 2002–January 2007.



GEORGETOWN UNIVERSITY

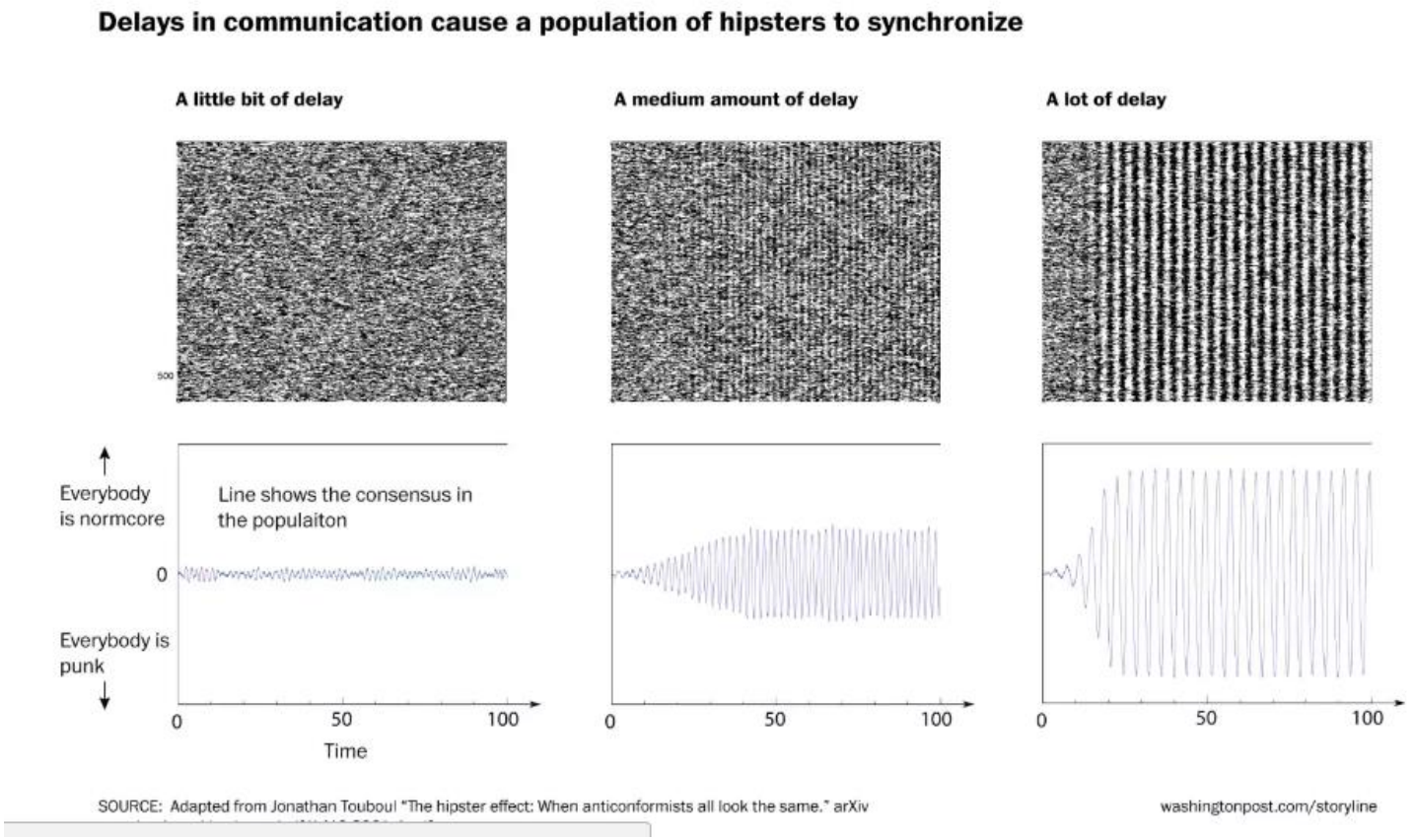
Machine Learning/Artificial Intelligence

- Machine learning is already improving on the predictive power of logistic regression.
- One approach is “random forests (RF),” which randomly samplse a set of features, randomly samples from the training data, builds decision trees for each point in these samples, and averages the outcomes.
- Limitations: RF does not work well with a few data points, and it works only for prediction but not for explanation/interpretation. Also, values on the outcome variable are restricted to those in the training set.

Another Method: Agent Based Modeling (ABM)

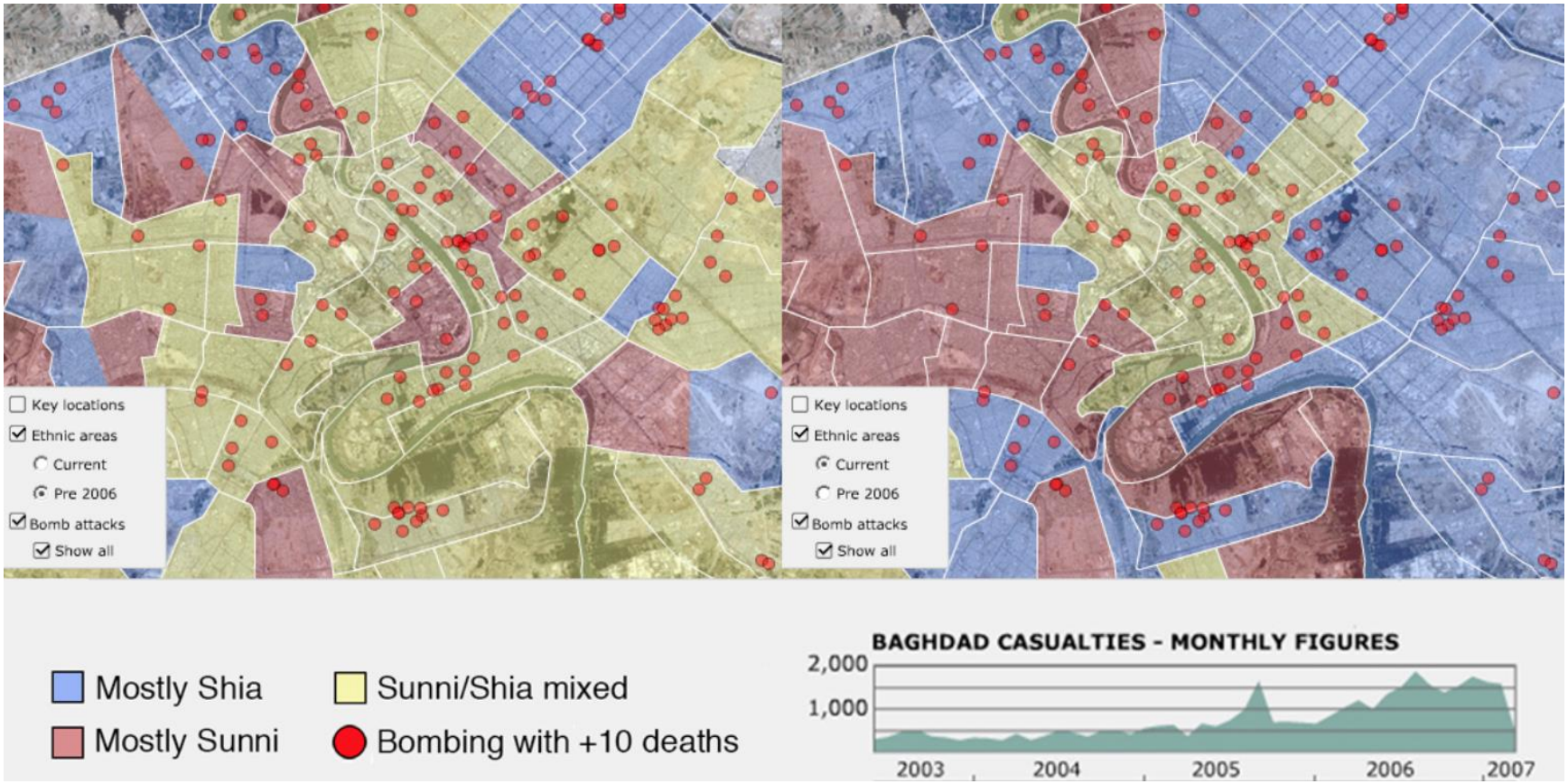
- ABM simulates actors in a system to study complex system-level dynamics.
- The computer gives each simulated actor a set of decision rules and constraints, including those on interactions with other actors and geographic features.
- The simulation tracks individual actors and population outcomes from the starting configuration forward.
- This allows modeling heterogeneous agents, geographic features, adaptation, learning, and path dependency.
- ABMs have been used to study infectious diseases, land use, ecosystems and resource management, electoral design, and ethnic and civil conflict.
- Recent work uses more complex rules and compares modeling results to empirical events.

A Trivial but Illuminating ABM Example: Why All Hipsters Dress Alike



GEORGETOWN UNIVERSITY

A Non-trivial example: ABM could be used to Model Refugee Flows from Ethnic Cleansing



GEORGETOWN UNIVERSITY

Experiments and Natural Experiments

- Political scientists have renewed their interest in experiments, especially on how individuals respond to media messages and other texts.
- These have uncovered (depressing) findings on such things as the persistence of impressions from fake news.
- Natural experiments have become popular as well, although it is hard to satisfy the “as if random” assumption.
- The situations in which experiments are ethically and fiscally do-able are limited, even outside of the IC and especially within it. The IC might draw on results of others’ experiments, and can do natural experiments, but should exercise great caution about doing any experiments of its own.

The Good Judgment Project(GJP) and Superforecasting

- Since 2011 the GJP, sponsored by IARPA and led by Philip Tetlock, Barbara Mellors, and Don Moore, has been doing forecasting tournaments
- Results show that with training and group information sharing, the best forecasters are better than computer algorithms or analysts with access to classified information
- These superforecastser achieve 80-85% accuracy, vs. 50-60% accuracy for typical forecasters
- Currently GJP is holding a hybrid forecasting competition, combining human analysis and computer algorithms
- Additional research could focus on how best to train people to become better forecasters, how to help the best forecasters become even better, and whether more explicit Bayesian or other forms of analysis can improve forecasts.

New Case Study Methods: Bayesian Process Tracing

- Case study methodologists have developed techniques of process tracing, or within-case causal inference, that are more fully and explicitly Bayesian
- This involves more explicit designation of the priors on alternative explanations and the likelihood ratios for pieces of evidence
- This kind of analysis can also be applied to do scenario building and forecasting, as it is in the GJP
- Explicit Bayesian analysis might have improved the 2003 assessment of Iraqi WMD
- More research is needed on whether explicit Bayesian process tracing improves explanations and predictions

New Case Study Methods: Typological Theorizing

- Typological theorizing categorizes cases into theoretical types, or combinations of variables
- This can address high-order interaction effects
- Typological theories can aid in case selection for process tracing and theory development
- This is in some ways analogous to matching approaches in statistical research, except that it uses coarsened exact matching
- Like matching, typological theorizing can help identify good and poor analogies for current policy cases
- This approach has been applied to alliance burden sharing, revolutions, and transnational actors in civil conflicts

Multimethod Research

- Researchers are increasingly combining quantitative and qualitative inferences in the same research project
- This allows analysis of both patterns in populations and mechanisms in individual cases
- Quantitative analysis can identify cases with unexpected outcomes, and process tracing on those cases can reveal omitted variables
- This is especially important for improving forecasts of rare events, as such forecasts usually have a high rate of false positives. Studying the false positives can reveal clues that differentiate the true positives.

Conclusion: How can the IC make use of these methods?

- The IC is already using many of these methods, and creating new ones as in the GJP
- There is still room for the IC to improve, especially in combining methods
- A best-practices package for intelligence analysis would use a multi-method combination of:

--machine learning

--group forecasting

--automated data collection

--case studies of poorly-predicted cases

- The key goal of the case studies would be to identify holes in the data collection, since machine learning can only optimally combine the kinds of data it knows about.

GEORGETOWN UNIVERSITY

Further Reading: The Replication Crisis

- "Why Most Published Research Findings Are False", John P. A. Ioannidis, *Public Library of Science Medicine*, 30 August 2005.
- "The Replication Crisis in Psychology" by Edward Diener and Robert Biswas-Diener, NOBA, 2016.
- "How science goes wrong: Scientific research has changed the world. Now it needs to change itself", *The Economist*, 19 October 2013.
- "1,500 scientists lift the lid on reproducibility" by Monya Baker, *Nature*, 25 May 2016 — "Survey sheds light on the 'crisis' rocking research."
- "Replication initiatives will not salvage the trustworthiness of psychology" by James C. Coyne at BioMed Central (peer-reviewed, open access), 31 May 2016.
- "Is Most Published Research Really False?", Jeffrey T. Leek and Leah R. Jager, *Annual Reviews*, March 2017.

Big Data Analysis

- **http://journals.sagepub.com/doi/abs/10.1177/0022002714553104 Van der Windt, Peter and Humphreys, Macartan. 2014. Crowdseeding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time. Journal of Conflict Resolution**

- Patrick Brandt, John Freeman, and Philip Schrodt. "Real Time, Time Series Forecasting of Inter- and Intra-State Political Conflict", *Conflict Management and Peace Science* 28:1 (2011)

Computer-assisted Content Analysis

- David Muchlinsky, David Siroky, Jingrui He, and Matthew Kocher, Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data, *Political Analysis* 24:1 (Winter 2016).
- **http://www.aclweb.org/anthology/D17-1194 uses automated text analysis to trace conflicts.**
- **http://www.aclweb.org/anthology/W17-2705: uses automated text analysis to trace conflicts. -- http://aclweb.org/anthology/W17-2700 -- conference with several relevant uses of full text analysis**

Bayesian Process Tracing and Typological Theorizing

- Alexander L. George and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences*. MIT Press 2005.

- Andrew Bennett and Jeffrey Checkel, eds., *Process Tracing: From Metaphor to Analytical Tool*. Cambridge, 2014.

- Andrew Bennett, "Using Process Tracing to Improve Policy Making: The (Negative) Case of the 2003 Intervention in Iraq," *Security Studies 22:4 (2015).*

- Andrew Charman and Tasha Fairfield, "Explicit Bayesian Analysis for Process Tracing: Guidelines, Opportunities, and Caveats," *Political Analysis* 25:3 (2017).

GEORGETOWN UNIVERSITY

Machine Learning/AI

- **http://journals.sagepub.com/doi/full/10.1177/0022343316684009 (predicting local violence in Liberia)**
- **https://asu.pure.elsevier.com/en/publications/comparing-random-forest-with-logistic-regression-for-predicting-c (predicting civil war onsets using random forests, which are a common machine learning method)**
- **http://www.springer.com/us/book/9781461453109 Handbook on Computational Approaches to Counterterrorism, 2013**
- **Robert Trappl. 2006. Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention. Dordrecht, Netherlands: Kluwer Academic Publishers.**
- **Michael Alvarez, ed. *Computational Social Science: Discovery and Prediction* (Cambridge: Cambridge University Press, 2016)**
- http://globalpolicy.gmu.edu/a-global-model-for-forecasting-political-instability/
- **http://journals.sagepub.com/toc/jpra/54/2 (entire issue of J of Peace Research on forecasting; March-2017: lots of great examples in there)**

Agent Based Modeling

- Paul K. Davis and Angela O'Mahoney, "A Computational Model of Public Support for Insurgency and Terrorism A Prototype for More-General Social-Science Modeling," RAND report 2013.
- Carlos Lemos, Helder Coelho, and Rui J. Lopes, "Agent-based modeling of social conflict, civil violence and revolution: state-of-the-art-review and further prospects." Available at: http://ceur-ws.org/Vol-1113/paper10.pdf
- Ross Hammond, "Appendix A, Considerations and Best Practices in Agent-Based Modeling to Inform Policy," in *Assessing the Use of Agent-Based Models for Tobacco Regulation*, Washington: National Academies Press, 2015. Available at: https://www.ncbi.nlm.nih.gov/books/NBK305917/

Natural Experiments, Group Forecasting, Multimethod Research

- Thad Dunning, *Natural Experiments in the Social Sciences: A Design Based Approach*. Cambridge, 2012
- Philip Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction*, Broadway Books, 2016.
- Jason Seawright, *Multimethod Sociel Science: Combining Qualitative and Quantitative Tools,* Cambridge 2016.

GEORGETOWN UNIVERSITY