Consensus and Volatility in Presidential Approval

by

Paul Gronke
Department of Political Science
Duke University
Durham, NC 27708-0204
email: gronke@acpub.duke.edu

Abstract

One of the central themes in research, teaching, and popular commentary about American politics is the power of the Presidency. This is no surprise, since American domestic politics and policymaking, and even many decisions on the world stage, rotate about the axis of the U.S. president. Presidential approval plays a particularly critical role in presidential power, as ample studies have demonstrated over the past twenty years. We must be concerned, then, over recent evidence of an increasingly fickle American public. Presidents have enjoyed new peaks but also suffered new lows in popular approval. Changes in approval seem to be more rapid and the diversity of opinions at any one time seems to be growing. If the American public has become more *volatile* and less *consensual* in their evaluations of the president, it will be difficult for presidents to build and maintain legislative coalitions, unify domestic actors behind foreign policy initiatives, and ultimately maintain their electoral coalitions. The president's success in swaying public opinion and the president's ability to use public support as a leadership tool may depend as much on the amount of agreement (consensus) and the lability of public support (volatility) as it does on the average level of approval. This paper takes an opening look at measuring and modeling volatility and consensus in presidential approval. Drawing on data from the New York Times/CBS polls from 1976–1993, the paper shows, via descriptive analysis, how mean approval, over time volatility, and cross sectional variability, are distinct processes. The paper then proposes some alternative ways to examine volatility and consensus.

Paper prepared for the annual meeting of the Southern Political Science Association, Atlanta, GA, November 1996. The author would like to thank John Aldrich, John Brehm, Lyle Scruggs, Renée Smith, Janet Box–Steffensmeier, and Jim Stimson for encouragement and help, and the ICPSR (and Suzanne DeBoef) for access to the NYT data. All remaining errors are certainly the fault of the author.
Draft Copy: Comments and criticism highly encouraged

Introduction

Presidential approval is a key political resource in an era of the public relations presidency. When approval ratings are higher, members of the president's party in Congress are less likely to be defeated in the midterm and the president's legislative proposals are more successful. The presidential agenda, both on the foreign and domestic side, is "bolder" under conditions of high approval.[1] And presidents enjoying high levels of approval are more likely to be reelected. The president's performance in this "perpetual election" (Simon and Ostrom, 1989) or "new referendum" (Brace and Hinckley, 1992) is clearly a key to understanding presidential power in the postwar era (Neustadt, 1990).

The conventional method for understanding changes in presidential approval has been to model movements in *aggregate mean approval* as a function of time in office, economic change, and "rally events."[2] Models of the mean, however, abandon an enormous quantity of information, namely, the attitudes of individual respondents. One recent piece (Brehm and Gronke, 1995) attempts to extract individual response variation via change in the variance portion of the time series. These results are briefly described below.

There is another, obvious way to "extract" individual level data: use the individual responses. While the move from aggregate to individual level data may seem obvious, there are many theoretical and technical difficulties associated with the change, many of which I have not surmounted by the time of this essay. Here, I present some preliminary ideas about modeling consensus and volatility using combined cross sectional surveys and aggregate political and economic data. I start by showing how individual and aggregate level variability are different theoretically, but how they are empirically difficult to disentangle. Next, I present some graphics which illustrate the basic differences which I hope to analyze. Finally, I present some suggestions about analytic strategies for these data, providing what I hope is sufficient grist for discussion and criticism.

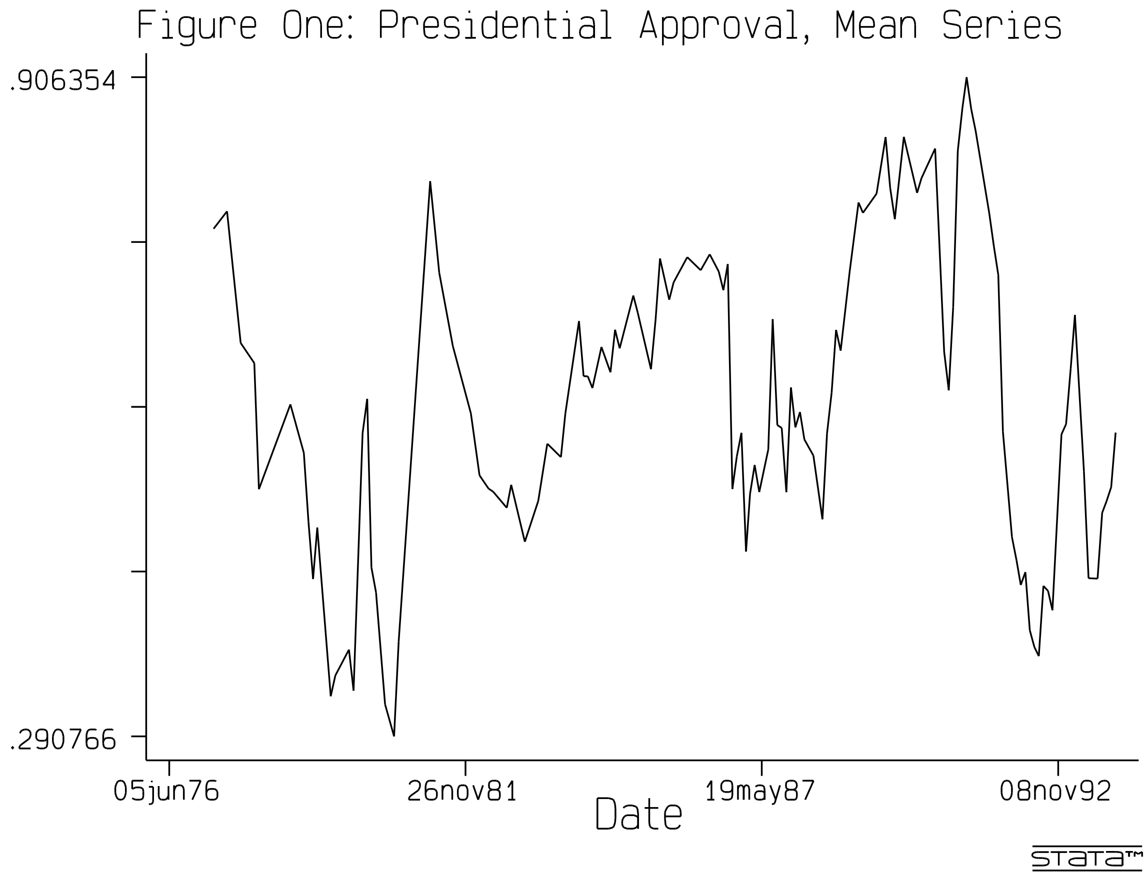
[1] These claims are well established in the literature on elections, presidential approval, and presidential leadership. The citations are too numerous to include here. Four recent volumes are Brody (1991), Brace and Hinckley (1992), Kernell (1993), and King and Ragsdale (1988). For evidence on legislative success, see Rivers and Rose (1985) and for evidence on "boldness" of agenda, defined as breadth of legislative proposals, frequency of vetoes, and use of military force, see Simon and Ostrom (1989).

[2] C.f. Jackman, 1995; MacKuen, Erikson, and Stimson, 1992; Simon and Ostrom, 1989; Ostrom and Simon, 1985; Brody, 1991; Brace and Hinckley, 1992; Kernell, 1983, 1978; Mueller, 1973.

Defining Volatility and Consensus

The presidential approval series is quite familiar to most political scientists. As shown in figure below, presidential approval has meaningful ebbs and flows, in response to the state of the economy, elections, wars, and presidentially relevant "events."[3] Thus, the figure below, plotted from 1976–1993, shows a relatively steady decline during Carter's term, with a "rally" after the Iran hostage crisis, followed by an even steeper fall; a honeymoon after Ronald Reagan's election, followed by a drop during the recession of 1981, followed by a long climb until the Iran-Contra hearings; another honeymoon after Bush's reelection, and a high point of popularity during the Gulf war, again followed by a free fall up to the election of Clinton. These patterns are well understood by the political science profession (see citations listed above). However, I would like to take another look at the approval series, and suggest that our understanding of what changes in approval mean have, until now, been missing a key element: the shape of the underlying distribution.

[3] The figure here plots aggregate presidential approval from the New York Times/CBS polls, conducted on approximately a monthly basis, from 1976 through 1993 (from the ICPSR archive). The item used is identical to that used by Gallup: "Do you approve or disapprove of the way (PRESIDENT) is handling his job as president?" The figures here correlate at .99 with the Gallup data.



Presidential approval is theoretically a continuous variable (see, for example, Ostrom and Smith, 1992; Ostrom and Simon, 1988), yet we only observe (except for a five point scale administered irregularly by the National Election Studies) dichotomous responses, disapprove or approve. While I do not have a special insight into the original justification of this item, its use through the years suggests that Gallup saw it as analogous to a "thermostatic" measure of public opinion about the president (MacKuen, Erikson, and Stimson, 1995). Some suggest, in fact, that the aggregate measure is all that is of interest, and we should ignore individual ratings of presidential job performance (Brehm and Gronke, 1994). Still, one could imagine a distribution of approval ratings, ranging from those who have a visceral reaction against everything the President says and does, to those who blindly endorse all presidential actions, where the bulk of Americans probably fall somewhere in the middle. If one accepts this premise, then one might also ask whether the shape of this theoretical distribution — peaked or flat, unimodal or multimodal — has any implications for public opinion about the president.

I suggest that it does, and the way to think about the distribution is via volatility and consensus. By *volatility*, I mean the degree to which presidential approval changes in response to roughly equivalent shocks across administrations. This is an aggregate, over time statistic. An administration where presidential approval decays more under one kind of shock than another administration under the same kind of shock reflects a more volatile public. Approval ratings that are likely to slip sharply in one direction or another are "volatile" ratings. Alternatively, volatility in approval can be conceptualized as the relative weight of the short term components of presidential approval (e.g. more rapid or steeper growth in the "coalition of minorities" (Mueller, 1973) or greater responsiveness to economic change), holding events constant.[4]

By *consensus*, I mean the spread of the distribution of presidential approval at any single time point. A less consensual public reveals a relatively wide variance in approval, regardless of the mean. Widespread agreement on the performance of the president, on the other hand, implies a tight variance about the cross–sectional mean.[5] Finally, consensus can be examined over time, in what I will call *consensus–building or dissensus–causing shifts* in approval. An example of consensus building events were the series of Soviet actions after World War II, which Page and Shapiro cite as a cause of the very high levels of consensus about the proper American position toward the Soviets (Page and Shapiro, 1992, Ch. 5). Examples of dissensus–building events were the U.S. actions in Vietnam in the early 1970s, which Zaller (1992) shows resulted in wide partisan disagreement on the U.S. role. This project moves beyond past studies by considering the shape as well as the central tendency of the distribution of presidential approval.

Should we care about volatility and consensus? In the most general sense, increases in volatility and decreases in consensus hurt the president's ability to build support in Congress and pass his legislative program. From the perspective of a rational political actor deciding whether to support or oppose a president, the possibility for rapid changes in approval introduces uncertainty into decisions. If a president's job approval is relatively stable, Congress, the media, and the public have a consistent image of the President. Political actors can make calculations without bothering

[4]From this perspective, events are "shocks" that we need to control for in order to gain leverage on the relative weight of the short and long term components.

[5]Because presidential approval is a binomial variable in the Gallup studies, there is no way to discriminate between different underlying variances at the individual level. However, as I will argue below, one can use group differences another measure of consensus.

to include uncertainty. A president with a wildly variable base of support, in contrast, introduces a host of uncertainties into the political system. Should I tie my fortunes to a popular President today, when he may be unpopular tomorrow? This dovetails with recent claims that modern presidents are highly sensitive to changes in opinion polls (Brace and Hinckley, 1992). Presidents delay certain actions because they cannot afford a drop in the polls or pursue other activities, especially foreign policy adventures, because they expect a boost in popularity. George Bush received criticism for attempting to protect his 1991 post Gulf War popularity surge into the 1992 election by purposely adopting a limited domestic agenda. What this research will help us answer is whether recent presidents have had to become more sensitive to public opinion because the public has become increasingly volatile. If change is important, the amount of change is as well.

Volatility allows us to test the hypothesis that presidential approval has become fundamentally less stable over time, with a resulting decline in the president's ability to act. Like many summary political judgements, presidential approval has both a long term and a short term component. The long–term, stable component of presidential approval includes party affiliation, strength of party attachments, and ideological beliefs (Mueller, 1973; Fiorina, 1981). Short term influences, those changes and events in the world which lead to perturbations in approval, include economic performance, wars, scandals, and what have come to be called "rally events" (Mueller, 1973; Kernell, 1978). If we conceive of volatility as the relative weight of short term components on approval, holding events to the side, then an increase in volatility implies that the president is no longer able to rely on a core base of support during his term in office. Presidential approval, and presidential power along with it, becomes a victim of events that are largely outside of the president's control.

If we conceive of volatility more broadly as the degree to which approval changes in response to comparable "shocks" and consensus as the level of intergroup agreement on the president's performance, then volatility and consensus will lead to a more nuanced picture of presidential "rallies." Extensive prior research, starting with Mueller (1973), has demonstrated the importance of key political events to presidential approval. Some events, such as wars, assassination attempts, and good economic news, boost the presidential stock while other events, such as scandals and civil discontent, slash away at presidential support. Conventionally, these events have been described narrowly as "rally points" and "negative events."[6] Yet, rallies and busts

[6]For example, see Simon and Ostrom (1989), Brace and Hinckley (1992), and Kernell (1978). Brody (1991) provides a nice critique of this approach.

contain implicit claims about consensus. In this respect, presidential rallies can be viewed like the stock market. The daily stock report tells us not only the rise or fall of the Dow (movement in the mean), but also whether gainers out paced losers (the cross sectional variance). A rally across the board is different than a rise marked by peaks and valleys — particularly if you've invested in the wrong stock! The same applies to presidential approval, and to public opinion more generally. Rather than leading only to a boost or drop in mean approval, we also want to know whether certain kinds of events might be associated with changes in consensus.[7]

Volatility and consensus, in tandem, provide us with the tools to explore changing patterns of presidential leadership. If the Neustadtian president was able to marshall generalized public support as a regular tool of leadership, the Kernellian president has to constantly appeal to a labile, fickle public (Neustadt, 1990; Kernell, 1993). If Kernell is right, and the combined impact of the Vietnam War, civil discontent, and the new style of press coverage led to a fundamental restructuring of the presidency, we ought to observe some evidence of this in presidential approval. These kinds of changes will not be evident in the mean, however, but in the new responsiveness of the public to events, i.e., volatility and consensus. Similarly, the different kind of political circumstances described in a historical study of presidential leadership (Skowronek, 1993) should be evident in the pattern of public response during those presidencies. A president during the "politics of disjunction," such as Carter, is unable to establish credibility and is "suddenly vulnerable;" while a president, like Roosevelt or Reagan, who practices a "politics of reconstruction" is able to rebuild the regime in a "new and more relevant way." (Skowronek, 1993, Ch. 2). These descriptions sound, without doing too much damage to this seminal work, like one president who encounters widespread dissensus, Carter, and two presidents who rebuilt consensus. This project will provide a quantitative lens to these qualitative accounts of presidential style and leadership.

Limits to Conventional Wisdom Regarding Volatility and Consensus

It remains to be seen, however, that the empirical representation of approval as a dichotomy provides any leverage at all on the shape of the theoretically continuous

[7]For example, scholars of international relations and public opinion and foreign policy describe a fundamental shift in the public's attitude toward U.S. involvement in foreign policy conflicts (e.g. Holsti, 1992; Jentleson, 1992). The implication is that the post–Vietnam public ought to react differently — less consensually – to foreign policy engagements. The data collected as part of this project will allow a direct test of this hypothesis.

distribution. At the aggregate level, of course, approval is measured in a more continuous fashion, bounded only as a proportion, at 0 and 1.0. This has led scholars to examine change in approval via the standard deviation in ratings. This evidence for volatility is spotty and inconclusive. One set of analysts suggest that "instability has not increased steadily over time" (Gallup and Edwards, 1990, p.122), while another set suggest that volatility increased during George Bush's term (Brace and Hinckley, 1992).

The technique of estimating the standard deviation of approval ratings over an administration compensates for unusual highs and lows, but does not recognize the turbulence of history. Some administrations have a high frequency of events (e.g., the Johnson administration) while others have a relatively low frequency of events (e.g., the Carter administration), yet the standard deviation of the two administrations are relatively close. Does the slightly higher standard deviation of approval under Johnson indicate a more "volatile" public than that under Carter, even if fewer consequential events took place? Similarly, how can one compare the standard deviation during Nixon's term with any other without accounting for Vietnam protests and the Watergate scandal? Comparing standard deviations alone is the wrong way to examine volatility. If the empirical referent for volatility is the variance in the approval equation, then this should be modeled directly.

Consensus, in contrast, has seldom been directly featured in any studies of presidential approval. I do not mean to suggest that disagreement has not been recognized. The "politics that presidents make," historically, has been a politics of taking advantage of partisan divisions and repudiating rising opposition (Skowronek, 1993). Kernell (1993), Brace and Hinckley (1992), and Brody (1991) all make reference to the level of public disagreement. But all the evidence is indirect, either newspaper accounts, measures of legislative opposition, or measures of support among party groups in the population. This study, to my knowledge, will be the first to examine subgroup variation within the cross sections directly. As such, it can build a new conventional wisdom about public consensus.

Measurement Issues: Time Series, Cross Sections, and Pooled Cross Sectional Time Series

In a time series, volatility can be measured, and modeled, in the stochastic portion of a time series. Measuring volatility and consensus in a cross section is substantially more problematic. While I can provide some suggestive ways to examine consensus, all involve significant assumptions about the structure of presidential approval.

Modeling Volatility in A Time Series[8]

Volatility can be conceptualized in two ways, as variance across a time series, and as changes over time in the weight that individuals place on certain considerations when evaluating the president. Each of these represent new approaches to the study of volatility of approval.

Scholars of the presidency know that we cannot understand *mean levels* of presidential approval without knowing the particular circumstances of each administration, whether this includes economic performance, time in office, and war casualties (Mueller, 1973), media coverage of the administration (Brody, 1991), or the "dramatic events of a term" (Brace and Hinckley, 1992, p. 10). Brehm and Gronke (1994, 1995) model the *variance* of presidential approval in a similar fashion. Typically, the analyst chooses an appropriate function to reparameterize the mean level of approval (μ) in terms of the substantive explanatory variables, while leaving σ^2 as a constant:

$$\mu = f(X, \beta) \tag{1}$$

What has been long neglected in studies of presidential approval is the opportunity to reparameterize σ^2 and to identify independent variables that help explain changes in the variance. Nor have there been any attempts to extend these models down to the individual level. Neither of these steps are merely methodological exercises. The aim is explicitly substantive: it allows the analyst to test hypotheses that implicate both over time and individual level indicators of political change on the president's ability to muster popular support.

In other work, my coauthor and I use an estimator which combines two classes of models of heteroskedasticity in time-series (Brehm and Gronke, 1994, 1995). One approach to modeling variance in a time series ("multiplicative heteroskedasticity") treats variance as a function of some set of explanatory variables, but is not conditional on variance in prior periods (Greene, 1993, p. 405–407; Harvey, 1990; for application to presidential popularity, see Beck, 1991). An alternative approach ("autoregressive conditional heteroskedasticity" or ARCH) treats variance as conditional on variance in previous periods, yet typically omits explanatory variables (Greene, 1993, p. 438-442; Harvey, 1990). In our work, we combine these two into an estimator we label ARCH-EDL. The autoregressive nature of the time series is retained, but

[8]This material in this section is presented in more detail in Brehm and Gronke, 1995

we add a series of explanatory variables. The parameterization for the variance thus becomes:

$$\sigma_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + Z\gamma \quad (2)$$

where α_0 and α_1 represent the ARCH parameters, Z is a matrix of explanatory variables, and $Z\gamma$ represents the impact of other explanatory variables on σ. This straightforward extension of Engle and Beck's work reveals a statistically significant increase in volatility over the post war era. The level of volatility is responsive to real world events.

Modeling Consensus and Volatility in a Time Series of Cross Sections

The cross sectional analogue of volatility is consensus: the degree of agreement or disagreement at any time point on presidential performance. What is substantially more problematic, however, is deriving an analogous measure for cross sections. Ideally, I would operationalize consensus as the variance within any single cross–section. This is impossible, given the binomial nature of the approval variable. In the cross section, there is no measure of the standard error that is not a simple transformation of the level of approval. Lacking access to a continuous distribution, there is another way to gain leverage on consensus. The underlying construct that I wish to represent is the level of disagreement on the president's performance, both within a cross section, and during significant historical events (thus comparing cross sections over time). An alternative way to think about consensus, and one which relates more directly to the problem of building and maintaining support coalitions, is to identify the positions of major groups within the U.S. population, and to think of consensus as a process of groups unifying and dividing on their level of approval. A more heterogeneous set of approval ratings across groups indicates lower consensus, which a unified set of ratings indicates high levels of agreement.

One quantitative technique for measuring heterogeneity on some characteristic across a set of groups has been developed elsewhere as a measure of population diversity (Gronke, 1993; see also Sullivan, 1973). Essentially, the statistic represents the probability that a pair of randomly selected individuals from a population will both "agree" (have the same value) on the characteristic under study, or, in our case, both approve of the president's performance:

$$A = 1 - (\sum_{k=1}^{p} Y_k^2)/V \quad (3)$$

where A = the probability of agreement, Y_k = the proportion of the population falling in a given category within each of the groups (e.g. the marginals), V = the number of variables (groups) which the population is broken down into, and p = the total number of categories across all variables (in our case, $2 * V$). The critical assumption underlying the technique is that the process of categorizing individuals into groups completely represents the underlying distribution of approval, thus each group member has an equal probability of "approving." If the groups are chosen carefully, this assumption does not do serious damage to our inferences. A consensus time series could be created out of cross sectional data, and be employed as a dependent variable in regression analyses.

The problem with this technique is that it treats all political cleavages equally, so that diversity among religious, economic, or racial groups receive identical weights. Second, among what categories do we calculate diversity? One possibility would be to calculate a measure heterogeneity or homogeneity within the group who approved of the president's performance. The argument would be that a more heterogeneous set of approvers indicates more consensual appeal. On the other hand, we may want to calculate diversity across the groups, calculating the likelihood that one "approver" and one "disapprover" (our sample pair) differs on the characteristics under study. My intuition is that these two ways of calculating diversity are simple transformations of each other, but I have not yet attempted to demonstrate this.

As a first cut at population diversity, I will instead rely on a more simple, straightforward measure of variance. I plot below the relative levels of agreement among politically consequential groups in the population. For example, how do Blacks and whites or Republicans, Independents, and Democrats differ in their evaluations of presidential performance? This procedure is no different than producing a set of time series on approval, aggregating from each randomly selected cross section, and pooling the time series. For each time point, an overall measure can easily be derived using the standard variance measure for pooled data, calculated across j groups, each with n respondents:

$$Variance_t = \frac{\sum(\bar{X}_j - \bar{X})^2 * n_j}{(\sum n_j) - 1}$$

This is a strictly descriptive measure, but may provide some insight into ebbs and flow in variance of approval.

Volatility may also be conceived of in another way, as the relative weight of short and long term considerations used by citizens when they evaluate the president. If the public has become more volatile over time, then regression estimates on, say, strength of partisanship ought to decline over time, while the weights attached to economic change and influential events should increase. Taken to a logical extreme, this representation implies a least square dummy variable (LSDV) specification of the approval model:

$$Y_{nt} = \alpha_{nt} + \lambda_t + \mu_n + X_{nt}\beta_k + \xi_{nt} \tag{4}$$

for $n = (1...N)$ cross sections over $t = (1...T)$ time points (Sayrs, 1989). Unit effects are assumed for each time point. LSDV, however, does not speak to the question of whether the β's vary over time. Furthermore, pooled cross sectional procedures within at least two statistical packages, SAS and STATA, assume *panel* data, e.g. time series on the same units pooled over time. This means that I cannot use the individual level data, but may proceed using the data aggregated into groups. While this strikes me as a serious loss of information, it has been suggested that, so long as I aggregate upon variables that are contained in the regression equation, going from cross sections to time series entails no loss of information. I have put off this step in the study for the future.

The Middle Ground: Assessing Causality in Pooled Cross Sectional Time Series

A final, promising direction for assessing the impact of individual level opinions and demographic characteristic as well as aggregate measures of important political events and economic changes lies somewhat in between purely aggregate analyses and purely cross sectional analyses. Three recent studies, mainly focussed on macropartisanship, touch on very similar issues to those raised here. Box Steffensmeier and Smith (1997) demonstrate that the macropartisanship series (MacKuen, Erikson, and Stimson, 1989) is fractionally integrated, meaning that the "aggregate measures of Republican and Democratic identifiers exhibit neither stationarity nor unit root behavior." This suggests both that there is behavioral heterogeneity in response to political events, or that the microfoundations of macropartisanship are heterogeneous across the population, but also that shocks to macropartisanship are relatively long lasting.[9] Box Steffensmeier and DeBoef (1996) both broaden the scope of the analysis, comparing movements in macroideology to macropartisanship, but more importantly for my purposes, move one level of analysis down, disaggregating the macropartisanship series into levels of political sophistication. They find, confirming the results of the

[9]The extent of responsiveness to system shocks in the macropartisanship series relates directly to changes in presidential approval in response to events.

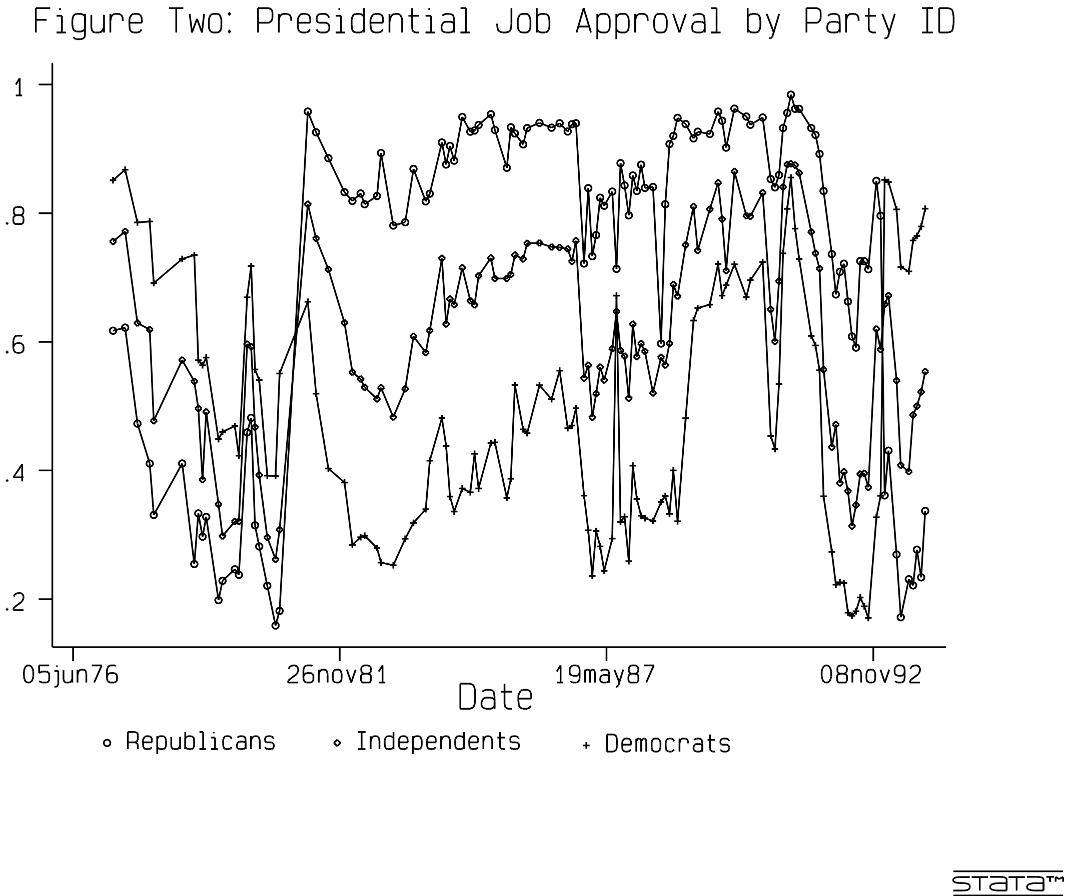
earlier piece, a stronger link between macropartisanship and macroideology among more sophisticated portions of the public. The question these studies raise for me is whether it is more productive to abandon attempts to model approval at the individual level, while pooling the cross sections, and instead move to a middle level of aggregation, and take advantage of the innovations of Box Steffensmeier, DeBoef, Simon, and Smith.

Suggestive Data and Future Directions

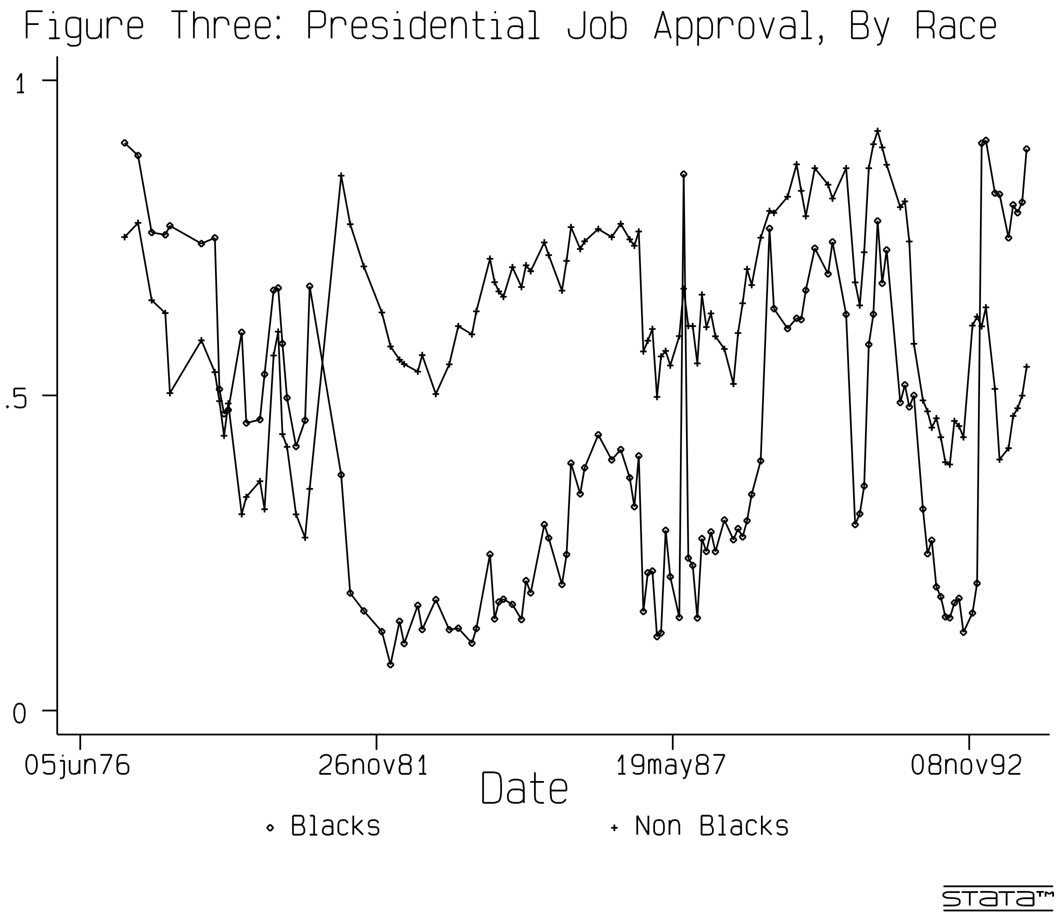
So far, I have presented a number of ideas about consensus and volatility, and listed a series of alternative ways to examine them. I close by presenting some preliminary data on variability in approval, mainly as suggestive directions for future work. I am open to comments and criticism on where I should turn next.

Do mean approval and variance in approval move in tandem? Or are there distinctive patterns? I examine these questions in two steps. First, I plot presidential approval, broken down by party ID and race. The patterns, shown in Figures Two and Three, are not unexpected. The starkest differences in approval are evident during the Reagan years, where the average level of approval among Democrats was approximately 40%, versus nearly 90% among Republicans. Carter's and Bush's partisan pattern of approval (distinct from their overall levels of approval, reported in the first figure), are somewhat similar. Both experienced the expected pattern, with inpartisans approving at higher levels than out partisans, yet the gap was seldom more than 30%. Similarly, Black/White differences were muted during the Carter years, as the average rating among both groups declined during the late 70s. Blacks and whites differed substantially on their rating of Reagan's presidency, varying by as much as 50 points, during the height of the 1981 recession. The pattern during the Bush years is something of a mix, neither as divided as during Reagan, nor as unified as during Carter.[10]

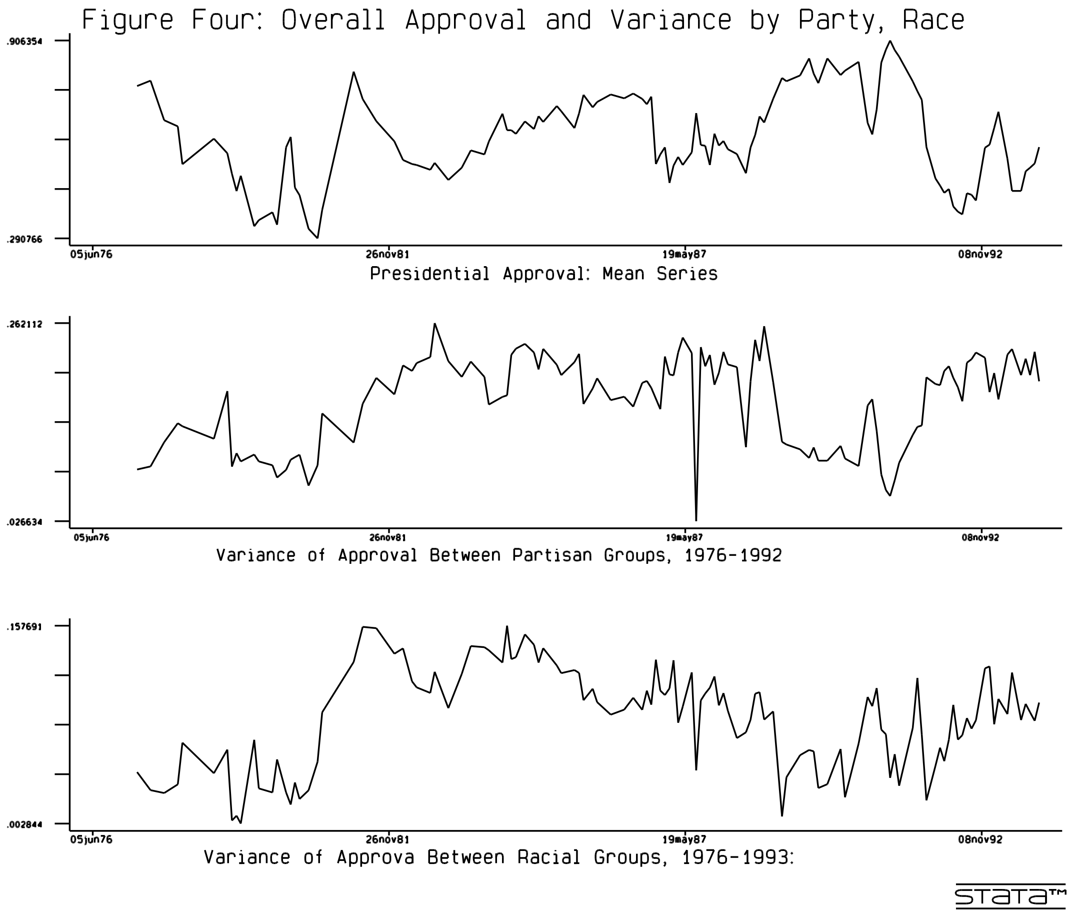
[10] I apologize to the reader for not removing the erroneous data point just after 19May87. I do not believe that Blacks were approving of Reagan's job performance at an 80% rate.



It is easy to jump ahead, and make an eyeball estimate of variability in these ratings. Partisan *differences* on presidential performance, expand and contract, and not always in tandem with rises or declines in average approval. For example, honeymoons (rises in mean approval) are clearly associated with consensus in approval. On the other hand, Reagan's approval rating during the 81-82 recession remained more stable among Republicans than among other partisan groups, and Bush's post Gulf War drop was much more severe among Democrats and Independents. And the gaps between Black and white expand and contract in dramatic fashion over this time period; at times (e.g. late in Carter's term, early in Bush's and Clinton's term), Blacks and whites were in agreement. During other periods, the differences are as great as we usually associate with the American racial cleavage. All this is not news, of course, but it reinforces the point that the spread of approval may vary in important ways, as does mean approval.



The next figure present these same data, but converted into an overall summary measure of variance. The top frame reproduces the mean approval series for ease of comparison, while the bottom frames shows the variance between partisan and racial groups. I will only point out a few features in these plots. First, notice, as was already evident in the previous plots, that Reagan's presidency was marked by long period of high variance in approval, regardless of the rises and falls in mean approval. Reagan's presidency, from these data, divided the public, on partisan and racial grounds, even though conventional interpretations of his presidency is one in which he enjoyed relatively high levels of public acclaim (e.g Brody, 1991, pg. 10–15). The uniqueness of the Reagan years in terms of trend – he failed to display a linear decline – is also apparent in terms of variability. George Bush, on the other hand, whose average level of popularity is described by Brace and Hinckley as more volatile than most other presidents (1994, pg. 144), actually experienced rather low variablity in ratings among partisan groups (not racial groups), with spikes and drops associated with the Gulf War conflict and the onset of the economic decline prior to the 1992 election.



Second, rapid increases (for example, during the three honeymoon periods covered by this series) and decreases (for example, during the 1991-2 recession and six months after the Gulf War) are associated with increases in variability. This is somewhat of a surprise, and contrasts with the findings of Brehm and Gronke (1995). Our supposition there was that honeymoons were "rallies" in every aspect, associated both with increases in mean approval and in consensus. On partisan and racial grounds, for these three presidents, this does not appear to be the case. Perhaps periods of rapid changes in approval are also periods of widespread uncertainty because of many conflicting interpretations of the presidential performance. When elite interpretations conflict, we ought to observe widespread disagreement among partisan groups (Zaller, 1992).

If this latter interpretation of the figures is correct, it should be reflected in a great amount of variability by party, and less so by other breakdowns. I have some indirect evidence on this point. The variability about the mean for racial groups is about half of that for partisan groups. On average, Blacks and whites varied 8.5 points (s.d. 0.4) from the mean, whereas partisan groups varied 16.9 points (s.d. 0.5) from the mean. Excluding the Reagan years does not reduce this gap, but controlling for party affiliation of whites (not done here) will reduce it. A nice comparison for these data would be plots broken down, again by party and race, during the 1960s, when elite messages were coded as often by race as they were by party.

Future Directions

This paper takes an exploratory look at volatility and consensus in presidential approval. In this paper, I take preliminary steps toward extending previous work on volatility to the individual and pooled cross sectional level. As a work in progress, there are no conclusions for this piece. I would like to reiterate some of my ideas, and forward some thoughts for the future.

I think it is clear that the dynamics of presidential approval change over time, whether we begin from simple descriptive plots or from ARCH models and cointegration tests on the aggregate series (Brehm and Gronke, 1995; Ostrom and Smith, 1992). Brehm and Gronke (1995) found that volatility is predictable: it increases at some points (during election years and after foreign policy events where the U.S. is perceived to be the aggressor) and decreases at others (over the course of an administration, during foreign policy events where the U.S. is perceived to be attacked). Ostrom and Smith do not address volatility per se, although they identify different kinds of presidential "events" in much the same spirit as Brehm and Gronke (and both following Ostrom and Simon, 1988). Still, Ostrom and Smith refer frequently to an "equilibrium level" of approval, with "errors" from that equilibrium (perturbations in response to events) that are corrected over time. It may be the case, if the general hypotheses presented here are true, that the size of these perturbations have increased over time, and the rapidity of return to equilibrium has decreased, as citizens have lost their partisan and ideological moorings.

The link between macro and micro approval, however, to borrow language from Box-Steffensmeier and DeBoef (1996), is only now being explored. The data portion of this project is daunting; I have accumulated economic and political time series (inflation, unemployment, war deaths), an event series, and pooled cross sectional surveys for the period 1976–1993. The eventual goal is to collect a longer set of cross sectional surveys and attaching aggregate level measure to them, but before taking that step, substantially more work needs to be completed on determining the particular quantitative models which will be used to analyze these data. This paper lists a series of alternative models which I hope will stimulate some reaction and discussion. A modest list includes:

- *Cointegrated time series and VAR regression*: employ the analytic techniques suggested by MacKuen and colleagues, Ostrom, Smith, Box Steffensmeier, and DeBoef. The time series are actually "pooled" in the way that Sayrs (1989) describes. Instead, the analyst tests for causality across the series. An excellent discussion of these methods are contained in the volume of *Political Analysis* in which the Ostrom and Smith piece appears (with application to presidential approval) . In the same volume, however, there is apparent controversy over the applicability and interpretation of the technology.

- *Pooled cross sectional time series analyses*: a variation on the cointegration tests, instead of testing for integration across the time series, treat each separate aggregation as an independent observation of the "average" respondent's placement on a 0-100 scale of approval. Thus, we are able to pool time series of whites times blacks times partisans. The problem is that once the sample is broken down into any moderately high number of categories, the sample size for each individual time point become too small.

- *Pooled cross sections*: another step down in aggregation, this time leaving each of the individual cross sections untouched, but still estimate pooled models. Markus (1988) followed this strategy with some success in order to address the "Kramer problem" in the politics and economics field.

- *Separate cross sectional analyses*: follow a primitive pooling strategy, testing for stability of coefficients over time. This should allow us to test for change in structural relationships (e.g. the declining influence of partisanship on approval), but doesn't allow a direct modeling of volatility.

This paper has only scratched the surface. I hope this essay, more fairly a work in progress, serves as a useful resource for ideas, and a target for reactions from others working in this area. Fire away!

Bibliography

Beck, Nathaniel. 1991. "Comparing Dynamic Specifications: The Case of Presidential Approval." *Political Analysis* 3: 51–88.

____________. 1989. "Estimating Dynamic Models Using Kalman Filtering." *Political Analysis* 1: 121–156.

Brace, Paul and Barbara Hinckley. 1992. *Follow the Leader: Opinion Polls and Modern Presidencies.* New York: Basic Books.

____________. 1991. "The Structure of Presidential Approval." *Journal of Politics* 53(4): 993–1017.

Brehm, John and Paul Gronke. August, 1995. "History, Heterogeneity, and Presidential Approval." Working Paper 197 in the Duke University Program in Political Economy, Durham, NC. Currently under review for publication.

____________. 1994. Modeling Volatile Approval: A Modified ARCH Approach to Presidential Approval Ratings. Paper presented at the Annual Meeting of the American Political Science Association, New York, NY.

Brody, Richard A. 1991. *Assessing the President.* Stanford, CA: Stanford University Press.

Engle, R. 1982. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflations." *Econometrica* 50: 987–1008.

Greene, William H. 1993. *Econometric Analysis, 2nd Edition.* New York: MacMillan.

Harvey, A. 1990. *The Econometric Analysis of Time Series, 2nd Edition.* Cambridge, MA: MIT Press.

Holsti, Ole R. 1992. "Public Opinion and Foreign Policy: Challenges to the Almond-Lippmann Consensus. Mershon Series: Research Programs and Debates." *International Studies Quarterly* 36: 439–66.

Jackman, Simon. 1995. "Re–thinking Equilibrium Presidential Approval: Markov–Switching Error–Correction." Paper presented at the 1995 meeting of the Political Methodology Section of the American Political Science Association, Bloomington, IN.

Jentleson, Bruce W. 1992. "The Pretty Prudent Public: Post Post–Vietnam American Opinion on the Use of Military Force." *International Studies Quarterly* 36 (1): 49–74.

Kernell, Samuel. 1993. *Going Public (2nd Ed).* Washington, D.C.: Congressional Quarterly Press.

____________. 1978. "Explaining Presidential Popularity." *American Political*

Science Review 72: 506–522.

Key, V.O. 1961. *Public Opinion and American Democracy.* New York: Alfred Knopf.

King, Gary. 1993. “The Methodology of Presidential Research.” In Edwards, Kessel, and Rockman (eds), *Researching the Presidency.* Pittsburgh, PA: University of Pittsburgh Press.

____________. 1989. *Unifying Political Methodology.* New York: Cambridge University Press.

King, Gary and Lyn Ragsdale. 1988. *The Elusive Executive.* Washington, DC: Congressional Quarterly Press.

MacKuen, Michael B., Robert S. Erikson, and James A. Stimson. 1995. “Dynamic Representation.” *American Political Science Review*, 89(4): 543-565.

____________. 1989. “Macropartisanship.” *American Political Science Review*, 83: 1125–1142.

Mueller, John. 1973. *War, Presidents, and Public Opinion.* New York: John Wiley.

Neustadt, R.E. 1990. *Presidential Power: The Politics of Leadership.* New York: John Wiley.

Ostrom, Charles M. and Dennis M. Simon. 1985. “Promise and Performance: A Dynamic Model of Presidential Popularity.” *American Political Science Review* 79: 334–358.

1in. 1988. “The President’s Public.” *American Journal of Political Science*, 32: 1096–1119.

Ostrom, Charles M. and Renée M. Smith. 1992. “Error Correction, Attitude Persistence, and Executive Rewards and Punishments: A Behavioral Theory of Presidential Approval.” In J. Williams (ed.), *Political Analysis 4th Ed.*, Ann Arbor: University of Michigan Press.

Page, Benjamin I. and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in American’s Policy Preferences.* Chicago: University of Chicago Press.

____________. 1983. “Effects of Public Opinion on Policy.” *American Political Science Review* 77: 175–190.

Rivers, Douglas, and Nancy L. Rose. 1985. “Passing the president’s program: Public opinion and presidential influence in Congress.” *American Journal of Political Science* 29:183–96.

Sayrs, Lois W. 1989. *Pooled Time Series Analysis.* Newbury Park, CA: Sage Publications.

Simon, Dennis M. and Charles M. Ostrom. 1989. “The Impact of Speeches and

Travel on Presidential Approval." *Public Opinion Quarterly* 53(1): 58–82.

Skowronek, Stephen. 1993. *The Politicsl Presidents Make.* Cambridge, MA: Harvard University Press.

Stimson, James. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29(4): 914–947.

____________. 1976. "Public Support for American Presidents: A Cyclical Model." *Public Opinion Quarterly* 40: 1–21.

Sullivan, John L. 1973. "Political Correlates of Social, Economic, and Religious Diversity in the American States." *Journal of Politics*, 35: 70–84.

Zaller, John. 1992. *The Nature and Origins of Mass Opinions.* New York: Cambridge University Press.