CHAPTER 43

Geographical Models of Crises: Evidence from ICEWS

Michael D. Ward, Nils W. Metternich, Christopher Carrington, Cassy Dorff, Max Gallop, Florian M. Hollenbach, Anna Schultz, and Simon Weschle

Duke University
Durham, USA
michael.d.ward@duke.edu

ABSTRACT

Developing political forecasting models is not only relevant for scientific advancement, but also increases the ability of political scientists to inform public policy decisions. Taking this perspective seriously, the International Crisis Early Warning System (ICEWS) was developed under a DARPA initiative to provide predictions of international crisis, domestic crisis, rebellion, insurgency, and ethnic violence in about two-dozen countries in the US PACOM Area of Responsibility. As part of a larger project coordinated by Lockheed Martin Advanced Technology Labs, a team at Duke University created a series of geographically informed statistical models for these events. The generated predictions have been highly accurate, with few false negative and false positive categorizations. Predictions are made at the monthly level for three months periods into the future.

Keywords: prediction, international crisis, security

1 INTRODUCTION

Predicting crises has been a research priority of the US intelligence and warning community for decades. For the past several years under DARPA funding, a large, multidisciplinary team of computer and social scientists from universities and small businesses developed the Integrated Crisis Early Warning System (ICEWS). ICEWS provides Combatant Command staffs (COCOMs) with highly accurate and timely forecasting of instability events of interest (EOIs) using an innovative combination of state-of-the-art computational social science models (O'Brien, 2010). ICEWS exploits dynamic, high-volume, heterogeneous data sources to estimate these models and provide operators with situational awareness of past and current events in countries of interest. Many components of this system are being transitioned into the Integrated Strategic Planning and Analysis Network (ISPAN) program of record at the United States Strategic Command by Lockheed Martin. ICEWS has also been deployed for user testing and evaluation at the Pacific and Southern Combatant commands in 2010 and 2011.

In political science, prediction is typically conceptualized as a conditional exercise, in which values on a dependent variable are calculated based on some estimated, or conditional, statistical model, and then compared with the actual observed values (Hildebrand, Laing and Rosenthal, 1976). But there is also a recent tradition of attempting to make political predictions about things that have not yet occurred. An early proponent of using statistical models for making such predictions in the realm of international relations was Stephen Andriole, a research director at ARPA in the late 1970s (Andriole and Young, 1977). In 1978, a volume edited by Nazli Choucri and Thomas Robinson provided an overview of the then current work in forecasting in international relations, much of which was done in the context of policy oriented research for the U.S. government during the Vietnam War. There were a variety of efforts to forecast or evaluate forecasting efforts, including Freeman and Job (1979), Singer and Wallace (1979), & Vincent (1980), and a few projects began to forecast internal conflict (Gurr and Lichbach, 1986). However, the median empirical article in political science (as well as sociology and economics) used predictions only in the sense of in-sample observational studies. Doran (1999) and others provided some criticism but most scholars avoided making predictions, perhaps because their models had enough difficulty in describing accurately what had happened. Still a few scholars continued to make predictions, including Gurr and Harff (1998), Pevehouse and Goldstein (1999), Schrodt and Gerner (2000), King and Zeng (2001), O'Brien (2002), Bueno de Mesquita (2002), de Marchi, Gelpi and Grynaviski (2004), Enders and Sandler (2005), Ward, Siverson and Cao (2007), Brandt, Colaresi and Freeman (2008), Bennett and Stam (2009), and Gleditsch and Ward (2010), among others. A summary of classified efforts was declassified and reported in Feder (1995) and a nice overview of the historical efforts along with a description of current thinking about forecasting and decision-support is given by O'Brien (2010).

The basic task of the ICEWS project is to produce predictions for five dependent variables, for 29 countries[1] in PACOM, for every month from 1997 through the present plus three months into the future. The variables in question are rebellion, insurgency, ethnic violence, domestic political crises, and international political crises. Each month we receive a drop of two sets of data. The first of these comprises the five dependent variables in this study, known as the ground truth

[1] The twenty-nine countries are Australia, Bangladesh, Bhutan, Cambodia, China, Comoros, Fiji, India, Indonesia, Japan, Laos, Madagascar, Malaysia, Mauritius, Mongolia, Myanmar, Nepal, New Zealand, North Korea, Papua New Guinea, Philippines, Russia, Singapore, Solomon Islands, South Korea, Sri Lanka, Taiwan, Thailand, & Vietnam.

data. In addition, we receive data for each event that transpires within or involving each of the 29 countries in the sample. These event data are gleaned from natural language processing of a continuously updated harvest of news stories, primarily taken from Factiva™ (Dow Jones), an open source, proprietary repository of news stories from over 200 sources around the world. The baseline event coder is called JabariNLP, and builds on the TABARI/KEDS software developed by the Philip Schrodt and colleagues (see http://eventdata.psu.edu/). It combines a "shallow parsing" technology of prior coders with a richer exploitation of syntactic structure. This has increased accuracy (precision) from 50% to over 70%, as demonstrated in a series of ongoing (informal) evaluations of its output by human graders (peak human coding performance is around 80% (King and Lowe, 2003)).

These data are augmented with a variety of other attribute and network data. In particular we use attributes, coded on a monthly or yearly basis from the Polity, MAR, and World Bank data set. We also include information about the election cycles (if any) in each of the countries. In addition, we use information about relations among the 29 countries, including geography, the length of shared borders, the amount of trade, the movement of people across borders, the number of refugees, as well as the number and types of events between each pair of the 29 countries (plus the US).

The next section provides a brief review of the theoretical motivations for each of our models, summarizes relevant variables and literature, and informs the reader of the type of model used to generate the results. A full explanation of mixed effects models is provided in the second section, followed by an overview of the model specifications for each dependent variable. Finally, we demonstrate the predictive power of all the models and explore the cumulative probabilities of predicting any event of interest.

2 MODELING EVENTS OF INTEREST

ICEWS focuses on five EOIs: Rebellion, Insurgency, Ethnic Violence, Domestic Crisis, and International Crisis. To model each EOI, we draw from relevant literature in political science to suggest a set of conditions under which the specific event is likely to occur.

Rebellion. Our model for predicting rebellion uses proxies for the level of latent conflict between the government and the opposition, and then models the circumstances under which this latent conflict will lead to rebellions. The proxies are directional measures of the number of conflictual words ("demand", "disapprove", "reject", "threaten") stated from the government towards opposition groups and vice versa. We suggest that the effect of conflict on the probability of rebellion depends on the number of ethnically relevant groups that are excluded from power. When there are no excluded ethnic groups, rebellion should be very unlikely, as disagreements can be solved in the political arena. However, if a large number of excluded groups exist, coordination problems arise, which also mitigate rebellions. Hence, rebellion becomes most likely when few excluded groups exist.

We also include proximity to elections, which can bring about an increase in violence. A recent example is the case of Kenya, where following the victory of incumbent President Mwai Kibaki, the opposition denounced the results and widespread protests led to violence. As Snyder argues, while elections and democracy are often seen as important mechanisms in the peace building process, they can actually increase the likelihood of violence (Snyder, 2000). Additional predictive factors are detailed in Table 1.

Insurgency. Access to power is a key variable to understanding the causes of insurgencies. Insurgencies involve groups attempting to wrest political power from the sitting government, and so groups without access to political power are especially of interest. The larger this excluded population, the more likely violence will be used to change the political landscape. Furthermore, evidence has shown that violence designed to undermine the government is faced with a collective action problem (Kalyvas and Kocher 2007). However, if anti-government groups observe attacks against the government, they may change their calculus. Thus, we include a measure of dissident groups' actions against the government because such actions can be used as a rallying force and recruiting tool, increasing the probability of insurgency. Similarly, it follows that insurgencies in nearby countries may update individuals' beliefs about who else will act against the government of their own country. For this reason we include a measure of insurgencies in nearby countries, lagged by three months. We also suggest that nearby insurgencies could potentially disrupt effective government repression, liberating sources of weapons, money, and information for would-be insurgents in the target country.

Ethnic Violence. While most quantitative studies focus on the effect ethnicity has on conflicts between rebels and the government, we are interested in inter-ethnic and inter-religious violence. Thus, our concept of ethnic violence matches ideas of non-state war (Sarkees and Wayman, 2010), non-state conflicts (Eck, Kreutz and Sundberg, 2010), or subnational wars (Chojnacki, 2006), where the primary actors are non-state. In line with recent work on ethnic conflicts (Cederman, Wimmer and Min, 2010), we argue that government policies play an important role in explaining these dynamics. Thus, in our models, we include the number of politically excluded ethnic groups in a country and the overall proportion of the excluded ethnic population. The existing literature also points to a polarization effect of political exclusion, which suggests including the squared term of the proportion excluded. In addition, we argue that periods of political transition increase incentives to lock in political power in future institutions. Hence, we include Polity and its squared term to model political transition periods (Hegre et al., 2001). Finally, we model the spatial component of ethnic conflict. An increasing number of scholars not only highlight the transnational dimensions of civil conflict (Gleditsch, 2007; Salehyan, 2009), but also its ethnic component (Cederman, Buhaug and Rød, 2009). Thus, our model takes into account possible spillover effects from neighboring countries.

Domestic Crises. Domestic violence and protests are frequently triggered by elections that were perceived to be unfair. We include proximity of elections in our model, with different effects depending on the level of executive constraints. We propose this approach because in countries with moderate levels of executive constraints, elections have meaningful implications regarding who holds office, but governments have the latitude to manipulate the elections and therefore domestic crises are more likely to center around elections. A second major factor that we believe affects the propensity of domestic crises onsets is a country's ethnic composition. When ethnic groups are excluded from political processes grievances are likely to arise. In authoritarian systems this effect is likely to differ from democracies, so in our model, the effect of the number of excluded groups varies by executive constraints. Hence, the likelihood of domestic conflict is conditional on different levels of executive constraints, with the coefficients for the proximity to elections and the number of excluded groups also varying by executive constraints. In addition to the random effects for proximity to election and number of excluded ethnic groups, we control for GDP per capita, population size, and a spatial lag of domestic crises.

International Crises. Our model of international crises tries to capture those situations when a leader is unable or unwilling to make the necessary concessions to avoid a crisis. A leader's incentives to avoid international crises will be conditional on domestic political institutions. Leaders in more democratic regimes may be less able to make concessions internationally due to threat of domestic audience costs. The costs of a crisis might be lower for leaders of more autocratic regimes since their constituency will not bear the brunt of any potential fighting (Bueno de Mesquita et al., 2003; Schultz, 2001). To account for systematic differences between the prevalence of crises under different regime types, the model includes a random intercept based on a country's democracy polity score. In addition, homogeneous populations impose few constraints on the bargaining of leaders in democracy. So, the model also includes a random effect for the number of politically relevant ethnic groups conditional on level of democracy. We also control for population size, international crises in politically similar states and include measures for both domestic political pressure and domestic conflict.

3 METHODS: PREDICTING EVENTS OF INTEREST USING MIXED EFFECTS MODELS

We model rebellion, domestic conflict, and international conflict using hierarchical models in which both the intercept and slope vary. This means that we group the data along an indicator, such as level of executive constraints, creating a different intercept for each group. Thus, the varying intercepts correspond to group indicators and the varying slopes represent an interaction between predictor variables x and group indicators:

$$\Pr(y_{it} = 1) = \text{logit}^{-1}(\alpha_{j[it]} + \beta^G_{j[it]} x^G_{it} + \beta^O_{j[it]} x^O_{it} + Z_{it}\gamma)$$

$$\begin{pmatrix} \alpha_j \\ \beta^G_j \\ \beta^O_j \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_\alpha \\ \mu^G_\beta \\ \mu^O_\beta \end{pmatrix}, \Sigma\right)$$

where i denotes the countries, t the month and j the grouping variable, α_j are the grouping variable's random intercepts. $x_{it}{}^{G}$ and $x_{it}{}^{O}$ are predictor variables; $\beta_j{}^{G}$ and $\beta_j{}^{O}$ are the associated random coefficients; γ is a vector of fixed effects associated with Z_{it}. Table 1 provides an overview over all model specifications.

For an illustrative example, this equation accurately presents our model of rebellion, where j denotes the tercile in which a country falls with respect to the number of excluded ethnic groups, α_j are the grouping-specific random intercepts, $x_{it}{}^{G}$ is the number of conflictual words from the government against the opposition in country i at time t, and $x_{it}{}^{O}$ is the number of conflictual words from the opposition against the government; $\beta_j{}^{G}$ and $\beta_j{}^{O}$ are the associated random coefficients. All models except ethnic violence take this form, which does not include grouping variables.

Table 1 Overview of variables in each EOI prediction model

	Grouping Variables	Controlled Effects
Rebellion	Excluded groups grouped by: Conflictual words opposition → government Conflictual words gov → opposition	Proximity to election Competitiveness of executive recruitment Executive constraints GDP per capita (log) Rebellions in surrounding countries
Insurgency	Country	Proportion of population excluded Number of excluded ethnic groups Exclusion interaction: population × number of groups High intensity actions: dissidents → government Insurgencies in nearby countries
Ethnic Violence	None	Number of excluded groups Number included groups Proportion of population excluded Squared proportion of population excluded High intensity actions: ethnic groups → government Polity score Squared polity score Violence in neighboring countries
Domestic Conflict	Level of executive constraints grouped by: Number of excluded groups Proximity to election	Population (log) GDP per capita (log) Crises in neighboring countries
International Conflict	Number of ethnic groups grouped by: Level of democracy	Population (log) Domestic EOIs Conflictual words: any domestic group → government International crises in politically similar countries

All effects are lagged three months, except proximity to election.

4 PREDICTIVE POWER

In this section we demonstrate the predictive power of our models and show how to use the individual predictive probabilities to create a cumulative predictive probability of any crisis event occurring. We use separation plots to visualize and assess the predictive power of our models. These plots provide a summary of the fit for each model by demonstrating the range and degree of variation among the predicted probabilities and the degree to which predicted probabilities correspond to actual instances of the event (Greenhill, Ward and Sacks, 2011). Red panels represent events and non-events are left white. The line through the center of the plot represents the expected probability for each model. Thus the probability increases from left to right in the plot and a good fit would be visualized with more red panels (events occurring) stacked at the right end of the plot.

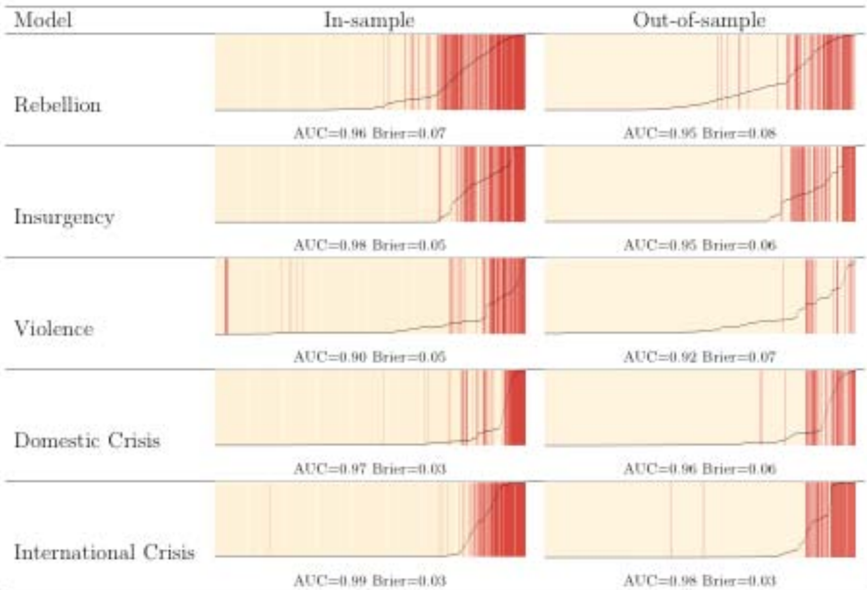

Figure 1 Predictive Capabilities of All EOI Models, including Area under the ROC curve (AUC) and Brier scores.

Another way to evaluate predictions that is employed here is the Brier score, defined as the average squared deviation of the predicted probability from the true event (Brier, 1950). It has been shown that the Brier score is one of the few strictly proper scoring rules for predictions with binary outcomes (Gneiting and Raftery, 2007). Brier scores closer to zero indicate better predictive performance. In Figure 1, we combine a visual and numeric interpretation of the models to display their predictive power. The separation plots and the provided statistics demonstrate a very good predictive performance of our models.

In addition to predicting specific crisis events, we can use the predictive probabilities derived by each of the models to create a cumulative predictive probability of any crisis event occurring. To do so we first create the new crisis variable that indicated the occurrence of any of the five crisis events above. To predict the experience of any crisis event we combine the individual predicted probabilities and calculate a cumulative probability of any crisis occurring. By simple probability theory, the cumulative probability is:

$$P(Y_{i,t}^C = 1) = [1 - P(Y_{i,t}^C = 0)] = 1 - \prod_{k=1}^{5}(1 - P(Y_{i,t}^k = 1))$$

where k ranges from one through five, enumerating the predicted probabilities derived for country i at time t for each of the individual EOIs discussed above. In words, the probability of a country to experience any crisis event in a given month is one minus the probability of no event occurring. This can be calculated as the product of the individual probability of non-occurrence for each of the individual events. Y stands for the occurrence of a crisis event in country i at time t, where the superscript indicates the type of crisis. We then use the cumulative probability to predict if a country experiences any of the above specified crises in a given month.

Figure 2 Separation plot for any crisis, indicating fit of the model, in and out-of-sample.

Figure 2 shows the separation plots of predicting any crisis using the cumulative predictive probabilities. There are very few actual crises that are missed, and relatively few false positives. Especially on the out-of-sample data, the cumulative prediction missed very few actual crises. A curious case is displayed on the in-sample separation plot, which shows a number of actual events on the left side with very low predicted probabilities of crisis occurrence. This is the case of the Solomon Islands in 2003, which experienced ethnic violence in five months during that year and was missed by the individual prediction models. Since the cumulative prediction is dependent on the individual models, the same case is missed in the in-sample prediction of the ethnic violence model. For the in-sample observations, the Brier score for predicting the occurrence of any crisis is 0.07, while the Brier score for the out-of-sample observations is 0.1. Thus, the model predicts quite accurate using the individual predictions to calculate a cumulative probability of any crisis occurring.

5 CONCLUSION

We have demonstrated the utility of creating forecasting models for predicting political conflicts in a diverse range of country settings, and have shown that this series of geographically informed statistical models is highly accurate, containing few false negative and positive predictions. These models can serve the public policy community and shed light on an array of critically important components of the political science literature on conflict dynamics. Moving forward, this project hopes to extend these models beyond their current geographical domains and onto a larger, worldwide data set.

ACKNOWLEDGMENTS

This project was undertaken in the framework of an initiative funded by the Information Processing Technology Office of the Defense Advanced Research Projects Agency to provide an Integrated Crisis Early Warning System (ICEWS) for decision makers in the U.S. defense community. The holding grant is to the Lockheed Martin Corporation, Contract FA8650-07-C-7749. For helpful insights we thank Scott de Marchi. All the bad ideas and mistakes are our own. We thank Mark Hoffman and Philippe Loustaunau for support and suggestions.

REFERENCES

Andriole, S.J. and R.A. Young. 1977. Toward the Development of an Integrated Crisis Warning System. *International Studies Quarterly* 21: 107–150.

Bennett, D.S. and A.C. Stam. 2009. Revisiting Predictions of War Duration. *Conflict Management and Peace Science* 26: 256-267.

Brandt, P.T., M. Colaresi and J.R. Freeman. 2008. The Dynamics of Reciprocity, Accountability, and Credibility. *Journal of Conflict Resolution* 52: 343-374.

Brier, G.W. 1950. Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review* 78: 1-3.

Bueno de Mesquita, B. 2002. *Predicting Politics*. Columbus: Ohio State University Press.

Bueno de Mesquita, B., A. Smith, R. M. Siverson and J. D. Morrow. 2003. *The Logic of Political Survival*. Cambridge: M.I.T. Press.

Cederman, L.-E., A. Wimmer and B. Min. 2010. Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics* 62: 87-119.

Cederman, L.-E., H. Buhaug and J.K. Rød. 2009. Ethno-nationalist dyads and civil war: A GIS-based analysis. *Journal of Conflict Resolution* 53: 496-525.

Chojnacki, S. 2006. Anything New or More of the Same? Wars and Military Interventions in the International System, 1946-2003. *Global Society* 20: 25-46.

Choucri, N. and T. W. Robinson, eds. 1978. *Forecasting in International Relations: Theory, methods, problems, prospects*. San Francisco: W.H. Freeman.

de Marchi, S., C. Gelpi and J. D. Grynaviski. 2004. Untangling Neural Nets. *American Political Science Review* 98: 371-378.

Doran, C.F. 1999. Why Forecasts Fail: The Limits and Potential of Forecasting in International Relations and Economics. *International Studies Review* 1: 11-41.

Eck, K., J. Kreutz and R. Sundberg. 2010. Introducing the UCDP Non-State Conflict Dataset. Uppsala University: Unpublished manuscript.

Enders, W. and T. Sandler. 2005. After 9/11: Is it All Different Now? *Journal of Conflict Resolution* 49: 259-277.

Feder, S. 1995. Factions and Policon: New Ways to Analyze Politics. In *Inside CIA's Private World: Declassified Articles from The Agency's Internal Journal, 1955-1992*, ed. H.B. Westerfield. New Haven: Yale University Press.

Freeman, J.R. and Brian L. J. 1979. Scientific Forecasts in International Relations: Problems of Definition and Epistemology. *International Studies Quarterly* 23: 113–143.

Gleditsch, K.S. 2007. Transnational dimensions of civil war. *Journal of Peace Research* 44: 293-309.

Gleditsch, K.S. and M.D. Ward. 2010. "Contentious Issues and Forecasting Interstate Disputes." Paper presented at the 2010 Annual Meeting of the International Studies Association, New Orleans.

Gneiting, T. and A.E. Raftery. 2007. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association* 102: 359-378.

Greenhill, B. D., M.D. Ward and A. Sacks. 2011. A New Visual Method for Evaluating the Fit of Binary Models. *American Journal of Political Science* 55: 991-1002.

Gurr, T.R. and M.I. Lichbach. 1986. Forecasting Internal Conflict: A Competitive Evaluation of Empirical Theories. *Comparative Political Studies* 19: 3-38.

Gurr, T.R. and B. Harff. 1998. Systematic Early Warning of Humanitarian Emergencies. *Journal of Peace Research* 35: 551-579.

Hegre, H., T. Ellingsen, S. Gates and N.P. Gleditsch. 2001. Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816-1992. *American Political Science Review* 95: 33-48.

Hildebrand, D. K., J. D. Laing and H. Rosenthal. 1976. Prediction Analysis in Political Research. *American Political Science Review* 70: 509-535.

King, G. and L. Zeng. 2001. Improving Forecasts of State Failure. *World Politics* 53: 623-658.

King, G. and W. Lowe. 2003. An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design. *International Organization* 57: 617-642.

O'Brien, S.P. 2002. Anticipating the Good, the Bad, and the Ugly: An Early Warning Approach to Conflict and Instability Analysis. *Journal of Conflict Resolution* 46: 791-811.

O'Brien, S.P. 2010. Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research. *International Studies Review* 12: 87-104.

Pevehouse, J.C. and J.S. Goldstein. 1999. Serbian compliance or defiance in Kosovo? Statistical analysis and real-time predictions. *Journal of Conflict Resolution* 43: 538-546.

Salehyan, I. 2009. *Rebels without borders: transnational insurgencies in world politics*. Ithaca: Cornell University Press.

Sarkees, M.R. and F. Wayman. 2010. *Resort to War: 1816-2007*. Washington, DC: CQ Press.

Schrodt, P.A. and D.J. Gerner. 2000. Using Cluster Analysis to Derive Early Warning Indicators for Political Change in the Middle East, 1979-1996. *American Political Science Review* 94: 803-818.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.

Singer, J.D. and M.D. Wallace. 1979. *To augur well: early warning indicators in world politics*. Beverly Hills: Sage Publications.

Snyder, J. 2000. *From Voting to Violence: Democratization and Nationalist Conflict*. New York: Norton.

Vincent, J.E. 1980. Scientific Prediction versus Crystal Ball Gazing: Can the Unknown be Known? *International Studies Quarterly* 24: 450-454.

Ward, M.D., R.M. Siverson and X. Cao. 2007. Disputes, Democracies, and Dependencies: A Reexamination of the Kantian Peace. *American Journal of Political Science* 51: 583-601.