

When Forecasts Fail: Unpredictability in Israeli-Palestinian Interaction

Charles Kurzman, Aseem Hasnain

University of North Carolina at Chapel Hill

Abstract: This article explores the paradox that forecasts may be most likely to fail during dramatic moments of historic change that social scientists are most eager to predict. It distinguishes among four types of shocks that can undermine the predictive power of time series analyses: effect shocks that change the size of the causal effect; input shocks that change the causal variables; duration shocks that change how long a causal effect lasts; and actor shocks that change the number of agents in the system. The significance of these shocks is illustrated in Israeli–Palestinian interactions, one of the contemporary world’s most intensely scrutinized episodes, using vector autoregression analyses of more than 15,000 Reuters news stories over the past three decades. The intervention of these shocks raises the prospect that some historic episodes may be unpredictable, even retrospectively.

Keywords: forecasting; prediction; unpredictability; Israel; Palestine

Editor(s): Jesper Sørensen, Delia Baldassarri; **Received:** March 7, 2014; **Accepted:** April 23, 2014; **Published:** June 23, 2014

Citation: Kurzman, Charles and Aseem Hasnain. 2014. “When Forecasts Fail: Unpredictability in Israeli-Palestinian Interaction.” *Sociological Science* 1: 239-259. **DOI:** 10.15195/v1.a16

Copyright: © 2014 Kurzman and Hasnain. This open-access article has been published and distributed under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited.

FOR millennia, experts have promised special insight into the future. This ambition was one of the original visions for the social sciences in the nineteenth century—“to enquire into the present, in order to foresee the future, and to discover the means of improving it” (Comte [1851] 1875:15)—and prediction remained one of the “principal tasks” of many social scientists in the mid-twentieth century (Schuessler 1968:418–19). In recent decades, this ambition has retreated somewhat, as “prediction has become almost a taboo word, connoting an embarrassing affiliation to vulgar positivism, scientism and technocracy” (Aldridge 1999). Still, prediction continues to flourish in a variety of social sciences, especially the applied wings of criminology, econometrics, and international relations, where government and business planning is heavily invested in forecasting (Land and Schneider 1987; Cooper and Layard 2002; Schneider, Gleditsch, and Carey 2010). Over the past generation, forecasters have established institutional venues for professional development, such as the International Institute of Forecasters, as well as specialized academic journals and distinctive statistical methods.¹

¹Some scholars distinguish between prediction and forecasting, though there is no uniform usage of either

One of the greatest challenges in forecasting is to develop models that anticipate historical discontinuities—sudden turns of fortune such as economic crashes, civil conflict, and revolution (Moore 1964; Sornette 2002; Bueno de Mesquita 2009; Goldstone et al. 2010). The more dramatic and counterintuitive the outcome, the more rewarding it is to “endogenize” the factors and dynamics that led to it. This is part of the appeal of “chaos theory,” whose models generate irregular trajectories based on complex processes of closed systems with fixed parameters (Smith 1998; Kellert 2008). Yet history is rarely a closed system—agents may enter and leave, preferences and priorities may be overturned, or the dynamics of interaction may change. These shifts have been given a variety of labels, including black swans (Talib 2007), contingency (Shapiro and Bedi 2007; York and Clark 2007), critical junctures (Collier and Collier 1991), punctuations (Baumgartner and Jones 2009), structural breaks (Chow 1960; Hansen 2001), and turning points (Abbott 1997). We refer to them here by the generic term “shock,” meaning parameter shifts

term. We use the terms interchangeably in this article to mean the application of observed evidence to unobserved circumstances.

that impact a system but do not appear to be generated by the system.²

Shocks are often invoked when forecasts fail. In the 1930s, for example, economist Alfred Cowles calculated that professional stock market forecasters performed no better than random (Cowles 1933; see also Friedman 2014). At the end of the twentieth century, despite vast improvements in forecasting techniques, economists were still unable to predict many recessions and crashes (Loungani 2001). Studies of civil conflict often report high levels of statistical significance for their models, but when examined as out-of-sample forecasts, these models are not nearly as successful (Tikuisis, Carment, and Samy 2012; Ward, Greenhill, and Bakke 2010). Indeed, some of the most important phenomena of recent history have proven inordinately resistant to prediction (Cerulo 2006; Harcourt 2007; Kurzman 2004; Tetlock 2005; for journalistic treatments of this subject, see Gardner 2011; Sherden 1997; Silver 2012).

Forecasters are aware of these challenges. They typically qualify their predictions with the caveat that the patterns they observe hold only so long as the parameters of their models remain stable, and they have cataloged numerous ways in which these parameters may shift (Clements and Hendry 2006:609–14). Forecasters have developed sophisticated methods to control for shocks through dummy variables, splines, and other techniques, in an attempt to identify the unvarying, underlying properties of their models.

This article proposes, by contrast, that shocks are not so readily tamed. Applying standard forecasting methods to Israeli–Palestinian interactions, one of the world’s most closely watched conflicts, this study finds that forecasts are most likely to fail during historic moments that social scientists are most eager to predict. The article then distinguishes four ways in which shocks may undermine prediction and identifies major episodes in Israeli–Palestinian interaction that illustrate these shocks. Each of these forms of shock corresponds to a distinct element in forecasters’ models: not just coefficients, which are

²This use of the term “shock” refers to actual shifts in parameters at particular historical moments, as distinct from the other usage of the term “shock” in vector autoregression analysis to refer to hypothetical shifts in parameters for the purpose of generating impulse-response functions.

the usual suspects in time series analysis of structural breaks, but also time-varying lag structures, shifts in the values of endogenous variables, and changes to the set of endogenous variables. The article concludes by extending the argument from forecasting to retrocasting, that is, to time series analyses and historical explanations more generally.

Israel–Palestine

Jews and Palestinian Arabs have lived in the Levant for ages, but only in the past century have they engaged in persistent conflict, capturing global attention. This conflict grew out of duplicate nationalisms that emerged in the last decades of the Ottoman Empire: political Zionism, beginning in the 1890s (Vital 1975), and Palestinian nationalism, beginning in the 1910s (Khalidi 1997; Nafi 1998). Both of these movements envisioned a nation-state between the Jordan River and the Mediterranean Sea (Biger 2004; Shelef 2010). For the past two decades, many Israelis and Palestinians³ have circumscribed their nationalist aspirations in support of a “two-state” solution that would divide the land into separate enclaves (Shamir and Shikaki 2010). Still, some Israelis and Palestinians reject such a compromise and aspire to control the entire territory (Gunning 2008; Milton-Edwards and Farrell 2010; Taub 2010). Violence has repeatedly undermined attempts to negotiate a lasting political accommodation.

This high-profile case has generated an interdisciplinary debate over the extent to which Israeli–Palestinian interactions may be forecast. Some observers emphasize the history of “unpredictable risks and limited rationality” (Dowty 2012:114). Others view this history as “an all-too-familiar pattern of confrontation and violence” (Tessler 2009:754). Recent quantitative analyses by economists, political scientists, psychologists, and sociologists are also split on this issue. Most emphasize predictability, finding statistically significant patterns through which Israeli and Palestinian actions follow from previous ac-

³“Israeli” and “Palestinian” are not mutually exclusive categories, because some self-identified Palestinian Arabs are Israeli citizens. This article follows common usage in referring to the two categories as distinct identities and combines the terms in alphabetical order.

tions (Beasley 2008; Braithwaite, Foster, and Sobek 2010; Brym and Andersen 2011; Brym and Araj 2006; Haushofer, Biletzki, and Kanwisher 2010, 2011; Kaplan et al. 2005; Maoz 2007). Other researchers, however, distinguish the predictability of Israeli actions and the limited predictability of Palestinian actions (Hafez and Hatfield 2006; Jaeger and Paserman 2006, 2008, 2009). Two recent analyses find periods of greater predictability alternating with periods of lesser predictability (Dugan and Chenoweth 2012; Golan and Rosenblatt 2011).

Golan and Rosenblatt (2011) suggest that it is “implausible” to assume a single impulse-response function for a phenomenon that bridges multiple “epochs,” and they call for further data gathering on additional “aspects of this multifaceted conflict.” Our study addresses both of these concerns. First, we analyze a greater palette of Israeli–Palestinian interactions than the narrow spectrum of events acknowledged in previous studies, which have focused primarily on fatal violence, sometimes combined with data on non-fatal attacks and imprisonment (see Table A1 in the Supplemental Materials). We incorporate many more forms of interaction, including negotiations, protests, and a host of other events that occupy much attention in the region and around the world. Second, we investigate the possibility that multiple forms of shocks may occur at irregular intervals, generating uneven, syncopated “epochs.” Previous studies involved dummy variables or time series end points corresponding with distinct periods of Israeli–Palestinian interaction, implying that transitions from one period to the next are exogenous to the interactions themselves. Yet, like many other time series analyses, these studies have not explored the ways in which patterns shift at break points of history. This article broaches the implications of multiple forms of shocks for the substantive understanding of the case and for the enterprise of forecasting and retrocasting more generally.

Data and Methods

To document Israeli–Palestinian interaction, we examine all 15,884 news reports of Israeli–Palestinian events from the Reuters news agency over the course of 11,219 days (April 15, 1979, to Decem-

ber 31, 2009). An algorithm developed by the Penn State Event Data Project (formerly known as the Kansas Event Data System) machine-coded the first sentence in each article to identify the actor, the action, and the target of action (Schrodt and Gerner 1994, 1997, 2000; Rasler 2000; Goldstein et al. 2001; Brandt and Freeman 2006; Brandt, Colaresi, and Freeman 2008; Brandt, Freeman, and Schrodt 2011; Dugan and Chenoweth 2012). Of the 139,025 articles in the Penn State “Levant” data set during this period (<http://eventdata.psu.edu/data.dir/levant.html>), we select all events with Israeli actors and Palestinian targets and all events with Palestinian actors and Israeli targets. Articles that report joint events, in which Israelis and Palestinians both engage in the same action toward one another on the same day, are recorded twice as separate directed dyads. For the analysis of *actor shock*, discussed later, we also included all 5,643 stories reporting Israeli and Palestinian interaction with Lebanese actors during this period.

The Levant data set sorts events into 20 categories of action, called CAMEO (Conflict and Mediation Event Observations) codes, with more than a hundred subcategories (Schrodt 2012). We sort the 20 root categories into an ordinal scale of seven categories, ranging from -3 (violent conflict) to $+3$ (cooperation):

-3 : Violent conflict

CAMEO code 18: Assault

CAMEO code 19: Fight

CAMEO code 20: Mass violence

-2 : Conflict

CAMEO code 14: Protest

CAMEO code 15: Force posture

CAMEO code 16: Reduce relations

CAMEO code 17: Coerce

-1 : Negative communication

CAMEO code 9: Investigate

CAMEO code 10: Demand

CAMEO code 11: Disapprove

- CAMEO code 12: Reject
- CAMEO code 13: Threaten
- 0: Neutral communication
 - CAMEO code 1: Statement
 - CAMEO code 2: Appeal
- +1: Positive communication
 - CAMEO code 3: Intent to cooperate
 - CAMEO code 4: Consult
 - CAMEO code 5: Diplomatic cooperation
- +2: De-escalation
 - CAMEO code 8: Yield
- +3: Cooperation
 - CAMEO code 6: Material cooperation
 - CAMEO code 7: Aid

We sum the events into a single score for each directed dyad for each day.⁴ We also include a dummy variable for Saturdays, the Israeli weekend, because there are 35 percent fewer reported events on this day than on other days. Table A2 in the Supplemental Materials presents descriptive statistics.

In addition, we performed parallel analyses with the articles coded according to the Goldstein (1992) scale of conflict cooperation, as applied to the Levant data set by Schrodtt (2007), and with net daily totals of material cooperation (CAMEO codes 6–9) and material conflict (CAMEO codes 15–20) used by Brandt et al. (2011). We also examined articles on fatal violence alone (Goldstein code –10), both as a daily count and as a daily

⁴Eighty-six percent of the articles in the Levant database have date phrases in the lead sentence such as “today,” “Monday,” “last night,” “next week,” or “April.” Of these, 92 percent refer to events that occurred the same day as the report. Two percent refer to the previous day, and another 3 percent refer to other days within the same week, some of them in the future (“negotiators will meet tomorrow”). Articles that have no identifiable date phrases in the lead sentence normally refer to developments on the same day as the report (“meetings have begun,” “leaders are trying”). Events clearly identified with a previous date are coded as occurring on that date; all other events are coded as occurring on the date of the report.

binary. And we analyzed a separate source of violent events, the B’Tselem list of fatalities in the Israeli–Palestinian conflict, both as a daily count and as a daily binary. The results of all these analyses are directly analogous to the findings reported here.

Our scale of daily events is an imperfect representation of relations between Israelis and Palestinians in a number of ways, among them the bias and filter of the news agency, its editors, and reporters in selecting what events to cover; the reduction of each news report to a single event category; the asymmetry of the actors’ decision-making processes, because the most frequently covered Israeli actions (such as military and police operations) are more likely to be dictated by a handful of government officials, whereas a large number of people may independently engage in the most frequently covered Palestinian actions (such as riots); and the use of this ordinal scale as a continuous variable. Notwithstanding these caveats, this data set is a much more nuanced proxy for Israeli–Palestinian interaction than the data in other recent time series analyses. Figure 1 displays the daily event totals for Israeli and Palestinian actions toward one another, smoothed with polynomial regression.

We adopt the methods of much recent quantitative analyses of Israeli–Palestinian interactions, and the advice of methods texts in forecasting and time series analysis (Box, Jenkins, and Reinsel 2008; Lütkepohl 2005; Montgomery, Jennings, and Kulahci 2008), by examining these data with vector autoregression (VAR) models, which compute simultaneous equations for each directed dyad. The general form of this analysis, using notation conventions from Lütkepohl (2005), is

$$y_t = \nu + A_1 y_{t-1} + \dots + A_p y_{t-p} + b_t x_t + u_t, \\ t = 0, 1, 2, \dots, p,$$

where $y_t = (y_{1,t}, \dots, y_{K,t})'$ is a vector of K time series variables; A_1 through A_p are matrices of coefficients; t designates the time unit from 0 through p lags; $\nu = (\nu_1, \dots, \nu_K)'$ is a vector of K intercept terms; b is a vector of coefficients for exogenous variable(s) x ; and $u_t = (u_{1,t}, \dots, u_{K,t})'$ is a K -dimensional white noise or innovation process. The specific implementation in the present analysis involves simultaneous equations with $K = 2$ time series variables (Israeli actions toward Palestinians and Palestinian actions toward

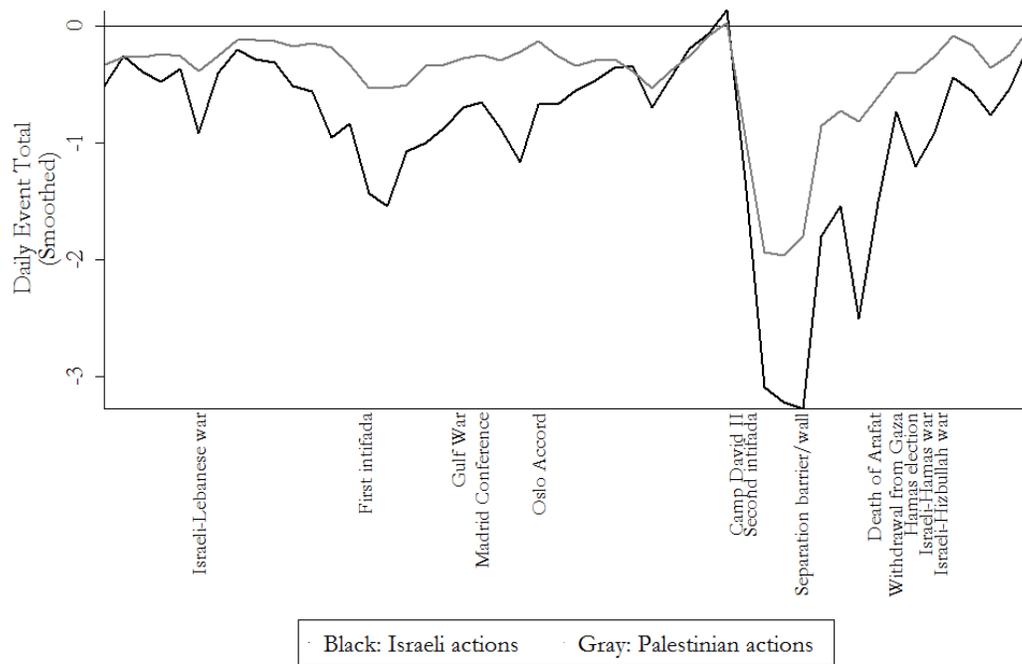


Figure 1: Israeli-Palestinian Interactions: Daily Events, 1979–2009. Daily event scores, smoothed with polynomial regression, reflect the deepening conflict during the First and Second Intifadas and the lessening of conflict during the period of the Oslo Accords.

Israelis—the section incorporating Lebanese actors involves $K = 6$ time series variables) and $p = 8$ days, incorporating 1-day through 8-day lags of both Israeli and Palestinian actions, based on the maximum optimal value of the Schwarz’s Bayesian Information Criterion (SBIC) (see the section on *duration shock* for a discussion of time-varying lag structures):

$$\begin{aligned}
 y_{1,t} &= \nu_1 + a_1 y_{1,t-1} + a_2 y_{2,t-1} + \dots + \\
 &\quad a_{15} y_{1,t-8} + a_{16} y_{2,t-8} + b_1 x_t + u_{1,t} \\
 y_{2,t} &= \nu_2 + a_{17} y_{1,t-1} + a_{18} y_{2,t-1} + \dots + \\
 &\quad a_{31} y_{1,t-8} + a_{32} y_{2,t-8} + b_2 x_t + u_{2,t},
 \end{aligned}$$

where y_1 represents Israeli actions toward Palestinians on a given day; y_2 represents Palestinian actions toward Israelis on that day; a_1 through a_{32} and b_1 through b_2 are coefficients; ν_1 and ν_2 are intercepts; x_t is an exogenous dummy variable for Saturdays; and u_1 and u_2 are white noise terms.

Granger causality tests confirm the significance of mutual effects of Israeli and Palestinian

actions at all lag structures from 1 to 81 days, the maximum that the Stata 12.0 software package allows. The results of Granger causality tests for selected lag structures are presented in Table 1; p -values less than 0.01 are evidence of the effect of past actions by one side on a given day’s actions by the other side. Nonstationarity of both time series—Israeli actions toward Palestinians and Palestinian actions toward Israelis—is ruled out with three versions of the augmented Dickey–Fuller unit root test (see also Haushofer et al. 2010, Table S2): a basic approach, a trend specification, and a drift specification (see Table 1). Negative Z -scores with p -values less than 0.01 are evidence for stationarity.

In-sample vector autoregression results with 1-day through 8-day lags are presented in Table 2. Of particular interest are the cells in the dark outlines, which indicate the effect of past Israeli actions on a given day’s Palestinian actions (upper right-hand cells) and vice versa (lower left-hand cells). The low p -values in these cells confirm the mutual impact of Israeli and Palestinian actions

Table 1: Mutual Effects of Israeli–Palestinian Actions

Lag structure	Test Statistic	Israeli actions toward Palestinians	Palestinian actions toward Israelis
Granger Causality Test			
1 through 8 days	χ^2	206.2*	170.1*
1 through 30 days	χ^2	199.4*	182.4*
1 through 80 days	χ^2	273.4*	299.0*
Augmented Dickey-Fuller Test			
Basic	Z	-77.4* (-3.4)	-84.4* (-3.4)
With trend term	Z	-78.0* (-4.0)	-84.8* (-4.0)
With nonzero drift	Z	-77.4* (-2.3)	-84.4* (-2.3)

Note: Based on 11,217 observations. 1% Critical Values for the Augmented Dickey-Fuller test are in parentheses.

* Denotes $(\text{Prob} > \chi^2) < .01$ for the Granger Causality Test and $p < .01$ for the Augmented Dickey-Fuller test.

on one another—six of eight lagged Israeli values are significant at the $p < 0.01$ level, and six of eight lagged Palestinian values are significant at this level.

A shift of 1 standard deviation in the Israeli or Palestinian daily scores generates small but statistically significant changes in the other's daily scores. The impulse-response function (IRF) in Figure 2 shows a long-lasting effect peaking at eight days after the shock, using in-sample analyses. The effect on Palestinian scores is half as large as for Israeli scores, just as the absolute scores of Palestinian daily actions are half as large as the absolute scores of Israeli daily actions. (The same pattern emerges with greater effects in models with orthogonalized IRFs.)

This model explains only a small portion of Israeli–Palestinian interaction over the past generation—17 percent of the variation in the in-sample Israeli daily scores and 12 percent of the Palestinian scores. However, the model performs as well as other recent time series analyses of Israeli–Palestinian interaction, whose median reported r^2 statistic is 12 percent (see Table A1). The model also performs well on the Theil inequality coefficient (TIC), which compares the root mean square errors (RMSE) of the in-sample predictions with the RMSE of a naive forecast of

“no change” in the event scores from one day to the next (Loungani 2001). According to this indicator, the model's prediction error is 23 percent lower than the naive forecast for Israeli scores and 25 percent lower for Palestinian scores (see Table 3).

Table 3 also presents additional robustness checks, including alternative data. If we abandon the ordinal scale of daily events and measure only a binary of fatal violence (Goldstein code of -10) and no fatal violence for each directed dyad for each day, the TIC performs just as well, beating the naive no-change forecast by 23 percent for Israeli actions and 25 percent for Palestinian actions. Parallel findings result from the net number of events of material cooperation and conflict for each directed dyad for each day, following the procedures of Brandt et al. (2011).

Similar findings emerge with a separate data source, a list of Israeli–Palestinian fatalities from the period September 29, 2000, through December 31, 2009, compiled by the peace group B'Tselem (<http://www.btselem.org/statistics>). Measured either as daily counts of fatalities for each directed dyad or as a binary of fatalities or no fatalities, the TIC for the B'Tselem data beats the naive no-change forecast by 25 to 28 percent (see Table 3). (Augmented Dickey–Fuller

Table 2: Vector Autoregression Results

	Israeli actions	Palestinian actions
Israeli actions:		
1-day lag	0.16* (0.01)	0.05* (0.01)
2-day lag	0.09* (0.01)	0.03* (0.01)
3-day lag	0.05* (0.01)	0.02* (0.01)
4-day lag	0.04* (0.01)	0.01 (0.01)
5-day lag	0.03* (0.01)	0.02* (0.01)
6-day lag	0.03* (0.01)	-0.003 (0.01)
7-day lag	0.05* (0.01)	0.02* (0.01)
8-day lag	0.03* (0.01)	0.02* (0.01)
Palestinian actions:		
1-day lag	0.10* (0.01)	0.12* (0.01)
2-day lag	0.08* (0.01)	0.08* (0.01)
3-day lag	0.05* (0.01)	0.05* (0.01)
4-day lag	0.05* (0.01)	0.05* (0.01)
5-day lag	0.05* (0.01)	0.03* (0.01)
6-day lag	0.03 (0.01)	0.02 (0.01)
7-day lag	0.003 (0.01)	0.03* (0.01)
8-day lag	0.06* (0.01)	0.03* (0.01)
Saturdays	0.11 (0.06)	0.07 (0.04)
Constant	-0.30* (0.03)	-0.11* (0.02)
R^2	0.17	0.12

Note: Based on 11,210 observations. Standard errors are in parentheses.

* Denotes $p < .01$.

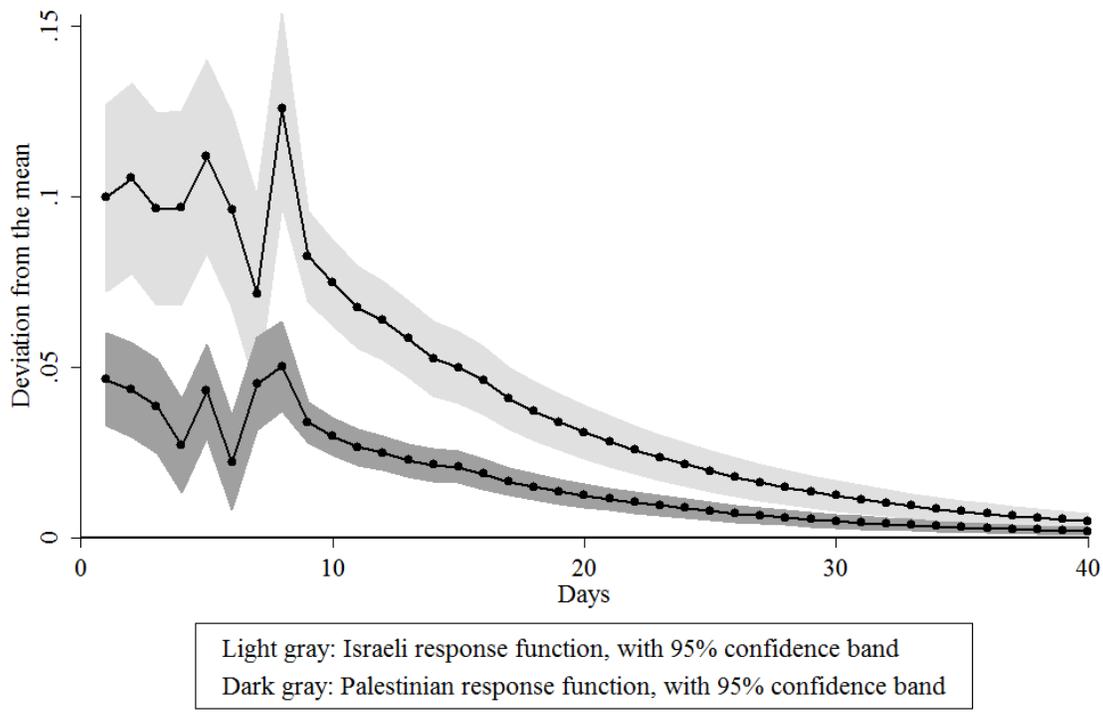


Figure 2: Impulse Response Functions for Israeli-Palestinian Interactions, 1979-2009. Israeli and Palestinian response functions reflect a statistically significant pattern of tit-for-tat interaction over the past generation.

tests indicate that all of these time series exhibit stationarity. The optimal SBIC for the binary measure of violence and the B'Tselem number of fatalities is 1-day through 10-day lags; the optimal SBIC for the binary measure of B'Tselem fatalities on a given day is 1-day through 6-day lags.)

All of the findings so far refer to in-sample analyses. However, the model is also robust to out-of-sample forecasting. Starting with a window of 90 days at the beginning of the time series, we use the coefficients from that period to forecast event scores on the ninety-first day; then we extend the window to 91 days and forecast event scores for the ninety-second day; and so on day by day to the final day in the data set. The root mean square of prediction errors for out-of-sample forecasts—using either the full event scale or the binary measure of violence—is virtually the same as with in-sample prediction for the Reuters events (see Table 3). The TIC statistics

are also almost identical for out-of-sample and in-sample predictions.

Similar findings result from Bayesian VAR models like the ones used in several recent papers (Brandt and Freeman 2006; Brandt et al. 2008; Brandt et al. 2011). As described in Appendix B, these models generate thousands of forecasts for each day for each directed dyad. Using the same extending-window approach as for non-Bayesian out-of-sample forecasts, the mean Bayesian out-of-sample forecasts generate root mean squared errors that are nearly identical—within 1 percent—of the root mean squared errors for non-Bayesian out-of-sample forecasts. Just more than three-quarters of the observed values for each day's directed dyad (76 percent for Israeli actions and 77 percent for Palestinian actions) fall within 1 standard deviation of the median Bayesian forecast—a 68 percent posterior probability interval used by Brandt et al. as an indicator of forecasting success.

Table 3: Robustness Checks

Dataset	Data	Forecasting method	Dependent variable			
			Israeli		Palestinian	
			RMSE	TIC	RMSE	TIC
Levant/Reuters	All events	In-sample	2.11	0.77	1.48	0.75
		Out-of-sample	2.12	0.78	1.49	0.76
	Binary of fatal violence	In-sample	0.13	0.77	0.15	0.75
		Out-of-sample	0.13	0.74	0.15	0.74
	Material coop./confl.	In-sample	0.74	0.76	0.47	0.74
		Out-of-sample	0.74	0.77	0.48	0.75
B'Tselem	Count of daily fatalities	In-sample	7.24 [†]	0.75	1.53	0.74
		Out-of-sample	8.17 [†]	0.84	1.65	0.79
	Binary of daily fatalities	In-sample	0.45	0.74	0.30	0.72
		Out-of-sample	0.46	0.76	0.30	0.74

[†] These figures are inflated by five days with particularly high numbers of Palestinian fatalities during the Israeli-Hamas War of 2008-2009. If those days' totals are capped at the next-highest number of daily fatalities (71 Palestinian fatalities on March 1, 2008), the RMSE drops by half.

This exercise confirms the reciprocal action-response dynamic of Israeli–Palestinian interactions over the past generation, using a more subtle daily indicator than in previous studies. Israelis and Palestinians have engaged in tit-for-tat interactions over the past 30 years, not just with regard to violence but also with regard to other forms of conflictual and cooperative actions as well.

Fluctuation in Prediction Error

We could stop here, as many time series analyses do, with the best models we are able to produce for the entire period under study. Hidden within our models, however, are dramatic swings in accuracy, which published time series analyses rarely mention. These models generate their worst prediction errors during the most important episodes in Israeli–Palestinian history, the interactions that we are most interested to explain.

As a threshold of historical importance, we take events during the period 1979–2006 that are mentioned in the chapter headings and subheadings of at least three out of seven recent English-language history books on Israeli–Palestinian interaction (Caplan 2010; Dowty 2012; Gelvin 2007;

Harms and Ferry 2008; Milton-Edwards 2009; Smith 2010; Tessler 2009). (Table A3 lists all events mentioned in the headings of at least two of the seven texts.) For eight of these 13 important events (two additional episodes cannot be dated precisely), prediction error as measured by RMSE is greater in the 90 days after the start date of the event than in the 90 days before the start date (shaded cells in Table 4). Prediction error increases during all of the six most important events, events that were headlined in five or more of the history books.

The fluctuation of prediction error is represented graphically in Figure 3, where the squared errors for out-of-sample predictions of each day's directed dyads have been smoothed with polynomial regression between each of the major historical events.⁵ Error is associated both with conflict and with cooperation—prediction error jumped after some of the major milestones in the peace process, including the Madrid Conference (especially for Israeli actions) and the first Oslo Accord (especially for Palestinian actions). However, the

⁵If the entire 30 years of prediction errors were smoothed as a single period, the line would appear to rise in the days leading up to historic moments such as the outbreak of the First Intifada, giving the misimpression that these events were anticipated by worsening prediction errors.

Table 4: Change in Prediction Error after Major Historical Events

Event	Start date	History textbooks	Change in prediction error	
			Israeli	Palestinian
Israeli-Lebanese war	June 6, 1982	5	+37%	+37%
Peace plans of the 1980s	No precise date	3		
First intifada	December 9, 1987	7	+147%	+102%
Gulf War	January 17, 1991	4	+1%	-45%
Madrid conference	October 30, 1991	5	+47%	+30%
First Oslo accord	September 13, 1993	7	+24%	+41%
Failure of Oslo accords	No precise date	3		
Camp David II	July 11, 2000	5	+34%	+111%
Second intifada	September 28, 2000	6	+212%	+259%
Separation barrier/wall	June 16, 2002	3	-22%	-17%
Death of Arafat	November 11, 2004	3	-15%	-14%
Withdrawal from Gaza	August 15, 2005	4	+53%	+58%
Hamas election	January 25, 2006	3	-23%	-30%
Israeli-Hamas war	June 28, 2006	3	+50%	+24%
Israeli-Hizbullah war	July 12, 2006	3	-16%	-25%

Note: Shaded cells indicate increases in the root mean squared error (RMSE) in the 90 days following the onset of a major historic event, compared with the prediction error in the preceding 90 days.

most dramatic increases in prediction error followed the outbreak of conflict, especially the First and Second Intifadas. These events were particularly poorly predicted by the interactions that preceded them, even if the forecast is based only on interactions directly preceding these events and ignores older interactions beyond a window of 30, 90, or 365 days. The pattern in Figure 3 is not simply an artifact of the scale of daily interaction scores—it emerges just as strongly with a binary indicator for violence, with similar leaps in squared prediction error after the First and Second Intifadas, the Madrid Conference (especially for Israeli actions), and the first Oslo Accord (especially for Palestinian actions).

These leaps in prediction error provide prima facie evidence of shocks in Israeli–Palestinian interactions. Our time series model passes the usual tests but is unable to forecast or retrocast historic episodes accurately on the basis of prior interactions.

These findings contradict the claims of previous time series analyses of Israeli–Palestinian interaction, which have not investigated time-varying errors in prediction. Like the bulk of time series analyses in general, these studies present models that best describe the time periods covered by their data. They imply, without present-

ing confirmation, that these models apply equally well to watershed moments and routine moments within each period. Studies that find tit-for-tat interactions draw conclusions about watershed episodes, not just routine ones. Studies that find no tit-for-tat interactions draw conclusions about routine moments, not just watersheds. We conclude that Israelis and Palestinians have been inconsistent in their responses to one another—their pattern of interaction appears stable during some parts of the past three decades but less stable during moments of historic change.

This pattern inverts the findings of Golan and Rosenblatt (2011), who concluded that predictable cycles of violence were visible only occasionally in Israeli–Palestinian interactions, “whereas in most periods, retaliation explains a minuscule portion of events, suggesting that the parties display no statistical regularity in their actions.” Using a longer time span (30 years instead of 8 years) and a richer set of interactions (not just killings and rocket attacks), we find that a consistent pattern of retaliation is less visible during periods of the greatest violence, while most periods exhibit greater statistical regularity. Next, we go further to identify four distinct aspects of interaction that may shift during moments of historic change.

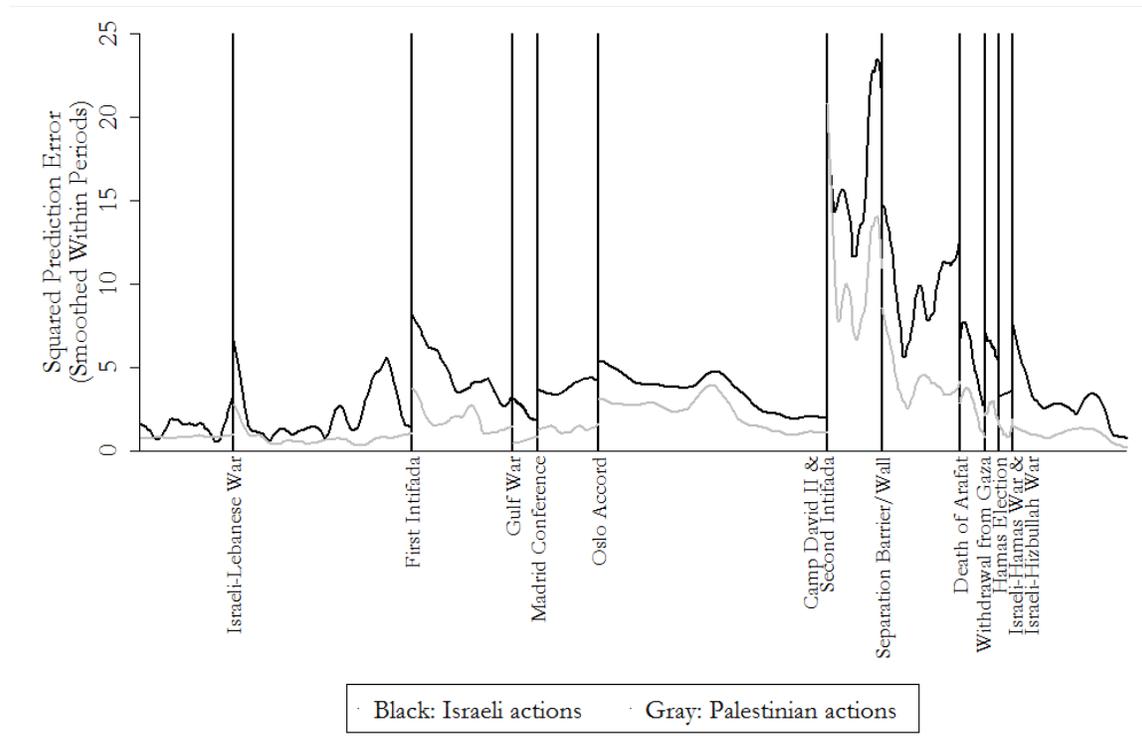


Figure 3: Israeli-Palestinian Interactions: Squared Prediction Error, 1979-2009. Out-of-sample prediction error leapt at many historic episodes in Israeli-Palestinian interaction. Note: Events within three months of each other are displayed in this chart as a single line: Camp David II and the Second Intifada in 2000; Israeli-Hamas War and Israeli-Hizbullah War in 2006.

Shock Absorbers

When faced with exogenous shocks, forecasters often attempt to identify and control for structural breaks, either by splitting the time series and calculating separate models for each period, or by introducing dummy variables corresponding to each period. These approaches face at least three challenges. First, on theoretical grounds, these approaches imply that historical trends ought to be understood net of their most important episodes, a point we will take up again in the concluding section of the article. Second, on empirical grounds, these approaches may not solve the problem of time-varying predictive accuracy. Third, on methodological grounds, these approaches fixate on structural breaks in coefficients while ignoring breaks in other parts of their models. Time series analysis of Israeli-Palestinian interaction illustrates all three of these challenges.

To examine whether fluctuation in prediction error is lower between structural breaks than across them, we chopped the time series of Israeli-Palestinian interaction into more than a dozen separate periods, based on the historic episodes identified by historians, and calculated separate vector autoregression models on each period. The resulting prediction error (Figure A4) rises and falls in a pattern almost identical to Figure 3, which forecasted interactions based on the entire 30-year sample. A similar pattern emerges with dummy variables for each period, using the first period as the reference category.

Instead of relying on historians' accounts to identify moments of epochal change, we might follow the inductive procedure proposed by time series texts such as Lütkepohl (2005) and search for the presence of structural breaks in the coefficients of the lagged dependent variables and the intercepts. We estimate the optimal number

of breaks, as indicated by the Bayesian Information Criterion, and the dates of these breaks, using econometric methods developed by Bai and Perron (2003) and implemented with the “breakpoints” function in the “strucchange” package in R (Zeileis et al. 2012a, 2012b). We set the minimum segment size at 90 days and limit the maximum number of breaks to 20 or fewer (results shown in Table A5, Model 1).

According to this inductive procedure, there are at least 20 structural breaks in the 30-year time series, but the optimal number of structural breaks in the time series is only 1. This break is estimated to occur on September 7, 2000, the same month that the Second Intifada began. Including a dummy variable for the period after this date scarcely improves the model, reducing prediction error by less than 2 percent (Table A5), and fails to smooth out the fluctuation in prediction error (Figure A6).

In empirical terms, then, forecasting methods designed to control for structural breaks may not control the problem we have identified: prediction errors remain highest at the moments of greatest historical importance. At the same time, these methods confirm the presence and substantive importance of breaks in Israeli–Palestinian interaction: coefficients for action–reaction models are significantly different before and after the break in September 2000. Before this time, impulse–response graphs display lesser and less consistent effects (Figure A7) than graphs calculated over the entire 30-year time series (Figure 2). (An alternative approach to structural breaks, using Chow tests on nonoverlapping segments of the time series, is described in Appendix C. The alternative approach identifies different break points but also fails to tame fluctuation in prediction error.)

Four Types of Shock

Structural breaks of the sort identified thus far, involving shifts in coefficients and intercepts, are the most common form of shock incorporated in forecasting analyses—indeed, much of time series analysis equates the concept of exogenous shock exclusively with this form of structural break. However, this is only one aspect of forecasting models that is vulnerable to shock (Bai 2010; Bai

and Perron 2003; Brandt et al. 2011; Castle, Fawcett, and Hendry 2011; Clements and Hendry 2006; Qu and Perron 2007). We contribute to this literature by documenting three additional types of shock, each type corresponding to a separate aspect of forecasters’ models:

- Effect shock: shifts in coefficients, or *how much* each action affects later actions
- Input shock: shifts in the values of endogenous variables, or *what* actions affect later actions
- Duration shock: shifts in the optimal lag, or *how long* each action affects later actions
- Actor shock: shifts in the number of endogenous variables, or *whose* actions affect later actions

The previous section of this article focused on what we call “effect shock”; this section presents examples of the other forms of shock and shows how recognition of these shocks, occurring at inconsistent intervals, complicates forecasters’ attempts to account for the past and predict the future.⁶

Input shock. We give the label “input shock” to shifts in actors’ behavior. In time series terms, the values of one or more endogenous variables break from their earlier pattern, as distinct from shifts in the coefficients for these variables. Forecasters rarely explore this form of shock.

We test for input shock over the full time series with the “strucchange” package in R, applied separately to the two univariate time series of Israeli actions toward Palestinians and Palestinian actions toward Israelis (Models 2 and 3 in Table A5). As with the joint model of interactions, this procedure finds at least 20 structural breaks in the 30-year time series both for Israeli actions and for Palestinian actions, but only 2 important breaks, according to the Bayesian Information Criterion. The first breaks are estimated to occur within days of each other in September 2000 (September 28 for Israeli actions and September 26 for Palestinian actions), at almost the same

⁶Technically, “duration shock” and “actor shock” are subsets of “effect shock,” because they can be expressed in terms of shifts in coefficients for particular lags or endogenous variables, but the following analysis suggests that these types of shock are substantively distinct.

time as the primary structural break in the joint model. The second breaks, however, differ for Israeli and Palestinian actions. Palestinian actions exhibit an input shock on August 3, 2002, around the start of construction of the Israeli separation barrier/wall, whereas Israeli actions exhibit an input shock more than two years later, on November 7, 2004, within days of Arafat's death.

These findings tell a complicated story that is not easily reconciled with the usual analyses of Israeli–Palestinian interactions or with the standard methods of time series analysis. The one consistent and unsurprising result is that Israeli and Palestinian actions shifted around the time of the Second Intifada in 2000, both in univariate and joint VAR models. However, we also find evidence of input shock without effect shock. Unexpectedly, input shock often appears at moments that are more commonly associated with actions on the other “side”—we see input shock in Israeli actions but not Palestinian actions after the death of Arafat, and input shock in Palestinian action but not Israeli actions after the construction of the separation barrier/wall. So far as we know, there is no readily available time series method to compare the importance of input shocks and effect shocks or to test for both forms of shock simultaneously. (The same problem arises with the alternative approach to breaks that is described in Appendix C.) The challenge for forecasting is that both types of shock may exist in the same time series data, sometimes coinciding at the same historic moment and sometimes not.

Duration shock. We give the label “duration shock” to shifts in how long an effect lasts. In time series models, duration shock is visible through changes in the optimal lag structure. To our knowledge, time series analysis has not developed criteria to identify multiple structural breaks in lags, but we will try to demonstrate the presence of duration shock in Israeli–Palestinian interactions using the three information criteria that are frequently used to determine optimal lag structures: the Akaike Information Criterion (AIC), the Hannan–Quinn Information Criterion (HQIC), and SBIC.

Beginning with a 365-day period at the beginning of the time series, we lengthened the window 90 days at a time and recalculated the optimal lag for the three information criteria. The values

of the optimal lags for each window are plotted in Figure 4. Each of the three indicators suggests that the optimal lag has risen as the time series lengthened—the longer the time series, the more long-lasting the effects of Israeli and Palestinian actions appear. Yet this rise was uneven, both across information criteria and across time. All three indicators suggest that lag structures lengthened during the Second Intifada in the 2000s. According to the AIC and SBIC, but not the HQIC, optimal lags lengthened during the First Intifada as well. In the mid-2000s, the AIC dropped, the HQIC rose, and the SBIC remained almost constant. Further inconsistencies ensue if we limit observations to rolling windows of one year or some other period. The sensitivity of these indicators to shifts in time period dampens hopes that lag structures will remain consistent for time series analysis of important historical episodes.

Shifts in optimal lag structure coincide with several of the historic events in Israeli–Palestinian relations identified by historians of the region: within 90 days after the outbreak of the Israeli–Lebanese War in 1982, the optimal lag indicated by the SBIC rose from 1 to 4; within 90 days after the outbreak of the Second Intifada in 2000, the optimal lag rose from 5 to 8. Other indicators of optimal lag structure also rose at this time or soon after. We may gauge the significance of these shifts with a Chow test that analyzes the 90 days before and 90 days after the outbreak of these two historic episodes. This test interacts the additional lagged dependent variables (the second, third, and fourth lags for the Israeli–Lebanese War; the sixth, seventh, and eighth lags for the Second Intifada) with a dummy variable equaling zero prior to the event and 1 afterward. The test is significant in both instances (chi-squared = 33.8 for the Israeli–Lebanese War and 49.5 for the Second Intifada, $p < 0.01$ for each), confirming that the lags experienced a structural break during these episodes. These tests are robust to different lengths of the periods before and after each outbreak.

Duration shock may involve substantive implications. If we examine only the period prior to the Second Intifada, for example, a five-day lag structure generates an IRF with smaller, shorter, and less statistically significant effects (shown in Figure A7) than for the full period of 1979 to 2009

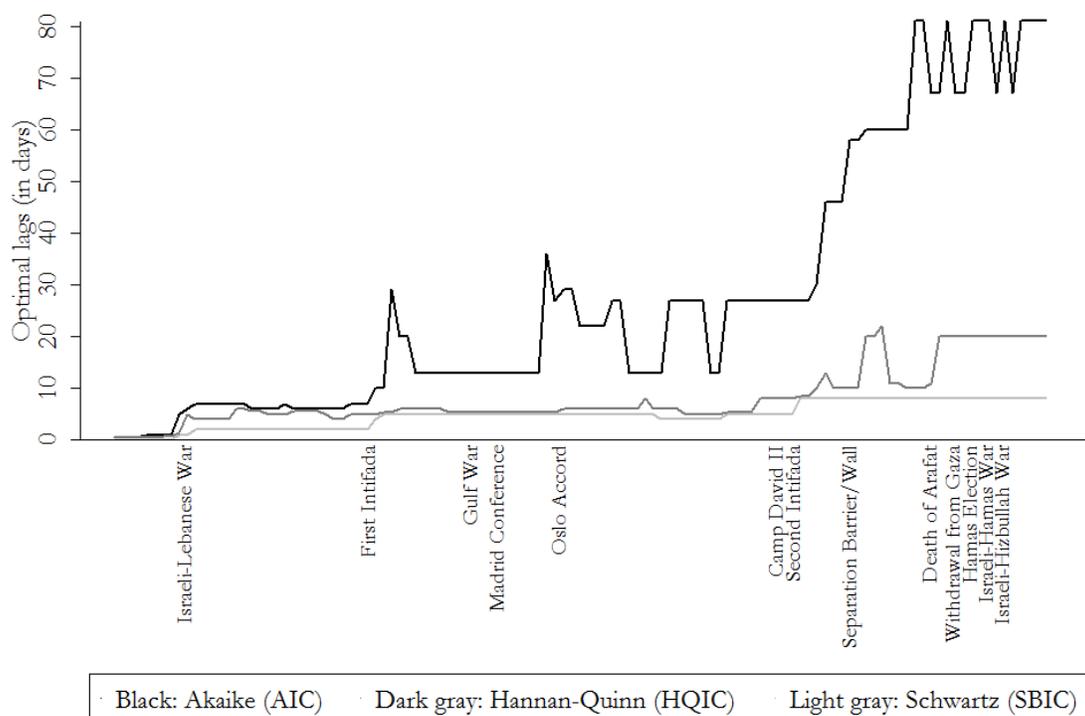


Figure 4: Duration Shock: Changes in Optimal Lag Structure, 1979-2009. Optimal lag structures fluctuate as the window of observations is extended over time.

(shown in Figure 2). Prior to the Second Intifada, in other words, Israeli responses appear less predictable than in the full time series; Palestinian responses, though lower in magnitude, appear to be more consistently predictable. The tit-for-tat pattern described in earlier time series studies of Israeli–Palestinian interactions only holds, it appears, for certain time periods and particular lag structures. Forecasting that relies on a single, unvarying lag structure may not be robust.

Actor shock. We give the label “actor shock” to shifts in the number of agents in the action-reaction system, for example, when an exogenous variable suddenly becomes endogenous, or vice versa. While time series analysis has long been concerned with omitted-variable bias, the concept of actor shock involves shifts in this bias over time.

With Israeli–Palestinian interactions, an example can be found in Lebanese actors—primarily Hizbullah but also the Lebanese government and other groups, as actors and targets of action in Reuters news stories. Directed dyads of Israeli–

Lebanese and Palestinian–Lebanese interaction are incorporated endogenously along with Israeli–Palestinian interactions in these analyses. For the full 30-year time series, the two primary tests for omitted variables contradict one another. The likelihood ratio (LR) test suggests that the inclusion of Lebanese actors improves the fit of the VAR model (chi-squared = 3,130.4, $p < 0.001$).⁷ Conversely, including Lebanese actors does not improve prediction errors—it slightly worsens RMSE, with an increase of less than 0.1 percent, for both Israeli actions toward Palestinians and Palestinian actions toward Israelis.

Hidden in these long-term findings are periods when the inclusion of Lebanese actors makes a

⁷The VAR model with endogenous Lebanese actors involves simultaneous equations for each of the following six dependent variables: Israeli actions toward Palestinians and vice versa; Israeli actions toward Lebanese and vice versa; and Palestinian actions toward Lebanese and vice versa. To compare fit, the reduced model requires the same six dependent variables, with the effects of lagged Lebanese actions and the effects of all parameters on Lebanese actions constrained to zero.

considerable difference in modeling Israeli–Palestinian interaction. Calculating RMSEs for the full and constrained models for every 90-day period in the time series, advancing day by day through all 30 years, we find moments in which the inclusion of Lebanese actors reduces prediction errors by more than 50 percent, and other periods when it increases errors by more than 20 percent (see Figure A8). The inclusion of Lebanese actors actually increases prediction errors in 62 percent of the windows. In other words, Lebanese actors are not uniformly exogenous or endogenous to Israeli–Palestinian interactions, and their inclusion or exclusion in models of Israeli–Palestinian interaction may matter a great deal—in different directions—at different points in time. Omitted-variable bias seems to be highly sensitive to the timing of the period of study.

As an illustration of actor shock, we examine prediction error during the Israeli–Hizbullah War of 2006, which lasted 59 days from July 12 through the end of the Israeli blockade on September 8, as well as the 59 days prior to the war and the 59 days following the war. The black bars in Figure 5 indicate the percentage change in RMSE for Israeli actions when Lebanese actors are included in the model; the gray bars indicate the percentage change in RMSE for Palestinian actions. During the war, including Lebanese actions reduces prediction error by 22 percent for Israeli actions and 40 percent for Palestinian actions. Just before the war, however, including Lebanese actions increases prediction error by 15 percent for Israeli actions and 39 percent for Palestinian actions. Just after the war, including Lebanese actions increases prediction error by 10 percent for Israeli action and reduces prediction error slightly, by 2 percent, for Palestinian actions. Over this six-month period, according to this standard measure of omitted-variable bias, Lebanese actors pop in and out of the model. (The RMSE analysis could not be confirmed with LR tests because nested VAR models would not run with these short time periods.) Nothing in the data just prior to the outbreak of the war would lead us to believe that Lebanese actors would soon become important for the forecasting of Israeli–Palestinian interaction. That is actor shock.

Several previous time series analyses of Israeli–Palestinian interaction incorporated a third party,

the United States (Rasler 2000; Goldstein et al. 2001; Brandt and Freeman 2006; Brandt et al. 2008). The first two of these studies found that interactions with America shifted in the early 1990s just as the Israeli–Palestinian peace process progressed. However, these studies do not explore the implications of discontinuity, and subsequent time series analyses have not followed suit. Our findings confirm that actor shock may be a crucial aspect of Israeli–Palestinian interaction at certain important junctures, illustrating how forecasting may have to contend with the sudden entry and exit of new endogenous factors.

Implications

Late in life, Auguste Comte wrote to an Ottoman statesman with a prediction (Comte [1853] 1876:xlili). Large states like the Ottoman Empire were doomed, he suggested, through the “normal application of the sociological law which everywhere restricts the territory of temporal dominions to its natural size.” Only “homogeneous” states, Comte concluded, would avoid “spontaneous dismemberment.”

Comte’s prediction was in general accurate: The Ottoman Empire was dismembered into successor states with more “homogeneous” populations. But this generally accurate forecast overlooked an important variety of shocks that Comte did not anticipate. Some dismemberment took the form of nationalist uprisings similar to the movements in Greece and Serbia on which Comte based his prediction. Elsewhere, however, nationalism emerged only after the dismemberment of empire, as in much of the Levant and the Arabian peninsula. Some of the new states, such as Greece and Turkey, achieved homogeneous populations only through ethnic cleansing. In the British Mandate of Palestine—the territory that is the subject of this article—dueling Israeli and Palestinian nationalisms cast doubt on Comte’s confidence in the existence of a “natural size” for each state. Comte’s forecast may seem prescient from a bird’s-eye view, but it looks more dubious when we examine the particulars.

The forecasting profession has developed far more sophisticated methods since Comte’s time. However, these methods continue to struggle with the problem of unanticipated shocks. This arti-

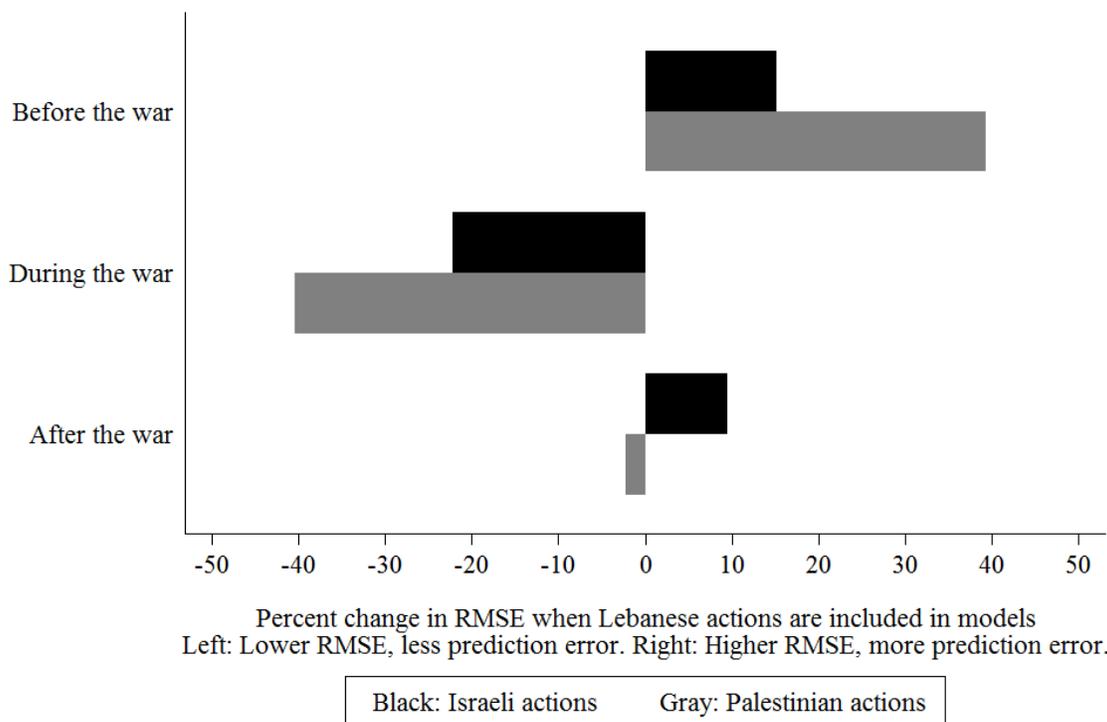


Figure 5: Actor Shock: Prediction Errors With a New Actor. Including Lebanese events reduces prediction error considerably during the 2006 Israeli-Hizbullah war, but not just before and after the war.

cle offers a statistical model that is adequately successful at predicting Israeli and Palestinian interactions over the past generation, according to the usual standards of the forecasting profession. On closer inspection, however, these models falter during the most important episodes of historic change. This article identifies four ways in which shocks have disrupted patterns of Israeli-Palestinian interaction at watershed moments, contributing to recent methodological work on the mechanisms of structural breaks: shifts in the duration of effects (duration shock), shifts in the scale of effects (effect shock), shifts in actions (input shock), and shifts in the number of agents in the system (actor shock). This typology is not intended to be exhaustive but to encourage research into additional forms of shock and more sophisticated tests for the presence of shocks of different sorts.

The identification of different types of shocks allows us to specify several unexpected aspects

of Israeli-Palestinian interactions, such as short-term input shock affecting Israeli actions, but not Palestinian actions, during the First Intifada, suggesting that the statistically significant change in behavior during that historic period did not occur on the Palestinian side but rather on the Israeli side. The Second Intifada, by contrast, involved input shock on both sides as well as effect shock scrambling the lag coefficients. These coefficients underwent a short-term effect shock several years later, with the death of Yassir Arafat. Israeli actions also shifted at that time, whereas Palestinian actions did not. Palestinian actions shifted with the construction of the separation barrier/wall in 2002, whereas Israeli actions did not. The duration of interaction effects increased abruptly at important moments, and at least one third party entered and exited Israeli-Palestinian interactions (the Lebanese in 2006).

Israeli and Palestinian leaders have frequently debated the subject of shocks as they struggled

to understand whether old patterns of interaction still held or new ones were forming. In the early days of the First Intifada in 1987, for example, the Israeli defense minister reassured a colleague that “the army will assert control very quickly,” while a foreign ministry official drew the opposite conclusion that “this was the beginning of something big.” Palestinian leaders were also disconcerted by the uprising. The head of intelligence for the Palestinian Liberation Organization recalled that “when the intifada broke out, we were at first afraid” that it would amount to nothing, because “nobody had been calculating on such an intifada, with its force and power. The one who was most in touch with the occupied territories was Abu Jihad [Khalil al-Wazir], but even he didn’t expect it to be like that” (Gowers and Walker 2003:358).

This article encourages the acknowledgment of surprise in the forecasting profession, as a check on the field’s eternal optimism. As a matter of professional “best practices,” forecasters should test for and report the presence of shocks of various sorts. Indeed, this practice should be adopted not just by forecasters but by retrocasters as well. The four types of shock identified in this article may apply not just to vector autoregression models but to other approaches to time series analysis as well, including models that focus on larger sets of variables with a smaller number of time points. Perhaps an extremely complicated model might account for syncopated shocks of different forms. That is the goal of time series methodology—“dealing with structural breaks” (Perron 2006) so that they will no longer obscure the underlying dynamics of the time series. Such a model would describe the process of Israeli–Palestinian interaction, net of the surprising and important events that have shaped this history over the past generation. This article makes the case, by contrast, that “dealing with” shock is not just methodologically difficult but substantively unsound. Shocks are not just disturbances to be controlled for statistically, in search of a stable underlying set of parameters. They are significant in their own right and a crucial feature of Israeli–Palestinian interaction over the past generation. Accounting for one form of shock may not help in anticipating the next one. Shocks can undermine forecasting accuracy at historic moments, without warning.

Shocks get to the heart of social-scientific explanation: the extent to which the past may be said to anticipate the future. Whether the analysis is conducted retrospectively with in-sample forecasts or prospectively with out-of-sample forecasts, whether it is grounded in quantitative or qualitative evidence, the intervention of shock may detach future from past. Shocks draw attention to discontinuities in history, while time series and other historical analyses are normally attuned to continuity. Explanations work best when the dynamics of cause and effect remain stable, but moments of historic change, when these dynamics shift, are less readily attributed to the dynamics that have gone before. When forecasts and retrocasts fail, they set bounds on social-scientific explanation.

References

- Abbott, Andrew. 1997. “On the Concept of Turning Point.” *Comparative Social Research* 17:85–105.
- Aldridge, Alan. 1999. “Prediction in Sociology: Prospects for a Devalued Activity.” *Sociological Research Online* 4, <http://www.socresonline.org.uk/4/3/aldridge.html>. <http://dx.doi.org/10.5153/sro.333>
- Bai, Jushan. 2010. “Common Breaks in Means and Variances for Panel Data.” *Journal of Econometrics* 157:78–92. <http://dx.doi.org/10.1016/j.jeconom.2009.10.020>
- Bai, Jushan and Pierre Perron. 2003. “Computation and Analysis of Multiple Structural Change Models.” *Journal of Applied Econometrics* 18:1–22. <http://dx.doi.org/10.1002/jae.659>
- Baumgartner, Frank R. and Bryan D. Jones. 2009. *Agendas and Instability in American Politics*. 2nd ed. Chicago: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226039534.001.0001>
- Beasley, Maya. 2008. “The Impact of Counterterrorism and Strain on Palestinian Terrorism.” Pp. 331–48 in *Armed Forces and Conflict Resolution: Sociological Perspectives*, edited by G. Caforio, G. Kümmel, and B. Purkayastha. Bradford, UK: Emerald.

- Biger, Gideon. 2004. *The Boundaries of Modern Palestine, 1840–1947*. London: Routledge.
- Box, George E. P., Gwilym M. Jenkins, and Gregory C. Reinsel. 2008. *Time Series Analysis: Forecasting and Control*. 4th ed. Hoboken, NJ: John Wiley. <http://dx.doi.org/10.1002/9781118619193>
- Braithwaite, Alex, Dennis M. Foster, and David A. Sobek. 2010. "Ballots, Bargains, and Bombs: Terrorist Targeting of Spoiler Opportunities." *International Interaction* 36:294–305. <http://dx.doi.org/10.1080/03050629.2010.502453>
- Brandt, Patrick T. 2012. "Markov-Switching, Bayesian, Vector Autoregression Models: Package 'MSBVAR.'" Version 0.7-2. Vienna, Austria: Comprehensive R Archive Network.
- Brandt, Patrick T. and John R. Freeman. 2006. "Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting, and Policy Analysis." *Political Analysis* 14:1–36. <http://dx.doi.org/10.1093/pan/mpi035>
- Brandt, Patrick T., Michael Colaresi, and John R. Freeman. 2008. "The Dynamics of Reciprocity, Accountability, and Credibility." *Journal of Conflict Resolution* 52:343–74. <http://dx.doi.org/10.1177/0022002708314221>
- Brandt, Patrick T., John R. Freeman, and Philip A. Schrodtt. 2011. "Real Time, Time Series Forecasting of Political Conflict." *Conflict Management and Peace Science* 28:41–64. <http://dx.doi.org/10.1177/0738894210388125>
- Brym, Robert J. and Robert Andersen. 2011. "Rational Choice and the Political Bases of Changing Israeli Counterinsurgency Strategy." *British Journal of Sociology* 62:482–503. <http://dx.doi.org/10.1111/j.1468-4446.2011.01375.x>
- Brym, Robert J. and Bader Araj. 2006. "Suicide Bombing as Strategy and Interaction: The Case of the Second Intifada." *Social Forces* 84:1965–82. <http://dx.doi.org/10.1353/sof.2006.0081>
- Bueno de Mesquita, Bruce. 2009. *The Predictioneer's Game: Using the Logic of Brazen Self-Interest to See and Shape the Future*. New York: Random House.
- Caplan, Neil. 2010. *The Israel–Palestine Conflict*. Chichester, UK: Wiley-Blackwell.
- Castle, Jennifer L., Nicholas W. P. Fawcett, and David F. Hendry. 2011. "Forecasting Breaks and During Breaks." Pp. 315–54 in *Oxford Handbook of Economic Forecasting*, edited by Michael P. Clements and David F. Hendry. Oxford: Oxford University Press.
- Cerulo, Karen A. 2006. *Never Saw It Coming: Cultural Challenges to Envisioning the Worst*. Chicago: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226100296.001.0001>
- Chow, Gregory C. 1960. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28:591–605. <http://dx.doi.org/10.2307/1910133>
- Clements, Michael P. and David F. Hendry. 2006. "Forecasting with Breaks in Data Processes." Pp. 605–57 in *Handbook of Economic Forecasting*, edited by Graham Elliott, Clive W. J. Granger, and Allan Timmermann. Amsterdam: Elsevier.
- Collier, Ruth Berins and David Collier. 1991. *Shaping the Political Arena: Critical Junctures, the Labour Movement, and Regime Dynamics in Latin America*. Princeton, NJ: Princeton University Press.
- Comte, Auguste. (1851) 1875. *System of Positive Polity*. Vol. 1. London: Longman, Greens.
- Comte, Auguste. (1853) 1876. *System of Positive Polity*. Vol. 3. London: Longman, Greens.
- Cooper, Richard N. and Richard Layard, eds. 2002. *What the Future Holds: Insights from Social Science*. Cambridge, MA: MIT Press.
- Cowles, Alfred, III. 1933. "Can Stock Market Forecasters Forecast?" *Econometrica* 1:309–24. <http://dx.doi.org/10.2307/1907042>
- Dowty, Alan. 2012. *Israel/Palestine*. 3rd ed. Cambridge: Polity.
- Dugan, Laura and Erica Chenoweth. 2012. "Moving beyond Deterrence: The Effectiveness of Raising the Expected Utility of Abstaining from Terrorism in Israel." *American Sociological Review* 77:597–624. <http://dx.doi.org/10.1177/0003122412450573>

- Friedman, Walter A. 2014. *Fortune Tellers: The Story of America's First Economic Forecasters*. Princeton, NJ: Princeton University Press.
- Gardner, Dan. 2011. *Future Babble: Why Expert Predictions Are Next to Worthless, and You Can Do Better*. New York: Dutton.
- Gelvin, James L. 2007. *The Israel–Palestine Conflict: One Hundred Years of War*. 2nd ed. Cambridge: Cambridge University Press.
- Golan, David and Jonathan D. Rosenblatt. 2011. “Revisiting the Statistical Analysis of the Israeli–Palestinian Conflict.” *PNAS* 108:E53–E54. <http://dx.doi.org/10.1073/pnas.1016378108>
- Goldstein, Joshua S. 1992. “A Conflict-Cooperation Scale for WEIS Events Data.” *Journal of Conflict Resolution* 36:369–85. <http://dx.doi.org/10.1177/0022002792036002007>
- Goldstein, Joshua S., Jon C. Pevehouse, Deborah J. Gerner, and Shibley Telhami. 2001. “Reciprocity, Triangularity, and Cooperation in the Middle East, 1979–97.” *Journal of Conflict Resolution* 45:594–620. <http://dx.doi.org/10.1177/0022002701045005003>
- Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2010. “A Global Model for Forecasting Political Instability.” *American Journal of Political Science* 54:190–208. <http://dx.doi.org/10.1111/j.1540-5907.2009.00426.x>
- Gowers, Andrew and Tony Walker. 2003. *Arafat: The Biography*. Revised ed. London: Virgin.
- Gunning, Jeroen. 2008. *Hamas in Politics: Democracy, Religion, Violence*. New York: Columbia University Press.
- Hafez, Mohammed M. and Joseph H. Hatfield. 2006. “Do Targeted Assassinations Work? A Multivariate Analysis of Israel’s Controversial Tactic during the Al-Aqsa Uprising.” *Studies in Conflict and Terrorism* 29:359–82. <http://dx.doi.org/10.1080/10576100600641972>
- Hansen, Bruce E. 2001. “The New Econometrics of Structural Change: Dating Breaks in US Labor Productivity.” *Journal of Economic Perspectives* 15:117–28. <http://dx.doi.org/10.1257/jep.15.4.117>
- Harcourt, Bernard E. 2007. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.
- Harms, Gregory and Todd M. Ferry. 2008. *The Palestine–Israel Conflict: A Basic Introduction*. London: Pluto Press.
- Haushofer, Johannes, Anat Biletzki, and Nancy Kanwisher. 2010. “Both Sides Retaliate in the Israeli–Palestinian Conflict.” *PNAS* 107:17927–32. <http://dx.doi.org/10.1073/pnas.1012115107>
- Haushofer, Johannes, Anat Biletzki, and Nancy Kanwisher. 2011. “Reply to Golan and Rosenblatt: Revisiting the Statistical Analysis of the Israeli–Palestinian Conflict Confirms that Both Sides Retaliate.” *PNAS* 108:E55–E56. <http://dx.doi.org/10.1073/pnas.1019420108>
- Jaeger, David A. and M. Daniele Paserman. 2006. “Israel, the Palestinian Factions, and the Cycle of Violence.” *American Economic Review* 96:45–49. <http://dx.doi.org/10.1257/000282806777212008>
- Jaeger, David A. and M. Daniele Paserman. 2008. “The Cycle of Violence? An Empirical Analysis of Fatalities in the Palestinian–Israeli Conflict.” *American Economic Review* 98:1591–1604. <http://dx.doi.org/10.1257/aer.98.4.1591>
- Jaeger, David A. and M. Daniele Paserman. 2009. “The Shape of Things to Come? On the Dynamics of Suicide Attacks and Targeted Killings.” *Quarterly Journal of Political Science* 2009:315–42. <http://dx.doi.org/10.1561/100.00009013>
- Kaplan, Edward H., Alex Mintz, Shaul Mishal, and Claudio Samban. 2005. “What Happened to Suicide Bombings in Israel? Insights from a Terror Stock Model.” *Studies in Conflict and Terrorism* 28:225–35. <http://dx.doi.org/10.1080/10576100590928115>
- Kellert, Stephen H. 2008. *Borrowed Knowledge: Chaos Theory and the Challenge of Learning across Disciplines*. Chicago: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226429809.001.0001>

- Khalidi, Rashid. 1997. *Palestinian Identity: The Construction of Modern National Consciousness*. New York: Columbia University Press.
- Kurzman, Charles. 2004. *The Unthinkable Revolution in Iran*. Cambridge, MA: Harvard University Press.
- Land, Kenneth C. and Stephen H. Schneider, eds. 1987. *Forecasting in the Social and Natural Sciences*. Dordrecht, Netherlands: Reidel. <http://dx.doi.org/10.1007/978-94-009-4011-6>
- Loungani, Prakash. 2001. "How Accurate Are Private Sector Forecasts? Cross-Country Evidence from Consensus Forecasts of Output Growth." *International Journal of Forecasting* 17:419–32. [http://dx.doi.org/10.1016/S0169-2070\(01\)00098-X](http://dx.doi.org/10.1016/S0169-2070(01)00098-X)
- Lütkepohl, Helmut. 2005. *New Introduction to Multiple Time Series Analysis*. Berlin: Springer. <http://dx.doi.org/10.1007/978-3-540-27752-1>
- Maoz, Zeev. 2007. "Evaluating Israel's Strategy of Low-Intensity Warfare, 1949–2006." *Security Studies* 16:319–49. <http://dx.doi.org/10.1080/09636410701547782>
- Milton-Edwards, Beverley. 2009. *The Israeli–Palestinian Conflict: A People's War*. London: Routledge.
- Milton-Edwards, Beverley and Stephen Farrell. 2010. *Hamas: The Islamic Resistance Movement*. Cambridge: Polity.
- Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulahci. 2008. *Introduction to Time Series Analysis and Forecasting*. Hoboken, NJ: Wiley-Interscience.
- Moore, Wilbert. 1964. "Predicting Discontinuities in Social Change." *American Sociological Review* 29:331–38. <http://dx.doi.org/10.2307/2091478>
- Nafi, Basheer M. 1998. *Arabism, Islamism and the Palestine Question, 1908–1914*. Reading, UK: Ithaca Press.
- Perron, Pierre. 2006. "Dealing with Structural Breaks." Pp. 278–352 in *Palgrave Handbook of Econometrics*, Vol. 1, *Econometric Theory*, edited by Terence C. Mills and Kerry Patterson. Basingstoke, UK: Palgrave Macmillan.
- Qu, Zhongjun and Pierre Perron. 2007. "Estimating and Testing Multiple Structural Changes in Multivariate Regressions." *Econometrica* 75:459–502. <http://dx.doi.org/10.1111/j.1468-0262.2006.00754.x>
- Rasler, Karen. 2000. "Shocks, Expectancy Revision, and the De-escalation of Protracted Conflicts: The Israeli–Palestinian Case." *Journal of Peace Research* 37:699–720. <http://dx.doi.org/10.1177/0022343300037006003>
- Schneider, Gerald, Nils Petter Gleditsch, and Sabine C. Carey. 2010. "Exploring the Past, Anticipating the Future: A Symposium." *International Studies Quarterly* 12:1–7. <http://dx.doi.org/10.1111/j.1468-2486.2009.00909.x>
- Schrodt, Philip A. 2007. "CAMEO Scale: Version 0.5B1 [07.03.21]." University Park: Pennsylvania State University, Event Data Project.
- Schrodt, Philip A. 2012. "CAMEO, Conflict and Mediation Event Observations: Event and Actor Codebook." Version 1.1b3. University Park: Pennsylvania State University, Event Data Project.
- Schrodt, Philip A. and Deborah J. Gerner. 1994. "Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982–1992." *American Journal of Political Science* 38:825–54. <http://dx.doi.org/10.2307/2111609>
- Schrodt, Philip A. and Deborah J. Gerner. 1997. "Empirical Indicators of Crisis Phase in the Middle East, 1979–1995." *Journal of Conflict Resolution* 41:529–55. <http://dx.doi.org/10.1177/0022002797041004003>
- Schrodt, Philip A. and Deborah J. Gerner. 2000. "Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant." *American Political Science Review* 94:803–17. <http://dx.doi.org/10.2307/2586209>
- Schuessler, Karl F. 1968. "Prediction." Pp. 418–25 in *International Encyclopedia of the Social Sciences*, edited by David L. Sills. Vol. 12. New York: Macmillan.
- Shamir, Jacob and Khalil Shikaki. 2010. *Palestinian and Israeli Public Opinion*. Bloomington: Indiana University Press.
- Shapiro, Ian and Sonu Bedi, eds. 2007. *Political Contingency: Studying the Unexpected, the Accidental, and the Unforeseen*. New York: New York University Press.

- Shelef, Nadav. 2010. *Evolving Nationalism: Homeland, Religion and Identity in Israel*. Ithaca, NY: Cornell University Press.
- Sherden, William A. 1997. *The Fortune Sellers: The Big Business of Buying and Selling Predictions*. Hoboken, NJ: John Wiley.
- Silver, Nate. 2012. *The Signal and the Noise: Why Most Predictions Fail—but Some Don't*. New York: Penguin.
- Smith, Charles D. 2010. *Palestine and the Arab-Israeli Conflict*. 7th ed. Basingstoke, UK: Palgrave Macmillan.
- Smith, Peter. 1998. *Explaining Chaos*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CB09780511554544>
- Sornette, Didier. 2002. "Predictability of Catastrophic Events: Material Rupture, Earthquakes, Turbulence, Financial Crashes, and Human Birth." *PNAS* 99:2522–29. <http://dx.doi.org/10.1073/pnas.022581999>
- Talib, Nissim Nicholas. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.
- Taub, Gadi. 2010. *The Settlers and the Struggle over the Meaning of Zionism*. New Haven, CT: Yale University Press.
- Tessler, Mark A. 2009. *A History of the Israeli-Palestinian Conflict*. 2nd ed. Bloomington: Indiana University Press.
- Tetlock, Philip E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.
- Tikuissis, Peter, David Carment, and Yiagadeesen Samy. 2012. "Prediction of Intrastate Conflict Using State Structural Factors and Events Data." *Journal of Conflict Resolution* 57:410–44. <http://dx.doi.org/10.1177/0022002712446129>
- Vital, David. 1975. *The Origins of Zionism*. Oxford, UK: Clarendon Press.
- Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "The Perils of Policy by p-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47:363–75. <http://dx.doi.org/10.1177/0022343309356491>
- York, Richard and Brett Clark. 2007. "The Problem with Prediction: Contingency, Emergence, and the Reification of Projections." *Sociological Quarterly* 48:713–43. <http://dx.doi.org/10.1111/j.1533-8525.2007.00098.x>
- Zeileis, Achim, Friedrich Leisch, Kurt Hornik, and Christian Kleiber. 2012a. "Package 'strucchange.'" Vienna: R Foundation for Statistical Computing.
- Zeileis, Achim, Friedrich Leisch, Kurt Hornik, and Christian Kleiber. 2012b. "strucchange: An R Package for Testing for Structural Change in Linear Regression Models." Vienna: R Foundation for Statistical Computing.
- Zeitoff, Thomas. 2011. "Using Social Media to Measure Conflict Dynamics: An Application to the 2008–2009 Gaza Conflict." *Journal of Conflict Resolution* 55:938–69. <http://dx.doi.org/10.1177/0022002711408014>

Acknowledgements: This project was funded by the National Science Foundation's program in Human and Social Dynamics (grant 0729405). We thank coprincipal investigators Mark Crescenzi and Robert Jenkins and research assistants Rebecca Best, Christine Carpino, Russell Grote, Ali Kadivar, Zafar Nazarov, Katrina Nowak, and Reed Wood, as well as colleagues Deborah Barrett, Lisa Feldman Barrett, and Philip Schrod, none of whom should be held responsible for the awkward findings that their assistance helped to generate. Direct correspondence to Charles Kurzman.

Charles Kurzman: Department of Sociology, University of North Carolina at Chapel Hill. E-mail: kurzman@unc.edu.

Aseem Hasnain: Department of Sociology, University of North Carolina at Chapel Hill. E-mail: ahasnain@unc.edu.