From Predictive to Causal Inference: The Use of Multivariate Generalized Autoregressive Conditional Heteroskedasticity-Derived Time Series of Covariance in the Study of Political Economic Systems

By

Nicholas Adams Judge

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Political Science)

at the

UNIVERSITY OF WISCONSIN–MADISON

2014

Date of final oral examination and acceptance without revision: 10/18/2013

The dissertation is approved by the following members of the Final Oral Committee:
Charles Franklin, Emeritus Professor, Political Science; Professor,
Law and Public Policy, Marquette University
Ken Mayer, Professor, Political Science
Janet Box-Steffensmeier, Professor, Political Science, Ohio State University
Alex Tahk, Assistant Professor, Political Science
Jon Pevehouse, Professor, Political Science

Contents

Chapter 1

An Introduction and Literature Review

1.1 Introduction

A profound growth in data has begun to change how questions throughout the social sciences are investigated. The origins of this change are mostly technological progress and concomitant reductions in the cost of data generation. This will only continue, and its effects on the social sciences will grow more profound.

There is a disjunction between the pace of technological change and the evolution of academic and intellectual tradition in all the sciences, and the empirical social sciences are no exception. Changes need to be made to align investigative patterns with the new data environment – never mind the data environment of ten or twenty years from now. To be clear, statistical methodology seems to be keeping pace with changes in the research environment, as thousands of social and ”hard” scientists, mathematicians and statisticians wait to pounce on the next uptick in computational power. *Research* methodology, however, has not moved as quickly at least, not in many fields.

How questions are investigated and the empirical forms debate takes seems to have been more refined than altered by the new data environment. That is a statement specific to political science and a matter of opinion, but many would agree that most literatures have moved down pre-existing veins of research more than they have been reworked to maximize their utility and cumulativeness for a new research environment.

In some areas, the new data environment has taken the form of supplementing the cross-sectional with time series. Literatures that once saw only sporadic generation of data now enjoy data generation on a weekly and even daily basis. To give some examples from American politics, public opinion is now surveyed either every day or quite nearly every day during American presidential campaigns. Federal Election Commission (FEC) data that used to be useful only as cross-sectional data now can be used as a rich set of time series. Media attention can be tracked by the hour on Google. Even social commentary, in the form of social media, can be tracked as time series.

This enhancement of the cross-sectional with the dynamic is incredibly promising for the social sciences. Variance that once had to effectively be treated as contemporaneous can now be put in a dynamic context. That's the difference between investigating the cause of a divorce knowing that both partners cheated on one another, and knowing that one partner broke their vows long before the other. It is far from a complete information set, but it is pretty darn important.

Furthermore, relationships that once had to be subjected to an assumption of time-invariance can now be allowed to vary over time in a realistic manner, thanks to models like the Dynamic Conditional Correlation (DCC) model (Engle 2002).[1] To characterize the importance of this, consider the growth in importance of semi and nonparametric techniques over the past two decades. Few would dispute that setting aside assumptions of global functional relationships has been a major source of development throughout the sciences. Nonparametric relationships are, often, linear relationships that are allowed to vary over the values of the variables in the model. DCC-derived relationships are, often, linear relationships that are allowed to vary over the values

[1]The DCC is one particularly convenient variant of multivariate Generalized Auto Regressive Conditional Heteroskedasticity (GARCH or MGARCH) models. The development of this class of models has spawned one of the largest literatures in econometrics.

of a particularly important variable, t, time.

Nonparametric methods are employed across the social sciences, while models like the DCC appear only very rarely outside their home field of economics. Only one article could be found employing it in political science (Lebo and Box-Steffensmeier, 2008), though several recent articles mentioned it as a possibility for future research. This to some extent must be a holdover from the previous rarity of time series data outside of market and macroeconomic contexts.

To rephrase the above, the dearth of time series in the non-economic social sciences is less and less ubiquitous. It is past time for the social sciences to adapt. This dissertation is part of that adaptation. One of its motivating questions is: *What would social scientific investigation look like if time series data and the methods that allowed for time- varying relationships had always been part of the debate?*

The answer that is given here: The over-time variance of the relationship between variables would be used to complete inferential tasks almost as regularly as the modeling of the variables themselves. That is a simple answer but, it is argued, its ramifications for the research process are not self-evident. After an introduction to the econometric developments that brought the social sciences to this point, Chapter 2 explores some of the ways the above answer can operationalize itself as a research methodology. There are two broad categories to that, reflecting the blurry bifurcation of the research process into data exploration and descriptive inference on the one hand, and the modeling process and causal inference on the other.

The first category involves simply observing how the relationship varies over time and coming to the conclusions that are rendered obvious by those observations. Because in political science most research problems have not enjoyed ample time series data, and because of assumptions of relationship time-invariance, large literatures have been built up over the years that can benefit from finally, simply observing relationships as they change over time. In fact, let's give that literary pattern its own name: A *dynamic bottleneck* has been built up in many literatures over the past decades, as shortages of data and inattention to the dynamics of the relationships between variables[2] have created difficult questions that would have been easy to answer without those problems.

Sometimes, empirical questions that are complex with a dynamic bottleneck present are relatively simple once the bottleneck is removed. Further, because very few scholars are working on removing these dynamic bottlenecks, it is likely that the most useful inferences that can be made at this stage come from the first category. That has been the experience of this dissertation. Once the hard work has been done to form time series of statistical relationships in this case time series of correlation matrices the argumentation and process of inference are actually quite simple. That's a strength of this methodological framework, not a weakness. It also means that the statistical workload is very much front-loaded, allowing readers who lack some statistical background to still enjoy the substantive findings, once they skip over a chapter or two.

The second category of methods explored in Chapter 2 are methods that enhance the traditional, modeling-based process of causal inference with dynamically conditioned statistical relationships. These relationships themselves can be used as (independent or dependent) latent variables in subsequent modeling. That is useful for several reasons. (1) The relationship e.g., a correlation at each time textitt itself may be the better operationalization of an underlying concept. (2) Combined with error correction models, the modeling and testing for equilibria around which relationships fluctuate can be completed straightforwardly.[3] Combined with cointegration models, the direct effects on the relationship of a set of exogenous variables can be taken into account. The equilibrium relationship can then be examined, almost as a counterfactual, with those effects removed. Perhaps most importantly, (4), Difficulties in numerical optimization mean that it can be more intuitive to explore the conditions that undergird a relationship by first modeling the time series of correlations and then exploring what variables affect that relationship.

[2] As opposed to the dynamics of a system of variables with assumptions of time-invariance built into the model of that system.

[3] Covered in Chapter 3.

In financial economics, where the wide availability of time series data has long been the case, the ability to conveniently model dynamically conditioned correlations is still fairly young. The statistical methods to do have existed in nascent form for at least three decades (Engle and Kramer 1983) but statistical and computational problems were significant to the point that the size of the econometric literature building these models rivaled the substantive applications of these models.[4] Only in the past decade has the applied literature blossomed to a truly massive scale. And only in recent years has the DCC been introduced into popular statistical software packages such as R (2009) and Stata (2011), and even then with notable limitations.[5]

Perhaps more importantly, the financial econometric literature nearly always emphasizes predictive, instead of causal, inference. Indeed, the name of Engle's book that brings together much of his research on multivariate GARCH models is entitled, "Anticipating Correlations" (Engle 2009). The financial advantages that comes from refining predicted values of covariances are so large that this it to be expected. That does not mean that the DCC is not also an incredibly powerful tool for causal inference; but it does mean that maximizing its potential for causal inference requires some explicit study from the perspective of the research methodologist. That is the purpose of Chapter 2.

The process of exporting MGARCH models to different data environments heightens the importance of certain statistical issues that are of relatively minor importance in the typical financial market setting. These are addressed in Chapter 3. In particular, sampling error; repercussions from working with shorter time series; creating time series of correlations for purposes of using the series as latent variables; and the empirical relationship between aggregate-level and cross-sectional correlations are covered.

Each of these statistical issues in their own way are greatly affected in practice by the

[4]Engle discusses the delayed response in economics to his original, univariate ARCH model in, New Frontiers for Arch Models, *Journal of Applied Econometrics* 17, no. 5 (2002): 425-446; the response to multivariate versions of GARCH (Bollerslev 1986) was further delayed by computational issues, among others.

[5]Particularly in *R*.

computational side of things. So, Chapter 3 also introduces two new R packages. One (Judge and Badanjak, forthcoming)[6] is written to facilitate the use of the DCC with survey data. It also emphasizes the flexibility of the modeling process, even while fitting very large numbers of time series of correlations. This is particularly important, because practical experience shows that the application of the DCC in political scientific settings entails fitting series that may behave in a manner more differentiated than assets in a financial market. It also includes other useful functions for time series analysis. It is by far the most flexible and powerful R package for multivariate GARCH modeling.

The other package (Judge, forthcoming) is a stochastic heuristic optimization function, designed to get around, in a computationally efficient manner, the local maxima problems that plague many models, particularly when applied to shorter time windows. Further, the process of fitting time series often involves fitting so many different models that scholars need computationally efficient solutions to the local maxima problem. So, the problem is particularly important to applications of the DCC in the political scientific setting. Both R packages are presented more thoroughly elsewhere (Judge and Badanjak, forthcoming; Judge, forthcoming). The presentation here facilitates the discussion of the underlying statistical issues. The stochastic heuristic optimization problem is laid out inparticular detail because the statistical issues that underlie it are not commonly covered in political science and are of heightened importance for applications of the DCC in the political scientific setting.

Fortunately, the issues in Chapter 3 need substantive examples. So, the data used in the bulk of this dissertation, the 2008 National Annenberg Election Survey (NAES) is introduced and used throughout Chapter 3.

Chapters 2 and 3 remove the bulk of the methodological, statistical and computational barriers to using the methodological framework presented here in a political scientific setting.

[6]The author programmed all the scripts used in this dissertation, which form the vast bulk of the R package presented here. Sanja Badanjak is working with the author on the process of readying the package for publication on CRAN.

With those out of the way, Chapters 4 and 5 can then take a more substantive focus, applying the framework to the study of American presidential campaigns.

Chapter 4 outlines, for the first time, the dynamic structure of public opinion during a presidential campaign. The investigation is completed systematically, putting the variables into four categories: voting behavior, party ID and ideology, politician evaluations, issue and non-politician evaluations. Each category is first characterized vis--vis its internal relations. When that is done, the relations between categories is examined.

The dynamics and trajectory of these correlations suggest the limits and capabilities of campaigns-as-organizations. They are, on the one hand, suggestive of a surprising level of campaign agency in affecting the correlation structure of opinion in some respects. On the other, they show how campaigns are surprisingly not capable of affecting other aspects of the structure of public opinion during presidential campaigns.

Chapter 4 covers a very broad topic. It is by design primarily a work of descriptive inference. Its findings, however, speak to a number of important subliteratures within thecampaigns literature. It is argued that the basic knowledge presented in Chapter 4 should be common knowledge among those studying campaigns.

With Chapter 4 in place, Chapter 5 can narrow the focus to the role campaigns play in activating latent dispositions, a process held by political scientists to be central to campaigns at least since Lazarsfeld, Berelson and Gaudet (1944). The findings in Chapter 5 are surprisingly clear cut. To the author’s surprise, they show that campaigns do little to activate latent dispositions. Instead, ”activation” is localized in its effects only to the candidates themselves. Nowhere else in the structure of public opinion does there appear strong evidence of latent disposition activation. Because of this, it is argued, the term ”activation” is a misnomer that has misled scholars about how campaigns engage the general electorate.

The remaining parts of this chapter do two things. First, what is meant by a ”methodological framework” is clarified. Second, the campaigns literature is given a general review, describing the larger patterns in the literature that motivate this dissertation. Some of the subsequent chapters review more specific questions.

1.2 A Methodological Framework

Theoretical frameworks organize diagnostic and prescriptive inquiry (Ostrom in Sabatier, ed., 2007, pg. 25). So, by methodological framework, it is meant here *a coherent organization of empirical methods into a strategy for social scientific inquiry.* This emphasis on organization is important. Statistical methods are regularly invented and introduced; less page space is given to the explicit organization of these methods into a broader mode of inquiry that maximizes their inferential potential. That is an odd reality, since most scholars are intuitively aware of the importance of the organization of their empirical investigation. So, this methodological framework does not so much emphasize the creation of new statistical methods as the organization of methods into a strategy for social scientific inquiry.

Throughout this study, statistical dependence is generally assumed to be linear; that is, well described by covariance. Chapter 4 discusses why this assumption is much safer when not imposing an assumption of relationship time- invariance. Covariances are conveyed nearly always as correlations, specifically, Pearson's *r*. The term correlation will generally be used this way, meaning the ratio of the covariance to the square root of the product of the variances. When discussing the methodological framework, however, correlation and covariance will be used as convenient shorthand, with the understanding that other measures of statistical association might better describe the structure of dependence among the variables under study.

The main point of the methodological framework is to leverage the over-time variance in covariance in order to characterize theoretically-important elements of a dgp. Consider the dgp of public opinion during the 2008 American presidential election: Some correlations between important variables evolve almost as if by a linear, deterministic trend not just during one point in the election, but from before the first nomination contest takes place, past the conventions and all the way through to election day. Others seem more affected by important campaign events. Some are prone to equilibria, while others undergo a random walk. Some relationships are surprisingly exogenous to any campaign effects, even while the variables that compose them swing wildly about in their aggregate values. These different changes in ”correlation series” convenient shorthand for time series of correlations are not obvious. Many of the findings alluded to above, and detailed in chapters four and five, were not expected *a priori*. They say important things about presidential campaigns and where in the causal chain actors and events can, and cannot, affect outcomes. They also speak to which parts of the process are inherent in the process itself and which are more conditionally present. Some of the findings have been found with more traditional methods, while others inherently cannot be investigated without elements of the methodological framework proposed here. In fact, some hitherto unused analytic leverage stems from inherent differences between many economic and political processes. For instance, many political processes, like elections, have conclusive or near-conclusive end dates. Whereas the vast majority of markets studied by time series scholars cannot be said to drive towards a final outcome value, campaigns always can. This, as will be shown, opens up inferential techniques to the political scientist which generally cannot be used in market settings. Similarly, inherent in fully functioning markets are mechanisms for the removal of predictable price patterns. This is not such a general rule for variables in non-market settings. Thus, political scientists may enjoy opportunities for inference that, at least theoretically, should not be common for students of markets.

1.3 American Presidential Campaigns

1.3.1 From Minimal to Significant Effects

Like much of American political science, Campbell et al. (1960) and the Columbia school (e.g. Lazarsfeld, Berelson and Gaudet 1944; Berelson, Lazarsfeld and Mcphee 1954) can be treated as the starting point for the modern study of the effects campaigns have on voters. Both sets of scholars found campaigns to have relatively minimal effects on the final vote, especially in comparison to the political and economic events that occurred in between campaigns.

These early findings would be reinforced by related theories that were receiving considerable attention at the time and were seen as plausible mechanisms in the minimal effects model. For instance, the political psychological work on selective exposure (e.g. Klappner 1960) suggested populations will actively filter the messages they receive to buttress pre-conceived notions.[7] Similarly, the classical conception of party identification as, in the words of Franklin (1984), pre-political meant that most voters would not be subject to persuasive campaign effects.

Over the past several decades, the minimal effects thesis also drew strength from the literature on election forecasting, which poses a natural question to those that emphasize the importance of campaigns: If general election outcomes are largely predictable from before the campaign begins, then how can campaigns be all that important? Erikson and Wlezien (2008) present findings that encapsulate much of the forecasting literature, not least of which is their long running series on pre-election forecasts (e.g. Erikson and Wlezien 1996). For the presidential elections 1952-2004, if the final vote total is regressed on trial heat poll results and leading economic indicators in June, prior to either party officially nominating a candidate, the adjusted R-squared is .68. By August, when the conventions are completed but most of the general election remains, that number climbs to a very impressive .88.

[7] McGraw and Hubbards chapter in Mutz, Sniderman and Brody (1996, ch 6) addresses the interplay of the early political psychology and political communication literatures.

Such high levels of explained variance might be taken as an open-and-shut empirical case for the minimal campaign effects thesis. There are several problems, however. Two in particular stand out. The first is an endogeneity problem. Trial heat polls taken during the summer reflect one or two large, quasi-national nomination campaigns that have already taken place. Both parties, furthermore, have been engaged in daily general election activity long before their nominations are officially sealed. The general election may have not taken place yet, but polls from early in the election year are by no means exogenous from the campaign effects and dynamics of that years specific race.

The second problem is one of external validity. The thirteen elections during these years are far from a large textitn, even if each election was somehow akin to unbiased draws from the population of potential political-partisan confrontations. Instead, each election involves groups of candidates and campaign operatives that change only incrementally. Every single election in Erikson and Wleziens sample involves at least one candidate who is on the ballot during subsequent elections. Indeed, one candidate, Richard Nixon, is on the ballot for nearly 40% of the entire sample. The limited inferential range and danger of over-fitting of such a sample is obvious.

The above two problems mean, in brief, that the election forecasting literature, while impressive, should not be taken as indicative of an ability to predict elections without reference to campaign events.

The early 1990s saw the unraveling of the minimal effects thesis. The attacks came on a number of fronts, but two stand out as particularly important. First, Zaller (1992) powerfully undermined the empirical basis of previous studies of media and campaign effects with a number of methodological criticisms. To give examples of some of his major points (see also Zaller and Price 1993; Zaller in Mutz, Sniderman and Brody 1996): Previous studies had assumed there to be a linear relationship between respondent information level and media effects, when in fact the relationship was convex; self-reported knowledge of campaign events is a terrible proxy for actual knowledge levels; and information flows from opposing campaigns were often not balanced in the real world, as many studies assumed they would be.

Second, Franklin (1991) makes the point that elections are exceptionally unstructuring institutions in terms of the campaign politics they beget, and so campaign effects are largely dependent on the actors involved. He demonstrates this empirically in his study of Senate campaigns, by comparing contemporaneous in-state perceptions of senators when one is up for reelection and the other is not, and then contrasting those differences with those that result from different campaign strategies. While elections alone increase clarity, these effects are small in comparison to the effect due to candidate campaign strategies (ibid.)

A recent study, one of a surprisingly modest number of true time series studies of campaigns, begins by characterizing the minimal effects thesis as largely overturned, saying, [w]e accept the fundamental finding of recent campaigns research: campaigns do matter (Box- Steffensmeier, Darmofal and Farrell 2009).

1.3.2 Contemporary Research and Open Questions

There is a divide in the literature that must be considered when asking how to advance our understanding of campaigns. There is, first, a large, fairly well-organized body of what can referred to as campaigns-driven literature: studies of how campaigns evolve, how they are affected, and the extent of their effects. Second, there is a less organized, though perhaps cumulatively larger, set of act-driven literatures: studies that are organized around the measurement of the impacts of particular types of campaign acts, such as television advertisements (e.g. Freedman, Franz and Goldman 2004; Huber 2007) or phone calls (e.g. Gerber and Green 2001).

This is a natural divide driven by different research agendas. It has, however, grown wider as scholars have taken to heart Shaws (1999) argument that studies with findings of minimal effects have mis-specified the dependent variable (by lumping campaign events together), as well as the dependent variable (by considering only the immediate impact). Both the act-driven and campaigns-driven literatures have grown more precise in their questions and operationalizations. As they have done so, they have come closer to answering their individual research questions. The literature as a whole, though, risks further atomization.

Whereas older work generally sought to characterize the effects of campaigns or campaign events on the election day vote, contemporary campaigns-driven work has seen a diversification of dependent variables. The dependent variable often remains the final vote. Increasingly, though, it may be vote intention at time *t*, fundraising (e.g. Christenson and Smidt 2011), or endogenous campaign effects on actors such as the media, individual campaigns and voters (Box-Steffensmeier, Darmofal and Farrell 2009). Meanwhile, as the act-driven literature has grown more refined, it has seen a greater level of specificity in its independent variables. For instance, after Gerber, Green and Larimer (2008) found social pressure to have impressively large effects on turnout, Political Behavior (2010) published an entire special issue dedicated to field experimental research on social pressures effect on turnout.

Such research constitutes genuine breakthroughs for behavioral political science in general. Vis-à-vis our understanding of campaign effects, however, the question of how to build a more cumulative literature remains unanswered. Act-driven findings are useful. By themselves, however, it is improbable that so many act-driven studies will be conducted in so many different contexts that the profession will be able to offer fully generalizable findings about campaign effects.

Similarly, the campaigns-driven literature has its own set of difficulties. Cross sectional studies of presidential campaigns are limited by the small number of elections. Most time series studies are hampered in their ability to speak to the role of the context of a given campaign act at time t because they have generally assumed constant parameters throughout the campaign. The question is how to build a literature that directly bridges the campaigns-driven and act-driven literatures. If scholars could travel seamlessly between these two bodies of understanding, the literature on campaigns would be fundamentally improved, and the stage would be set for a more cumulative literature.

Just how to bridge those literatures is a difficult question. The new data environment alone will not suffice. Sides and Vavreck (2013) leverage a mass of data - media, public opinion and campaign resources data - to tell the tale of the 2012 presidential election. Save for the first debate, they emphasize the stability of opinion in the face of major campaign events, nodding to the redundancy of political information (Rahn et al in Ferejohn and Kuklinski 1990). Sides and Vavreck's book exemplifies the richer picture of the campaign process that can be gained from the new data environment. It embeds many act-specific investigations in the context of a single campaign. It is difficult, however, to see how a similarly structured study could be brought bear on a larger set of campaigns.

Box-Steffensmeier, Damofal and Farrell advance the theory of campaigns. The key to understanding campaigns," they write, "as democratic instruments, we argue, rests in examining endogeneity not as a methodological nuisance, but instead, as the critical substantive feature of campaigns (Box-Steffensmeier, Damofal and Farrell, 2009).

Almost inherent in that perspective is the need to allow for relationships to change over time. In the pages that follow, some tools and methods are refined for a political scientific setting, combined creatively and organized into a methodological framework. That framework points the research process in the direction of generalizability and cumulativeness. It does so by leveraging the over-time change in covariance to characterize the aspects of the dgp that are endogenous and exogenous to the process being studied; an examination of the nature of the change that does occur situates particular actors within the dgp, delimiting their influence to particular roles, while evidencing the manner in which they can claim agency. Practically speaking, an approach of this type is not a sufficient condition for bridging the gap between the two literatures; it is, however, a necessary one.

Chapter 2

The Basis for a Framework

2.1 Introduction

This chapter introduces the working material of the proposed methodological framework, the time series of correlations that are generated by Dynamic Conditional Correlation (DCC) models. After a short discussion of the use or lack thereof of time series of correlations in political science, a few important terms are defined. The next section presents the DCC and its development in the literature. Some familiarity with the basics of time series analysis is assumed, but this section is also a general refresher on GARCH models, the broader family of models of which the DCC is a part. Section IV discusses time series of correlations broadly from the perspective of the political research methodology literature. Section VI presents some of the ways that time series of correlations may be used as research tools. Before the conclusion, two in-depth examples are given that highlight important possibilities for the framework proposed here.

2.1.1 The DCC and Political Science

The creation of the Dynamic Conditional Correlation model (Engle 2002), and MGARCH models more broadly,[1] was an important step forward for the social sciences. Despite the regularity with which they are used in financial economics, there has been little discussion of their use outside that field. The DCC was developed to facilitate predictive inference. The possibilities it helps open up for causal inference is likely what political scientists will find useful, and this is the topic explored here. It will be argued that MGARCH models and the accurate time series of correlations they generate constitute a novel toolkit for social scientific investigation.

Lebo and Box-Steffensmeier introduced the DCC to political science in the July 2008 issue of AJPS. Five years on from that exceptionally lucid article, not a single published article or working paper by a political scientist was found employing the model.[2] This is political science operating far from its ideal as a creative, energetic field that incorporates useful methodological advances in a timely manner.

There are a set of possible reasons why the DCC has not yet been picked up by political science. Political scientists have made the pursuit of the most accurate time-invariant parameters a priority for the better part of a century. Perhaps after so much time that investigative pattern has burrowed its way into our analytic intuitions.

More concretely, as with many young methodological innovations, it is common for scholars to stop their efforts at a technical understanding, leaving novel applications to the substantive field that generated the new method.[3] Similar to the advent of univariate ARCH modeling, it may take years of incremental, article-by-article expansion of practical examples and model alterations until scholars grasp the full potential of a particular methodological approach.[4]

[1] Tsay (2006) offers a thorough review of the MGARCH family of models

[2] Though one published article (Conraria, Magalhaes and Soares 2012) and one working paper (Mattiacci, 2011; `http://polisci.osu.edu/conferences/vim/Mattiacci_VIM.pdf`) were found that briefly discuss the DCC or its application.

[3] 10 This despite Lebo and Box-Steffensmeiers (ibid.) innovative use of the DCC.

[4] 11 Robert Engle discusses the delayed response of economics to his original ARCH model in, New Frontiers for Arch Models, Journal of Applied Econometrics 17, no. 5 (2002): 425-446.

This chapter aims to leap-frog some of that slow development by exploring the ways that social scientists can utilize time series of correlations to investigate causal connections. This is particularly important because there are a number of possible applications that have yet to be seen in economic, econometric or political scientific journals. Together with applications of the DCC already found in economic and finance journals, they constitute a promising new mode of investigation one that combines some of the richness of cross-sectional data with the aggregate-level variance of time series data.

The DCC has had to contend with another important bottleneck. As Greene (2011, p. 1089) points out, one of the interesting aspects of the development of econometric methodology [is] that the adoption of certain classes of techniques has proceeded in discrete jumps with the development of software. In the summer of 2011, the first version of Stata that facilitates the use of the DCC was released.[5] The first R package with code for analyzing dynamically conditioned correlations was published on CRAN in late 2008.[6] Still, shortcomings in the practical application of statistical software to DCC modeling remains, particularly in R. That is why Chapter 3 presents two new R packages, written by the author. The first facilitates DCC modeling, particularly in the political scientific setting. The second helps with the numerical optimization problems that become especially important when dealing with data common in political science. With these practical but important barriers to ease-of-use out of the way, the time is ripe to begin developing a fuller understanding of the research potential of dynamically conditioned correlations.

Like many questions that arise in the exploration of time series data, exact statistical criteria for some of the methods proposed here are not always clear. Like the development of multivariate GARCH models in general, substantive research concerns have pushed the development and use of these models faster than a full explication of their statistical properties would allow (McAleer

[5]Stata 12 facilitates the DCC model and another model that produces dynamically conditioned correlations, that of Tse and Tsui (2002).

[6]Tomoaki Nakatani (2010). ccgarch: An R Package for Modeling Multivariate GARCH Models with Conditional Correlations

et al. 2008, pg. 1555). This pattern is replicated here. Suggestions for creative applications of realistically modeled time series of correlations are put forward to illustrate the full research potential of an important class of models. A larger research agenda is needed to answer precisely some of the questions that are raised along the way.

2.2 A Note on Terminology

One definitional point is particularly important because it reflects a larger theoretical argument. As used in this paper, a correlation relationship does not mean simply the correlation between two variables, whether conceptualized as time-variant or static. Though the time series of correlations themselves or correlation series, for short are typically the most interesting part of the correlation relationship, there are a number of other characteristics of that statistical relationship.

The higher moments of the time series of correlations may be important. For instance, two variables with the same mean correlation to a third variable but with very different variances of their respective correlation series likely have different relations correlative and causal to the third variable.

Similarly, the dynamics of the correlation series are important. A correlation series that is clearly stationary suggests a relationship between the variables that differs from an integrated correlation series in theoretically important ways. So, I define here a correlation relationship as the correlations themselves, together with the moments and dynamics of their time series. Of course, this richer notion of a correlation relationship is also not causality. It may, however, allow for far more causal inference than the correlations by themselves.

Realistic correlation series is a phrase used here as convenient short hand for correlations that are modeled in a manner that does a reasonable job of capturing the correlations and the dynamics of their series. As will be shown below, besides MGARCH models, cross-time change in correlations has been modeled in various ways. Some simpler methods, such as rolling averages, do not attempt to accurately model the dynamics of the correlation series, and are only rough approximations of the actual correlation at a given time t. Correlations from properly specified MGARCH models are realistically modeled correlation series, as are other reasonable approximations, such as the daily values of correlations based on intra-day data in financial markets, or correlations from a survey that is repeated regularly enough to form a time series of correlations.

2.3 Review of Methods: Motivations, Background and Estimation

2.3.1 Motivations

The thought that realistic correlation series are a fairly young feature of the social sciences might be surprising. Surely weve been able to easily model changing correlations for some time now? In fact, the methods of measuring time-varying correlations previously deemed adequate in political science are rudimentary to the point of generally, fundamentally misrepresenting how correlation relationships change over time.

Two traditional techniques are a rolling average (or moving window) of correlations, and exponential smoothers. These methods have numerous problems.[7] Consider the time- conditioned covariance matrix, $H_t \equiv E_{t-1}x_{t-i}x^{'}_{t-i}$, where x_t is a vector of random variables at time t. For the rolling average, where n is the length of the moving window,

$$H_t = 1/n \sum_{i=1}^{n} x_{t-1}x^{'}_{t-1}$$

[7] See Lebo and Box-Steffensmeier (2008)for a lengthier comparison of time-varying correlations obtained from the DCC and other methods.

For the exponential smoother, where λ is a parameter, $0 < \lambda < 1$

$$H_t = \lambda(x_{t-1}x_{t-1}^{'}) + (1 - \lambda)H_{t-1}$$

While these functional forms do not determine the direction of change, they strongly predetermine the dynamics with which covariances change over time, and are generally rough estimators of present covariances.[8] Their popularity was enhanced by their use in non-academic reports and data generation, e.g. by financial firms producing popular reports or risk firms selling pre-boxed software to other financial firms.

Realistic correlation series are of pressing importance in the world of finance, however. And it may be that the particulars of data in that field delayed the arrival of computationally convenient ways to generate realistic correlation series. In settings that involve market values, daily correlations could be found using intra-day data. This method is not without its own problems,[9] but is a fairly effective way to represent a correlation on a particular day. Its superiority to rolling averages or exponential filters may have reduced the demand for better models of how correlations change over time. Of course, the option of using intra-day-derived correlations almost never exists outside of market settings.

2.3.2 Background

Many articles and time series textbooks offer a good review of models that produce time series of correlations.[10] This review is organized by how these models developed chronologically.

[8] Engle (2002a) should be credited with this illustration.

[9] Engle, Robert. 2002. New Frontiers for Arch Models. Journal of Applied Econometrics 17(5): 425-446.

[10] Enders (2010, Ch. 3) has a good break down. Greene (2012, pgs. 930-937 treats well the multivariate GARCH family of models but does not mention dynamically conditioned correlations. Engle and Sheppard (2001) lay out the motivation and specifics of the DCC in great detail. Bauwens et al (2006) provide perhaps the most thorough review of multivariate GARCH models. Tsay (2006) is also an excellent source. Poon and Granger (2003) focuses only partly on GARCH models, but contains an especially useful appendix, where many conditional volatility models are laid out side by side. Two articles that stand out for the clarity with which they present the DCC are Engle (2002) and Lebo and Box-Steffensmeier (2008).

Realistically modeled correlation series were, in retrospect, the product of a natural progression in the literature. Understanding that development will aid in the comprehension of the models themselves.

George Box and Gwilym Jenkins published their seminal work on time series analysis in 1970.[11] Like that book, much of the work that their research helped spawn kept to a focus on the practical questions of predictive inference. When models are created to predict future movements in asset prices and thus might be worth quite a bit of money they tend to be formed not within the confines of formal theory but instead with a singular focus on predictive accuracy.

ARCH models, as their name suggests, model an assets time-varying variance. The variance (volatility)[12] of an assets price is intimately connected to its value and along with its liquidity helps determine the riskiness of holding that asset. When Engle (1982) introduced ARCH modeling, he was seeking to model and predict changes in volatility. The framework he invented Engle is also the creator of the DCC is perhaps the most important part of Engles stunningly influential work.[13]

At their most basic, ARCH models take autoregressive modeling and apply it to the series' second moment. So, whereas we model the future value of a variable as, say, a first-order autoregressive process,

$$Y_{t+1} = \beta_0 + \beta_1 Y_t + e_t$$

an ARCH model seeks to predict the error structure along the general lines of,

$$Y_{t+1} = \beta_0 + \beta_1 e_t^2 + \epsilon_t \tag{2.1}$$

[11] Box, G.E.P., G.M. Jenkins, and G.C. Reinsel. 1970. Time series analysis

[12] Market volatility is at times modeled differently than the variance of price data. For example, it has be operationalized as the log of the absolute price difference. Most of the time, however, it is indeed the price variance, and for our purposes here, like in most of the literature, the volatility and variance will be treated as synonyms.

[13] Diebold, Francis X. 2004. The Nobel Memorial Prize for Robert F. Engle. The Scandinavian Journal of Economics 106(2): 165-185

Where ϵ_t is an unpredictable white noise element with mean of zero. In actuality, variances cant be allowed to take a negative value. So, the formula used for the simplest first-order ARCH process is,

$$e_t = \epsilon_t \sqrt{\beta_0 + \beta_1{}^2_{t-1}} \tag{2.2}$$

Here the white noise element, ϵ_t, is standardized with a mean of zero and a variance of unity.

The properties of an ARCH model are very convenient and make it exceptionally easy to use and interpret. From the above, $E_{t-1}e_t = 0$ and $E_{t-1}(e_t e_{t+i}) = 0 \; \forall i > 0$, but $E_{t-1}e_t^2 = \beta_0 + \beta_1 e_{t-1}$. The unconditional variance (homoscedastic) term is $\beta_0/(1 - \beta_1)$ while β_1 represents the previous shock's effects on the current shock. Let's make the model look simpler by calling everything under the square root sign h_t. So, in this case, $h_t = \beta_0 + \beta_1 e_{t-1}$. Here we arrive at the very important,

$$e = \epsilon_t \sqrt{h_t}. \tag{2.3}$$

The above paragraph is the key building block for the literature that follows. *All* of the progression of the models covered in this article can be understood as a progression of how to model h_t, or in the multivariate context, the matrix of h_t's, H_t. The next major milestone (Bollerslev, 1986) serves as a good example. Bollerslev allowed the variance to follow not just an autoregressive but a full autoregressive moving average (ARMA) process. So, in the first-order GARCH(1,1) process, using Bollerslevs notation,

$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1} \tag{2.4}$$

Models with greater flexibility can take into account features of the dgp under study e.g. asymmetric volatility effects or threshold effects. They grew rapidly in the literature duringhe 1990s. Acronyms like GARCH-M, QARCH and TARCH began to appear regularly in the top econometric journals. Perhaps the best illustration of the size of this literature is the YARCH model, which stands for Yet another ARCH model.[14]

Bollerslev, Engle and Wooldridge (1988) use a variant of a multivariate GARCH, (MGARCH) model. In MGARCH models, the t h from the univariate model is replaced by the matrix of h_ts, H_t. For example, in the bivariate ($N = 2$) case, $H_t = \begin{smallmatrix} h_{11t} & h_{21t} \\ h_{21t} & h_{22t} \end{smallmatrix}$. H_t is determined by the multivariate equivalent of a GARCH(p,q) process, with some restrictions necessary. Bollerslev, Engle and Wooldridge lay out one such model,

$$vech(H_t) = C + \sum_{j=1}^{p} B_j vech(H_t - j) + \sum_{i=1}^{q} A_i vech(e_{t-i} e_{t-i}^{'}) \tag{2.5}$$

where $vech$ is an operator that stacks the lower triangle of a matrix, columb by column, into a single $1/2N(N + 1)$ length vector; A_i and B_j are the square matrixes of the α and β parameters from (4), and are of dimension $1/2N(N + 1)x1/2N(N + 1)$; C is a vector of length $1/2N(N + 1)$; p and q are the autoregressive and moving average lag lengths; and e_t is a vector of innovation terms at time t. The details of that particular model are not vital at this point. The important point to notice now is the presence, the second we switch to the multivariate framework, of the covariance term, h_{12t}, in the H_t matrix. As a reminder, recall that the correlation of two variables is usually expressed as their covariance divided by the square root of the product of their variances. With both variables holding a mean of zero and expressed in the time-varying context that becomes,

$$\rho_{ijt} = \frac{E_{t-1}(x_{it} x_{jt})}{\sqrt{E_{t-1}(x_{it}^2) E_{t-1}(x_{jt}^2)}}$$

Hence, as soon as we model variances in a multivariate context, we have the time-varying covariances and variances - the two necessary ingredients for a time series of correlations.

The problem with a time series of correlations just dropping into out laps like this, though, is

[14]Engle, New Frontiers for Arch Models. The YARCH model is from a conference presentation by Figlewski (1995)

one estimation. For instance, an unrestricted version of the model used by Bollerslev, Engle and Wooldridge would require $(N+1)+N(N+1)+N^2(N+1)^2(p+q)$ parameters to be estimated. Even in very simple models, then, estimation can be difficult. Bollerslev, Engle and Wooldridge solution is to model each covariance with the restriction that it is affected only by its past values and innovations, with $p=q=1$ lags.[15] This solution is far from perfect, however, in part due to issues of numerical optimization. Chapter 3 will go into more detail.

The question of which restrictions are the most appropriate has framed much of the literature that followed the introduction of the multivariate GARCH model. The goal is to find the best balance between the number of parameters, other numerical optimization issues,[16] and the accuracy of the model. Bollerslev (1990) proposed an interesting model, the constant conditional correlation (CCC) model, which proved to be the predecessor to the DCC. The bivariate example is representative,

$$H_t = \begin{matrix} h_{11t} & \frac{\rho_{12}}{\sqrt{h_{11t}h_{22t}}} \\ \frac{\rho_{12}}{\sqrt{h_{11t}h_{22t}}} & h_{22t} \end{matrix} \tag{2.6}$$

where ρ_{12} is the correlation between the two variances.

To see what is going on in the H_t of the CCC more clearly, it is helpful to break it down into its variance and correlation components,

$$H_t = D_t^{-1} R D_t^{-1} \tag{2.7}$$

where D_t is the diagonal k x k standard deviation matrix, with $\sqrt{h_{1t}}, \sqrt{h_{2t}} \ldots \sqrt{h_{kt}}$ on the diagonal and zeros elsewhere, and R is the time-invariant matrix of correlations.

The simplicity of the CCC means that estimation is considerably easier. Much of the

[15]They accredit French, Schwert, and Stambaugh (1986) and Poterba and Summers (1986) with a similar idea.

[16]Other issues arise. Speaking broadly, the likelihood of MGARCH models are notoriously difficult to maximize. For instance, one benefit of the CCC is that it greatly reduces the number of matrix inversions that must take place during estimation, often from thousands to dozens. For models with large numbers of variables or high orders, this can be especially important. The DCC retains much of the relative computational ease of the CCC.

financial data that scholars like Engel and Bollerslev were concerned with, however, does not display constant correlations. So, Engle took Bollerslevs model and allowed for a t subscript on ρ. So in the DCC,

$$H_t = D_t^{-1} R_t D_t^{-1} \tag{2.8}$$

Repeating the bivariate example,

$$H_t = D_t^{-1} R_t D_t^{-1} = \begin{matrix} h_{11t} & \frac{12}{\sqrt{h_{11t}h_{22t}}} \\ \frac{12}{\sqrt{h_{11t}h_{22t}}} & h_{22t} \end{matrix} \tag{2.9}$$

ρ_{ijt} is based on the time-varying covariance matrix. Adopting the notation found in the literature,

$$\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{iit}q_{qqt}}} \tag{2.10}$$

Engle (2002) considers various ways to model the q_{ijt}'s[17] and tests the accuracy of the resulting models with Monte Carlo experiments and real-world data. One of the best performing and intuitively accessible methods is to allow for a GARCH estimation of the changing covariances. Taking the (1,1) model as an example,

$$q_{ijt} = \rho_{ij} + \alpha(\epsilon_{i,t-1}\epsilon_{j,t-1} - \rho_{ij}) + \beta(q_{ij,t-1} - \rho_{ij}) \tag{2.11}$$

Where ρ_{ij} is the time-invariant, or unconditional, correlation between ϵ_{it} and ϵ_{jt}; that is, the same ρ_{ij} from the CCC. Engle and Mezrich (1995) propose estimating unconditional covariance directly as the sample covariance a practice known as "variance targeting" thereby reducing the estimation burden. A small literature has popped up around this idea, and practical experience finds this works well for the unconditional correlation.

The specification of parameters in (11) must guarantee positive definiteness the correlation

[17] Because of his alternative specifications of q_{ijt} and the many alterations proposed in the subsequent literature, one can speak of the DCC as a larger family of models.

matrix. (11) is known as the ”two parameter” variant of the DCC. Ding and Engle (2001) lay out a number of more flexible MGARCH models, which can also be used for the DCC. The two parameter is vastly more popular than any other variant, but it does suffer some important drawbacks. Chapter 3 explores some of these issues further. Engle (2002) and Engle and Sheppard (2001) give the log-likelihood of the DCC. Enders (2010, Ch. 3, sec. 8 and ap. 3.1) provide a good review of MGARCH MLE more generally, and the DCC specifically.[18]

The log likelihood of the DCC is,

$$l = 12\sum_{t=1}^{T}(klog(2) + 2log(|D_t|) + e_t D_t^{-1} D_t^{-1} e_t - \epsilon_t \epsilon_t + log(|R_t|) + \epsilon_t^{'} R_t^{-1} \epsilon_t) \tag{2.12}$$

Where ϵ_t is the vector of time-conditioned standardized residuals.[19] Since D_t and R_t enter the equation seperately, (12) can be treated as a two-step estimator, wherein first the D_ts are estimated, and then the series of R_ts. Engle and Shephard (2001) give a proof of the two stage MLE consistency and asymptotic normality based on Newey and Mcfaddin (1994) proof of two stage GMM estimators.

Drawbacks

The DCC is not the conclusion to the literature related to time-varying correlations. A number of scholars have offered further generalizations. For instance, Chan, Hoti and McAleer (2003) introduce a generalized autoregressive conditional correlation (GARCC) model where the correlations are derived from a random coefficient VAR of the standardized residuals.[20]

Bauwens, Laurent and Rombouts (2006) point out the shortcomings of (11), the two parameter variant. They echo a number of authors, noting that, with large matrices, the imposition

[18]Though Enders description of DCC modeling of the correlations as smoothing could not be found elsewhere in the literature.

[19]That is, the residuals standardized vis-à-vis the standard deviation at time t.

[20]McAleer et al (2008) is another source for an in-depth comparison of the GARCC to other models, the DCC included.

of a single scalar acting on each element of the covariance matrix can be overly restrictive. Caporin and McAleer (2013) unpublished work is very critical of Engle, Shephard and Sheppard's (2008) proofs of estimator consistency, and Aielli (2013) notes that the consistency and asymptotic normality of the second step estimator is based on the first step's accuracy though Engle (2002) explicitly mentions that second-stage consistency is based on that of the first stage estimation. Aielli offers a "corrected" (cDCC) estimator that outperforms the DCC in forecasting under some circumstances.

There are three good reasons to stick with the DCC for now, though. First, the DCC holds up well to less restricted MGARCH models,[21] and has been tested many times on simulated data. Second, given the delayed response of political science to the DCC, it seems unlikely that enough scholars will take up additional, more complex models, where there are such declining marginal returns of accuracy for increased complexity, and no pre-existing software packages to facilitate modeling.

Third, and most important, any scholar that has spent much time actually using multivariate GARCH models for research, as opposed to focusing on their theoretical properties, appreciates that the most important limits imposed by model type involve numerical optimization, especially when exogenous variables are involved. By breaking the modeling process down into two steps, the DCC facilitates modeling with higher order models or more flexible model variants where other GARCH models would fail. This is the motivation behind the CCC and then the DCC's creation. Works such as Caporin and McAleer (2013) that do not put such issues front and center are simply not in touch with the actual process of model fitting when real-world data requires the fitting of complex models.

If the DCC can be taken as the standard approach to modeling time series of correlations, then it is time to turn to questions of application. The burst of creativity that came from the scholars who produced the literature that led to the DCC has for the most part not been matched

[21] See, for instance, Engle 2002

by those that have applied the model to answer substantive questions. To be sure, there are still a number of open methodological issues concerning MGARCH models, the DCC included.[22]

These issues are important and need continued attention. It is time, though, for a research methodology literature on time-varying covariance to compliment the econometric literature.

2.4 Thinking about Time-Varying Correlations

Consider Jackmans definition of latent variables: quantities that are not directly observable. [23] He describes the inferential problem involved with latent variables as, Conditional on observable data y, what should we believe about latent quantities x?[24] Using the definition of correlation used here, Pearsons r, the sample correlation is directly observable from y, as the ratio of the covariance and the square root of the product of the variances.

A correlation series, by definition, though, defines correlation as a characteristic of the dgp at each time t. In most research settings in time series analysis, however, the researcher is left with only one observation of each variable at each discrete point in time. So, unless there is a sample of y at each point t large enough to observe the sample correlation, the correlations that are derived should be considered latent variables, whereas their atemporal cousins, the sample correlation, need not be.

[22]Bauwens, Laurent and Rombouts (2006) provide a list of outstanding issues. It is reproduced here in full here as an indicator of the issues found in the contemporary econometric literature:
1. Improving software for inference (this is a prerequisite for progress in applications).
2. Comparing the performance and assessing the financial value of different specifications in applications.
3. Implications of stability or not of a model class with respect to linear transformations.
4. More flexible specifications for the dynamics of correlations of DCC models.
5. Unconditional moments of correlations/covariances, marginalization and temporal aggregation in DCC models.
6. Development of a copula tool for specification and inference.
7. Impact of choice of the square root decomposition of Ht on statistical procedures.
8. Conditions for two-step efficient estimation (MGARCH on residuals of the mean model).
9. Asymptotic properties of MLE (in particular low level, easy to check, sufficient conditions for asymptotic normality when it holds).
10. Further developments of multivariate diagnostic tests.

[23]Jackman in Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier, eds. 2008. *The Oxford Handbook of Political Methodology*, pg. 119.

[24]*Ibid.*

Conceptualizing the correlation series as a time series that is latent in the dgp is helpful in structuring our thoughts. The questions that jump into the mind of those that first hear of the DCC, though, are likely to be questions of accuracy and biasedness.

DCC-derived correlations at time t are, by the design of the model, the forecast of the correlations at time t+1, based on information up to time t. Despite the popularity of the DCC, precise statistics concerning standard errors of the correlations at a given time are difficult to come by. While this is not ideal, DCC-derived estimates tend to be quite accurate, especially as the time window grows.[25]

Tests on simulated data and tests of forecast error using real-world data are revealing of the accuracy of the DCC. In financial settings, the DCC performs very well, even against less restricted MGARCH models.[26] The same may be said for Monte Carlo experiments, where the DCCs summed mean forecast errors over a series of experiments generally outperformed other even computationally more intensive MGARCH models.[27]

Ignoring issues of measurement error in the variables themselves, DCC estimates of correlations should be considered a valid measure of correlations, assuming the correct functional form.[28] In particular, if the correlation series is suspected of being integrated, one should use Engles (2002) integrated DCC model.

Assuming the correct functional form is used, questions concerning the measurement validity of time-varying correlations can be stated in more generalized terms, and the standard literature that addressed measurement broadly and latent variables in particular applies (e.g. Adcock and Collier 2001; Jackman in Box-Steffensmeier, Brady and Collier, eds. 2008; King Keohane and Verba 1994; Gerring 2001, forthcoming).

[25] Lebo and Box-Steffensmeier 2008, pg. 696

[26] Engle and Sheppard 2001

[27] Engle 2002

[28] Validity is roughly analogous to the notion of unbiasedness in the context of parameter estimation. Jackman in Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier. 2008. *The Oxford Handbook of Political Methodology*, pg. 121

Given the emphasis on time-variant relationships, issues of contextual specificity (Adcock and Collier 2001, pgs. 534-6) should be theorized in great detail by the researcher beforehand. Especially when modeling correlations with a transfer function or similar model (see below), the use of event-specific dummy variables in particular should be determined by theory before any models are run. The appropriate use of event dummies, even in the same sample, may vary by the question being asked,[29] but post hoc reasoning will likely undermine the scientific nature of the modeling process.

One point that has not yet been made in the context of correlation series, and is especially important for data used by political scientists, is that measurement error in the underlying variables will accentuate (increase in absolute value) or, more likely, attenuate (decrease in absolute value) the values of the correlation series. With survey data, for instance, the variance introduced from sampling error may be a considerable portion of the overall observed variance. To put it more plainly, DCC-derived correlations are not robust against errors in the data on which they are based.

In the cross-sectional context, assuming the measurement errors in two variables, x and y, are random and independent of one another as well as the variables themselves, measurement error reduces the measure of correlation from the actual correlation by,

$$\rho - \frac{\rho}{\sqrt{(1 + var(e_x)var(x))(1 + \frac{var(e_y)}{var(y)})}} \tag{2.13}$$

Where ρ is the true correlation and e_x and e_y are the measurement errors. With survey data, expected variation from sampling error is generally known, to a rough approximation. Issues with weighting, and measurement issues more broadly complicate matters (see, for instance, Traugott and Wlezien, 2009). The simplest solution would be to strip that approximated sampling error-derived variance from the qs in the denominator in (10). Alternatively, one could plug the

[29] context-specific indicators and adjusted common indicators are not always a step forward, and some scholars have self-consciously avoided them. (Adcock and Collier 2001, pg. 536)

relevant values into (13), using the initial estimates of ρ_t in place of rho, and then add the resulting T x 1 vector of (13)-derived values to the original estimates of the correlation series to calculate a new, more accurate series of correlations.[30]

Heteroskedasticity that comes about from changing sample size is covered in Chapter 3.

2.5 Uses of Correlations Series

The MGARCH literature has used time series of correlations in several ways: As a measure of market volatility spillover (contagion) ; as characteristics of the price behavior of assets or entire markets; and as parameters in models focused on predictive inference. There are three broad categories of additional, largely unexplored uses of realistic correlation series in the social sciences. Each of the three, overlapping categories entail different techniques and possible strategies for inference. They will be explored below. The material presented in the following section is the result of an attempt to answer the question, What is the unexplored causal inferential potential of realistic correlation series?

(1) Examining the entirety of correlation relationships and the correlation structure of the dgp. This is similar to the process of data exploration, but focusing on the relationships between variables, instead of the values of the variables itself. Practical experience has shown this to be a very useful way to learn about a dgp.

(2) Using transfer functions, vector auto-regressions (VARs) and cointegration models to directly model the conditions that underlie a correlation relationship, or the impediments to that relationship. There are several reasons to adopt this approach. One of the less obvious but most important: While the inclusion of exogenous variables in the original DCC equation is one way to reach the same modeling goals, in practice it greatly complicates the optimization process. The mechanics of including exogenous variables and the ramifications of doing so are active topics in

[30] This topic has not been addressed in the econometric literature. For instance, in Bauwens, Laurent and Rombouts (2006) lengthy treatment of MGARCH methodological issues, the phrase measurement error does not appear once.

the literature (e.g. Iglesius and Phillips, 2013). Practically speaking, the process greatly restricts other model specification choices. Creating a time series of correlations that are then explored with subsequent modeling is one way to get around these issues while simplifying the modeling process by breaking it down into discrete steps.

Finally, (3) using correlation series themselves as latent variables in models to increase the realism, parsimony and explanatory power of models. For example, when a model calls for a variable that represents the strength of a given relationship at the aggregate level, a correlation series may better operationalize that concept than would the inclusion of the variables themselves in the model.

Save for some of (2), in Lebo and Box-Steffensmeier (2008), a treatment of these additional possible uses of dynamically conditioned correlations could not be found. This chapter seeks to illustrate the research potential that these methods offer, and so encourage political science to turn to the DCC.

The novelty and wide range of possible uses of dynamically conditioned correlations in political science necessitates a general approach to the subjects below. For many of the uses of correlation series proposed below, more traditional research or modeling alternatives may work just as well or better than those that incorporate the methodological arguments made here. The argument is that investigation of correlation series may supplement, not replace, some of the more traditional techniques.

2.5.1 Investigating, Exploring and Comparing Time Series of Correlations

When the data allows us to do more, we waste a profound amount of information when we focus solely on the time series of the variables themselves. A dgp with n observed variables has $n(n-1)/2$ correlation series. A researcher with eight observed time series actually has twenty-eight correlation series on her hands. Viewed this way, the volume of wasted information is striking. The parameters of the DCC equation that produces the correlation series are another source of important information. Whether or not and to what extent correlations series are stationary reveals important information about the dgp. An integrated series of correlations might indicate structural changes in the correlation process (Engle 2002), whereas an obviously stationary process suggests the correlation-generating process is robust against the variance in the conditions that occur during the time window.

Another basic point is that variables may share similar mean correlations to a third variable but very different variances in their correlation series. The relationship with the higher- variance correlation series is therefore more conditioned on other elements of the dgp. Say we are interested in forecasting election results (e.g. Erikson and Wlezien 2008), and are considering two different indexes of economic activity as a dependent variable. We are probably interested in the index with the highest correlation to the vote. We may be willing, though, to sacrifice a small amount of unconditioned correlation for a correlation series with a smaller second moment. Visual plots of the correlation values over time are useful, both for the researcher and the reader. In addition, three dimensional scatterplots are an efficient way to communicate information about correlations. With correlations on the y axis, the time window on the x axis, and a conditioning variable as the third dimension, the underpinnings of the correlation relationship are efficiently consumed.

Visual aids and descriptive statistics give the researcher a general sense of the correlation relationships within a dgp. They serve as a means to explore, investigate and communicate basics of the data. Still, because a research methodology that focuses on the over-time variance of relationships is so new, some of the most exciting findings may come from this basic, explorative stage of analysis.

2.5.2 Transfer Functions, VARs and ECMs

Modeling the conditions under which correlation relationships strengthen and weaken offers considerable promise. Lebo and Box-Steffensmeier (2008) use a transfer function the time series equivalent of a simple multivariate regression with a correlation series as the dependent variable to model the effects of various conditions on time-varying correlations. This is the first time in political science that an MGARCH-derived correlation series was used as a variable in another model, and similar examples were not found in the econometric literature this despite a large literature that has not developed truly satisfactory ways to include exogenous variables into the correlation modeling process.

Such transfer functions allow the researcher to directly model how other variables structure a correlation relationship. For instance, one of the more interesting findings from Lebo and Box-Steffenensmeier (ibid.) is that presidents receive diminishing marginal returns from economic growth in their approval rating. They come to this finding by regressing the correlation between economic data and presidential approval on a number of variables, GDP growth among them. GDP growth had a strongly negative relationship on the value of the correlation series.

Endogeneity will often be an issue with transfer functions, like with traditional transfer functions. The researcher may turn to VARs and cointegration models to deal with problems of endogeneity. Cointegration and correlation series mix particularly well: Finding the presence of an equilibrium in correlation values once exogenous forces are taken into account is a useful way to disconfirm hypotheses. The second example in Section VI provides an example.

2.5.3 Using Time-Varying Correlations as Variables in Models

Several general advantages of using correlations as time series of latent variables are important First, realistic correlation series may operationalize a concept better than other variables. An example can be taken from economics, where indexes of volatility (e.g. the Chicago Board of Exchanges Volatility Index (VIX)) are often used as a measure of the impact of crises on markets, but some analysts prefer as a measure the mean intra-day correlation of the SP 500. For another example, an interesting way to characterize the strength of partisanship at the aggregate level is by its time-varying correlation with vote intentions.

Both these examples speak to the most general potential of correlation series as latent variables. Correlations are a measure of the strength or degree of a (linear) statistical association. So, as a time series of latent variables, correlation series can be useful as an operationalization of concepts that involves the strength or degree of connectedness. Next, a correlation series may allow for more parsimonious models than otherwise possible. There are two reasons for this. First, including one correlation series in the model is, in one sense, more parsimonious than including the two constituent variables. Second, correlations between two variables may serve as a representation of a net effect of a broad array of causal forces.

A common mistake in the social sciences is the confusion of the main causal pathway between variables for the causal relationship itself. For example, Hetherington (1996) treats perceptions of the economy as a necessary link in the causal chain connecting the economy and the presidential vote. In effect, he ignores the many causal connections between the economy and presidential vote that do not flow through voter perceptions of the economy. Correlation series can be representative of a larger set of causal mechanisms, allowing us to create a variable that leap-frogs questions of causal mechanisms while retaining the measure of association. Of course, there are disadvantages to including a correlation series in a model as a latent variable. Running the DCC or other MGARCH models themselves, choosing between different models of the correlations and developing an understanding of MGARCH models takes time. Like many latent variables, the truly appropriate conclusion when findings are close to the line of significance is not clear.

Whether these disadvantages outweigh the advantages will for the most part be determined by the specifics of the research environment. Given the novelty of realistic correlation series in political science, though, the potential for many useful applications seems noteworthy.

2.6 Two Examples

The following examples help illustrate the usefulness of MGARCH-derived correlation series in political science. Each example concludes with important generalizations about the inferential potential of mixing certain time series models and correlation series.

2.6.1 Example 1: Disaggregating the Effect of Primary Outcomes and The Primary Itself on Campaign Fundraising

The question of the drivers of campaign fundraising at the presidential level has been investigated by a number of political scientists. Hinckley and Greens (1996) study was especially important. They found that at least in 1988 the effects of developments in the campaign were largely secondary to a candidates initial fundraising organizational capacity. They note, though, that problems with using Federal Election Commission (FEC) data as time series could explain some of their findings. Campaign fundraising has changed, though, from a process of checks being physically mailed and eventually deposited, to a process heavily affected by online, instant contributions. As such, scholars have found more success lately in using the still noisy FEC data as daily time series (e.g. Christenson and Smidt 2011).

There has been a larger literature on the effect of primary outcomes on multi-stage presidential nomination contests (e.g. Bartels 1988). Thinking about the development of FEC data in light of this literature, it is natural to ask, What is the immediate effect of primary victories on fundraising?

A difficulty in answering this question lies in the simultaneity of the effects of the primary election itself and the effects of the outcome of that election. That is, the increased attention the race gets will exact an effect on fundraising in a manner that confounds our attempts to measure the effects of the strategic outcome of the primary.

The worry, then, is that a model that measures the influence of primary outcomes on fundraising while not accounting for the influence of the primary itself will suffer from omitted variable bias.

Is this the case? One convenient method to test this is to turn to a transfer function where the correlations stem from the fundraising totals of the competing candidates. The correlations between fundraising totals will likely increase as increased attention from the primary benefits fundraising of all competitive campaigns. If the strategic nature of the outcome typically has a significant immediate impact on fundraising, though, the pre and post primary effects on the fundraising correlation series will differ.

The more recent the presidential nomination contest, the greater the role of internet fundraising. So, the 2008 struggle between Sens. Clinton and Obama was chosen as a case study. The daily correlations of the two campaigns FEC fundraising figures[31] from a DCC (1,1) model are the dependent variable. *Fig. 1* presents the correlations as they vary over time. Visual inspection suggests a significant sensitivity of the correlation series to the primaries. *Fig. 2.2* presents the correlations between fundraising reported by the two campaigns, over both candidate's fundraising totals. *Table 2.1* presents the correlation relationship. Due to the noisiness of the data, the absolute value of the correlations is compressed towards zero. The integrated nature of the correlations is table 2.1s most theoretically useful information. Recall that an integrated series means that the correlation structuring process changes significantly over time. This is evidence that dynamic elements of the campaign have grown in importance since the 1988 race studied by Hinckley and Green (1996). Two variables are included as independent variables in the transfer function, one measuring the proximity of the decisively-won primary,[32] from three days before the primary, and one the proximity for three days after the primary. Both take a null

[31]The data is highly seasonal, with the day of the week strongly affecting numbers. The FEC data is thus first deseasonalized. Similarly, as Christenson and Smidt (ibid) cover in greater detail, the days before filing deadlines are outliers, as campaigns attempt to display strong numbers by rushing to elicit as many donations as possible and report them before the deadline. The three days before each filing deadline are simply dropped from the sample as outliers.

[32]Those are the days of the primaries of IA, NH, NV, SC, then the primaries of Feb. 9, 12, and 19, March 4 and 11, April 22 and May 6.

Figure 2.1

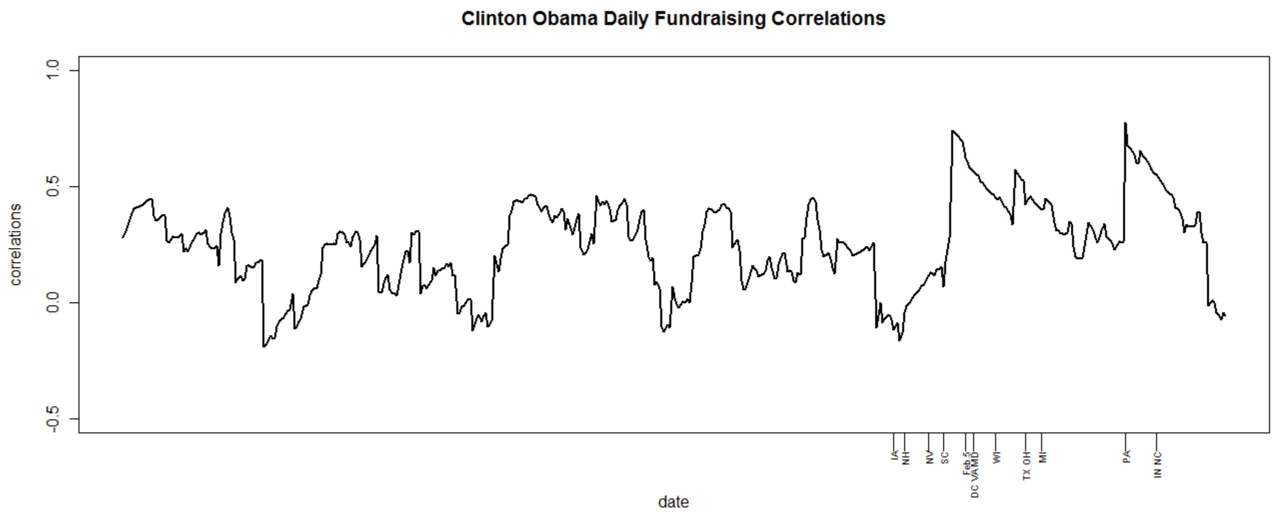
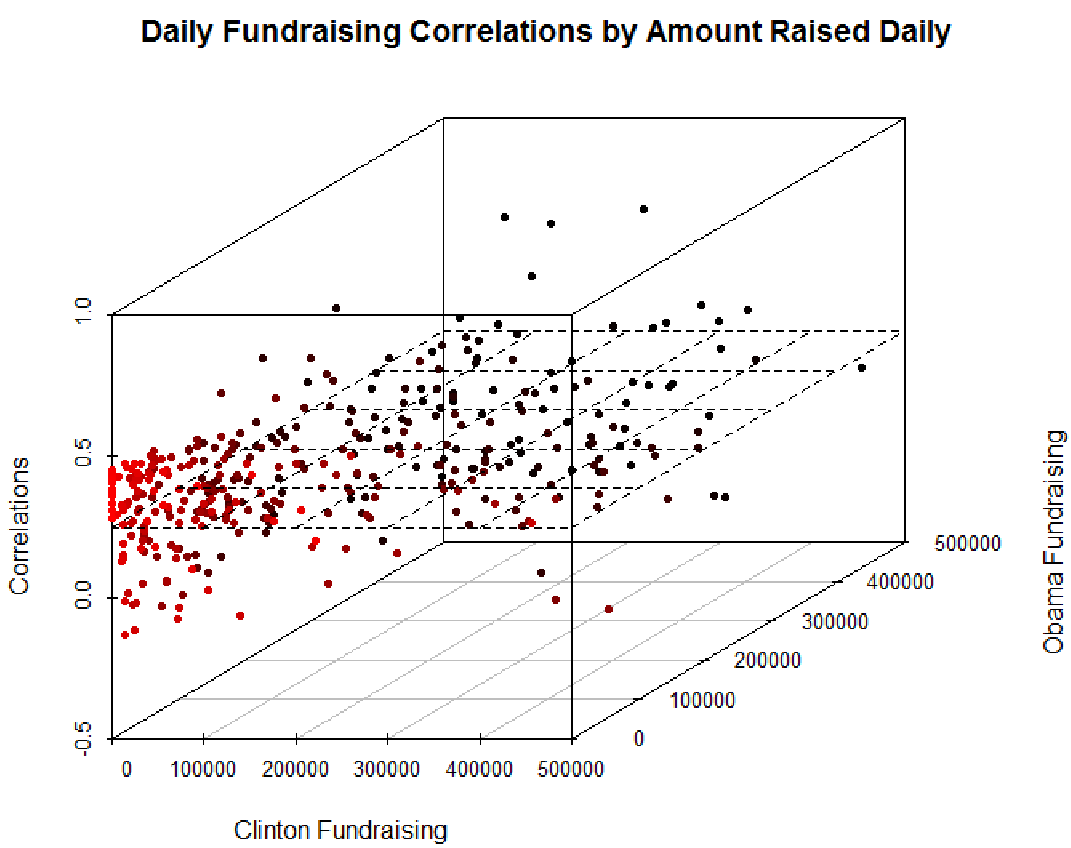


Table 2.1: Correlation relationship between Clinton and Obama Daily Fundraising

min	-.1910
Mean	.2460
1st Quartile	.1173
3rd Quartile	.3906
Var	.0367
Skew	-.0140
Kurtosis	2.651
P(stationarity)	.0235*

*augmented Dickey-Fuller test w/ 12 lags


Daily Fundraising Correlations by Amount Raised Daily
Correlations
1.0
0.5
0.0
-0.5
0
100000
200000
300000
400000
500000
Clinton Fundraising
Obama Fundraising

value at all other times.[33]

Again, the point is to get a sense of the significance of the impact of outcome of the primary, juxtaposed to the impact the primary itself has on individual campaign coffers. So, we are left with the model,

$$\delta\rho_t = \lambda\delta\rho_{t-1} + \theta_1 x1t - 1 + \theta_2 x_{t-1} + \epsilon_t \tag{2.14}$$

where ρ_t is the correlation at time t, δ is the difference operator, the two x_ts are a measure of pre and post-primary proximity, respectively, ϵ_t is the error term at time t, and λ and the θs are time-invariant parameters. Such a simple model is chosen for illustrative purposes. The model does, though, provide preliminary evidence of an immediate impact of the strategic impact of decisively-won campaigns. Table 2 presents the results of (13). The effects are not huge. Given, though, the noisiness of the data and that two independent variables vary in less than 10% of the sample, this is to be expected.

Table 2.2: $\delta\rho_t = \lambda\delta\rho_{t-1} + \theta_1 x_{1,t-1} + \theta_2 x_{2,t-1} + \epsilon_t$

Correlations	Coefficient	S.E.	z-score	$P(\rho < 0)$
t-1	-.0266048	.0447454	-0.59	0.552
preprimary	*.0188556	.0055695	3.39	0.001
postprimary	-.0037896	.0055405	-0.68	0.494

$RMSE$: .062344; R^2 : 0.0233; χ^2 : 11.84852; $P = 0.0079$

Notice the differently signed coefficients on the pre and post-primary variables. Though the latter differs insignificantly from zero, the former is easily significant. The coefficients are significantly different from one another and in the expected directions. At least between Sens. Clinton and Obama, primaries increase the similarity of campaigns fundraising as they approach, but do not exert the same influence as media coverage focuses on the results of the primary. This is especially important because campaigns are counter-intuitively effective in their use of primary

[33]They are lagged to allow for filing time. Larger lag values were far from significance at any level, suggesting campaigns are very efficient in processing donations raised over the internet.

losses as appeals for fundraising (Christenson and Smidt 2011). That strategy may aid their overall fundraising success of a primary losers campaign, but it is apparently not do enough to mitigate the damage to the losers campaign, as the fundraising correlations diverge immediately from their pre-primary high.

2.6.2 Example II: Cointegration Models and Correlations

I have investigated in greater detail elsewhere (Judge Working Paper, 2011) claims recently made that the changed asset price correlations of gold and the dollar are evidence of the beginnings of an active erosion of the dollars unique position in global finance. Augmented Dickey-Fuller tests and KPSS tests are used to ascertain that, like many currency-related correlation series, the dollar-gold correlation series are indeed non-stationary. As noted above, this does suggest a shifting correlation generating process.

The story of the euro crisis, though, has dominated global markets since 2009. The question naturally arises: Is the dollar-gold correlation process cointegrated with a few basic measures of the single currency sovereign debt crisis? If it is, then that means the correlation process would likely be stationary absent the direct effects of a euro crisis. If that is the case, it disconfirms the use of the changing gold-dollar correlation as evidence for any recent changes involving the dollar. It is, after all, hard to use a trendless, stationary process as evidence of systemic change!

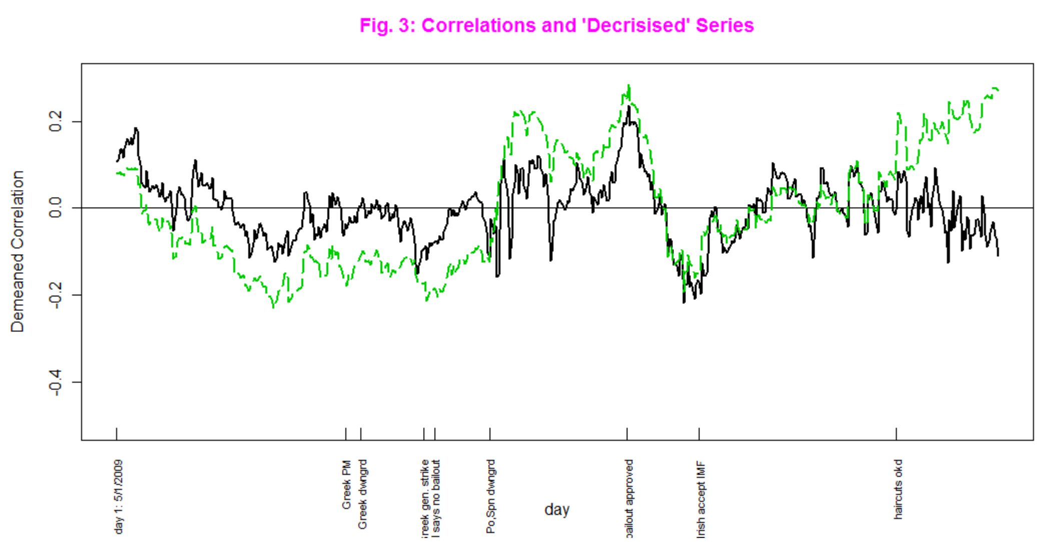
I use a vector of a handful of European bond yields and one measure of market volatility to represent the euro crisis. Table 3 presents the model, the results and several tests from the first step in the Engle-Granger two-step process cointegration modeling.[34] A test for stationarity of the residuals and another test for cointegration are presented in that table. Clearly, the two series are cointegrated, suggesting the integrated nature of the gold-dollar correlations during this era is largely the result of the euro crisis. Absent a claim that changes in the dollar played a strong

[34] As can be seen by table 3, that test is a simple regression of the variable of interest on the suspected cointegrating variables; though, in the presence of cointegration, that regression has its own special statistical properties. See Ch. 6 in (Enders 2010) for more details.

causal role in the euro crisis, this cointegration model is strong evidence against any notion of the dollar-gold correlation breaking down in a manner suggestive of a changed global status of the dollar.

Fig. 3 presents the demeaned correlation series alongside the residual series from the first step of the Engle-Granger two-step Cointegration model. Though not readily apparent from visual inspection, the variance of the correlation series (.018) is over three times that of the cointegration disequilibrium values (.005). To generalize from this example, a cointegration framework seems

Figure 2.2



particularly promising for the investigation of the conditions that underlie changing correlation generating processes. If the correlation generating process leads to integrated correlations, then this is evidence of structural change in the relationship between the constituent variables. If the correlation series is cointegrated with a set of variables entirely causally disconnected from the purported source of that change, however, that is powerful falsifying evidence concerning the causal connections between that variable and the underlying structural change.

2.7 VII. Conclusion

Correlations are one of the basic building blocks of empirical social scientific analysis. The development of computationally convenient, parsimonious models of how correlations change over time is no small matter.

During the same general time period these models have developed, there has been a rapid growth in the number of politically-relevant time series. Finally, as of the summer of 2011, all of the most popular sophisticated statistical software environments facilitate DCC modeling. With the methods, data and software in place, it is time for political science to develop a rich research methodology literature on the inferential possibilities of correlation series. Section V, where many of the potential uses of correlation series are laid out, is necessarily incomplete. It is a first-pass exploration intended as a suggested set of starting points for the creation of a larger research methodology literature.

The benefits of adopting a framework that involves correlation series varies by the question being asked. In some contexts, it may constitute just an interesting mode of data exploration. In others, it may provide more fundamental insights or improvements in model performance that could not be achieved without such an approach.

Chapter 3

Aggregate Variance-Derived Covariance and the Political Scientific Study of Survey Data

3.1 Introduction

Practical and theoretical issues arise when applying mGARCH models to survey data that do not come up during their application to the study of financial markets. So, there is no published work specifically on the application of mGARCH models to survey data. This chapter begins with the practical, statistical side of things; sampling error and optimization issues, for example. For reasons that will be explained, they are often of heightened importance in political scientific research settings, and they make up roughly the first half of the chapter.

Beyond the statistical issues that need to be resolved before substantive study gets under way, there is another research barrier that, historically, has greatly affected the direction of the social sciences: the availability of software that allows researchers to easily implement and test models. To that end, two new R packages are introduced (Judge, forthcoming; Judge and

Badanjak, forthcoming). The first provides a computationally efficient way around the local maxima and flat likelihood problems that often plague GARCH models.

The second, larger package is designed to facilitate univariate GARCH, a variety of different (multivariate) DCC models, and related useful functions for time series analysis.[1] The flexibility of a range of different forms of DCC-GARCH models, inclusion of a variety of optimization routines from other packages, and the ability to include exogenous variables in various functional forms makes this package by far the most flexible and powerful DCC package currently available in *R*. In addition, several features are designed specifically to facilitate the DCC-based modeling of survey data. The package is laid out in full elsewhere (Judge, forthcoming; Judge and Badanjak, forthcoming). Here, reference is made to these packages only when it facilitates the discussion of problems already being discussed in this chapter.

The statistical and practical quickly become the empirical, however; and it is those issues that motivate this chapter. First, characteristics typically found in time series of public opinion demand that variants of the DCC must, for the first time, be introduced to political science. Second, an important element of time series studies of political behavior is the relationship between individual-level, cross-sectional covariance and the estimated covariance derived from those variables' time series. Without knowledge of that relationship, time series studies can "speak to" and "inform" the traditional cross-sectional approach but cross-sectional and time series findings cannot be directly compared.

This is a particularly important question because a major characteristic of the mass political behavior literature over the past two decades has been the informing of debates traditionally based in individual-level data by time series, methods and data. *The Macro-Polity* in general, Stimson's "policy mood" work, and Box-Steffensmeier's contributions to the party ID literature are all prominent examples.

[1]The author wrote all of the code for this dissertation in R. Sanja Badanjak is assisting with transforming the body of scripts written for this dissertation into two separate, CRAN-ready packages.

Ideally, aggregate-level variance derived estimates of parameters could be used as a convenient stand-in for the more traditional cross sectional sample parameter. To clarify, consider the relationship between a parameter derived from cross-sectional variance e.g. a survey of public opinion observed at a particular point in time, ρ_{ct}, and the parameter derived from its aggregate-level, time series variance, ρ_{at}.

$$\rho_{at} = f(\rho_{ct}) \tag{3.1}$$

If the practical hopes of researchers ruled the world, $f()$ would be an identity, $\rho_{at} = \rho_{ct}$. For a variety of reasons, though, that seems unrealistic. Characterizing (1) as it exists for a particular topic is a useful, preliminary step if time series and cross-sectional findings are to be made more directly comparable.

The functional form in (3.1) seems like too vague a question one whose answer likely varies over the topics being studied. It is, though, an interesting question to ponder. For each research topic, it should be asked, are aggregate-level variance-derived measures of covariance close approximations of cross-sectional covariance? And, failing such a degree of accuracy that we can consider discrepancies a form of measurement error, can the relationship between the two types of covariance be characterized, perhaps as a probability distribution?

The second half of this chapter begins by presenting findings that show the tightness the relationship between aggregate-level variables and their individual-level counterparts. This is done over a wide array of variables representing in a broad sense the major components of political public opinion during presidential campaigns. The distribution of differences among the forms of correlations is presented, and contrasted with the effectiveness of the DCC in modeling that aggregate-level covariance.

A third benefit of this chapter is that it speaks to the GARCH literature in a unique way. The benchmarks previously used to measure the accuracy of GARCH models are each imperfect in their own way. In Monte Carlo simulations, the actual correlation is known and model-derived correlations can be compared to the true correlation values. However, simulations often rely on assumptions that simplify the dgp to something rarely found in the real world.[2] With real-world data, however, there are the problems with defining the true correlation at time t that were covered in Chapter 2. Time series of survey data give us an interesting, alternative way to investigate the performance of GARCH-derived correlation models. They offer the benchmark for GARCH-derived correlations that is most intuitive: real-world sample correlations at each time t.

Not coincidentally, the data used in this dissertation is particularly well-suited to address the above issues. So, it is with its description that the main body of the chapter begins.

3.2 Data

The 2008 National Annenberg Election Survey (NAES) was the third quadrennial edition of a survey conducted during the American presidential campaign by the Annenberg Public Policy Center. The telephone portion of the pre-election survey, based on thirty-minute phone interviews, consisted of 57,967 interviews, conducted from December 17, 2007 until election day, November 3rd, 2008, for a total of 316 days. This leaves a daily mean of roughly 183 responses a day. The survey was directed by Richard Johnston, Diana Mutz and Kathleen Hall Jamieson.

There is a standard deviation of roughly 70 responses a day for questions that were asked of all respondents. This high standard deviation is due in part to a two-month period during the early summer when response numbers to major questions had a daily mean of roughly just 70. Excepting this period, the mean daily response level was a bit above 200 and the standard deviation around that mean was roughly 50.

A number of questions were asked during the whole time window, and it is to those questions this dissertation turns most regularly. The first section of *Appendix I* to this chapter lists

[2] See the "Stochastic Optimization" section below.

the question wording and response options for all the questions used in this dissertation. The second section of the appendix gives the operationalizations of the questions as daily time series. To clarify, when a researcher is forming a daily time series from the data, she must first choose the form the daily value will take.

This operationalization conundrum, to give an example, is perhaps why macropartisanship is defined in the manner that it is, as the portion of party identifiers that identify with the Democratic party. That macropartisanship does not incorporate the portion of the electorate not associated with either party is a particularly weighty example of the trade-offs researchers face when converting cross-sectional data into time series.

Most of the other *time series operationalizations*, to coin a term, are of a more mundane nature. For instance, each day respondents were asked to rate Senator McCain's trustworthiness on a scale of 0 10.[3] The researcher has the choice of choosing the mean ranking, median or some other suitable alternative as the daily value. For instance, a researcher might wonder about the portion of people that give Senator McCain a positive rating on trust, and so time series operationalize his trustworthiness as the portion of voters giving him a 5 or above.

While choices of time series operationalizations can be important, such as in the macropartisanship example, the choices that the 2008 NAES presented were fairly obvious. Reasonable alternative operationalizations produced negligible differences in most series.

Including four different operationalizations of the party ID variable, time series were used, giving ( ) different correlation series. This is a good example of the rich tapestry of relationships made available by a methodology based on time series of correlations.

[3]The NAES found no significant difference between the usual 1-100 rankings and the 0-10 scale. Similarly, measures of Pearsons r are not generally not significantly attenuated when continuous survey data is recategorized into ten ordinal groups. In experiments where continuous survey data is recategorized into progressively rougher groupings, the attenuation of correlation measures becomes statistically significant around seven categories and, depending on the distribution of the data and the placement of the cutoffs between categories, may become considerably attenuated as the categorizations become rougher (cite those old articles on the from the 70s).

3.3 Optimization of Time Series Models in Political Science

GARCH models are notoriously difficult to optimize. There are a number of reasons for this. For instance, if a large, discrete change occurs in the middle of the time window, the dgp might be a mixture of two distinct probability distributions. More fundamentally, there is a conundrum common to the dgps studied with GARCH and other models popular in time series, such as Markov switching models: If the data are cleanly generated by the same, relatively simple probability distribution, we probably would not be using these models in the first place. *Fig. 2* presents a stylized version of the problem. A time series is being studied. The two black lines are values from the likelihood function over the first and second halves of the time window. Perhaps a large structural break occurred right in the middle of the window. The red line, though, is that of the entire distribution, and that is the multi-model distribution that actually greets the researcher and her numerical optimization routine.

Here is the point rarely made but particularly important for political scientists: *Ceteris Paribus*, local maxima in the likelihood function are more likely to be found with shorter time windows. There are several reasons for this, but chief among them is that shocks that affect the data for a short period of time may constitute a substantial portion of the data set and so, perhaps, their own little bump in the llf. *Fig. 2b* presents the same dgp as *fig. 2*, but with a time window five times longer and with five times the structural breaks than *fig. 2*. Notice the nice, unimodal llf. It is the central limit theorem working its magic on the likelihood function.

Unfortunately, political scientists generally are forced to work with relatively short time windows. A typical time series for an economist might consist of the better part of a decade's worth of daily price data. A political scientist, on the other hand, feels blessed to have a T of several hundred. This makes the numerical approximations of the maximum likelihood estimates all the more important for our field.

The instinctive solution is to turn to the Markov chain Monte Carlo (MCMC) methods regularly employed in Bayesian settings to fully flesh out the distribution of the likelihood over the parameter values. The process of fitting a GARCH model, however, involves comparing so many models that this would be quite consuming of time and computer power. Typically, it is not prohibitively time-consuming. But the requisite time and computing power is further increased by the methods advocated here, where the very point is to create a rich tapestry of $\binom{k}{2}$ time-varying relationships from just k series.

Stochastic optimization algorithms are an enticing solution, though they are not well known in political science. The basic logic is similar to deterministic routines, like the popular Newton-Raphson-type methods. Deterministic routines at each step try to find a higher point in the likelihood function. Once no higher likelihood value can be found, the algorithm concludes it has converged to the parameter values that generate the maximum likelihood. Stochastic optimization routines work similarly, but inject a stochastic element into the direction and distance it jumps in the parameter space from step to step; and/or into the decision to accept the new point in the parameter space.

Threshold accepting algorithms (Dueck and Shueuer 1990) are a useful class of stochastic optimization routines. Instead of the algorithm demanding the likelihood be increased with each step, some small reduction in likelihood is permitted. The size of the tolerable reduction itself may be stochastically chosen. Commonly, the allowable reduction is fairly large at the beginning of the algorithm and is shrunk as the algorithm progresses. Often the allowable reduction in likelihood effectively disappears by the end of the algorithm meaning the stochastic optimization algorithm finishes off behaving similarly to a deterministic algorithm.

The value of such procedures is that they are often capable of reaching out beyond local maxima, but at a computational burden a fraction of that of standard MCMCs. Both types of explorations of the likelihood space share some similar underlying ideas. The relative utility is purely in computational efficiency.

Winker and Maringer recently proposed such a stochastic threshold-accepting algorithm.[4] I base the central algorithm in the forthcoming R package, ”stochOptim” (Judge) on theirs, though with some significant modifications. stochOptim takes the same arguments as R’s optim procedure[5] and returns the same values, plus significance values based on individual Wald tests of each parameter. Winker and Maringer’s algorithm works as follows,

1. compute the likelihood at the initial parameter values λ.
2. choose a neighboring parameter value $N(\lambda)$, by altering one parameter. A parameter is selected at random, and altered by a value $\rho : -u > \rho < u$
3. Calculate L_{i+1}, the new likelihood
4. If $L_{i+1} > L_i - \tau$, accept L_{i+1}; if not, revert to L_i
5. reduce u and by a very small amount
6. repeat 2 - 6.

This algorithm can run for an arbitrary amount of time. Its effectiveness is sensitive to the initial value of u, tau, and the rate at which they decline, as well as the overall number of iterations (though experience has shown that there is generally little need to tweak the default values of stochOptim). rho is drawn from a uniform distribution with boundaries at u and u.[6]

Winker and Maringer design their algorithm to make a point about the convergence properties of stochastic optimization routines. I noticed three inefficiencies in their design when the algorithm is applied in practical research settings. First, the starting values are chosen randomly from the parameter space. Second, allowing only one parameter to change value at a

[4]Winker, Peter, and Dietmar Maringer. 2009. The Convergence of Estimators Based on Heuristics: Theory and Application to a GARCH Model. Computational Statistics 24(3): 53350.

[5]Indeed, part of the algorithm relies on calls to Rs optim command. There are several additional commands that can be called. For instance, a specification of the number of iterations. The default is 5,000. In addition, if the default function (a GARCH function, with or without exogenous variables) is optimized, starting values need not be supplied.

[6]More complex distributions make the process of tuning the values of u, τ and their decline rates less intuitive

time can keep the algorithm from converging. For instance, it was noticed while watching the algorithm converge[7] for a set of GARCH models that, especially early on, the algorithm might choose a poor value for one parameter which would then render the alterations to the other parameters pointless until that parameter jumped out of the space that created the valley in the likelihood. This slows the algorithm down and can even can keep it from converging, if the values are poor enough; e.g. inappropriately in negative territory.

Third, giving a role at appropriate points for a more efficient deterministic algorithm reduced the overall number of necessary computations. Once the general "hill" of the global maxima was found, reverting to the Nelder-Mead method[8] proved more efficient.

To address the first criticism of Winker and Maringer, stochOptim begins by choosing starting values for GARCH models with a three-step process.

1. A first guess of the AR values are found by least squares and then attenuated slightly.

2. Reasonable starting values for the MA parameters are derived algebraically by the equation for the mean variance. Using a $(1, 1)$ model as an example, $Var(y) = \alpha_0/(1 - \alpha_1 - \beta_1)$.

3. The Nelder-Mead method was run to find a local maximum, which more often than not proved to be the global maximum. The final values of the Nelder-Mead algorithm were used as the starting values for StochOptim. This has the additional advantage of checking the local maximum with the largest jump distances.

For the second criticism, Winker and Maringer's steps 2-6 were altered by allowing the algorithm to, every twelfth run, "reach out" and shift the candidate parameter values in more than one

[7]For two or three dimensional parameter spaces, or for problems when two or three parameters are of particular interest, the author realized that in R one can record the values from any optimization routine and then "play" them in a loop of hundreds or thousands of plot commands. This is much more efficient then scrolling through a large matrix of parameter values. When there are tens of thousands of dots plotted in the parameter space, it is best to use a new plot for each step in the loop (though the image is that of a moving dot, since the plots change so quickly). When there are less, one can add a new point at each step of the loop. Either way, the sensation is of watching the optimization routine explore the parameter space. If the points are kept on the screen, they can be color coded by step number so as to record the progress of the algorithm

[8] *R*'s default optimization algorithm.

dimension simultaneously. The number of dimensions, which dimensions and the euclidian distance from the previous point were chosen from uniform distributions. In this step, the maximum distance the algorithm would reach out from the starting point was set at $u * .5*$ (the number of parameters stochastically chosen). So, every now and then, the algorithm would experiment with odd angles in the parameter space, and, less often, reach further outside its neighborhood than is typical.

Finally, once the algorithm was run for a given number of iterations, the Nelder-Mead algorithm is called, using as starting values the maximum likelihood estimate found thus far. The following list gives the pseudo code for stochOptim.

1. Estimate AR initial values
2. Estimate MA values algebraically
3. Run the Newton-Raphson algorithm with starting values from (1) and (2) to find the point in the likelihood space, L_i,
4. Choose an element of the parameter vector, λ; every 12th step choose multiple elements in λ.
5. Choose ρ: $-u > \rho > u$; every 12th step select multiple 's
6. Form the neighboring parameter value $N(\lambda) = \lambda + \rho$
7. Calculate L_{i+1}, the new likelihood
8. If $L_{i+1} > L_i - \tau_i$, accept L_{i+1}; if not, revert to L_i
9. Reduce u and by a very small amount
10. Repeat steps 4-9 a large number of times

11. Choose the value with the highest likelihood found thus far and use it as a starting point for the Netwon-Raphson algorithm.

The messiness of real-world data, particularly political scientific data, relative to the simple simulated data that is commonly found in published work was well illustrated by the testing of this algorithm. A Monte Carlo of 1,000 simulated GARCH(1, 1) series, with true parameter values chosen from a $N(.35, .15)$ distribution, constrained by $1 > \alpha_1 - \beta_1$. The mean gain in likelihood scores of stochOptim over the Nelder-Mead algorithm - with starting values supplied in steps 1 and 2 - was not remotely significant. For the vast majority of series, the deterministic and stochastic numerical algorithms settled on the same parameter values.

The gains from stochOptim come, however, when the numerical optimization task isn't trivial that is, when multiple local maxima exist. The first-stage GARCH estimation of the forty-three NAES series were used as trial balloons. A set of GARCH models of different orders, and some with exogenous variables and some without were run using the forty three series that are the focus of this dissertation. The gains were, in some cases, notable. The likelihood is improved in roughly a quarter of all the models. In roughly ten per cent, the likelihood is improved by at least one percent of the total likelihood.

3.4 Sampling Error

Two facts raise the importance of sampling error when studying public opinion time series. Both are straightforward but have received little attention in the literature. First, as a practical matter, time series of public opinion usually have sampling sizes that vary over time. Were the sampling size to be constant, parameter estimates would be biased but at least in a consistent, non-confusing fashion. With notable variance in sample size, however, what may look like genuine change in the series is purely an artifact of changes in the sample size. Things start to look nefarious when one considers that sampling size and the phenomena under study may not be independent - as is the case with the NAES data. Important periods during a presidential campaign, for instance, tend to receive much more polling attention than less eventful periods.

Second, time series are often differenced before the real research begins. This may increase the portion of the observed variance that is generated by sampling error. To explain, consider an observed series, Y_t , that is the sum of an underlying process, y_t, and some sampling error, e_t.

$$Y_t = y_t + e_t \tag{3.2}$$

Sampling error is assumed to be a white noise process. That is, $E(e_t) = E(e_t e_{t-k}) = 0 \,\forall\, k = 0$; but, with a variance greater than zero. In a case common with survey data,

$$E(e_t^2) = \sigma_e = p(1-p)/n \tag{3.3}$$

where p is the population proportion and n the sample size. The variance of a differenced series is the typical formula for the variance of a difference,

$$Var(\Delta Y) = Var(Y) + Var(LY) - 2Cov(Y, LY) \tag{3.4}$$

where L is the lag operator, and $\Delta = (1 - L)$ is the difference operator. Another way to describe the left hand side of (3) is,

$$Var(\Delta Y) = Var(\Delta y) + Var(\Delta e) - 2Cov(\Delta e, \Delta y) \tag{3.5}$$

Since $Cov(\Delta e, \Delta y) \approx 0$ by assumption, the right-hand side of (5) can just be expressed as, $Var(\Delta y) + Var(\Delta e)$. Now, $Var(\Delta y) = Var(y) + Var(Ly) - 2Cov(y, Ly)$; and, $Var(\Delta e) = Var(e) + Var(Le) - 2Cov(e, Le)$. Generally, for lengthy series, $Var(e) \approx Var(Le)$ and $Var(y) \approx Var(Ly)$. So, we can simplify these equations further to

$Var(\Delta y) \approx 2Var(y) - Cov(y, Ly)$; and $Var(e) \approx 2Var(e) - Cov(3, Ly)$. By the definition of white noise, though, $Cov(e, Le) = 0$, and so $Var(\Delta e) \approx 2Var(e)$ However, the same cannot be said of the covariance of y with its lagged values. That is, $Var(\Delta y) \neq 2Var(y)$. So, finally, (5) can be expressed as,

$$Var(\Delta Y) \approx 2Var(y) + 2Var(e) - 2Cov(y, Ly) \tag{3.6}$$

In short, the portion of the variance of the differenced series is amplified because the noise element lacks any autocorrelation that would be subtracted from the differenced series' variance, while the same cannot at all be said about the autocorrelation of the underlying process. Since there is no covariance in the white noise process, $Cov(Y, LY)$ should be a close approximation of $Cov(y, Ly)$. Because of this and (6), the researcher can easily calculate the magnitude of differencing's amplification of the role of sampling error in a series as,

$$Var(\Delta Y)Var(\Delta Y) - Cov(Y, LY) \tag{3.7}$$

When political scientists do turn to time series analysis of public opinion data, they are perhaps most familiar with using Kalman filtering to deal with sampling error. Jong and Penzer (2004) show the steady state space representation of ARMA and ARIMA models that is possible when the summed coefficients of the ARMA model lie on or within the unit sphere. When already working in the ARMA framework, using a Kalman filter is vastly less parsimonious and at best carries little benefit to adjusting for sample error within the ARMA framework.

GARCH processes are those whose squared residuals follow an ARMA processes. Granger and Morris (1976) explore the nature of ARMA processes that are the sum of multiple series. One such sum is of particular interest for our purposes, since (2) often takes such a form. They show that an ARMA (p, q) series added together with white noise can be represented as an ARMA(p, p) series if $p > q$, and simply a (p, q) model if $q > p$. The act of adding additional MA parameters to deal with sampling error makes good intuitive sense: The additional moving average parameter(s) picks up the white noise component and moves the expectation for time $t+1$ back towards its mean by an amount proportional to the expected sampling error.

So, if the sampling error is largely constant throughout the series, the additional parameters in the univariate GARCH series will do a good job of filtering out the extra, sampling error-derived variance by themselves. Unfortunately, even with the NAES, one of the most impressive surveys of public opinion, the daily sampling size varies considerably. Fig. 3 shows the sampling size for a high-response question throughout the lifespan of the 2008 NAES.

To deal with the problem of the over-time variance of the sampling error-derived variance, an exogenous variable along the lines of,

$$x_t = \frac{mean(p)(1 - mean(p))}{n_t} \tag{3.8}$$

can be constructed for each GARCH series.[9] As a simple illustration of the importance of changing sampling sizes on the heteroskedasticity of time series of survey data, some simple bivariate regressions are run. The squared residuals from the 43 NAES series used here are the dependent variables in 43 equations, their respective x's from (8) are the independent variable. *Fig.* 3 reports the r–squareds from those regressions. The mean r-squared was .15, while the maximum was .25. The sample size variable was included in all of the first-stage GARCH equations, and, not surprisingly, was nearly always highly significant.

3.5 Estimating the First-Stage Models

DCC modeling can be – and usually is, for computational reasons – broken down into two steps. The first models the volatility of each univariate series. The second takes as data the

[9] At first glance, using the $p(1-p)$ from each time t may be tempting, but then an independent variable would be constructed that is functionally derived in part from the contemporaneous dependent variable.

normalized that is, the demeaned time series divided by the time series of the square root of the dynamically conditioned variances derived from the first step - and models the correlation series.

This bifurcation of the modeling process allows for greater flexibility over which model to use. This is flexibility is particularly useful for data from political processes, which, speaking broadly, enjoy less homogenous dynamics than the financial time series upon which the DCC is typically applied.

With the 43 NAES series, the bulk of the GARCH series were stationary or trend stationary. Those whose dynamically conditioned second moment were indeed integrated were modeled with an integrated GARCH (IGARCH) model, with grid searches approximating the parameter values for the exogenous variables. These IGARCH models provided remarkably good likelihood values, for series that otherwise were very difficult to numerically optimize, with *llf* values that were comparable to the optimized stationary variables. *Appendix II* goes into more detail.

Questions of differencing, over-differencing, trend-stationarity, fractional differencing and model accuracy in the context of mistakenly differenced or non-differenced series are covered by a very large literature (e.g. Box-Steffensmeier and Smith 1996; Baillie 1996). That literature need not be rehashed but there is one factor that is rarely mentioned that is of importance here. The volume of modeling that needs to take place when creating a correlation-based narrative of mass political behavior is very large.

That point is pertinent for two different reasons. First, and most obviously, for the researcher to complete her work, some cautious mass-production of model-fitting should be tolerated. Second, and perhaps more importantly, there is an inferential benefit to, when appropriate, a preference for uniformity of decision-making regarding differencing. If some variables are differenced while others are not, while others still are fractionally differenced with different values, interpretation of findings becomes fairly tricky. The substantive interpretation of a correlation at time *t* among two variables-in-levels or variables-in-differences straightforward to the point of triviality. Substantive interpretation becomes less parsimonious if one series is differenced and the other isn't, or if fractional differencing is introduced.

The correlation series produced by DCC models are effectively one-step-ahead forecasts of the correlation matrix within the dgp at time *t*. Bonnie Ray (1993) examines estimation of ARMA models on non-differenced data in the presence of fractional integration. She uses a battery of simulations to assess the increase in forecasting error for 1 to 20 step-ahead forecasts that occurs when fractional differencing is ignored. The length of her time window, 400, is similar to the T of the NAES, 316, and is useful for daily time series of public opinion during elections.

Speaking of the longer step-ahead forecasts, she concludes,

> Using an AR(p) model of moderate order (e.g. $p = 8, 9, 10$) to describe a fractional noise process does not cause much loss of long-range forecasting accuracy, indicating that the ease of estimating an AR model would seem to outweigh the extra effort required to estimate the fractional model in samples of moderate length.

Those findings are considerably stronger for shorter-range forecasts, such as those of concern here.[10] At the same time, poor model fits can lead to misleading results. If IGARCH models in the first stage do not successfully normalize the data,[11] some solution, such as differencing or fractional differencing, should certainly be introduced. The point here is merely that overfit DCC models can give some wiggle room to the researcher looking to maximize the inferential convenience of correlation series.

In terms of the mass-production of fitting the univariate first-stage models, fortunately, the parameter values from the volatility stage of DCC estimation is of little inferential value when the goal is creating a data set of a $\binom{2}{k}$ correlation series. If the time window is long enough, the best efficiency-accuracy trade-off seems to be to overfit the models and then check the normality of the

[10] See *Table I* in Bonnie Ray (1991, pgs. 515 517) for a related distribution of increases in forecast error over AR(p) models and values of *d*.

[11] To recall, a purpose of the first-stage estimation is to normalize the residuals by dividing them by their time-variant standard deviations.

residuals. If certain series present problems, those can be investigated, and, for the rest, the overfit models can be accepted. This convenience should be weighed against the loss of observations that higher-order models produce. The two-stage estimation of the DCC requires $p + q$ lost observations from the (p, q) used in the first stage modeling, and then, again, another $p + q$ lost observations from the second stage model order.

Finally, there is the question of how to assess the fit of the first-order models. From the standpoint of fitting the correlation series, the important factor is the normality of the residual series that is produced by dividing the original de-meaned series by its (time varying) standard deviation. This is the proof-is-in-the-pudding standard of the differencing and volatility modeling choices the researcher makes in the first stage of estimation. So, I prefer comparisons of the distribution of standardized residuals against a normal distribution to check for errors in the modeling process. If the approximation to the normal is only a rough approximation, however, the maximum likelihood estimate in the second stage still does not lose its useful properties. Instead, it retreats to the status of a Quasi-Maximum Likelihood estimate (Engle 2002, pg. 342).

3.6 Estimating the Second Stage Models

The second stage of the modeling process models the correlation component of the maximum likelihood estimate of the time-varying covariance matrix. In the multivariate setting, the correlation matrix must be positive definite. This is a particularly tricky problem because a single non-positive definite matrix produced throughout the entire time series of matrices, during just one of the iterations of a numerical optimization routine, will provide difficulties for the simple optimization algorithms.

Different forms of the DCC offer different ways to abide by the positive definiteness requirement. Ding and Engle (2001) cover the various forms of MGARCH models on which one can base the DCC.[12] The simplest and most common is the scalar model, often referred to as the two-parameter model. This is the model found in most software packages. For sake of simplicity, consider the first order model,

$$H_t = (1 - \alpha - \beta) + \alpha e_{t-1} e_{t-1}^{'} + \beta H_{t-1} \tag{3.9}$$

Where α and β are scalar parameters. Here, the parameter values must be optimized via constrained optimization. The summed value of the parameters must be at or within the unit sphere, and, effectively, the parameters cannot be negative.

Forcing each element of the covariance matrix to be affected by the same, scalar parameter makes sense if you are modeling small-to-medium sized covariance matrices of, say, equity prices from the same industry, or some other grouping that is known to create similar dynamics among each element. If, however, the researcher seeks to model a covariance matrix whose elements are without *a priori* reason for enjoying similar dynamics, then this assumption is dangerous. For instance, is it realistic that the covariance between party ID and ideology reverts back to its mean during presidential campaigns at the same rate as the covariance between perceptions of the candidates and voting intention? The answer may be yes, but it may very well not be.

Similarly, especially in higher order models, the assumption that the parameter values will be positive may be inappropriate. For example, perhaps candidates who benefit from the current structure of public opinion during a campaign tend to react strategically to changes in that structure by attempting to reverse yesterday's shocks to the structure. If their reactions tend to be effective, then this might be picked up as a negative value for α or β. This may or may not be true, but the point is that to restrict the parameters on the off-diagonal elements of the covariance matrix to above zero without some theoretical reason to do so seems dangerous.

[12] Variance targeting, a method assumed here for deriving the mean correlation values, is discussed in Engle and Mezrich (1996)

The vector-diagonal model offers a particularly elegant solution to this problem. It is,

$$H_t = (ii^{'} - \alpha\alpha^{'} - \beta\beta^{'}) + \alpha\alpha^{'} * e_{t-1}e_{t-1}^{'} + \beta\beta^{'} * H_{t-1} \tag{3.10}$$

here, *i* is a 1 x k vector of ones, α and β are vectors of parameters, and * is the Hadarmard, or element-by-element, product. The outer product of the parameter vectors form matrices of rank one. This has two major advantages. First, each element of the covariance matrix isn't forced to have the same coefficients. The ijth element of the $e_te_t^{'}$ matrix, for instance, is multiplied by $\alpha_i\alpha_j$. Second, off-diagonal values of the matrix are allowed to be negative. So long as $(ii^{'-\alpha\alpha^{'}-\beta\beta^{'})}$ is positive definite which in practice is effectively a requirement that the series are stationary in their second moments (9) will be positive definite. Finally, there is the BEKK (Baba *et al.*, 1991) model,

$$H_t = (ii^{'} - A - B) + Ae_{t-1}e_{t-1}^{'}A^{'} + BH_{t-1}B^{'} \tag{3.11}$$

The BEKK model offers the greatest flexibility of the models listed here, and guarantees positive definiteness, assuming stationarity. In practice, though, it uses so many parameters that anything beyond first-order models on small or medium sized covariance matrices is difficult to numerically optimize.

All of these models can be used in the R package put forth here. There is an inherent tradeoff between model flexibility and the ease of convergence. Depending on the size and nature of the data being studied, each may be appropriate in different circumstances. For instance, larger parameter values may entice the numerical optimization routine to explore the parameter space beyond the unit sphere and so derail the optimization routine, even if their maximum likelihood estimate lies within the unit sphere.[13] In this case, the two-parameter model might be able to

[13]The package presented here has a few safeguards against a small portion of the T covariance matrices traveling into non-positive definite territory, but no routine can effectively handle a significant portion of all matrices being non-positive definite, and the final outcome of the optimization routine must produce positive definite matrices for all T matrices. For more complex DCC models, the constrained optimization routines in R have difficulty constructing numerical approximations of the gradient near the boarders of the parameter space. Supplying analytical gradients for

converge where more complex models fail to do so.

Because the decision of which model to use relies heavily on idiosyncrasies in each time series, which order the modeler prefers, the size of the data set, and still other issues, no software package can automate the decision making process for the researcher. Some trial-and-error is be necessary.

In practice, the vector-diagonal model was found to usually offer the best trade-off between flexibility and convergeability. This is particularly true with the inclusion of exogenous variables, which is an active topic in the DCC literature; and is allowable in several different forms in the DCC package put forth here.

Especially with the two-parameter model, it is easy to produce reasonable-seeming models that converge properly. Without some theorizing and, in most cases, testing of the appropriate model, though, inference based on such models are built on shaky ground. Fortunately, the practice of choosing which DCC variant to implement forces the researcher to think about the nature of the dgp in a manner that often enhances her substantive, theoretical thinking.

The next issue to consider is the size of the correlation matrix. The researcher has the option of modeling the entire $nxnxT$ array of correlation series in one fell swoop. Alternatively, she may choose to run a number of models with smaller correlation matrices among subsets of the variables, filling the *ij*th element of the larger matrix in with values taken from each matrix. The latter method allows for greater flexibility in parameter values, but is considerably more labor intensive.

The author is not aware of formal methods designed particularly to test for the appropriateness of the size of the covariance matrix. One can, however, think of the size of the covariance matrix as a form of model restriction. This is literally true since, in all of the above models, the maximum likelihood estimate of one parameter acting on a given covariance series is restricted in some form by its need to contribute to the maximum likelihood estimate of the

the many different model orders and sets of exogenous variables is possible, but a prohibitively laborious task.

parameter acting on another element of the covariance matrix. This reasoning suggests an LR test.

3.6.1 A Proposal for a Likelihood Ratio Test for the Size of the Covariance Matrix

In the first step, the full covariance matrix is run with n variables. In practice, convergence issues prevent models other than the two-parameter model from going much beyond twenty variables. In the next step, the same model is run with some smaller number of variables, nk.

Third, the likelihood score is calculated for the smaller matrix, but at the parameter values from the larger model. Finally, in the fourth stage, the likelihood scores of the two models are compared, with $n - (n - k)$ degrees of freedom. Notice also the outcome of this test will often depend on whether (8), (9), or (10) is used. The DCC package presented here provides a simple command to implement the LR test, once the two models are fit.

3.6.2 Tweaking typical Modeling Concerns

The researcher must answer a fundamental question before estimating the second stage models. Is the goal of the second-stage modeling process inference or the creation of the most accurate possible time series of correlations, which will then be used for some subsequent purpose? If the latter, there are several repercussions minor revisions to the textbook modeling approach wherein the typical set of modeling concerns is weighted differently. For instance, the parsimony of models that are not referred back to once they generate their correlation series is much less important.

Similar to the first round fitting of the variance equations, overfitting is a concern insofar as it makes the optimization routine more difficult, and shortens the time series being studied. Overfitting does not run up against the parsimony bulwark in the same way that it does when the model itself is at the center of the inferential task.

In the DCC package written for this dissertation, structural breaks and time trends, when included in the final model, can be first estimated by OLS on each element of the covariance matrix. This keeps the computational burden for larger models manageable. It also facilitates assessing the significance of structural breaks and deterministic time trends. The package retains the option of estimating deterministic time trends and structural breaks in the larger MLE equation.

Survey data rarely undergoes the entirety of a structural break in a single day. Information diffusion and political change takes time. For instance, it's hard to identify a precise day in 2008 when the Democratic presidential primary season effectively ended and the general election began. Therefore, the package put forward here allows the user to specify the length and shape of the structural break. The default length is for a one-week linear break, but any day or shape can be easily specified; e.g., a "hard" break of a single day, or a smooth, polynomial break over a period of weeks.

3.7 Aggregate-Level and Cross Sectional Variance

Before the 43 NAES series were fit with DCC-GARCH models, time series of the daily cross-sectional correlations were created. This was initially done to create a benchmark against which the DCC-derived correlation series could be compared. The NAES is so huge that large sample properties applied to both the survey as a whole and the vast majority of the 316 daily observations. This allows for the very rare opportunity to compare time series of daily cross-sectional parameters and their aggregate variance-derived DCC counterparts using real-world data.

Importantly, with many hundreds of correlation series lasting hundreds of days each, a comparison can be made between the time-invariant cross-sectional covariance (roughly 58,000 respondents) and the covariance of the time series operationalized daily values (309 observations). So, the comparison between dynamic and cross-sectional correlations can be neatly disaggregated into two, distinct parts. (1) The relationship between cross-sectional covariance and aggregate-level, over-time covariance. (2) the relationship between DCC-derived correlations and their daily cross-sectional counterparts.

With the answers to (1) in hand, (2) is a straightforward question of assessing model fit. A particularly important benefit of addressing (1) first is that we can see the change in behavior of a variable that comes from its time series operationalization. There are a number of other benefits. Perhaps most important, the dialogue between cross-sectional research and time series research in political science is, as noted in the introduction, an important feature of recent political science. As time series data grows in abundance, this dialogue will likely become all the more important.

The relationship between cross-sectional and aggregate-level covariance is easy to characterize. *Fig.* 4a presents $\binom{39}{2} = 741$ correlation series from the 2008 NAES.[14]) correlation series from the 2008 NAES. 55 The y-axis is the cross-sectional correlations from the roughly 58,000 respondents. The x-axis is the time-invariant correlations of the time-series operationalized versions of the variables, with $n = 316$, the number of days in the sample. The correlation between the correlations and those of their time series operationalized counterparts is fairly tight, but with room for disjunction among any given two sets of correlation: .809. The cross-sectional correlations are, in their mean, .03 less than their time series counterparts, a value that is not statistically significance.

The sources of difference between the two versions of correlation can be broken down into two categories. First, there are a set of measurement issues. Correlations between ordinal variables with few categories will be attenuated. When a time series operationalization changes, say, a binary variable to a 0 100 variable, it removes that attenuation, pushing the correlation of the two types of correlation away from 1. Next, the time series operationalization itself will alter slightly the nature of the variance. Finally, the time series operationalization effectively weights

[14]The 39 series are from the total of 43, less four alternative operationalizations of party ID.

each day the same, despite changing sample sizes.

The second type of culprit is the (theoretically more interesting) substantive differences between the over-time, aggregate-level variance and its cross-sectional counterpart. There is no general rule saying that aggregate level covariance reflects identically the cross-sectional relationship among voters. That is, recall equation (1) is $\rho_{at} = f(rho_{ct})$, where ρ_{at} is is the aggregate level correlation at time t and ρ_{ct} is the cross sectional counterpart. Taking a closer look at *fig.* 4a., there is a set of correlations that cluster around zero at the cross-sectional level, but whose correlations stretch out as far as -.5 and .5 at the aggregate. These are, by definition, relationships that covary over time considerably but not at the time- invariant individual level. They enjoy little relationship at the individual level but they develop at the aggregate level over the course of the campaign with some statistical significance.

So, how much of the aggregate-cross-sectional disjunction stems from measurement issues and how much from (1) being something other than an identity? A simple way to approximate the maximum role for measurement error is to observe the relationship between the cross-sectional correlations and the correlation of the differenced time series. The measurement error issues change little, but most of the trend is removed from the overall covariance. This relationship is plotted in fig. 4b. The overall correlation of correlations increases from .809 to .934, closing in on a level of disjunction between the aggregate and cross-sectional that could be regarded as an acceptably small amount of measurement error. Indeed, gone in fig. 4b. is the subset of variables that cluster around zero in the cross-sectional correlations but not for their aggregate-level counterparts. 95% of all time series correlations can be found within .18 of their corresponding cross-sectional correlations.

In sum, most variables closely mimic the cross-sectional in their aggregate behavior, in a time-invariant sense. There are some that do not, and for the 2008 NAES, that disjunction seems to explain roughly twelve per cent of the overall correlation between time-invariant aggregate and cross-sectional level correlation; while the remaining six or seven percent can be attributed to various measurement issues. This is a rough approximation, but gives a good idea of the scale of the underlying sources of disjunction between the cross sectional and the aggregate. It serves as rough but useful evidence of the precision with which time series methods can directly speak to cross-sectional findings.

Question (1) concerned the relationship between cross-sectional and aggregate level variance, both conceptualized as time-invariant. Now that that relationship has been laid bare, at least insofar as the NAES, question (2) should be addressed, which is effectively question (1) but with the assumption of time-invariance removed. That is, the question of the relationship between the DCC-derived correlation series and the time series of daily cross sectional correlations.

Correctly–specified DCC models have been shown to be very accurate representations of correlations. That changes little in this circumstance. Further, the accuracy increases as the over-time variance of the correlation series increases: As always, the greater the variance, the greater the variance that can be explained, and the DCC functions best when we need it to. The DCC perform the worst when the over-time variance of correlations is minimal, leaving the fluctuations in the DCC to be largely white noise. Such instances are obvious when visually inspecting the correlation series.

The correlations that vary the most over time can correlate with their DCC-derived counterparts by an impressive .9. Further, the mean of the daily correlations is close to perfectly captured by DCC model, thanks to the variance-targeting design (Engle and Mezrich, 1996) of the models proposed here.

In sum, the vast majority of the difference between DCC-derived correlation series and their cross-sectional counterparts come from underlying differences between aggregate and cross-sectional variances, not measurement error induced by the modeling process.

fig. 2. likelihood function of a series with a large break in the conditional mean

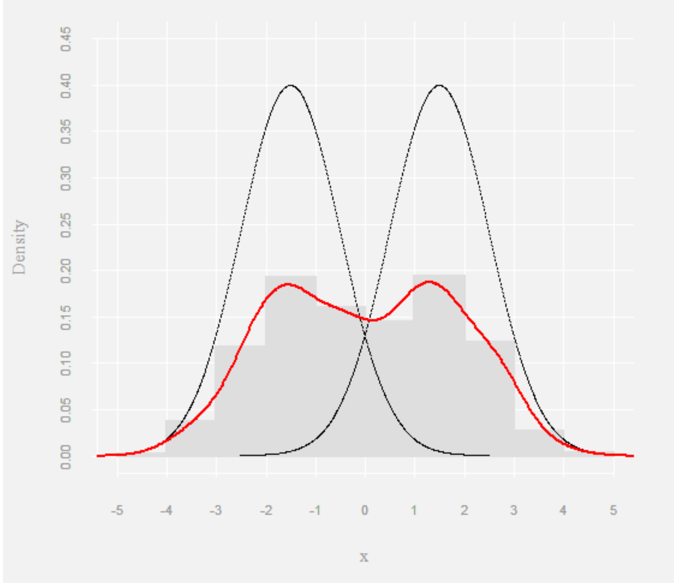
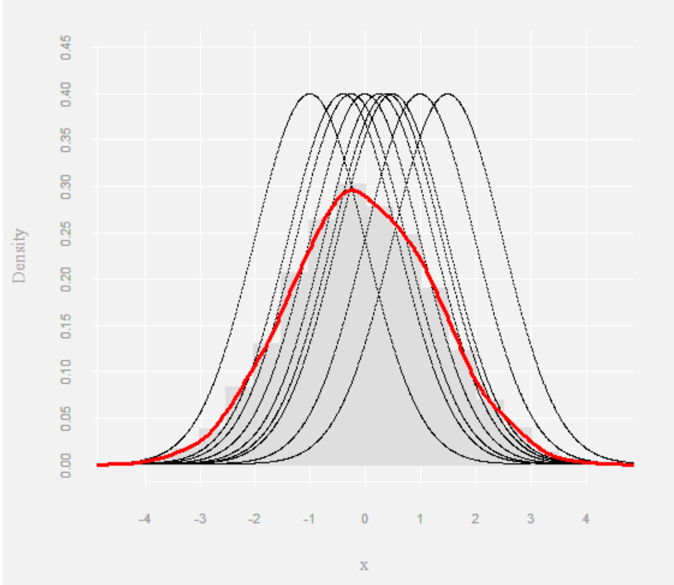


fig. 2b. likelihood function of a series with ten large break in the conditional mean



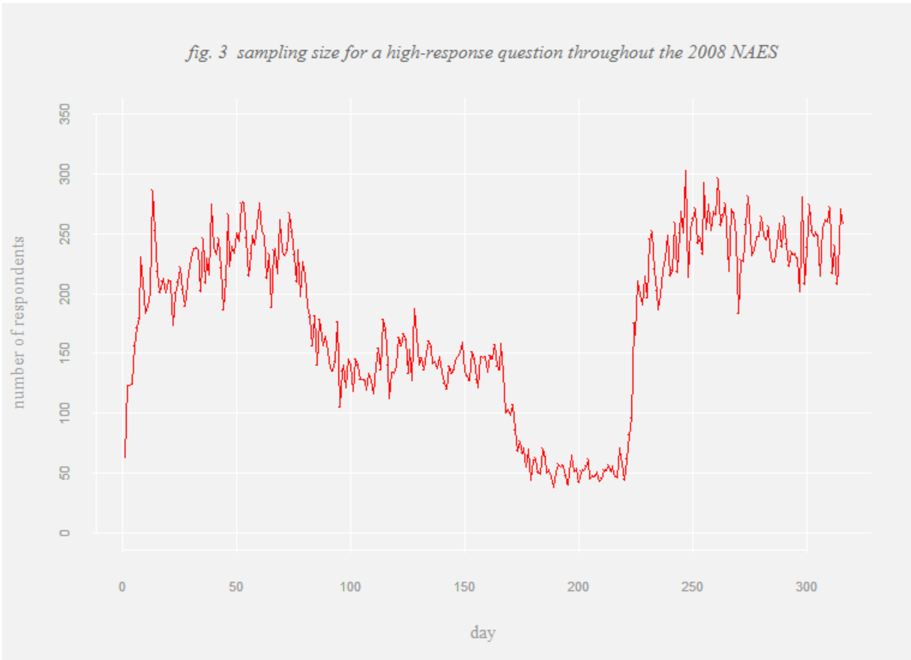


fig. 3 sampling size for a high-response question throughout the 2008 NAES
number of respondents
0
50
100
150
200
250
300
350
0
50
100
150
200
250
300
day

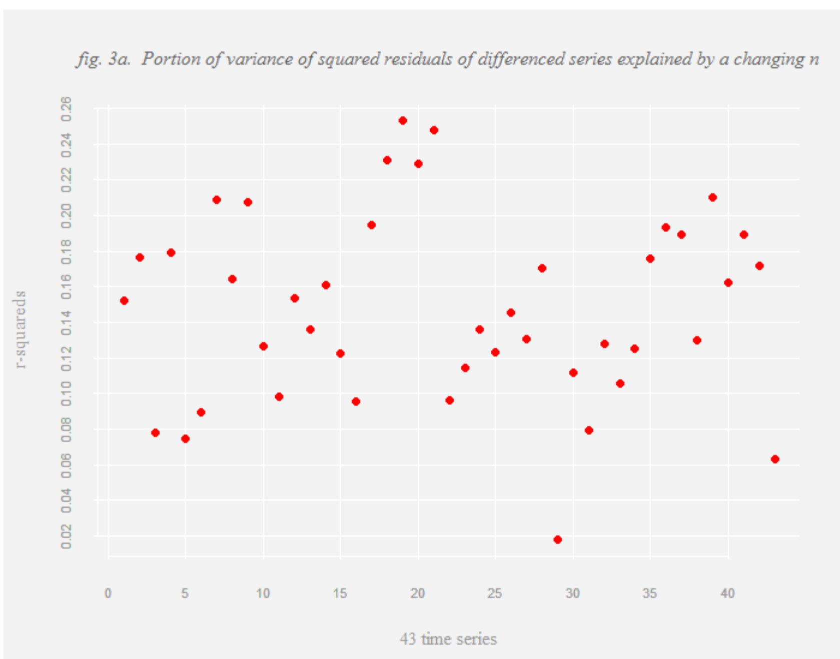
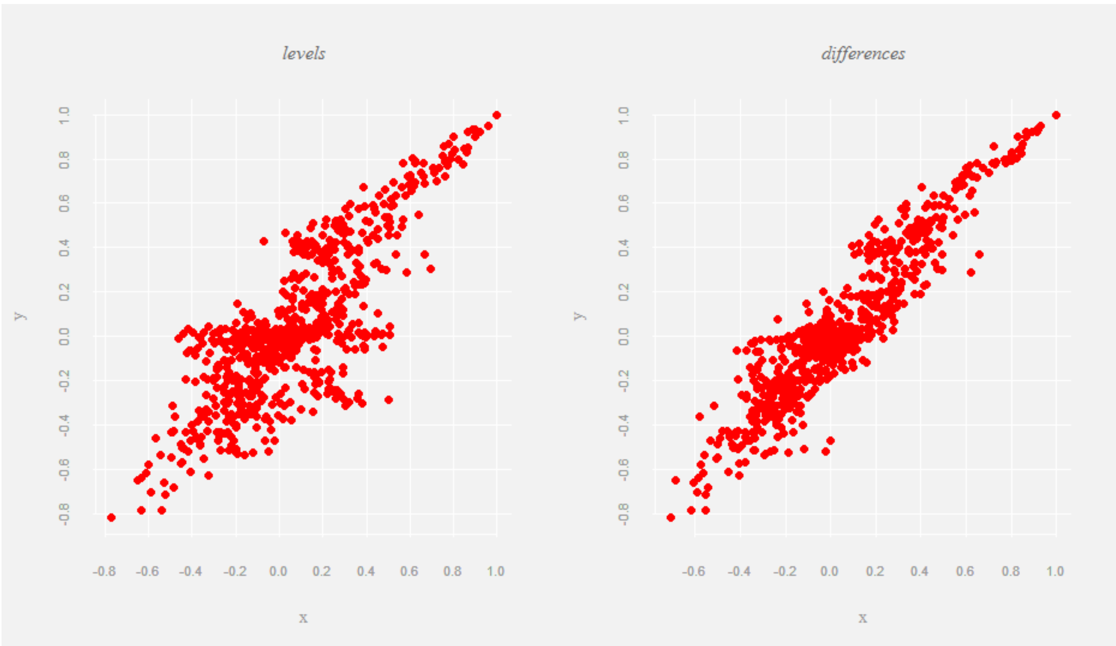


fig. 3a. Portion of variance of squared residuals of differenced series explained by a changing n
r-squareds
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0.16
0.18
0.20
0.22
0.24
0.26
0
5
10
15
20
25
30
35
40
43 time series

fig. 4a and 4b



Chapter 4

The Structure of Public Opinion During American Presidential Campaigns

4.1 Introduction

The literature on American presidential campaigns is one of the largest in political science. Still, the field lacks an understanding of some of the basics of the evolution of public opinion during the course of the campaign. Since the earliest studies of voting behavior, limitations of methods, computating power and, above all, data have prevented political scientists from describing some of the elemental aspects of how the structure of public opinion changes because of presidential campaigns.

During the first decade of the new millennium, that research environment changed not completely, but fundamentally. Daily time series of many of the dimensions of political behavior are now often available, and the methods now exist to straightforwardly analyze not just daily time series of aggregate-level values but the time-varying structure of covariance that undergirds those values, as well.

No one, however, has yet stepped back and simply described the changes the structure of public opinion undergoes during the course of the campaign. The literature has advanced considerably as data and statistical methods have grown in sophistication, but it has mostly moved further down preexisting veins of research. The basic, descriptive inferential task of just what happens to the structure of public opinion during campaigns has yet to be completed. It is as if explorers and cartographers made much progress mapping out a new land, but didn't demand of their fellow practitioners a new set of common understandings once satellites were invented.

Perhaps a reboot is in order. Four points of research methodology come together to form the appropriate pushing-off point. All three are pertinent only for some questions and in some research environments. First, when daily time series of public opinion are available, overly discrete notions of change entail tossing aside a tremendous amount of information. Change in a system does not need to be modeled and/or broken up into chunks when it can be directly witnessed as a time series.

Second, and relatedly, assumptions of the time-invariance of relationships or effects can be discarded. If relationships can be shown to be time-invariant, then that is a very useful finding, once shown. The more assumption-less, scientific starting point, though, is that relationships are subject to change. Describing relationships with parameter values that lump the year before the election together with the days leading up to the election is a pitfall in the research process that in many circumstances is now antiquated. Often, it turns out, it *is* appropriate but rarely is it appropriate by assumption. Further, delineating where assumptions of time-invariant relationships are appropriate turns out to reveal some of the most important characteristics of a system.

The third point concerns research when the object of study is a broader social system, like a campaign or markets, as opposed to a particular dependent variable, like vote totals or price behavior in a particular market. The proper perspective concerning the significance of findings should be not only statistical significance but the size of the findings *relative to change that is witnessed elsewhere in the dgp.*

One of the lessons taken from this dissertation is that the process of social scientific research lends itself to losing the forest for the trees. If a theory is popular, confirmatory findings that are just barely significant in a statistical sense can be granted a role of central importance in the literature, even if those findings are a tiny fraction the size of findings that support alternative theories. The methods proposed here begins with a widening of our lens and broadening our focus, as a necessary preliminary to digging into specific questions. There are many benefits to beginning with the descriptive inferential task of characterizing the changing structure of the dgp during the time window being studied. Not least among them is that the scale of findings is, by the design of the research process, given an intuitive, substantive component, along with the more common statistical component.

Fourth, the levels of variables are of course important, but the structure of relationships seems like the more generalizable and interesting question. Of course, applied research will always have an eye on explaining the levels of important variables, e.g. presidential approval or party identification. Literatures that ask questions broader than a single dependent variable, though for instance, how do campaigns affect public opinion all too often read like a pasting together of atomized sub-literatures, each centered on modeling different dependent variables in a system. This is a natural byproduct of multiple research agendas, but it is far from the most efficient way to accumulate a social scientific understanding of a given social system. Instead, focus should be on characterizing the structure of the system.

There are two concepts in that statement that need clarification, ”system” and ”structure.” A system can be usefully called a system when the structure of relationships maintains some consistency over cases. Values of variables can be expected to vary plenty, but there should be some consistency in the set of functional relationships that describes the system. For example, in the post war party system, unemployment might change considerably from election to election. If the same function can be used to describe the relationship between unemployment and incumbent party vote totals, then that's fine, from a methodological standpoint. The system is a system, taking different inputs along the way. If those (dynamic) relationships maintain little in the way of consistency, though, the usefulness of conceptualizing the set of relationships as a system is more or less gone. Each instance should be taken on its own and theory derived from one case will not help much to predict outcomes in another.

In short, defining a ”system” delimits the inferential range of a study. That’s a vague point of background interest. The notion of the ”structure” of public opinion or any other system is very important for the research presented here, however, and needs to be absolutely clear. Imagine a world composed purely of continuous variables and linear relationships between them. Whatever the value of variables, the structure of relationships is perfectly described by a covariance matrix at time t, re-formatted as a correlation matrix for interpretive convenience. A shock to the ith variable will affect the jth by an amount described by the ijth entry of the correlation matrix. This literal structuring of variance is why it is appropriate to consider the correlation matrix at a point in time as the operationalization of the structure of a system, as opposed to the set of top-line values of those variables.

With that, a notion of a social system and the operationalization of its structure is in hand. So, it can be seen why taking a step back from the values of variables to look at the correlation structure that undergirds their values is akin to stepping away from an investigation of the inputs into a system in favor of investigating the systemic structure itself. This is the argument for an explicit focus on the time series of correlation matrices being more likely to produce generalizable findings capable of building more cumulative literatures.

That argument is a sufficient condition for pursuing the methodology laid out here, not a necessary one. If perfectly true, then clearly this approach is warranted. If not, the relative novelty of such a focus may still well produce important insights. In practice, the operationalization of structure as a correlation matrix can fail for a limited number of reasons. (1) imperfections in the operationalization and measurement of variables; (2) non-linearity in effects; (3) mischaracterization of the correlation matrix; e.g. assuming the correlation matrix at time t equals the correlation matrix at time $t+1$; and (4) the correlation matrix can present a flawed picture due to omitted variables.

The imaginary world of continuous variables and linear relationships is surprisingly close to the reality of the research environment that confronts students of mass political behavior during presidential elections, if one adopts the research framework recommended here. The simplicity, uniformity and ease of communication of findings that results is one of the major benefits of this framework.

(1) and (4) are pitfalls in nearly all research. As regards (2), a world of continuous variables is accomplished easily by considering variables in their time series operationalized form. For instance, binary variables like agree-disagree questions become 0 - 100 scales measuring the portion of the population that takes a given opinion. This also changes the question from relationships at the individual level to relationships at the aggregate, systemic level. Some of the important characteristics of the relationship between the cross-sectional and the aggregate were covered in the previous chapter.

Finally, there is (3). Linearity as the functional form of relationships is an assumption that often breaks down. However, assumptions of linearity exist at some level in most research projects; even if just as local linearity, such as some approaches to semi and nonparametric modeling. In a similar fashion, doing away with the assumption of time-invariance makes an assumption of linearity much more tenable. Thinking this way, one can see a parallel between methods like the DCC that discard strong assumptions of time-invariance and semiparametric methods: Semiparametric methods allow the (often linear) relationship to vary over the values of the variables. Doing away with assumptions of time-invariance allows the (often linear) relationship to vary over the values of *t*. The mechanics are somewhat different, but the concerns are similar.[1] With that parallel in mind, one can feel much safer about assuming a world of

[1]In fact, the DCC is flexible enough to incorporate non-linear and nonparametric elements into its modeling of the correlation matrix; threshold effects being perhaps the most straightforward example. For an excellent review of Nonparametric time series methods in general though only briefly touching upon GARCH-related material see Hardle, Lutkepohl and Chen (1997).

linearity when assumptions of time-invariance are removed.[2]

4.1.1 Organizing the Exposition

This chapter's task is that of descriptive inference. With about forty variables in the mix, there are roughly nine-hundred relationships to be described. If organized well, reading this chapter will result in an understanding of the basics of what happens to the structure of public opinion during a presidential campaign. If organized poorly, it could be useless or even counterproductive. One way that organization could collapse on itself is if too many questions of causal inference are put into the mix. The ability to cover such a breadth of material demands that more detailed questions be left for subsequent research. The focus here must be on descriptive inference.

Even with such a broad brush, the possibilities are exciting. To give some important examples, the limits of campaigns-as-organizations will be clear, in terms of how far into - and where in - the structure of opinion they exact their influence. The characteristics of the change that does occur in the structure incremental and cumulative, or large event-driven is a telling element of presidential campaigns that is relevant to many different questions of political behavior. In short, the basic characteristics found here are substantively important unto themselves and they also form a very useful frame for future research.

With that in mind, the question becomes a practical one of just how to organize the exposition of such a large number of time series of correlations, or correlation series, for short. Blocking the variables off into appropriate groups seems a natural way to start. That way, the internal relations of each group can first be characterized, followed by the relationships between different groups.

Four categories or substructures suggest themselves naturally. These categories are chosen to

[2]Indeed, one advantage of Lebo and Box-Steffensmeier's (2008) use of correlation coefficients as dependent variables in a transfer function is that it also functions as a *de facto* test of the appropriateness of semi/non- parametric modeling of the relationships characterized by those correlation coefficients.

group together the major components of mass political behavior during an election. The separation of particular components of political behavior is not meant to imply any theoretical point but merely to give the organization of this chapter an intuitive feel. Those four categories are,

- Issue space: positions (e.g. abortion) and (non-politician) evaluations (e.g. state of the economy)
- party identification and ideology
- political evaluations
- voting behavior

With those four substructures, there are ten category relationships to be described: the relationships within the four categories, plus the six relationships between them.

The four sets of within-category relationships will be evaluated first. This is an important part of the description in its own right, and will also serve as a test of the usefulness of organizing the roughly forty variables into their respective categories.

In describing each of the ten sets of relationships, the first step is to find which relationships are exogenous to the campaign. This ”relationship exogeneity” will help simplify the analysis, narrow the focus, and, hopefully, reveal telling patterns. Next, the endogenous, time-varying relationships will be investigated.

Two points should be made about relationship exogeneity. First, the variables themselves need not be exogenous, just the relationship between them. Feasibly, two variables could appear largely exogenous from the campaign, perhaps varying only slightly within a tight band, but their relationship (correlation) could be profoundly endogenous to the campaign. Thus, relationship exogeneity is not a sufficient condition for exogeneity as it is typically used, but it is a necessary condition.

Second, the investigation below, being one primarily of descriptive inference, does not confirm that variance observed during the campaign is caused by the presidential campaign and the actors in it. For instance, some variables, particularly Iraq-related variables, seem likely caused by events causally exogenous from the campaign. This being a work primarily of descriptive inference, ”exogeneity from the campaign” employs the typical definition of exogeneity, but ”the campaign” is meant in this context as a period of time, as opposed to an institution or process.

Throughout, the use of graphical presentation of data is used to communicate the data and relationships. This done for the sake of efficiency. With so much information to present, it is necessary. As the previous chapter covers, some forms of the DCC are overly restrictive when an assumption of similar dynamics across correlation series is inappropriate. This danger grows with the size of the correlation matrix. Further, if some of the series undergo significant structural breaks while others do not, the problem becomes all the more serious. Models will usually converge, but parameter values will be too high for some series and too low for others.

The result is an awful lot of noise in each series, making some series appear volatile when they are not and some appear to change less than they really do.[3]

The package written for this dissertation presents an option for a three-step, automated solution for large correlation matrices. The aim is to provide as flexible a solution as necessary, while preserving enough automation to allow for the estimation of hundreds or thousands of correlation series at a time. First, the vector-diagonal model is used (Ding and Engle 2001), as outlined in Chapter 3. When estimating large matrices, however, the vector-diagonal model is still not all that more flexible then the vastly more common two-parameter model. This is because each parameter plays a role in its own variance series, and then, as an interaction effect with the other series' respective parameter, in the parameter acting on the covariance series of which that variable is a part. (see Ch. 3, pgs. 18-19). So, to communicate the problem imprecisely: With,

[3]Though in practice, DCC models tend to be noisier than we would like.

say, dozens of variables, each parameter will be pulled towards the mean of the values that would be the best-fit value for each covariance series. Hence, practically speaking, there will be little advantage to the harder-to-fit vector-diagonal model than the two- parameter model.

The second step, therefore, is to fit the entire correlation matrix through permutations of smaller matrices. In our case, the $43x43xT$ correlation array is fit by fitting a set of $8x8xT$ arrays. The $8x8$ matrix size was chosen because in practice it easily passed the likelihood ratio test proposed in Chapter 3; it seemed to be a computationally efficient size; and in validating the results of the LR test, the correlation series generated appeared very similar to, but a bit less noisy than, the correlation series derived from individual runs of $2x2$ models. The method of fitting a correlation matrix with larger matrixes presumably does away with the positive definite restriction for the larger correlation matrix and is far from truly computationally efficient, but allows for more accurate fitting of any given correlation series than other realistic alternatives.

In step three, each individual correlation series can be chosen among two different models, using a penalized likelihood function such as the AIC. The penalized likelihood is based on the likelihood of the $2x2$ matrix at the parameter values derived from the $8x8$ matrix. This choice between two different models greatly increases the model flexibility for each correlation series. It is particularly useful when some series undergo structural breaks or are trend-stationary while others are simply stationary.

For the NAES, both second-stage models were fit with an order of $(2, 1)$. One model was fit without any time trend or structural breaks, and another with a time trend and three structural breaks representing the end of the Republican primary season, the end of the Democratic primary season and the beginning of the post-convention season. The first break was a hard, one day break. This was important because the filter on some NAES questions changed the day McCain sealed the Republican nomination. The next two breaks were week-long linear breaks. That is, two variables were created that took the value zero until the break, and changed linearly over the course of a week to a value of 7.

Typically, each individual correlation series is chosen based on a penalized likelihood function of its 2 x 2 correlation matrix. However, with the NAES there is the luxury of using the time series of cross-correlation series as a benchmark. Thus, the decision to include structural breaks and time trends in the final model is based on the correlation of the cross-sectional and DCC-derived correlation series. For the NAES, the correlation series generated by the two models usually differed only slightly. However, a substantial minority would have been fundamentally mis-characterized without some nine hundred series getting the choice between two different models.

In double-checking the three-step, automated procedure of the package presented here (ibid.), none of the sample of 50 series from among the roughly 900 series appeared to be mis-chosen among the two different models. In short, the *R* package put forward here allows for the automated fitting of a large number of correlation series indeed, infinite, if a computer were to be run forever in a manner that gives the modeler a very large degree of flexibility in the functional form of the models and in parameter values for each ijth element of the correlation series. It can then automatically choose between two different models. With a benchmark of cross-sectional correlations, the automatic selection process appears to be extremely accurate. Further work needs to be done to refine the penalty on the likelihood function to generate the best automated choices in the absence of cross-sectional correlations as a reference. It is quite possible that no truly automatic procedure can be a substitute the painstaking work of fitting hundreds of models without some objective referent, such as cross-sectional correlation series. A battery of Monte Carlo experiments should easily answer this question, though is beyond the scope of this paper.

Many series still appeared too noisy. Optimizing higher order models may have produced less jumpy series. However, the process of numerically optimizing models of higher order than those used here proved to be too much for larger models that also took exogenous variables. For some of the reasons already outlined, the task of fitting many disparate public opinion series proved to be considerably more difficult in practice than fitting the time series of financial data with which the author had previously worked. This was one of the reasons the R package put forward here was written. To deal with the noise, some correlation series were smoothed with a moving average, usually of three days, never more than one week, i.e. three days on each side of t.

4.2 Within-Category Relationships

4.2.1 Issues and Non-Candidate Evaluations

The first category consists of ten variables. As is true for all the variables used in this dissertation, Appendix 3.*A*. outlines the details of each question, while 3.textitB. presents their time series operationalizations. Table I. presents the variables in this category.

Table 4.1

VARIABLE	CONTENT
econ1	sociotropic economic well-being
econ2	personal economic well-being
tax	taxes too high?
abortion	under which conditions should it be allowed?
citizenship	favor a path to citizenship?
fence	favor building a US-Mexico fence
marriage	favor gay marriage
Iraq	Withdraw as soon as possible?
Iraqworth	Was Iraq Worth it?
Iraqter	Leaving Iraq increase terrorism?

The defining feature of the structure of the issue space throughout the campaign is that of exogeneity from the campaign. From the first primaries until election day, the campaign leaves the structure of voter issue positions almost entirely unchanged. Candidates, campaigns and parties may or may not see a strategic imperative to rally voters into groups based on their issue positions, but the underlying structure of the issue space does not grow more cohesive or change in any notable fashion.

Fig. 1 provides a sample of four of the relationships. *Fig. 1C.* is particularly interesting in that even as the economy underwent historic shocks affecting the population and the nation at different rates, reflections on sociotropic and personal economic well-being kept to a very tight relationship.

Fig. 1B is included partly as an example of the danger of low-variance variables mixing with the heteroskedasticity caused by sampling error. The inclusion of a sample size variable proved to be very effective at dealing with changes in sampling size. Of course, it did not perfectly remove all the sampling error. This imperfection had a negligible impact most of the time. When the underlying variables barely move over the course of a process this leaves room for inaccuracy in any modeling process. The black and blue lines in *fig. 1B* are the underlying marriage and border fence questions, whose time series operationalizations (see appendix 3.b.) move very little over the course of 2008. The small portion of heteroskedasticity that remains from changing sampling size seems to briefly force the correlations upwards, though they return to previous levels soon thereafter.

Regardless of these methodological issues, though, the point should be emphasized: The structure of the issue space during the campaign, as recorded by NAES questions, does not change it is surprisingly exogenous from the campaign. Correlation series fluctuate only briefly and rarely, and quickly revert to the value around which they tightly equilibrate. As discussed in the introduction, this does not mean that variable values aren't affected by the course of the campaign, but that the relationship between them is not.

4.2.2 Party Identification and Ideology

This category is the smallest but, of course, quite important. It is composed of self- reported ideology on a 1-5 liberal-conservative scale and self-reported party ID, on the traditional 1-7

Fig. 1.

scale.

A defining trait of the 2008 campaign appears to be, again, the exogeneity of some of the fundaments of public opinion over the course of a presidential campaign. In this particular case, partisan identification and ideology. Indeed, the relationship between the two appear to be so entirely exogenous from the campaign that it would be hard to imagine this as a campaign-specific null finding.

These findings confirm on a shorter time horizon the study of Box-Steffensmeier, Knight and Sigelman (1998) that found no relationship among party ID and ideology at the aggregate level over the course of several decades. One interesting course for future research would be to see if there are similar temporal patterns in daily time series of the ”sophisticated electorate” as those found by Box-Steffensmeier and De Beouf (2001).

The findings are still surprising, though, since one might expect there to be at least some change as campaigns move from contesting primaries to seeking the presidency as a party’s national nominee. *Fig 2* presents the DCC-derived correlation series of ideology and the portion of the electorate that identifies with the Democratic party. The (imperfect) cross- sectional correlations the dotted black line are included in the background along with the more appropriate DCC-derived series. Alternate time series operationalizations of party identification made only slight differences in the mean level of the correlations and nearly none at all on the dynamics and direction or lack thereof of the series.

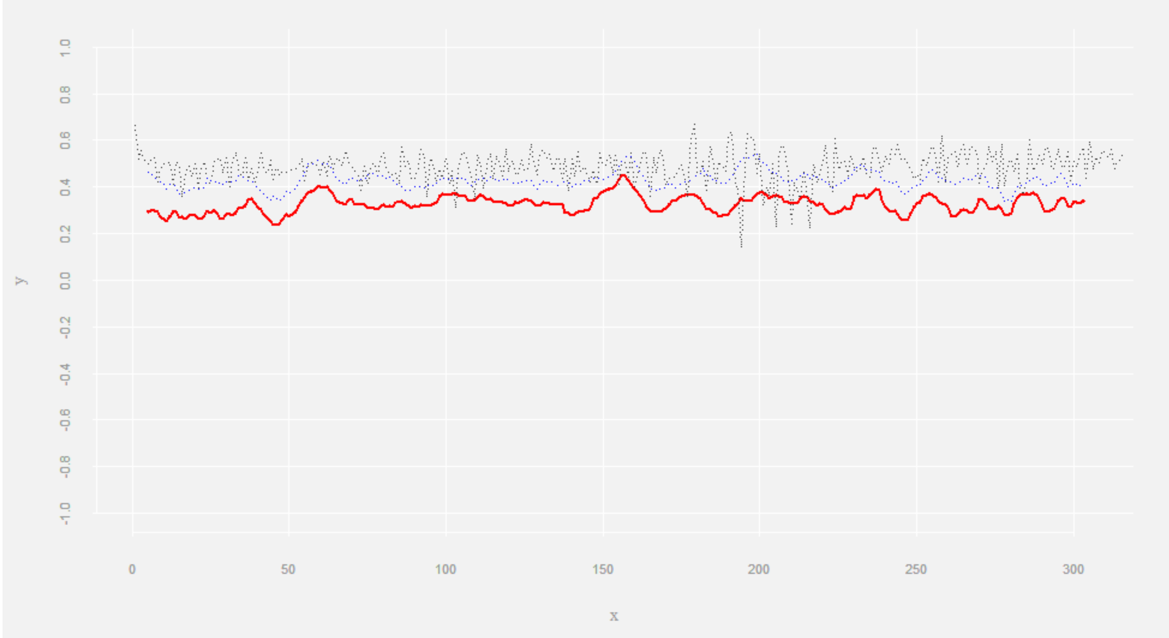


Fig. 2. Ideology and Party ID

4.2.3 Politician Evaluations

The variables included in this category are outlined in *Table II*. If the internal structures of issue positions and of the ideology and party ID category seem to be off limits to campaigns, the power of campaigns appears to be located in their influence over the structure of voter perceptions of candidates. The disjunction is profound: It does not appear to be a difference in degrees but a

Table 4.2

VARIABLE	CONTENT
bushfav	Bush Favorability
bush	Presidential Approval
cong	Congressional Approval
hillfav	Sen. Clinton Favorability
obamafav	Obama Favorability
obamaideol	Perceived Obama Ideology
mccainfav	Mccain Favorability
mccainideol	Perceived Mccain Ideology
mccainslIs	Mccain a strong leader?
mccaint	Is he truthful?
mccaine	Does he have the experience to be president?
mccainj	Does he have the judgement to be president?
obamasl	Is Obama a strong Leader
obamat	Is he truthful?
obamae	Does he have the experience to be president
obamaj	Does he have the judgment to be president
Otraits	Obama's Average of above four traits
Mtraits	McCain's Average of above four traits

qualitatively different ability to affect the structure of this category.

That being said, much of the internal structure of campaign and politician evaluations is, like the previous categories, exogenous from the campaign. Perhaps the most substantively interesting case of campaign exogeneity is between two major figures, Senator Clinton and President Bush. The relationship between the two does not change throughout the campaign, even when Senator Clinton is actively seeking her party's nomination for the presidency. The correlation between the two politician's approval ratings maintains a value close to its mean of - .69.

Voters were asked to evaluate four personal traits of the major party candidates throughout the campaign: were they a "strong leader," were they truthful, did they have the judgment to be president, and did have the experience to be president. The relationship between those four is surprisingly steady for both candidates. "Within-candidate " will refer to the structure of relationships among perceived traits, favorability and other the other variables from *Table* II. for a particular candidate, and "between-candidate" to the set of relationships between perceptions of different candidates.

Fig. 3a. gives an example of within-candidate campaign exogeneity, that of McCain's truthfulness and his judgment. For Senator McCain, there is close to no change throughout the campaign. For Senator Obama, there is a slight upward trend, from roughly .8 in January to roughly .9 by the end of the campaign. An example is given in *Fig. 3*. In this category, most cross-sectional correlations can be treated as something more scientific than convenient shorthand, since most variables are on a 0-10 scale. Speaking just of the cross-sectional correlations, the relationship exogeneity or very minor change of Figs. 3 a and b holds for all of the within-candidate structure of opinion.

The DCC-derived within-candidate correlation series, however, exhibit an exception to that campaign exogeneity for just one trait for each candidate. For Obama, it is his truthfulness, while for McCain it is his experience. This odd disjunction between the aggregate and cross- sectional correlation series was robust against many modeling specifications. *Fig. 3c* and *3d* illustrates this

Fig. 3.

strange juxtaposition.

It was hypothesized that these two traits, generally assumed to be strengths of the candidates, remained steady while their other traits varied at the aggregate level. This did not appear to be the case, however, as the variance of these traits shrunk no less than the others when the last ninety days of the campaign to the prior 226 days. Nor was the overall correlation from those days in the raw data significantly different during the period in question. Nor would that theory address the cross-sectional, aggregate-level disjunction.

The decline in aggregate-level correlation for these two traits should not be over-stated. It is a decline of a bit more than .2 in the correlation series. With roughly 900 series, it is likely that at least some of these series will exhibit noisy behavior that looks substantively interesting, at least for a portion of their series. It is quite strange, though, that this would appear in one trait for each candidate. In short, while it may be noise, no explanation could be found, and further research is necessary.

To summarize the within-candidate subset of the structure of the politician evaluation space, there was a single, trait-specific question mark for each candidate, and the correlation series among candidate Obama's perceived traits showed a very mild, incremental strengthening over time. Otherwise, the structure of within-candidate opinion structure were exogenous to the campaign, even as their aggregate values changed notably.

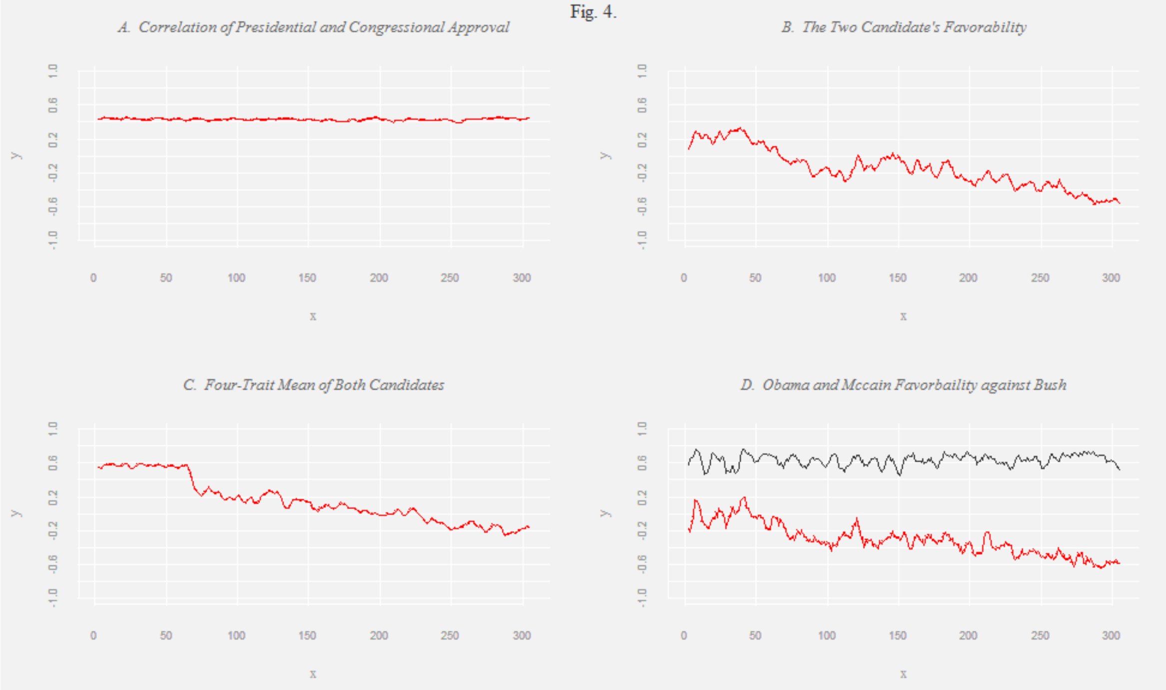
The correlation between presidential job approval and that of Congress' is a noisy series, as Congress' approval is fairly steady, though on a slight downward trajectory, whereas Bush's mean approval rating is more volatile, both in its daily movements and its downward trajectory. The correlation between the two institutions seems to weaken, mildly and noisily, over time, as *Fig. 4a*. On the one hand, different parties control both branches, but on the other they both are perceived in poor light. Interestingly, this results in an easily positive overall correlation, at .38. This suggests that the performance evaluation of the two dominates the ideological or partisan evaluations at the aggregate. Finally, as the slide in approval for both accelerates with the onset of the crash, their correlations jump significantly, regaining the heights at the beginning of the campaign.

It is in the between-candidate relationships that the campaign exacts its largest effects. *Figs. 4.b-d* give the change between the correlations of Obama's and McCain's favorability; the mean of their four trait-ratings; and the two candidate's correlation series with the president's favorability. *Fig. 4.b* is particularly important. *Fig. 4.c.* shows a larger decline, but some of that comes from a switch in who is asked each question, form within-party to the general electorate, on day 79 of the sample.

The contrast between between-candidate evaluations and the previous categories is striking. The extent and nature of the change between particular candidate traits is nearly identical to the approval and averaged trait series, though a small bit noisier. Campaigns are fundamentally capable of restructuring public opinion vis-à-vis the candidates running, even as they might just as well not have happened when it comes to the structure of issue positions. Finally, Senator

Clinton and McCain's nearly time-invariant correlation series with President Bush contrasts sharply with that of the initially-lesser-known Senator Obama. The latter two correlation series are displayed in *Fig. 4.d.*

Fig. 4.



4.2.4 Voting Intentions

Four variables are in this category, presidential vote intention, partisan house vote intention, and voting behavior for the two in the previous elections, 2004, and 2006, respectively. Unfortunately, the NAES did not ask about general election voting intentions for the Presidency until only 228 days before the election, and 118 days for the House. So, much of the dynamics of the two campaigns shifting from their party's nomination battles to the general election is not captured by the NAES.

The correlation between house and presidential vote intention is, perhaps surprisingly, largely exogenous from the campaign, hovering around a mean of .75. Similar to the correlation between the (self-reported) previous presidential vote and 2008 vote intentions, there is a sizable dip about two weeks before the election, though it quickly adjusts back to its previous mean. It is

caused by Obama's support falling while the Democratic House vote share increases, and then both series returning, partially, to the pre-dip levels. The correlation series are shown in *Fig. 5a* and *b*.

There seems to be a small uptick in the correlation between past and present presidential voting behavior as the race shifts into the general election, perhaps as some Clinton supporters return to the Democratic ticket.

Table 4.3

Category I. Voting Behavior	Category II Politician Evaluation
Vote Intention	Bush Favorability
House Vote Intention	Presidential Approval
Pres. Vote 2004	Congressional Approval
	Sen. Clinton Favorability
	Obama Favorability
	Perceived Obama Ideology
	Mccain Favorability
	Perceived Mccain Ideology
	Mccain a strong leader?
	Is he truthful?
	Does he have the experience to be president?
	Does he have the judgement to be president?
	Is Obama a strong Leader?
	Is he truthful?
	Does he have the experience to be president?
	Does he have the judgment to be president?
	Obama's Average of above four traits
	McCain's Average of above four traits

4.3 Between-Category Relationships

4.3.1 Politician Evaluations and Voting Behavior

The category most proximate to voting behavior is the perception of candidates and politicians. Let us start here and see if there is much shift between these two tightly knit categories, outlined in *Table III*.

The correlation between elements of the vote intention bloc and politician evaluation category is dominated by a pattern of incrementally increase, either negligibly or up to .2 throughout the campaign. Interestingly, unlike the internal structure of politician evaluation structures, this increase is not differentiated by whether or not the politician in question is actively seeking office. The same pattern emerges in the relationships between the four candidate traits

Fig. 6.

and the vote, again with the curious exception of Obama's truthfulness and McCain's experience.

4.3.2 Ideology and Party ID and politician evaluations.

Table 4.4

Category I. Party ID and Ideology	Category II Politician Evaluation
Ideology	Bush Favorability
Party ID	Presidential Approval
	Congressional Approval
	Sen. Clinton Favorability
	Obama Favorability
	Perceived Obama Ideology
	Mccain Favorability
	Perceived Mccain Ideology
	Mccain a strong leader?
	Is he truthful?
	Does he have the experience to be president?
	Does he have the judgement to be president?
	Is Obama a strong Leader?
	Is he truthful?
	Does he have the experience to be president?
	Does he have the judgment to be president?
	Obama's Average of above four traits
	McCain's Average of above four traits

Interestingly, for non-candidates Bush and Senator Clinton, the structure of their perceptions vis-à-vis PID and ideology is campaign exogenous. Similarly, at the aggregate level, both ideology and Party ID correlate with candidate perceptions in a manner that is exogenous from the campaign. *Fig. 7.a* and *b* present those two series, respectively, while *c* and *d* present the same for McCain. The cross sectional correlation pattern differs, however, and show a clear upward trend of varying degrees of strength. This appears to be the first clear-cut case of aggregate level variance leading to depictions of campaign dynamics that differ fundamentally from that of their cross sectional counterparts. The conclusion seems inescapable: Party ID and ideology are, in the

aggregate, unmoved by the campaign. *Fig. 8* displays their raw values, unsmoothed and unfiltered for sampling error, over the course of the campaign. Simply put, the variance observed is sampling error and other white noise, to an eyebrow-raising level. It is not that these variables equilibrate tightly around a certain value: it is that they are completely unmoved.[4] So, it is no surprise that the aggregate-variance derived correlations are unmoved. As with any measure of covariance, it takes two to tango. As was already shown, though, the structure of candidate perceptions changes dramatically over the course of the campaign, and that change is incremental, steady and, as in *Fig. 7*, at least partly through partisan and ideological filters. All the more so because the cross- sectional correlations are attenuated by the ideology scale of 1–5, the Party ID scale of 1–7; in addition, the statistical problems that come from being forced to pretend that these are evenly spaced, ordinal scales increases the noise that might reduce the value of the

[4]So much so that one might think the NAES weighted its sample by ideology and party ID, but they practiced random digit dialing and made no mention of any method that would weight their respondents, on any daily time frame.

Fig 8. PID and Ideology Throughout the Campaign

cross-sectional correlations.

What is surprising is the strength of the dichotomy of the cross-sectional and aggregate-level variance. The old party ID adage can be extended to ideology and specified: Party ID and ideology are unmoved in the aggregate and movers at the individual level.

4.3.3 Voting Behavior and Party Identification and Ideology

Table 4.5

Category I. Party ID and Ideology	Category II Voting Behavior
Ideology Presidential Vote 2004 House Vote 2006	Presidential Vote Party ID House Vote

It would be very telling of underlying dynamics if the same pattern holds between the party ID/ideology category and categories other than perceptions of politicians. In *fig. 9* the aggregate variance-derived correlations are the dashed lined, since the very low over-time variance of the ideology and correlation series means the series is very noisy, and not much should be read into their trends. The cross-sectional correlations are attenuated more severely than in the previous category, since major party vote choice is a two category variable.

Fig. 9.

With those methodological caveats in mind, it appears the same pattern holds, with a weakening in correlation as the vote variable begins to move more profoundly during the events of October. It should also be noted, though, that the NAES' polling of the final forty days of the two party vote share shows no discernible trend Obama's rise and fall occurs earlier in the sample and that this lack of movement in October seems to contrast with pollster.com's estimates.

In sum, the data in this sections is the noisiest and shortest combination of the ten sections. All that can be gleaned from the methods put forward here is that the disjunction between the cross-sectional and aggregate variance-derived covariance is consistent across party ID and ideology's relationships with both voting and candidate evaluations.

4.3.4 Issues Space and Party Identification and Ideology

The issue and non-politician evaluation variables are listed in *Table 1*. The structure of these variables is entirely exogenous from the campaign. Over the twenty relationships (between party ID and ideology and the variables listed in Table 1) only the slightest perturbations occur. Like the other party ID and ideology series, the aggregate level variables display more noise than they should, due to the high ratio of sampling error to substantive movement, but none of the correlation series, aggregate or cross-sectional, show discernible trends or developments.

It's worth taking note of the profundity of this campaign exogeneity. It is across-the-board: not even a major economic meltdown shakes the relationship - or lack thereof - between economic evaluations and ideology. The same may be said of party ID. It is as if we are looking for campaign effects in places very, very far from where we might find them. One might expect a campaign to rally a more ideological or party-cohesive issue space. At least in 2008, it does no such thing - anywhere.

4.3.5 Issues Space and Politician Evaluations

If the structure of the issue space is so time-invariant as regards ideology and partisanship, even as political actors are spending billions of dollars to affect election outcomes, it is worth asking if the same holds for the relationship between the campaign's issue space and evaluations of politicians. We know that campaigns do not do much to change the structure of the issue space, nor the PID and ideology relationship. Conversely, we know that the campaign changes the structure of the candidate evaluation space, and that much of that change travels through the cross-sectional partisan and ideological groups. One might expect then, to see a similar

Table 4.6

Category I Issue Space	Category II Political Evaluation
Sociotropic Economic Well-Being	Bush Favorability
Personal Economic Well-Being	Presidential Approval
Taxes too High	Congressional Approval
Under Which Conditions Should it be Allowed	Sen. Clinton Favorability
Favor a Path to Citizenship	Obama Favorability
Favor Building US-Mexico Fence	Perceived Obama Ideology
Favor Gay Marriage	Mccain Favorability
Withdraw as Soon as Possible	Perceived Mccain Ideology
Was Iraq Worth It	Mccain a strong leader?
Leaving Iraq Increase Terrorism	Is he truthful?
	Does he have the experience to be president?
	Does he have the judgement to be president?
	Is Obama a strong Leader
	Is he truthful?
	Does he have the experience to be president
	Does he have the judgment to be president
	Obama's Average of above four traits
	McCain's Average of above four traits

relationship between issue positions and candidate evaluations.

Interestingly, though, *Fig. 10.a* and *b* show the two major candidates correlation series against two issues, marriage and taxes, that did not see a lot of movement during the campaign. The contrast between evaluations and party ID and ideology is notable. *Fig. 10.c* also presents an interesting picture. McCain's favorability, as the incumbent party candidate, is much more tightly linked to the economy. As the economy enters crisis mode, the correlation with McCain's favorability tightens. Viewing the correlation series involving Obama's favorability does not make much sense without first seeing the series involving his opponents.

A major feature of these series is a mild trend in the cross-sectional data, and seemingly direction-less movement in the aggregate variance-derived series. Some social issues, abortion and immigration issues, exhibit a very noisy but discernible trend. Oddly, marriage does not, for Senator McCain, yet does so at the aggregate level for Senator Obama. Taxes exhibit no trend whatsoever, perhaps their stand-in as a small government issue getting swamped by the financial crisis. Economic issues show movement but no trend until mid-September, when they strengthen rapidly, even while their cross sectional correlations show no trend whatsoever. Some of these developments are included in the last graph in the chapter, *Fig. 11*.

Perhaps this movement in the aggregate and lack of movement in the cross-sectional is evidence that ”the economy” exerts its profound electoral influence less as an issue unto itself and more through its effect on perceptions of competence, direct effect on people's quality of life, and through other avenues. This suggests that analyzing the difference between cross- sectional and aggregate variance-derived correlation series is a particularly useful way to detangle the large web of causality in economic voting. However, further investigation needs to be conducted before this one, suggestive finding can be said to speak seriously to such a large literature as that of economic voting (e.g. Carey and Lebo 2006).

Fig. 10d presents the most dramatic example of the opposite cross-sectional-aggregate disjunction, that of no development at the aggregate and trend at the cross-sectional: the correlation series between Obama's approval and opinions on Iraq. The latter question allowed for only three responses, and so correlations and the change between them will be considerably attenuated from a variant that would have asked respondents to place themselves on a 1–100 scale. Still, the campaign re-aligns much of the electorate, from .2 to .5 by the end. However, that

does not filter up to the aggregate level in any discernible way. Other correlation series exhibit more mild trends in the cross-sectional usually a 2 increase with a similar level disconnect between that trend and the trendless, noisy development of the aggregate level variance. The pattern is widespread enough that its implications should be investigated further.

4.3.6 Issue Space and Voting

Fig. 11b documents how, even as Congress' approval becomes linked to the nose- diving economy, the economy's correlation with the vote intention for the party in control of Congress strengthens in the opposite direction. Certainly, the Republican party was held responsible for the financial crisis far more than the Democratic congress.

Fig. 11.

The structure bridging the variables in *Table I.* and *Table II* is very similar to that of the issue space's relation to candidate evaluations. *Fig. 11* presents four representative cases. In each case, the red line is the primary, DCC-derived correlations series. The black line is the cross-sectional correlations, and the dashed blue line is the related correlation series, with the vote series switched out for the appropriate approval series. For the presidential vote, it is Senator Obama's approval rating; for the House it is congressional approval.

To summarize the final two sections, voting behavior and candidate perceptions are, as would be expected, knit to the issue space in very similar ways. However, that connection varies from issue to issue, and is typically not nearly as profound as the change in correlation series between candidate evaluations. Social, ideological issues appear to have a propensity to sort voters at the individual level, but this seems to only be loosely suggestive of their behavior in the aggregate. Lastly, candidate approval is so tightly mapped onto voting behavior at the presidential level only that future substantive research that leverages the disconnect between the cross-sectional and the aggregate may be wise to focus on the continuous approval variable over the binary vote variable.

4.4 Conclusion

This investigation goes a long way towards laying out the elemental facts of public opinion's evolution during a presidential election. These basic facts, in turn, speak directly to the campaigns literature, particularly the campaign effects literature. What is found is a world of defined, qualitative limits, one where campaigns can fundamentally restructure some of the structure of public opinion, but are completely incapable, or unwilling, to move other parts of the structure.

The partisan, ideological, and issue structure of the electorate seems almost casually unimpressed by the efforts of elites to affect it over the course of a single campaign. Top-line values of variables in these categories might be moved, but to do so meaningfully, you need to move the others by the amount the structure of opinion dictates, and that is a much, much heavier lift. Perhaps it is something that requires the muscles of generational replacement or major political economic developments, not some comparatively puny billion-dollar campaign organization. If that is the case, then that is a mooring of democracy that's a lot more stable than Zaller's balanced information flows by themselves.

Conversely, though, candidates, though not politicians in general, see their relationships in public opinion change dramatically over the course of the campaign, and the development of the within-candidate structure does not seem especially nuanced. At the cross-sectional level, particularly, some of the dramatic change stems from mapping candidates onto the issue positions they have long had. Though the public grows no more partisan or ideological in its approval of institutions or non-candidate politicians than it was at the beginning of the race. But they do grow a small bit more partisan and ideological in their assessment of candidates.

That change unto itself seems fine. But ten months before the election, the approval of the likely nominees after they've been campaigning for many months, for opposing parties' nominations were still positively correlated. The change that followed may be the public learning about candidates, but it is learning of a disconcertingly late and basic nature. In short, ten months before the election, the public is not at all apolitical in its beliefs, but is apolitical in its politics. The journey from that place seems very republican, of competing factions alerting the republic to legislator's stances. The normative concern is not in the nature of the journey but that the starting point is so far from where it needs to be change in correlation of .7, from positive to negative that if the trajectory is off by just a bit, the destination is wholly missed. Of course, just one election was studied here, but the qualitative nature of the differences between the changes witnessed from substructure to substructure and between substructures suggests that the change witnessed here isn't so much the product of a particular strategy or circumstance. Rather, they speak to what campaign organizations are - and are not - capable of doing.

The research plan employed in this chapter should be duplicated on other elections before one can truly talk about generalizability. Still, there is another argument that speaks to the generalizability of these findings: The change that stands out, at least when one is painting with as broad a brush as this chapter has done, is one of incremental, day-in-day out campaign effects. When elements of the structure of opinion finish off the race in very different circumstances than they started, more often than not it is a slow, almost linear change that occurs.

Incremental change could mean many things, but it certainly seems to suggest that change is driven more by the day-in, day-out efforts of parties, campaigns, and groups than by a small set of important, national political events. Of course, this is excepting the financial crisis, which certainly shook things up; but is also quite a different event from a debate or a major policy speech.

Methodologically, some of the major findings point to developing theory further to allow for the leveraging of differences between cross-sectional and aggregate level-derived covariance. Chapter three explored the distribution of time-invariant correlations that were based on the two different types of covariance. The relationship between the two was fairly tight, but far from universally so. It turns out, from this chapter, that those differences are not just noise and measurement error, at least as they play out in a dynamic context. Something deeper and more telling is at work.

It is the substantive findings that drive this chapter, though. Combined, they provide the basics of how public opinion and voting behavior evolve over the course of the campaign. Important unto itself, it is hoped that these findings will be considered the descriptive inferential background against which causal inferential work will operate. The next chapter, for instance, will use the information presented here as a starting point for an investigation of activation effects. The methods and preliminary findings concerning economic voting may be a useful basis for further research. Similarly, the long-run relationship between ideology and partisanship in the aggregate was mirrored perfectly in the short run. It is important to find out if the same long run–short run relationship holds among the ”sophisticated” electorate. Finally, the pattern of incremental, nsteady change that was witnessed in much of the campaign-endogenous portion of the structure of mass opinion should be investigated further. What does it say about the role of campaign actors and major campaign events in shaping opinion?

Chapter 5

Activation Effects? What the Dynamic Structure of Public Opinion Says

5.1 Introduction

Chapter 1 presented the underlying logic of the proposed methodological framework and situated the reader in the literature on American presidential campaigns. Chapter 2 reviewed the econometric developments that brought us to the point of using aggregate level variance to model correlations at each specific point in a time series. It presented some useful definitions and examples, and went on to lay out the logic behind the methodological framework used here.

Chapter 3 covered the application of multivariate GARCH models to survey data. It presented some of the problems that come up when applying MGARCH models in political scientific research settings and presented some ways around them, aided by two R packages, which it also introduced. Finally, it examined the relationship between cross-sectional and aggregate-level parameters.

Chapter 4 mapped out the basic characteristics of the structure of public opinion during presidential campaigns. Which relations between different elements of public opinion were exogenous to the campaign and which were endogenous was surprisingly obvious. With the background understanding given by Chapter 4 of how the structure of opinion evolves during a campaign, more specific questions can be answered. Activation effects constitute one area in which the methods advocated here are capable of contributing to some of the oldest, most important issues in the literature.

5.2 The State of Activation Effects Literature

Political scientists have put activation at the heart of the causal narrative of presidential campaigns for seven decades now. A literature review of campaign effects in general was given in Chapter 1. Numerous other reviews exist (e.g. Hillygus in Leighley, 2010). This review situates activation effects in the larger campaigns literature and updates the reader on the most recent developments.

Lazarsfeld, Berelson and Gaudet (1944) declared ”Political campaigns are important primarily because they activate latent dispositions” (emphasis in original). Since that oft-cited statement (e.g. Hillygus in Leighley, 2010; Kaplan, Park and Gelman, 2012), campaigns’ ability to activate feelings of partisan, ideological and group identities has been central stage in the campaigns literature.

That the Columbia school did not just point to activation as an important part of campaigns but gave it primacy is a major reason for activation effects’ prominence in the literature. Another is that activation offers a resolution to one of the puzzles that has helped organize the political behavior literature: the predictability of election outcomes in spite of the observed volatility of public opinion during the campaign. Indeed, a modern ”classic” (Enns and Richman, 2013) on activation effects carries the title, ”Why Are American Presidential Election Campaign Polls so Variable When Votes Are so Predictable?” (Gelman and King, 1993). Hillygus and Jackman (2003), explain activation effects’ capability of resolving the tension: ”..[C]ampaigns might help to activate latent preferences, but the preferences are thought to be in place before the election period begins. So, again, individual votes and election outcomes can be predicted without accounting for the campaign."

Kaplan, Park and Gelman (2012) emphasize that non-campaign specificity of the mechanisms by which campaigns activate "the fundamentals" in an interesting manner. They model the growth of the role of fundamentals in predicting vote choice during three recent presidential elections. Then they fit the model for each year with the coefficients from the other years' models, and find little in the way of loss of fit.

A major reason Gelman and King (1993) has been so influential is that the literature has adopted their operationalization of activation effects. A set of "fundamentals" such as party ID, ideology and demographic variables determine a portion of the electorate's vote choice. That portion increases over the course of the campaign; and it is that increase that is considered activation. Along with sampling error (Erikson and Wlezien 2012), it is widely considered a major source of over-time variability in polls.

Mclurg and Holbrook (2009) evidence activation effects along these lines. They compare battleground states and other states during two presidential elections and find that voters in states where campaigns are most active "behave in a more predictable fashion." (ibid. pg. 502) They also find that this pattern took different forms in 1988 and 1992, probably the consequence of different campaign strategies.

Finkel (1993) looks at NES panel data in 1980. He argues not just for activation effects, but that campaigns "mainly" serve to generate activation effects; and that this generation is in line with the minimal effects thesis.

Which variables are appropriate to treat as fundamentals is a matter that lacks consensus, though it does not appear to receive much active debate. The answer given generally relies more on the substantive concern of the given research then on an in-depth theoretical argument. Issue and economic variables, presidential approval, party ID, ideology and demographic variables are some options. Of those, the latter three are nearly always used. Chapter 4 presents novel, confirmatory evidence or the treatment of PID and ideology as exogenous from the campaign over short (less than a year) time windows: Not only is PID and ideology exogenous from the campaign in their aggregate values, the relationship between the two shows no discernible development over the course of the campaign.

Most of the evidence concerning activation effects is cross-sectional in nature. Shaw (1999), however, provides aggregate-level evidence for the role of fundamentals, if not the activation thereof, when he shows that events have a greater impact when they help the candidate that is behind or hurt the one that is ahead the further from some equilibrium value, the greater the pull back towards that equilibrium, at least when some event does occur.

Enns and Richman (2013) harbor skepticism of the growth-in-fundamentals thesis. They turn to NAES data to show that, among those who (self-reportedly) care about the outcome of the campaign and those that do not, there is no growth throughout the campaign in the explanatory power of "the correctly weighted fundamentals."[1] What does change in their data, though, is that roughly twelve percent of respondents go from the category of not caring about the election outcome to caring about it, the latter category displaying a tighter relationship between vote choice and fundamentals.

Enns and Richman's central point is that much of the apparent growth in the fundamentals may actually just be a reduction in "survey satisficing" (Krosnick 1991) during the interview process. Enns and Richmann further evidence this by using NES data to show that the fundamentals do a better job of predicting the vote during face-to-face interviews than during phone interviews. When the costs of satisficing are raised either by interest in the campaign, the social cost of giving the inappropriate answer, or other reasons respondents are more likely to

[1]They use as fundamentals PID, ideology and demographic variables, plus presidential approval and issue mood and economic perceptions. As regards the "correctly weighted fundamentals:" They follow Gelman and King (1993) by weighting those variables throughout the campaign by the weights voters attach to them in the final week of the campaign, arguing that this is a better approximation of full information preferences.

give an answer in line with their true feelings. They note that similar satisficing arguments can be made about survey completion and response bias. They conclude not by disavowing campaign effects but by arguing that much off what appears to be activation effects are actually artifacts of changing survey error.

Huckfeldt et al. (2007) emphasize the role of information diffusion. They study House campaigns in 2002, and find that, absent significant spending, even knowledgeable voters are less likely to know a race is competitive. Put another way, campaign spending is a necessary condition even for high-information voters to assess the basics of the race. Questions can be raised about whether this part of their evidence lines up with their causal story, but their article in general points strongly towards the activity of the campaigns themselves, as opposed to the general, election-infused political environment as the mechanisms that do the bulk of the activating.

That emphasis on information diffusion is, to some extent, frictional with the notion of the primacy of activation effects. There is need for caution about such statements, though. As Hillygus puts it, ”...[t]he lines between the standard typology of campaign effects are quite blurry” (Hillygus in Leighley 2010). The dividing lines between activation, learning, priming, framing and persuasion are not clear cut. Nor should they be: These categories represent substantively different phenomena, but ones that play roles in the others’ causal chains.

Still, these concepts must be assessed against each other in terms of the role they play in altering the vote, if we are to characterize just what it is campaigns do to public opinion. That has been for some time an entire research agenda. The goal of this article is particular to activation effects.

5.3 The Argument

The central argument that is presented does not challenge the empirical tale that Gelman and King (1993) and Kaplan, Park and Gelman (2012) have built. The fundamentals do indeed grow as the campaign progresses at least in their cross-sectional impact on the vote. What is demonstrated is that the growth-in-fundamentals operationalization of the activation thesis is inappropriate.

The argument of this articles is that if ”activation” is a reasonable moniker for the increasing role of the fundamentals in explaining vote choice during a campaign, then activation effects should show up *somewhere* in the structure of opinion besides vote intention and candidate assessments. It is hard to picture attitudes and identities that lie latent in voters, and then when activated only apply to two particular people. To emphasize, activation does not need to appear everywhere, or even in most of the places one might expect. However, somewhere in the structure of the relationship between parties, other politicians and/or issue positions, there needs to be similarly patterned change. Otherwise, ”activation” is really describing something other than the activation of latent dispositions and so should be set aside so that the field does not mis-characterize the type of change that goes on during campaigns.

As it turns out, there is not any such evidence of activation effects elsewhere in the structure of public opinion. In the chapter that follows, this is first demonstrated. Next, reasonable alternative causal narratives are discussed. Finally, the consequences for how we picture campaigns and, indeed, politics are addressed. Several very important conceptions of campaigns and the relation between elites and voters is at stake.

5.4 Evidence

2008 is a particularly good year to critically examine the appropriateness of using ”activation” to describe change in public opinion during a presidential election. The growth in the role of the fundamentals’ in predicting the vote is roughly twice that of the previous two elections. Kaplan, Park and Gelman find that, ”in the 2000, 2004, and 2008 campaigns, the

Fig. 1.

fundamentals account for approximately .08, .08, and .15 percentage points (sic) more of the variance in vote choice by the end of the campaigns than at their beginnings." (Kaplan, Park and Gelman, 2012, pg. 856) . If activation effects were to show up elsewhere in the structure of opinion, it should be especially easy to find in 2008.

5.4.1 Where in the structure of opinion does activation occur, and where does it not?

The most direct, damning argument against activation effects is what the campaign does to the structure of opinion as it pertains to politicians other than the candidates. If PID, ideology and other variables were to some extent latent in voters' hearts and minds, and then activated, this would show up in assessments of politicians in general, not just the candidates. At the least, it would show up for politicians that are active figures and topics of debate during the election. 2008 is a good year to look for this effect because two politicians, President Bush and Senator Clinton loomed over the race, regularly mentioned in often even the focus of the news cycle. Fig. 1

shows the daily correlations of ideology and approval for the two major party nominees, steadily growing at the cross-sectional level, though lessening at the aggregate during the onset of the financial crisis. The DCC-derived correlations (the red line) have been smoothed with a one week moving average. On the cross sectional level, at least, even in the face of the financial crisis, activation seems to work its magic.

Consider, however, the last two cells in *Fig. 1*. The same relationship between ideology and approval for two of the country's more prominent, partisan politicians seems completely unaffected by the campaign. Both were very visible, either in appearances or campaign ads, throughout the campaign, yet absent is any additional pull of the slowly activated latent dispositions.

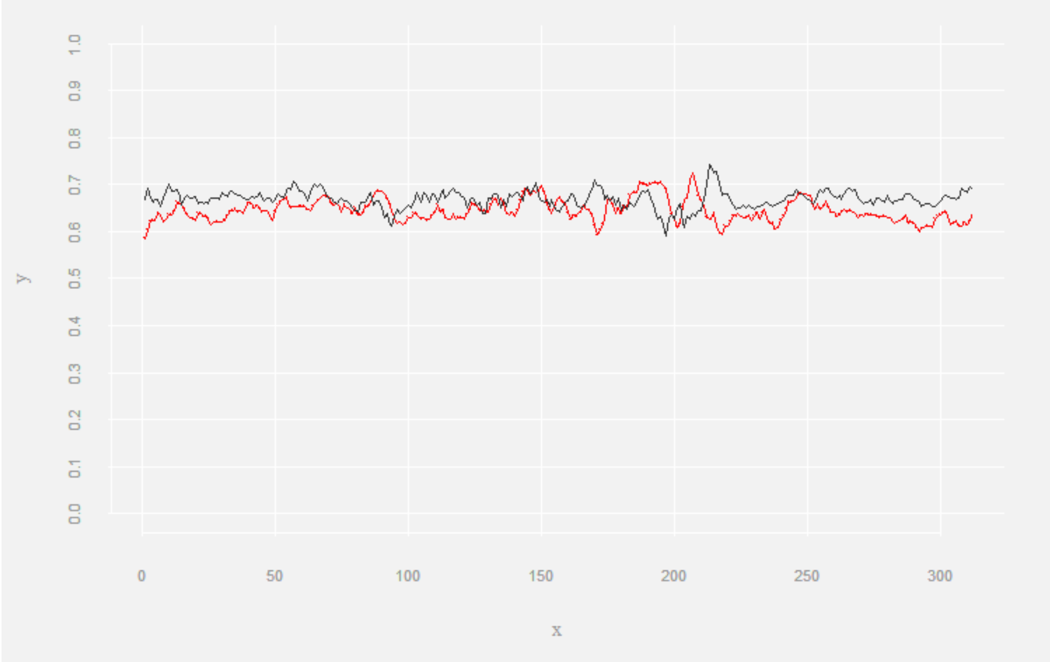
A pattern similar to what we see in ideology and candidate favorability is apparent with about half the issue positions the NAES asked about throughout the length of the study. For partisanship and some issues, the dichotomy exhibited in *Fig. 1* repeats itself. *Fig. 2* samples some of these relationships. The difference being that growth in explanatory power of issues vis-à-vis the candidates is smaller. The complete lack of growth via non-candidates is the same.

An oddly-ignored question proponents of the activation thesis must answer is what happens to activation at the aggregate level. *Fig. 3.A.* shows the levels of partisans and ideologues[2] in the electorate, smoothed over a five day period to minimize the raw data's sampling error. Simply put, there is no increase whatsoever. The brief jumps come from sampling error during the days that saw only a few dozen respondents surveyed.

There is also no tightening in the relationship between Macropartisanship and Ideology. *Fig. 3.b.* shows how minimal in the aggregate and time-invariant the relationship is in a daily time series

[2]Ideologues are counted as those that list themselves as " liberal" or " conservative." The results do not change if the definition is restricted to the subset that are "very liberal" or "very conservative." Partisans are those that identify with either party. If one restricts the definition to just strong identifiers, there is a roughly 5% increase in strong partisans as the primaries end, and then, from more than 200 days before the election, there is no increase.

Fig. 3.a. Ideological and Partisan Voters

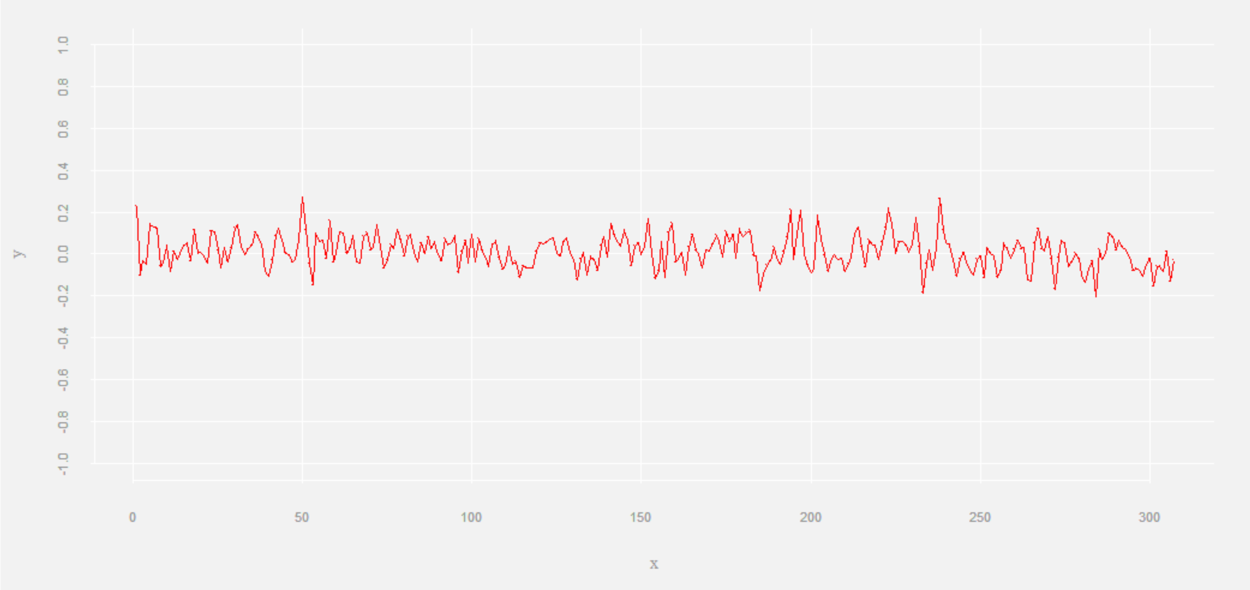


lasting 316 days during a presidential campaign. The cross-sectional relationship is of course not zero. It is, however, time invariant over the course of the campaign. Both cross- sectionally and in the aggregate, the presidential campaign does not activate ideologically motivated partisanship, and vice-versa.

To return to non-candidate politicians, perhaps the non-activation of latent attitudes towards Senator Clinton and President Bush is an effect of the exceptional stature of Senator Clinton and President Bush. Perhaps the activation of opinion via president Bush and Senator Clinton happened elections cycles ago. This would quite clearly delimit the activation thesis to only less well-known candidates. Parenthetically, it would also require of the electorate a memory for which it is not known.

The next place we should look for activation effects is between issues, as opposed to between issues and politician favorability. *Fig. 4* shows the relationship between some of the

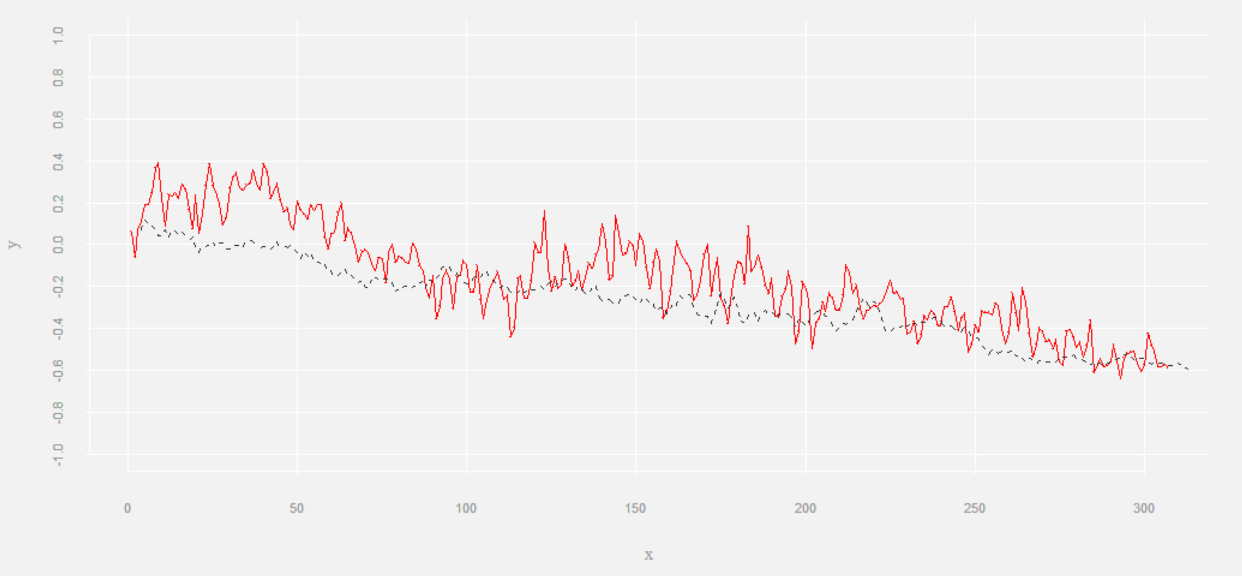
Fig. 3.b. The Correlation between Ideology and Macropartisanship



issues that each show the greatest level of ”activation” vis-à-vis the candidates. *Fig. 4.c.* and *4.d.* also show representative relationships between these issues and the parties. The correlations with the macropartisanship and ideology series are noisy because their standard deviation is so low that the small degree of changing sampling error-derived heteroskedasticity that is not filtered out by the first round GARCH models plays an exceptionally prominent role in the series. Again, though, the change that is witnessed in the relationship of these issues and the candidates is not found in the relationships between the parties and issues. The issue space does not grow more ideologically cohesive, nor does it grow more partisan, at least insofar as the NAES data allows.

The only significant issue-party or issue-ideology change occurs vis--vis Iraq. At the aggregate level, begins the series 316 days before the election not correlated with macropartisanship, grows in correlation to roughly -.4, and then returns to zero by election day. The Iraq-ideology series does not change in any directional manner. Both the party-Iraq and ideology-Iraq series do not change in their cross sectional relationship over the course of the campaign. Given its unique

Fig. 5. Obama and Mccain Favorability



behavior in the time series of public opinion, it seems likely that the aggregate-level change is the result of developments in Iraq more than developments in the campaign.

5.5 The scale of activation effects

Another important element of the argument against the activation thesis is the scale of change that is witnessed. Compare the change in *Fig. 1* to the differentiation between the candidates that takes place. *Fig. 5* shows the changed correlation between the favorability of candidates Obama and McCain. The DCC-derived time series of favorability correlations (the red line) begins in low positive territory and ends just shy of -.6. The dashed line, the cross sectional correlations, begins in positive territory, though just barely, and ends in nearly the exact same place.

The change in candidate differentiation is roughly four times that which is observed in most relationships that could serve as an example of activation effects at work. The profundity of the impact of the campaign on candidate differentiation is further evidenced by how the relationship switches signs, from mildly positive to negative. It seems appropriate to say that the relationship between the candidates changes from being apolitical to being highly political.

5.6 Tying it Together

So we know that Party ID, ideology and issues become more important in predicting individual-level voting patterns. Yet, if this is because latent concerns, values and identities are activated by the campaign, why does it not show up elsewhere in the structure of public opinion during presidential campaigns? Four facts stand out. 1. The change is specific to the candidates, not even to prominent politicians. It does not affect the other politicians that the NAES asked about throughout its sample. Partisan figures George W. Bush and Hillary Clinton evince no activation effects. 2. PID and ideology do not meaningfully change throughout the campaign, nor does their relationship change, at the aggregate or cross-sectional level. 3. The relationship between issues does not move in any directional way. Even issues like gay marriage and abortion, that could be expected to become ideologically intertwined or more closely associated with partisan identities do not do this over the course of the 2008 campaign. 4. The scale by which issue candidate, ideology candidate and party candidate tighten their relationship is dwarfed by the scale of candidate differentiation that takes place.

Given the strength of evidence, it seems clear that political science have mis- characterized the mechanisms by which opinions about the candidates are aligned with the ”fundamentals.” The process of activation presumably does occur for activists and those most involved in politics. That’s an important process, but it does not show up in public opinion polls, at least not in the NAES in 2008. It is also a very different story from that of elites, parties and groups reaching down into the public’s psyches and activating concerns that had fallen latent.

5.7 Possible Alternatives

The most intuitive alternative is that what had been cast as a process of activation is actually a process of candidate-specific learning. The profundity of the difference between how candidates and prominent non-politician candidates relate to the rest of the structure of public opinion during the election suggests that change involves, in some fashion, candidate-specific information. What this study does not speak to is the other, non-activation mechanisms could describe how that candidate-specific information gets treated by voters.

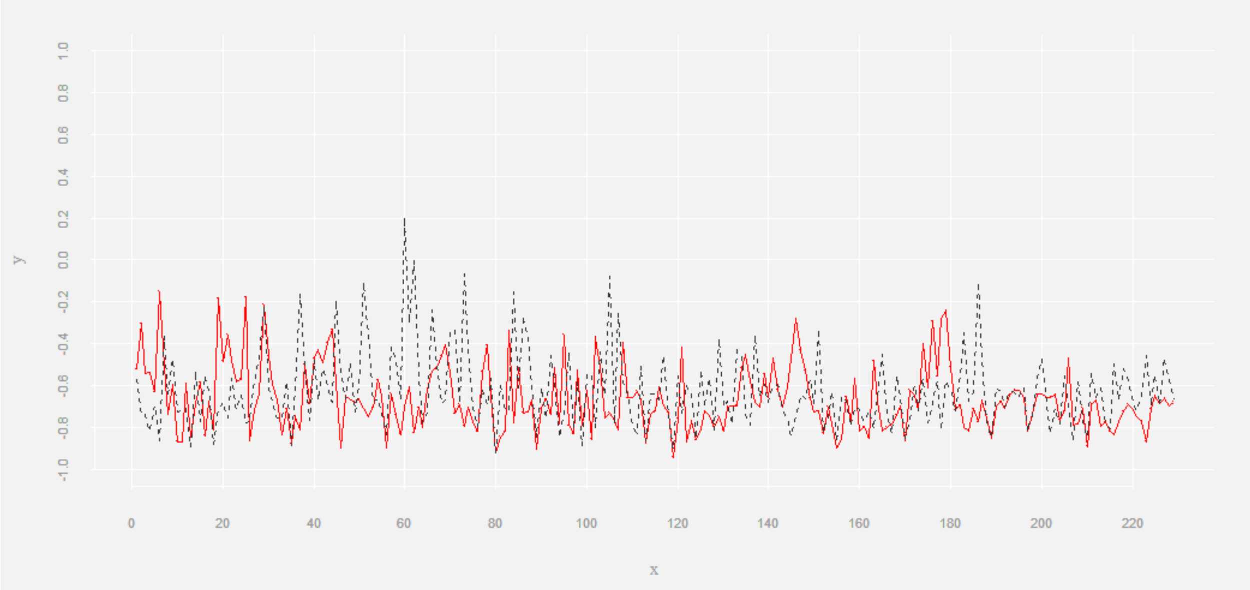


Fig. 6. Obama and Mccain Approval and the Vote

Priming also is a possibility, insofar as it is a notion more easily localized to just the candidates and vote choice. If that is the case, though, theories of priming must navigate the fact that the candidate favorability-vote choice relationship shows very little development over the course of the campaign; what change there is is fleeting and non-directional. *Fig. 7* shows that relationship. It undergoes brief periods of attenuation, but never for very long, ending the election where it began. If issues were primed over the course of the campaign in a manner consistent with the gradual change in the structure of opinion observed in the relationships previously used as evidence of activation one might expect that priming to affect vote choice more than candidate favorability. This is, though, only suggestive evidence.

This study also says little about persuasion. Candidate-specific appeals might well persuade voters. The evidence strongly suggests that it is those whose issue, partisan and ideological stances are at odds with their vote choice that are most persuadable by the campaign (e.g. Hillygus and Jackman 2003).

At any rate, an investigation of the other, non-activation mechanisms is beyond the scope of this paper. What is clear is that activation is a mischaracterization of how voters align their beliefs and their candidate preferences. The process is not one of the structure of beliefs, or some part thereof, being activated, it is the placement of candidates in that structure of beliefs.

Even the broadly interpretable alternative mechanism of ”learning” is too specific for the argument presented here. What does occur is some process of candidate differentiation. What does not occur is an activation of latent attitudes, beliefs and identities, at least not on a scale picked up by the NAES, a survey of roughly 57,000 voters.

5.8 Conclusion

This investigation was started with the expectation of finding strong, confirmatory evidence of activation effects. Clearly, though, activation is not the right causal story for the observed statistical tale. What is at stake is fundamental for political scientists' perception of democracy in that it positions voters and campaigns in very different positions of power in relation to one another. In the traditional narrative, voters have latent attitudes that lie dormant until activated by elites. In the other, more accurate, narrative, voters place candidates into their preexisting structure of beliefs.

One narrative for this process of placement is that voters learn about candidates. The fundamental variables show no evidence of differing latency before, during or after the campaign. Whatever the alternative causal narrative, though, it seems the ball is certainly back in the court of those that claim the evidence suggests activation is a first-tier effect of presidential campaigns.

The simplicity and directness of the argument presented here is possible because we can observe daily time series of the structure of public opinion during the campaign, and not restrict their relationship to being time-invariant. Even without these advantages, though, one has to ask why the field was so willing to settle on a consensus such as they exist in political science without finding much evidence of activation effects beyond voter opinion of the candidates and the closely related variable of vote choice. Activation is certainly an intuitively appealing, convenient idea, but those are not qualities that can be allowed to drive the development of a social science.

That political science was so willing to accept the activation thesis is part of a larger pattern. Most science studies systems that are complex. Yet the capacity to conduct research is limited by constraints cognitive, computational, financial and professional. So, scientists must choose which components of the system get investigated and in what depth.

Humans are humans, however, and the choices they make are often selective and self-serving. The forest doesn't just get lost for the trees. The trees get sought out while pretensions of forest-hunting remain. The methodological framework proposed here disciplines the researcher against that impetus. Minor evidence for a theory was found to be minor precisely because the variance that was witnessed in one portion of the system was, first, sought out in other, closely related areas, and it did not show up. Next, it was compared to the variance witnessed elsewhere and, in that light, was seen to be very minor.

Both of these insights were not just possible but obvious because the methodological framework put forward in this dissertation demands of the researcher the construction of a three dimensional array of covariances, and, then, the exploration of that rich structure. Findings from one part of the system's structure could be compared and contextualized by other parts.

There are a number of other advantages to this framework but that is the most fundamental. There are many difficult modeling decisions that must be made at the beginning, and then there is much work to be done to investigate and communicate. But at some point in the middle, the structure, the forest, stands out in front of the researcher.

Abramowitz, Alan I. 1978. "The Impact of a Presidential Debate on Voter Rationality." *American Journal of Political Science* 22(3): 680-90.

—. 1988. "An Improved Model for Predicting Presidential Election Outcomes." *PS: Political Science and Politics* 21(4): 843-47.

—. 1989. "Viability, Electability, and Candidate Choice in a Presidential Primary Election: A Test of Competing Models." *The Journal of Politics* 51(4): 977-92.

—. 1995. "It's Abortion, Stupid: Policy Voting in the 1992 Presidential Election." *The Journal of Politics* 57(1): 176-86.

—. 2006. "Just Weight! The Case for Dynamic Party Identification Weighting." *PS: Political Science and Politics* 39(3): 473-75.

Abramson, Paul R., John H. Aldrich, Phil Paolino, and David W. Rohde. 1992. "'Sophisticated' Voting in the 1988 Presidential Primaries." *The American Political Science Review* 86(1): 55-69.

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *The American Political Science Review* 95(3): 529-46.

Adkins, Randall E., and Andrew J. Dowdle. 2002. "The Money Primary: What Influences the Outcome of Pre-Primary Presidential Nomination Fundraising?" *Presidential Studies Quarterly* 32(2): 256-75.

Aguiar-Conraria, Lus, Pedro C. Magalhaes, and Maria Joana Soares. 2009. "Dynamical Decomposition of Political Time-Series: An Application of Wavelet Analysis to Electoral Cycles in the United States Presidential Elections." *SSRN eLibrary*. http://papers.ssrn.com.ezproxy.library.wisc.edu/Sol3/papers.cfm?abstract_id=1450048 (July 3, 2012).

Aguiar-Conraria, Lus, Pedro C. Magalhaes, and Maria Joana Soares. 2012. "Cycles in Politics: Wavelet Analysis of Political Time Series." *American Journal of Political Science* 56(2): 500-518.

Aielli, Gian Piero. 2011. *Dynamic Conditional Correlation: On Properties and Estimation*. Rochester, NY: Social Science Research Network. SSRN Scholarly Paper. http://papers.ssrn.com/abstract=1507743 (October 14, 2013).

Aldrich, John H. 1980. "A Dynamic Model of Presidential Nomination Campaigns." *The American Political Science Review* 74(3): 651-69.

Aldrich, John H., and R. Michael Alvarez. 1994. "Issues and the Presidential Primary Voter." *Political Behavior* 16(3): 289-317.

Altman, Micah, Jeff Gill, and Michael P. McDonald. 2003. *Numerical Issues in Statistical Computing for the Social Scientist*. 1st ed. Wiley-Interscience.

Alvarez, R. Michael, and Charles H. Franklin. 1994. "Uncertainty and Political Perceptions." *The Journal of Politics* 56(3): 671-88.

Audrino, F. 2011. *Forecasting Correlations During the Late 2000s Financial Crisis: Short-run Component, Long-run Component, and Structural Breaks*. School of Economics and Political Science, Department of Economics, University of St. Gallen.

Baillie, Richard T. 1996. "Long Memory Processes and Fractional Integration in Econometrics." *Journal of econometrics* 73(1): 5-59.

Bartels, Larry M. 1985. "Expectations and Preferences in Presidential Nominating Campaigns." *The American Political Science Review* 79(3): 804-15.

—. 1988. *Presidential Primaries and the Dynamics of Public Choice*. Princeton University Press.

Bartels, Larry M., and John Zaller. 2001. "Presidential Vote Models: A Recount." *PS: Political Science and Politics* 34(1): 9-20.

Bauwens, L., S. Laurent, and J.V.K. Rombouts. 2006. "Multivariate GARCH Models: a Survey." *Journal of applied econometrics* 21(1): 79-109.

Blisle, Claude J. P. 1992. "Convergence Theorems for a Class of Simulated Annealing Algorithms on Rd." *Journal of Applied Probability* 29(4): 885-95.

Bollen, Kenneth A., and Kenney H. Barb. 1981. "Pearson's R and Coarsely Categorized Measures." *American Sociological Review* 46(2): 232-39.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of econometrics* 31(3): 307-27.

—. 1987. "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return." *The review of economics and statistics*: 542-47.

Bollerslev, Tim. 1990. "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model." *The Review of Economics and Statistics* 72(3): 498-505.

—. 2008. "Glossary to ARCH (GARCH)." *SSRN eLibrary*. http://papers.ssrn.com.ezproxy.library.wisc.edu/sol3/papers.cfm?abstract_id=1263250 (February 9, 2012).

Bollerslev, Tim, Robert F. Engle, and Jeffrey M. Wooldridge. 1988. "A Capital Asset Pricing Model with Time-Varying Covariances." *Journal of Political Economy* 96(1): 116-31.

Box, G.E.P., G.M. Jenkins, and G.C. Reinsel. 1970. *Time Series Analysis*. Holden-day San Francisco.

Brandt, P.T., and J.R. Freeman. 2009. "Modeling Macro-political Dynamics." *Political Analysis* 17(2): 113-42.

Campbell, James E., Helmut Norpoth, Michael S. Lewis-Beck, Charles Tien, Alan I. Abramowitz, Robert S. Erikson, Christopher Wlezien, Thomas Holbrook, Brad Lockerbie, Alfred G. Cuzn, Charles M. Bundrick, and Carl E. Klarner. 2009. "Forecasting Recap: Assessments of the 2008 National Elections Forecasts." *PS: Political Science and Politics* 42(1): 19-25.

Caporin, Massimiliano, and Michael McAleer. 2013. *Ten Things You Should Know About the Dynamic Conditional Correlation Representation* . Kyoto University, Institute of Economic Research. KIER Working Paper. http://ideas.repec.org/p/kyo/wpaper/870.html (October 14, 2013).

Cappiello, L., S. Manganelli, and B. Gerard. 2005. "Measuring Comovements by

Regression Quantiles." *ECB Working Paper No. 501* .

Chan, F., S. Hoti, and M. McAleer. 2003. "Generalized Autoregressive Conditional Correlation." *unpublished paper, School of Economics and Commerce, University of Western Australia*.

Chan, W.C. 2011. "The Correlation Analysis of Remuneration for Precious Metals Related to the Trending in US Dollar."

Christenson, D.P., and C.D. Smidt. 2011. "Riding the Waves of Money: Contribution Dynamics in the 2008 Presidential Nomination Campaign." *Journal of Political Marketing* 10(1-2): 4-26.

Christenson, Dino. "Correlation." In *International Encyclopedia of Political Science*,.

Christos S. Savva, Denise R. Osborn, and Len Gill. 2006. *Periodic Dynamic Conditional Correlations Between Stock Markets in Europe and the US*. Economics, The Univeristy of Manchester. http://ideas.repec.org/p/man/cgbcrp/77.html (October 27, 2011).

Crato, Nuno, and Bonnie K. Ray. 1996. "Model Selection and Forecasting for Long-range Dependent Processes." *Journal of Forecasting* 15(2): 107-25.

Diebold, Francis X. 2004. "The Nobel Memorial Prize for Robert F. Engle." *The Scandinavian Journal of Economics* 106(2): 165-85.

Ding, Zhuanxin, and Robert F. Engle. 2001. "Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing." http://papers.ssrn.com.ezproxy.library.wisc.edu/sol3/papers.cfm?abstract_id=1294569 (October 14, 2013).

Dueck, Gunter, and Tobias Scheuer. 1990. "Threshold Accepting: a General Purpose Optimization Algorithm Appearing Superior to Simulated Annealing." *Journal of computational physics* 90(1): 161-75.

Durr, R.H. 1994. "An Essay on Cointegration and Error Correction Models." *Political Analysis* 4: 185-227.

Enders, Walter. 2009. *Applied Econometric Times Series*. 3rd ed. Wiley.

Engle, R. 2007. "High Dimension Dynamic Correlations." *NYU Working Paper No. FIN-07-045*.

Engle, R. F., and V. K. Ng. 2012. "Measuring and Testing the Impact of News on Volatility." *The Journal of Finance* 48(5): 1749-78.

Engle, R.F., and C.W.J. Granger. 1987. "Co-integration and Error Correction: Representation, Estimation, and Testing." *Econometrica: Journal of the Econometric Society*: 251-76.

Engle, Robert. 2002a. "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models." *Journal of Business & Economic Statistics* 20(3): 339-50.

—. 2002b. "New Frontiers for Arch Models." *Journal of Applied Econometrics* 17(5): 425-46.

Engle, Robert F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50(4): 987-1007.

—. 1984. "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics." *Handbook of econometrics* 2: 775-826.

Engle, Robert F., and C. W. J. Granger. 2001. "Co-integration and Error Correction: Representation, Estimation, and Testing." In Cambridge, MA, USA: Harvard University Press, 145-72. http://dl.acm.org/citation.cfm?id=781840.781849 (February 25, 2012).

Engle, Robert F., and Kenneth F. Kroner. 1995. "Multivariate Simultaneous Generalized Arch." *Econometric Theory* 11(1): 122-50.

Engle, Robert F., and Kevin Sheppard. 2001. "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH." *National Bureau of Economic Research Working Paper Series* No. 8554. http://www.nber.org/papers/w8554.

Engle, Robert F., Mark W. Watson, Tim Bollerslev, and Jeffrey R. Russell. 2010. *Volatility* *and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Oxford University Press.

Enns, Peter K., and Brian Richman. 2013. "Presidential Campaigns and the Fundamentals Reconsidered." *The Journal of Politics*: 1-18.

Erikson, Robert S. 1989. "Economic Conditions and The Presidential Vote." *The American Political Science Review* 83(2): 567-73.

Erikson, Robert S., Joseph Bafumi, and Bret Wilson. 2001. "Was the 2000 Presidential Election Predictable?" *PS: Political Science and Politics* 34(4): 815-19.

Erikson, Robert S., Michael B. Mackuen, and James A. Stimson. 2002. *The Macro Polity*. Cambridge University Press.

Erikson, Robert S., Costas Panagopoulos, and Christopher Wlezien. 2004. "Likely (and Unlikely) Voters and the Assessment of Campaign Dynamics." *The Public Opinion Quarterly* 68(4): 588-601.

Erikson, Robert S., and Christopher Wlezien. 1996. "Of Time and Presidential Election Forecasts." *PS: Political Science and Politics* 29(1): 37-39.

—. 1999. "Presidential Polls as a Time Series: The Case of 1996." *The Public Opinion Quarterly* 63(2): 163-77.

—. 2008a. "Are Political Markets Really Superior to Polls as Election Predictors?" *The Public Opinion Quarterly* 72(2): 190-215.

—. 2008b. "Leading Economic Indicators, the Polls, and the Presidential Vote." *PS: Political Science and Politics* 41(4): 703-7.

—. 2012. *The Timeline of Presidential Elections: How Campaigns Do (and Do Not) Matter*. University of Chicago Press.

Ferejohn, John A., and James H. Kuklinski. 1990. *Information and Democratic Processes*. Univ of Illinois Pr.

Franklin, Charles H. 1984. "Issue Preferences, Socialization, and the Evolution of Party Identification." *American Journal of Political Science* 28(3): 459-78.

—. 1991. "Eschewing Obfuscation? Campaigns and the Perception of U.S. Senate Incumbents." *The American Political Science Review* 85(4): 1193-1214.

—. 1992. "Measurement and the Dynamics of Party Identification." *Political Behavior* 14(3): 297-309.

—. 1993. "Senate Incumbent Visibility over the Election Cycle." *Legislative Studies Quarterly* 18(2): 271-90.

Franklin, Charles H., and John E. Jackson. 1983. "The Dynamics of Party Identification." *The American Political Science Review* 77(4): 957-73.

Freeman, John R. 1983. "Granger Causality and the Times Series Analysis of Political Relationships." *American Journal of Political Science* 27(2): 327-58.

Gardner, E.S., and others. 2006. "Exponential Smoothing: The State of the art-Part II." *International Journal of Forecasting* 22(4): 637-66.

Gardner Jr, E.S. 1985. "Exponential Smoothing: The State of the Art." *Journal of Forecasting* 4(1): 1-28.

Gelman, Andrew. 2007. "Rejoinder: Struggles with Survey Weighting and Regression Modeling." *Statistical Science* 22(2): 184-88.

Gelman, Andrew, and Gary King. 1993. "Why Are American Presidential Election Campaign Polls so Variable When Votes Are so Predictable?" *British Journal of Political Science* 23(4): 409-51.

Gerber, Alan, and Donald P. Green. 1998. "Rational Learning and Partisan Attitudes." *American Journal of Political Science* 42(3): 794-818.

Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. Cambridge University Press.

Gordon C. R. Kemp. 1999. "The Behavior of Forecast Errors from a Nearly Integrated AR(1) Model as Both Sample Size and Forecast Horizon Become Large." *Econometric Theory* 15(2): 238-56.

Granger, C. W. J., and M. J. Morris. 1976. "Time Series Modelling and Interpretation." *Journal of the Royal Statistical Society. Series A (General)* 139(2): 246-57.

Granger, C.W.J. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14(2): 227-38.

Granger, C.W.J., and R. Joyeux. 1980. "An Introduction to Long-memory Time Series Models and Fractional Differencing." *Journal of time series analysis* 1(1): 15-29.

Granger, Clive WJ, and Paul Newbold. 1974. "Spurious Regressions in Econometrics." *Journal of econometrics* 2(2): 111-20.

Green, Donald P., Alan S. Gerber, and Suzanna L. de Boef. 1999. "Tracking Opinion Over Time: A Method for Reducing Sampling Error." *The Public Opinion Quarterly* 63(2): 178-92.

Green, Donald, Bradley Palmquist, and Eric Schickler. 1998. "Macropartisanship: A Replication and Critique." *The American Political Science Review* 92(4): 883-99.

Green, Donald Philip, and Bradley Palmquist. 1994. "How Stable Is Party Identification?" *Political Behavior* 16(4): 437-66.

Greene, William H. 2011. *Econometric Analysis*. 7th ed. Prentice Hall.

Hamilton, James Douglas. 1994. 2 *Time Series Analysis*. Cambridge Univ Press. http://journals.cambridge.org.ezproxy.library.wisc.edu/ production/action/cjoGetFulltext?fulltextid=3183396 (October 14, 2013).

Hrdle, Wolfgang, Helmut Ltkepohl, and Rong Chen. 1997. "A Review of Nonparametric Time Series Analysis." *International Statistical Review* 65(1): 49-72.

Harvey, A.C. 1987. "Applications of the Kalman Filter in Econometrics." In *Advances in econometrics: fifth world congress*, , 285-313.

Hillygus, D. Sunshine. 2005. "Campaign Effects and the Dynamics of Turnout Intention in Election 2000." *The Journal of Politics* 67(1): 50-68.

Hinckley, Katherine A., and John C. Green. 1996a. "Fund-Raising in Presidential Nomination Campaigns: The Primary Lessons of 1988." *Political Research Quarterly* 49(4):

693-718.

—. 1996b. "Reply: 'We Can't Believe He Ate the Whole Thing'." *Political Research Quarterly* 49(4): 727-28.

Holbrook, T. M. 2009. "Economic Considerations and the 2008 Presidential Election." *PS: Political Science and Politics* 42(3): 473-78.

Holbrook, Thomas M. 1996. "Reading the Political Tea Leaves A Forecasting Model of Contemporary Presidential Elections." *American Politics Research* 24(4): 506-19.

—. 2008. "Incumbency, National Conditions, and the 2008 Presidential Election." *PS: Political Science and Politics* 41(4): 709-12.

Hosking, J. R. M. 1981. "Fractional Differencing." *Biometrika* 68(1): 165-76.

Huckfeldt, Robert, Edward G. Carmines, Jeffery J. Mondak, and Eric Zeemering. 2007. "Information, Activation, and Electoral Competition in the 2002 Congressional Elections." *Journal of Politics* 69(3): 798-812.

Iglesias, Emma M., and Garry D. A. Phillips. 2011. "Small Sample Estimation Bias in GARCH Models with Any Number of Exogenous Variables in the Mean Equation." *Econometric Reviews* 30(3): 303-36.

Johnston, R., and E. Thorson. 2009. "The Economy, the Candidates, and the 2008 Campaign." http://papers.ssrn.com.ezproxy.library.wisc.edu/sol3/papers.cfm?abstract_id=1451215 (July 4, 2012).

De Jong, Piet, and Jeremy Penzer. 2004. "The ARMA Model in State Space Form." *Statistics & probability letters* 70(1): 119-25.

Kenski, Kate, Bruce W Hardy, and Kathleen Hall Jamieson. 2010. *The Obama Victory: How Media, Money, and Message Shaped the 2008 Election*. Oxford; New York: Oxford University Press.

Lau, Richard R. 1985. "Two Explanations for Negativity Effects in Political Behavior."

American Journal of Political Science: 119-38.

Lau, Richard R., and David P. Redlawsk. 1997. " Voting Correctly." *American Political Science Review*: 585-98.

—. 2001. " Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science*: 951-71.

Lazarsfeld, Paul Felix, Bernard Berelson, and Hazel Gaudet. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. Duell, Sloan and Pearce.

Le, and Chang. 2010. " Oil and Gold: Correlation or Causation?" *Working Paper*.

Lebo, Matthew J., and Daniel Cassino. 2007. " The Aggregated Consequences of Motivated Reasoning and the Dynamics of Partisan Presidential Approval." *Political Psychology* 28(6): 719-46.

Lebo, Matthew J., and Janet M. Box-Steffensmeier. 2008. " Dynamic Conditional Correlations in Political Science." *American Journal of Political Science* 52(3): 688-704.

Ledoit, Olivier, Pedro Santa-Clara, and Michael Wolf. 2003. " Flexible Multivariate GARCH Modeling with an Application to International Stock Markets." *The Review of Economics and Statistics* 85(3): 735-47.

Li, Qi, and Jeffrey Scott Racine. 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton University Press. http://books.google.com.ezproxy.library.wisc.edu/books?hl=en&l r=&id=BI_PiWazY 0YC&oi=fnd&pg=PP2&dq=li+and+racine+non parametric+econometrics&ots=eX45lbmj5y&si g=8mgArbrglN4jW4H1lOdKpq2VioE (October 14, 2013).

Linton, O., and E. Mammen. 2005. " Estimating Semiparametric ARCH (?) Models by Kernel Smoothing Methods." *Econometrica* 73(3): 771-836.

Lumley, Thomas. 1996. " Generalized Estimating Equations for Ordinal Data: A Note on Working Correlation Structures." *Biometrics* 52(1): 354-61.

MacKuen, Michael B., Robert S. Erikson, and James A. Stimson. 1989.

" Macropartisanship." *The American Political Science Review* 83(4): 1125-42.

Maddala, G. S., and In-Moo Kim. 1999. *Unit Roots, Cointegration, and Structural Change*. Cambridge University Press.

Marzo, and Zagaglia. " Gold and the U.S. Dollar: Tales from the Turmoil." *Working Paper*.

Mayer, William G. 2003. " Forecasting Presidential Nominations or, My Model Worked Just Fine, Thank You." *PS: Political Science and Politics* 36(2): 153-57.

McAleer, M., F. Chan, S. Hoti, and O. Lieberman. 2008. " Generalized Autoregressive Conditional Correlation." *Econometric Theory* 24(06): 1554-83.

Mills, T.C. 2004. " Statistical Analysis of Daily Gold Price Data." *Physica A: Statistical Mechanics and its Applications* 338(3): 559-66.

Mills, Terence. 2004. " Exploring the Relationship Between Gold and the Dollar." *Significance* 1(3): 113-15.

Mitchell, Dona-Gene. 2012. " It's About Time: The Lifespan of Information Effects in a Multiweek Campaign." *American Journal of Political Science* 56(2): 298-311.

Newey, Whitney K., and Daniel McFadden. 1994. " Large Sample Estimation and Hypothesis Testing." *Handbook of econometrics* 4: 2111-2245.

Niemi, Richard G., and Larry M. Bartels. 1984. " The Responsiveness of Approval Voting to Political Circumstances." *PS* 17(3): 571-77.

Otranto, E. 2010. " Identifying Financial Time Series with Similar Dynamic Conditional Correlation." *Computational Statistics & Data Analysis* 54(1): 1-15.

Perron, Pierre. 1989. " The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 57(6): 1361-1401.

—. 1990. " Testing for a Unit Root in a Time Series with a Changing Mean." *Journal of Business & Economic Statistics* 8(2): 153-62.

Perron, Pierre, and Cosme Vodounou. 2005. " The Variance Ratio Test: An Analysis of Size and Power Based on a Continuous-Time Asymptotic Framework." *Econometric Theory* 21(3):

562-92.

Peterson, David A. M. 2009. "Campaign Learning and Vote Determinants." *American Journal of Political Science* 53(2): 445-60.

Poon, Ser-Huang, and Clive W. J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41(2): 478-539.

Ramirez, Mark D. 2013. "The Policy Origins of Congressional Approval." *The Journal of Politics* 75(01): 198-209.

Rapach, David E., and Jack K. Strauss. 2008. "Structural Breaks and Garch Models of Exchange Rate Volatility." *Journal of Applied Econometrics* 23(1): 65-90.

Ray, Bonnie K. 1993. "MODELING LONG-MEMORY PROCESSES FOR OPTIMAL LONG-RANGE PREDICTION." *Journal of Time Series Analysis* 14(5): 511-25.

Ridout, Martin S., Clarice G. B. Demtrio, and David Firth. 1999. "Estimating Intraclass Correlation for Binary Data." *Biometrics* 55(1): 137-48.

Lee Rodgers, Joseph, and W. Alan Nicewander. 1988. "Thirteen Ways to Look at the Correlation Coefficient." *The American Statistician* 42(1): 59-66.

Romer, Daniel. 2006. *Capturing Campaign Dynamics, 2000 and 2004: The National Annenberg Election Survey*. Philadelphia: University of Pennsylvania.

Santos, Manuel S., and Adrian Peralta-Alva. 2005. "Accuracy of Simulations for Stochastic Dynamic Models." *Econometrica* 73(6): 1939-76.

Shaw, Daron R. 1999. "A Study of Presidential Campaign Event Effects from 1952 to 1992." *The Journal of Politics* 61(2): 387-422.

Sides, John, and Lynn Vavreck. 2013. *The Gamble: Choice and Chance in the 2012 Presidential Election*. Princeton University Press.
http://books.google.com.ezproxy.library.wisc.edu/books?hl=en&lr=
id=gwBPySWcAx4C&oi=fnd&pg=PP2&dq=sides+vavreck+gamble&ots=IwzPf1uuDA
sig=QrnHbhRSCe0Y1x1GQsDMzE6QQNA (October 15, 2013).

Silvennoinen, A., and T. Tersvirta. 2009. " Multivariate GARCH Models." *Handbook of Financial Time Series*: 201-29.

Skintzi, V.D., and A.N. Refenes. 2006. " Volatility Spillovers and Dynamic Correlation in European Bond Markets." *Journal of International Financial Markets, Institutions and Money* 16(1): 23-40.

Box-Steffensmeier, J. M. 1996. " A Dynamic Analysis of the Role of War Chests in Campaign Strategy." *American Journal of Political Science*: 352-71.

Box-Steffensmeier, J. M., D. Darmofal, and C. A. Farrell. 2009. " The Aggregate Dynamics of Campaigns." *The Journal of Politics* 71(1): 309-23.

Box-Steffensmeier, Janet M. 1996. " A Dynamic Analysis of The Role of War Chests in Campaign Strategy." *American Journal of Political Science* 40(2): 352-71.

Box-Steffensmeier, Janet M., and Anand E. Sokhey. 2007. " A Dynamic Labor Market: How Political Science Is Opening up to Methodologists, and How Methodologists Are Opening up Political Science." *PS: Political Science and Politics* 40(1): 125-27.

Box-Steffensmeier, Janet M., and Suzanna De Boef. 2001. " Macropartisanship and Macroideology in the Sophisticated Electorate." *The Journal of Politics* 63(1): 232-48.

Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier, eds. 2008. *The Oxford Handbook of Political Methodology*.

Box-Steffensmeier, Janet M., Kathleen Knight, and Lee Sigelman. 1998. " The Interplay of Macropartisanship and Macroideology: A Time Series Analysis." *The Journal of Politics* 60(4): 1031-49.

Box-Steffensmeier, Janet M., and Renee M. Smith. 1998. " Investigating Political Dynamics Using Fractional Integration Methods." *American Journal of Political Science* 42(2): 661-89.

Box-Steffensmeier, Janet M., and Rene M. Smith. 1996. " The Dynamics of Aggregate Partisanship." *The American Political Science Review* 90(3): 567-80.

Traugott, Michael W., and Christopher Wlezien. 2009. " The Dynamics of Poll Performance

During the 2008 Presidential Nomination Contest." *Public Opinion Quarterly* 73(5): 866-94.

Tsay, Ruey S. 2005. *Analysis of Financial Time Series (Wiley Series in Probability and Statistics)*. 2nd edition. Wiley-Interscience.

—. 2006. "Multivariate Volatility Models." *Lecture Notes-Monograph Series* 52: 210-22.

Tse, Y.K., and A. Tsui. 2000. "A Multivariate Garch Model with Time-varying Correlations."

Weber, Enzo. "Structural Dynamic Conditional Correlation." *Working Paper*.

Weide, Roy van der. 2002. "GO-GARCH: A Multivariate Generalized Orthogonal GARCH Model." *Journal of Applied Econometrics* 17(5): 549-64.

Winker, Peter, and Dietmar Maringer. 2009. "The Convergence of Estimators Based on Heuristics: Theory and Application to a GARCH Model." *Computational Statistics* 24(3): 533-50.

Wlezien, Christopher. 2003. "Presidential Election Polls in 2000: A Study in Dynamics." *Presidential Studies Quarterly* 33(1): 172-87.

Wlezien, Christopher, and Robert S. Erikson. 2002. "The Timeline of Presidential Election Campaigns." *The Journal of Politics* 64(4): 969-93.

—. 2004. "The Fundamentals, the Polls, and the Presidential Vote." *PS: Political Science and Politics* 37(4): 747-51.