

# Bayesian Methods in Political Science: Introduction to the Virtual Issue

Jeff Gill

Washington University, St. Louis

jgill@wustl.edu

## 1 Introduction and Motivation

In episode 96 (season 5, number 15, 1994), of *The Simpsons*, Homer calls NASA to complain about their “boring space launches.” The crew for the impending mission is described as: “a mathematician, a different *kind* of mathematician, and a statistician.” While most of the world is unaware that there are two principle types of *statisticians*, I would guess that the bulk of our colleagues *do* know about this distinction. While many political scientists understand that there exists a contrast in empirical work between Bayesians and Frequentists, this is actually not correct. There are almost no Frequentists in political science because the core tenet of Frequentism is the availability of an infinite stream of independent, identically distributed data that the researcher can draw from. So in this unrealistic setting there really are 19 more identical experiments from which to consider a confidence interval, and p-values do eventually become probabilities. It turns out that this is completely inappropriate for a discipline that relies almost exclusively on situational observational data that will never, ever, be collected again. This, by the way, is why we are a *hard* science and physics, chemistry, and engineering are *soft* sciences. Actually, the appropriate statistical contrast for us is between Bayesians and Likelihoodists.

Fisher created (or discovered, depending on your view of epistemology) maximum likelihood estimation in the 1920s (1922, 1925a, 1925b) to find the fixed  $\theta$  “most likely” to have generated a single set of data  $\mathbf{X}$  (Stigler 1986). Furthermore, he considered the null hypothesis as merely something to be nullified when the evidence for an hypothesized effect is substantial (Gill 1999). In fact, Fisher loathed the mechanistic Frequentist approach of Neyman and Pearson (1928a, 1928b, 1933a, 1933b, 1936a, 1936b), in which one hypothesis was rejected and the other one was accepted(!). Of course it is drilled-into us in graduate studies in political science that we never accept an hypothesis because there are an infinite number of alternatives that were not tested. And rightly so. Except that we kind-of do accept the alternative hypothesis when you carefully read text in the paragraphs following a regression table (my least-favorite part of any article).

So the real contrast in empirical political science is between Bayesian practitioners and Likelihood practitioners. Or is it? Both approaches create a likelihood function from the joint distribution of the observed data. The two approaches are asymptotically equivalent: the data subsumes any reasonable prior in the limit for a Bayesian model. Actually, a likelihood model is equivalent to a Bayesian model with the appropriately chosen uniform prior. So wait, doesn't that make Likelihoodism a special case of Bayesianism? The answer is yes. All of us are Bayesian, some of us are aware of it. This is even more true when you consider that most Bayesians in political science use flat priors on all of their model parameters. This leads to the question of why would we care about the difference? There are two principle reasons to prefer to do Bayesian work in the discipline, and neither one of them are philosophical or need to draw from the acrimonious history of Bayesian versus Frequentist statistics.

First, in Bayesian inference all unknown quantities are treated probabilistically. This includes: the right form of the model specification, the true parameter values, and the missing data. It also means that the results are treated probabilistically. So I can say, for instance, that there is a 94% probability that some explanatory variable has a positive

effect on the outcome variable (holding other model quantities constant), if 0.94 of the density of the corresponding coefficient posterior distribution is to the right of zero. Substantively, this is normally considered a strong finding and I, for one, would be willing to bet money that there is a positive effect, conditional on trusting the whole model enterprise of course. Note that this would fail reach typical significance levels and would be unworthy of “stars.” The point is that it is not only more convenient to discuss results in probabilistic terms, and avoid dancing around “confidence” or overly-arbitrary testing, it is also more intuitive to readers since humans like to think probabilistically (Gigerenzer and Murray 1987), even if we are not very good at it (Tversky and Kahneman 1974, 1981). Thus there is great value in keeping all uncertainty on the probability scale and discussing results in this fashion.

Second, the post-1990 Bayesian estimation engine is the most powerful vehicle for obtaining model results available in statistics. Markov chain Monte Carlo (MCMC) was introduced into the general statistical literature by Gelfand and Smith in a 1990 review article that appeared in the *Journal of the American Statistical Association* after lurking undetected in statistical physics and image-restoration for decades. Bayesian stochastic simulation replaces analytical solutions and numerical mode-finding with a computational process that describes multidimensional posterior distributions, which may be impossible to integrate, by exploring them using a Markov chain. Since each step of the chain is a multidimensional position, marginalizing the joint posterior is simply equivalent to looking at the history of each dimension individually. Marginalizing is what we want since a row of the regression table is just a marginal summary of a particular coefficient estimate. Of course I am glossing-over a whole host of challenges (Gill 2008), and much work has been dedicated since 1990 towards making this process work more efficiently across a wide range of models and data types. The important point is that MCMC, either Gibbs sampling or the Metropolis-Hastings algorithm, is more powerful than maximum likelihood, we just do not know it yet. After all, it took about 40 years from Fishers important papers on MLE until the full set of properties were revealed by Birnbaum (1962). Yet the reasons are clear why MCMC is more powerful: it gives the same information as MLE (the mode and curvature around the mode), it gives full information about the posterior so that quantities like quantiles and Bayes Factors can be determined, and the *process* can reveal information on the way (especially in Bayesian nonparametrics).

## 2 Where We Are Now

The use of Bayes’ Law in political science as a manipulator of probability statements is an old practice, and many of these works use decision-theoretic, psychological updating, or rational choice arguments. However, Bayesian regression-style models did not really appear in political science until the mid-1990s with the appearance of works like Bartels (1997), Gill (1999), Gelman and King (1994), Katz and King (1999), Quinn, Martin and Whitford (1999), Western (1998), and Western and Jackman (1994). An important exception to this wave is Chris Achen’s 1978 paper, although he restricts most of the Bayesian discussion to an appendix. See also Sidney Ulmer’s 1975 critical essay, which contains no data analysis but was way ahead of its time. After 2000, Bayesian models were regular features of prominent political science journals, and a search for “Bayesian” in the quarterly issues of *Political Analysis* (since 1999) gives 176 articles.

There is not much controversy amongst the more quantitatively-oriented political scientists about the use of Bayesian models, and even the least likely to use these methods see them as a principled way to incorporate prior information (quantitative or qualitative), make probabilistic claims from regressions, or to conveniently specify hierarchies. Regretfully, a non-trivial proportion of the discipline still regards Bayesian models as exotic or perhaps even sinister. On the other hand, one can find entire panels of political scientists at the APSA meeting that regard *regression*

as exotic and sinister. If there is any appreciable level of distrust, I believe that it comes from not fully understanding the role of the prior distribution. Nearly all Bayesian papers in political science seek to minimize the influence of the prior specifications. This stems from a general lack of interest in fleshing-out principled prior forms from the literature, and a desire to quell journal reviewers. Neither of these are good philosophical reasons to minimize the discussion of priors but both are strongly vocational. My hope is that this changes. Perusing the back cover of any issue of the *Journal of the American Statistical Association* over the last decade or two shows that there is nothing controversial about Bayesian approaches in general statistics research.

In the dramatic increase in Bayesian methods in political science we see applications to GLMs, causal inference, time-series, change-point problems, ideal point estimation, expert elicitation, missing data imputation, genetics analysis, textual analysis, nonparametrics, ecological inference, neural networks, structural equation models, and factor analysis. This list is important because it demonstrates that Bayesian approaches are not just another “tool” in the standard sense, but are instead a general philosophical way of thinking about data and uncertainty. This discussion appears in many places and I will not repeat it here (see Samaniego [2010] for a recent detailed look). Critically, the Bayesian approach will continue to gain in popularity because it is well-suited for the type of data we deal with (observational) and the types of theories that we care about. Almost no political scientist believes that the phenomenon they care about is fixed and unyielding over time and circumstance. We tend to care about quantities such as the likelihood that two nations go to war, the probability that a certain type of voter will pick a particular party, the tendency for legislators to vote in patterns, and so on. These are, by definition, varying quantities and therefore best described with distributions.

### 3 Articles In this Virtual Issue

Of the 176 articles in *Political Analysis* that address Bayesian methods in some fashion I was asked to pick a relatively modest number. This small  $p$  binomial choice is regretful since many political methodologists have contributed excellent work over the last 12 years of the quarterly release of the journal, which I took to be my sampling frame. The six works chosen are a mix of papers that I believe to be fundamentally important, and papers that I personally enjoyed. Most are both. While two papers are from 2010, I tried to have a range of dates to reflect the genesis of Bayesian political methodology. I also tried to vary the types of methods that Bayesian inference is used to address. There is also a nice range of seniority in the discipline reflected in these authors with three of them contributing solo-authored works as pre-tenured scholars. One paper (Quinn 2004) is from the Bayesian special issue of *Political Analysis*, and one paper (Martin and Quinn 2002) is of sufficient age and quality that it has collected 450 citations. After briefly introducing this set of articles, I will retain the convention in these introductions of suggesting areas for future research.

**Andrew Martin** and **Kevin Quinn**’s (2002, Vol 10(2), pp. 134-153) article won the 2001 Harold Gosnell Prize by the Political Methodology Section of the American Political Science Association. They are principally concerned with estimating ideal points for justices on the U.S. Supreme Court and how they change over time. Using data from The United States Supreme Court Judicial Database (Spaeth, 2001), which covers 1953 to 1999, they create a dynamic spatial model built on item response theory (IRT) foundations. Their challenges are formidable relative to standard legislative settings: the number of subjects is small for any given court, the institution is secretive rather than open in its deliberations, the standard identification problem is not easy to solve, and the interest is in dynamic behavior rather than constant over a single term. Unlike standard IRT models that assume quantities of interest like student aptitude are fixed, Martin and Quinn assume that justice ideal points are variable. Clearly a model addressing these challenges

cannot be cleanly implemented without a Bayesian distributional approach. Therefore their prior is built on a random walk strategy that conditions justice ideal points on past estimates. This model requires a customized MCMC solution for inference. The authors not only coded their sampler from scratch in C++, they made it available as a general resource for others. Later versions developed an R interface and other model specifications. Their substantive results are not only interesting, but also contradict important previous findings in the literature.

Factor analysis is a popular tool in the social sciences because it is extremely easy to run and greatly simplifies many multidimensional problems. Unfortunately standard approaches cannot viably mix ordinal and continuous variables into the same factor, although this routinely violated in practice. **Kevin Quinn** (2004, Vol 12(4), pp. 338-353) tackles this problem with a new model that unifies the two data cases by assuming that they are both determined, at varying levels of quality, by an underlying latent continuous measure. Therefore the true data generating process is completely continuous in multidimensional space, but the observed manifestation differs due to some intervening process. The resulting factor analysis specification is sufficiently complex to estimate that a Metropolis-Hastings algorithm is required. Quinn provides a customized software solution, made available through the R package `MCMCpack` (Martin, Quinn, and Park 2011). While the application focuses on the index of political-economic risk, this paper gives a solution that is universal across data-oriented disciplines, and is therefore an extremely important scientific contribution.

Time-series analysis is a small cottage industry among political methodologists, and Bayesian time-series is an even smaller subset branch of the field. The problem is that Bayesian time-series can be very *hard*. **Patrick Brandt** and **John Freeman** (2006, Vol 14(1), pp. 1-36) review the current state of time-series, and in particular Bayesian time-series, in political science. Except for other work by Brandt and Freeman (Brandt, Colaresi and Freeman 2008, Brandt and Freeman 2009, Brandt, Freeman, and Schrodtt 2011), there has not been much Bayesian work in political science since this article appeared. However, Freeman and Brandt have greatly influenced how we think about dynamic models including prediction in general and they have encouraged our continuing movement away from econometrically-driven specifications that do not fit longitudinal political data very well. Surprisingly, in the review part of this article they find that political scientists who use time-series methods often provide no measures of uncertainty for their causal claims, and no error bands on many reported quantities. Brandt and Freeman then introduce the Bayesian vector autoregression (BVAR) specification of Doan, Litterman, and Sims (1984) where the data are assumed to be first-order integrated with a drift, or that the classic first differences of each series cannot be predicted. They then contrast the well-known “Minnesota” prior with a new reference prior based on the research of Sims and Zha (1998). It turns out that the latter gives a more detailed, and theoretically driven form of the structural model for Bayesian forecasting. In addition, they give a rigorous procedure for evaluating the sensitivity of the priors about the dynamics, which differs from standard prior sensitivity analysis.

Change points are common in political data. We study lots of phenomenon where some change to a regime, an institution, a set of voters, causes a stream of data (usually measured over time) to shift noticeably. This is an easy problem when we can point to an event that has occurred at some known time: a coup, a constitutionally required change, a macro-political event that alters voters preferences, and so on. Unfortunately there are occasions where we know that there has been a fundamental shift but we cannot point to the exact time that it occurs. Change point models are constructed to estimate the time of this shift are ideal applications of Bayesian methods since the assertion of the shift is often distributional rather than as a fixed single point. **Arthur Spirling** (2007, Vol 15(4), pp. 387-405) looks at Bayesian change point models in political science for nonlinear outcome models. By constructing a series of useful link functions (log-linear, logit, exponential duration) Spirling adds to our Bayesian toolbox in a very useful

way since these specifications accommodate covariate inclusion, which is often not provided in similar work. This article has a wonderful “workshop” feel to it in that Spirling provides enough detail, and `bugs` code, so that the reader can immediately begin developing their own related models.

Text analysis is at an early but very exciting stage in political science. Unfortunately the modeling demands are high for producing useful substantive conclusions since political communication comes in institutional contexts that have different word-to-phrase contexts. So the word “sentence” may be used differently structurally by a court than by a campaigning politician. **Justin Grimmer** (2010 Vol 18(1), pp. 1-35) takes-on this problem, in an article that won the 2011 Miller Prize by the Political Methodology Section of the American Political Science Association, looking at one institutional setting: press releases by Senate offices. The problem with “unsupervised learning methods,” which generally just assign words or phrases to topic categories, is that they do not account for the relative emphasis that the speaker or the author places on different passages. Naturally these tools are becoming more sophisticated with time, but it is still difficult to address Grimmer’s problem where there exists a hierarchy in the data: press releases are grouped within senators, and different senators see word-usage differently. The resulting Bayesian hierarchical model developed in this paper is innovative in that it gets at the “expressed agenda” for each senator by specifying a multinomial individual senator draw from a Senate-wide Dirichlet distribution, which then provides conditioning for a von Mises-Fisher distribution for the individual press releases. This is a classic Bayesian multilevel model approach that is tailored to a specific political setting. There is also a lot going on in this paper besides the model specification: an immense amount of data collection and conditioning, construction of a variational estimation algorithm, and sophisticated model-checking.

Time-series cross-sectional (TSCS) data has been a strong interest of political methodologists for decades. A large amount of political data arrive for us with a longitudinal component across cases. As noted above, there is an unfortunate paucity of sophisticated Bayesian time-series work in political science, including cross-sectional models. **Xun Pang**’s article (2010 Vol 18(4), pp. 470-498) is clearly an exception. She produces a order-p autoregressive error process for unbalanced binary TCSC data. This specification is also a multilevel Bayesian generalized linear model in the conventional sense. This is an important model specification that has not been developed in any statistics journal article to date. A key problem that Pang faces is the common occurrence of heterogeneity across individual units *and* over time. The multilevel specification is necessary to handle these simultaneous issues, and the autoregressive structure is built to both correct serial correlation and improve fit. This model is also a contribution in that it handles nonlinear outcome variables in the TSCS context in addition to the other features. All of these steps forward (which could have been multiple models across multiple published papers) mean that estimation is especially tricky. So Pang develops a customized MCMC procedure based on data augmentation and a Cholesky decomposition of the error matrix that results from modeling the serial correlation. But that is not all. The off-diagonal correlation structure means that a naïvely constructed chain will mix very poorly through the sample space. So she borrows a tool from statistical physics and Euclidean quantum physics that adds a coarsened auxiliary grid over the fine grid of the original problem. This provides a means of temporarily “jumping” to a faster moving grid but staying on the same target sample space. This too could have been a completely separate *Political Analysis* article. Finally, the examples demonstrate substantially better fit and prediction with a range of data commonly encountered in political science.

## 4 Future Research Agenda

There is plenty of Bayesian work to be done, both methodologically and in applied settings. In Bayesian time-series, there is a need to add more structural features and simultaneity in both the error structures and the hierarchical components. The works described here demonstrate that this is difficult, but rewarding, both in terms of theory development and in terms of resulting computation. So far change point models are fairly basic in political science (though very useful!), and the problem of an unknown number of multiple change points has not been adequately addressed. The state of the art in statistics is the paper by Girón, Moreno, and Casella (2007) using intrinsic priors, but only for the homoscedastic normal linear model. Since language is naturally hierarchical, models for text analysis can clearly be improved by extending Bayesian multilevel specifications, as suggested by the single specification given by Justin Grimmer. Bayesian nonparametrics is another exciting area, especially now that the computational challenges are mostly under control (Kyung, Gill, and Casella 2009, 2010). This family of tools based on Dirichlet process priors can account for latent clustering that regular specifications ignore. Another general area of Bayesian modeling that can use more attention is the specification of priors with contextual information or with desirable mathematical properties. In the first case, some disciplines (notably medicine), have been successful in incorporating defensible prior knowledge into prior distributions as a way to improve the quality of the posterior. In the second case, the so-called “objective Bayesian” group promotes the development of (possibly complex) alternatives to flat priors for low-information specifications.

## 5 Concluding Remarks

Political scientists have increasingly embraced Bayesian methods as helpful ways to address empirical and methodological challenges. Over the last two decades, any sense of controversy has receded from the general field of statistics. With a wide range of available MCMC tools, estimation challenges are now manageable, even under difficult circumstances. This leads to an environment whereby political scientists have few impediments in developing useful and principled Bayesian models for their empirical questions. It is clear that *Political Analysis* has played an important role in getting to this current state.

This essay is not to suggest that Bayesian methods are a panacea. I have seen plenty of evidence that it is easily as possible to construct flawed Bayesian specifications as it is to construct flawed non-Bayesian specifications. To be fair, the bulk of this evidence is from conference presentations and review manuscripts. Yet the Bayesian paradigm gives a uniformly more principled approach to describing uncertainty from data and models. As Ed George observes, “All good procedures are Bayesian, but not all Bayesian procedures are good” (personal communication).

## 6 Funding

Financial support from the National Science Foundation (grant SES-1120976), and the National Institutes of Health (National Cancer Institute Transdisciplinary Research on Energetics and Cancer (TREC) Center is gratefully acknowledged.



## 7 About the Author

Jeff Gill is a Professor in the Department of Political Science (College of Arts & Sciences), the Division of Biostatistics (School of Medicine), and the Division of Public Health Sciences (School of Medicine), at Washington University in St. Louis. His current areas of research interests include: Bayesian hierarchical models, Markov chain Monte Carlo theory, Bayesian meta-analysis, the psychological effect of war on children, linkages between obesity and cancer, and Bayesian nonparametrics.

## 8 References

- Achen, C. H. (1978). Measuring Representation. *American Journal of Political Science* **22**, 475-510.
- Bartels, L. M. (1997). Specification Uncertainty and Model Averaging. *American Journal of Political Science* **41**, 641-674.
- Brandt, P. T. and Freeman, J. R. (2006). Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting, and Policy Analysis *Political Analysis* **14**, 1-36.
- Brandt, P. T. and Freeman, J. R. (2009). Modeling Macro-Political Dynamics. *Political Analysis* **17**, 113-142.
- Brandt, P. T., Colaresi, M., and Freeman, J. R. (2008). The Dynamics of Reciprocity, Accountability, and Credibility. *Journal of Conflict Resolution* **52**, 343-374.
- Brandt, P. T., Freeman, J. R., and Schrodtt, P. A. (2009). Real Time, Time Series Forecasting of Inter- and Intra-State Political Conflict. *Conflict Management and Peace Science* **28**, 41-64.
- Birnbaum, A. (1962). On the Foundations of Statistical Inference. *Journal of the American Statistical Association* **57**, 269-306.
- Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and Conditional Projection Using Realistic Prior Distributions. *Econometric Reviews* **3**, 1-100.
- Fisher, R. A. (1922). On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Statistical Society of London A* **222**, 309-360.
- Fisher, R. A. (1925a). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- Fisher, R. A. (1925b). Theory of Statistical Estimation. *Proceedings of the Cambridge Philosophical Society* **22**, 700-725.
- Gigerenzer, G. and Murray, D. J. (1987). *Cognition As Intuitive Statistics*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Gill, J. (1999). The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly* **52**, 647-674.
- Gill, J. 2008. Is partial-dimension convergence a problem for inferences from MCMC algorithms? *Political Analysis* **16**, 153-178.
- Gelfand, A. E. and Smith, A. F. M. (1990). Sampling Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association* **85**, 398-409.
- Gelman, A. and King, G. (1994). A Unified Method of Evaluating Electoral Systems and Redistricting Plans. *American Journal of Political Science* **38**, 514-554.

- Grimmer, J. (2010). A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis* **18**, 1-35.
- Katz, J. N. and King, G. (1999). A Statistical Model for Multiparty Electoral Data. *American Political Science Review* **93**, 15-32.
- Kyung, M., Gill, J. and Casella, G. (2009). Characterizing the variance improvement in linear Dirichlet random effects models. *Statistics and Probability Letters* **79**, 2343-2350.
- Kyung, M., Gill, J. and Casella G. (2010). Estimation in Dirichlet Random Effects Models. *Annals of Statistics*, 38, 979-1009.
- Martin, A. D. and Quinn, K. M. (2002). Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999. *Political Analysis* **10**, 134-153.
- Martin, A. D., Quinn, K. M., Park, J. H. (2011). MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software* **42**, 1-21. URL: <http://www.jstatsoft.org/v42/i09/>.
- Neyman, J. and Pearson, E. S. (1928a). On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference. Part I. *Biometrika* **20A**, 175-240.
- Neyman, J. and Pearson, E. S. (1928b). On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference. Part II. *Biometrika* **20A**, 263-294.
- Neyman, J. and Pearson, E. S. (1933a). On the Problem of the Most Efficient Test of Statistical Hypotheses. *Philosophical Transactions of the Royal Statistical Society, Series A* **231**, 289-337.
- Neyman, J. and Pearson, E. S. (1933b). The Testing of Statistical Hypotheses in Relation to Probabilities. *Proceedings of the Cambridge Philosophical Society* **24**, 492-510.
- Neyman, J. and Pearson, E. S. (1936a). Contributions to the Theory of Testing Statistical Hypotheses. *Statistical Research Memorandum* **1**, 1-37.
- Neyman, J. and Pearson, E. S. (1936b). Sufficient Statistics and Uniformly Most Powerful Tests of Statistical Hypotheses. *Statistical Research Memorandum* **1**, 113-137.
- Pang, X. (2010). Modeling Heterogeneity and Serial Correlation in Binary Time-Series Cross-sectional Data: A Bayesian Multilevel Model with AR(p) Errors. *Political Analysis* **18**, 470-498.
- Quinn, K. M. (2004). Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses. *Political Analysis* **12**, 338-353.
- Quinn, K. M., Martin, A. D., and Whitford, A. B. (1999). Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models. *American Journal of Political Science* **43**, 1231-1247.
- Samaniego, F. J. (2010). A Comparison of the Bayesian and Frequentist Approaches to Estimation. New York: Springer-Verlag.
- Sims, C. A. and Zha, T. A. (1998). Bayesian Methods for Dynamic Multivariate Models. *International Economic Review* **39**, 949-968.
- Spaeth, H. J. (2001). United States Supreme Court Judicial Database, 1953-2000 Terms [Computer File] . Ann Arbor, MI: Inter-University Consortium for Political and Social Research, 16th edition.
- Spirling, S. (2007). Bayesian Approaches for Limited Dependent Variable Change Point Problems. *Political Analysis* **15**, 387-405.
- Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA:



Harvard University Press.

- Tversky, A. and Kahneman, D. (1974). Judgment Under Uncertainty: Heuristics and Biases. *Science* **185**, 1124-1131.
- Tversky, A. and Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice *Science*, New Series, **211**, 4481, 453-458.
- Ulmer, S. S. (1975).  $H_0$ : Post Hoc Con-Straw-Man Con = 0. *American Journal of Political Science* **19**, 565-570.
- Western, B. (1998). Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach. *American Journal of Political Science* **42**, 1233-1259.
- Western, B. and Jackman, S. (1994). Bayesian Inference for Comparative Research. *American Political Science Review* **88**, 412-423.