

UC Berkeley

Working Papers

Title

The Exposure Thesis Revisited: Aggregate Seat Change and Individual Vulnerability

Permalink

<https://escholarship.org/uc/item/7jj8f2m0>

Author

Jarvis, Matthew G.

Publication Date

2005-08-05

The Exposure Thesis Revisited: **Aggregate Seat Change and Individual Vulnerability**

Matthew G. Jarvis
mjarvis@berkeley.edu
University of California, Berkeley

Abstract

There are numerous models predicting the partisan distribution of seats in the U.S. House of Representatives using national-level variables. The partisan distribution in Congress is simply an aggregation of 435 separate elections. This paper is an attempt to formulate measures of differential partisan vulnerability in elections from individual election results. The measures constructed are compared to models in the literature. Findings indicate that either the measures are not indicative of the exposure thesis or that, similar to the literature on sociotropic versus “pocketbook” economic voting, the district-level causal mechanism is complex and unclear.

Numerous models exist purporting to accurately predict the aggregate results of congressional elections.¹ As noted by Lewis-Beck and Rice (1992), however, some of these models are not truly predictive in nature, but explanatory, using data that is generally not available until after an election has occurred. Thus, the number of models that can be used to generate predictions about what the results of an upcoming election will be is somewhat smaller. Within both types of models exists another possible flaw. Various models utilize measures as independent variables that have problems of endogeneity. Stated simply, many models attempt to predict what will happen by looking at measures of what will happen, leaving the reader wondering what has actually been explained. This same phenomenon exists within the literature and folklore of predicting presidential elections.² Furthermore, many modelers seem obsessed with maximizing the predictive power of their models, often at the expense of theoretical tractability. This is not to say that all models of congressional elections are either theoretically unjustified, mis-specified or ill suited to prediction. However, these problems plague many of them to various degrees.

Perhaps the area of greatest theoretical difficulty for scholars of congressional elections is how the majority of these models predict the aggregate results of congressional elections using aggregate data. This is tantamount to assuming that “all politics is national.”³ Research into this question has usually found the opposite to be true. Brady, D’Onofrio and Fiorina (2000) study the phenomenon in various ways over the 20th century, and find that local effects greatly

¹ ‘Aggregate results’ could refer to either of two distinct concepts. The first concept is the final vote tally for a given congressional district (either over time or for a single election)—the understanding of ‘aggregate’ in this case being of the individual voters. The second understanding of the term—and the one that will be used herein—is the aggregate result of the 435 separate elections, namely, the partisan composition of a given Congress.

² The most famous aphorism of this sort is: “As goes Maine, so goes the nation.” Between 1860 and 1928 Maine only ‘mis-predicted’ the President 3 times: 1884, 1892, and 1916. Since then, its ‘accuracy’ has gone down considerably (correlated at only 0.06)

predominate over national effects. Researchers must recognize that 435 separate elections occur every other November, and that the partisan balance in Congress comes from an aggregation of these elections. This approach has a couple of major advantages. The first is that, as a discipline, we know a lot more about individual congressional elections than aggregate ones. More research has been done into them, and as a result, there are some easily measured, district-specific traits that make predicting the results of a given election much more certain. Second, there is the law of large numbers. There have been 12183 congressional elections since World War II, as opposed to 28, which is barely enough cases to be able to run simple regressions on, much less analyze with any degree of complexity. However, it is the simple obvious truth of the system of elections that should lead scholars to pay attention to the method with which our data are generated. The result of a national congressional election is the sum of 435 simultaneous, linked (but separate) elections. With that in mind, we can proceed to an analysis of models that have been previously expounded upon.

Previous Literature

Lewis-Beck and Rice (1992) summarize a number of major models of national elections, including congressional elections. In evaluating these models, they pay particular attention to the predictive accuracy (and by extension, whether the model is for midterm or all elections) and the amount of time prior to the election that predictions can be made. If election forecasting is the only goal of the researcher, these are appropriate goals. However, the guiding principle for

³ Another possible interpretation of these models is that they assume that while congressional elections do have “local” components, the distribution of these components is random with respect to the individual national effects. However, this author has never seen evidence of either this fact or even its explicit assumption.

this project is that a model that predicts accurately without an adequate theoretical basis is not a good model. Thus, another assessment of these models is in order.

One of the simplest models (and one created by a pioneer in the forecasting field) is that of Edward Tufte (1975). Tufte's model treats midterm elections as referenda on national circumstances, predicting congressional vote share from a presidential approval and an economic status variable. There is a great deal of theoretical tractability to this model. First, presidential approval is one of the two variables that Tufte theorizes affects the magnitude of the national vote loss by the President's party, which in turn leads to seat loss. James Piereson's (1975) work with presidential approval and vote choice in congressional and other elections confirms that this theory operates at the level of the individual voter. This is more than just a passing concern; in studying as few elections as there have been in the post-war period, it's quite possible for us to find a correlation by chance, and regression models are ill suited to working with so few degrees of freedom. Tufte's measurement of his second variable, change in the economy, is relatively straightforward; he simply uses yearly change in real disposable income (RDI) per capita. The macro level relationship between the economy and elections has been established (although there is a rich literature on how this relationship functions on the level of the individual). There is an important characteristic of Tufte's model that must be discussed. Tufte standardizes the midterm vote loss by using the difference between the midterm results and the average of the last 8 elections.⁴ This type of 'adjustment' for what the relative party strengths are at a given time is common within the literature.⁵ Such adjustments must be made very carefully, however. Any

⁴ Tufte explains that he does not use the previous 8 elections in all cases (notably 1938, as he recognizes the New Deal realignment as changing what the 'default' environment was (Tufte 1975, fn. 14). The choice of 8 as a particular reference number is not made immediately clear. Nor is there a clear guideline for how future researchers should implement his methodology. He claims that the 1932 shift is 'clear' in his data. However, Jacobson (2001 and personal communication) includes a post-1992 variable in his analyses to account for what he sees as a shift in national voting behavior. For a treatment of the post-1992 era, see footnote 32.

⁵ Indeed, as shall be noted, it is often confounded with separate causal variables.

adjustments made to a variable based on other scores of that variable risk introducing autocorrelation into the data. Adjustments for ‘relative party strength’ made by looking at vote shares risk polluting the data, because those observations of the dependent variable are not guaranteed to be representative of any ‘rest’ tendency.⁶ However, in the end, Tufte presents a very parsimonious model that makes simple assumptions that seem to be borne out by the data.

It is unfortunately true, however, that we must question the measures used. The measures of seat change (ignoring for the moment standardization of those measures) are very reliable and pose no problems for any research in this field. The same cannot be said of the measures of presidential approval and the state of the economy. While examining a different question of the Gallup Poll, Erikson and Sigelman (1995) come to the conclusion that the Gallup Poll is biased, but systematically so. In fact, the bias is related to the party of the president, so any statistical model that includes this as a variable can eliminate it. Moore and Saad (1997) confirmed this result, but with an interesting methodological caveat; the Gallup Poll increases in accuracy when only the responses of “likely voters” are examined. At first glance, this creates no problem. However, attempts to get the likely voter responses are not always successful.⁷ Nor is there agreement on *when* the approval of the president should be measured. It ranges from June of the election year to the last poll taken days (or sometimes hours) before the election. Thus, it is quite

⁶ For example, if one adjusted expectations of seats in the 1990s based on the 1980s (as being indicative of recent trends in partisan disposition), one would be conveniently ignoring the well-known bias towards the Democrats that the 1981-1982 redistrictings had, which partly explain why the Democrats did as well as they did. This example leads to the general problem of how large a ‘window’ should be drawn to calculate the ‘normal’ partisan distribution. Furthermore, it is unclear how to accurately predict what the *current* distribution is from the *previous* distributions. The average of the series 1,2,3,4,5,6,7,8 is 4.5; yet perhaps what this series demonstrates is that by the 9th point in the series, a better sense of the ‘average’ might take into account trends as well as history. Weighted averages could be used, but those run into the two problems of specifying what kind of weights to use as well as the possible exaggeration of a recent data point that is an outlier in some sense. c.f. Erikson, 1990.

⁷ Using the data resources of a major research university, this author was unable to get the Gallup Poll generic ballot responses for likely voters with any sense of completeness. The responses of registered voters are archived quite thoroughly by the Roper Center at the University of Connecticut. An additional fee service is available that may or may not contain the likely voter data. What further complicates the investigation of previous models using Gallup numbers is that it is rarely (if ever) specified what sample the estimates are based on.

possible that our measurement procedures for a relatively simple variable to generate radically different results.⁸

Unfortunately, such problems also plague our other main independent variable. After the 2000 presidential election, where most of the popular models of presidential vote share were significantly off, Bartels and Zaller (2001) remarked that the models were wrong due to modeling errors and the fact that the economy changed at that last minute (after it had been measured by the modelers). However, Erikson, Bafumi and Wilson (2001) contend that Bartels & Zaller's Bayesian models should have included presidential approval as a predictor. For the purpose of this paper, what is interesting is what Erikson, Bafumi and Wilson term "an interesting subplot" of the Bartels and Zaller piece—

“whether a choice could be made as to which measure of the economy provided the best specification. They tentatively concluded that it is best to measure prosperity by per capita growth in disposable personal income rather than in GNP and to measure it (discounted over the span of the presidential term rather than as a flat growth rate over the last year of the administration of the last year before the election...Our analysis brings us to the same conclusion, only more strongly.”
(p 818)

This leaves the congressional election modeler with a difficult decision to make. If voters in presidential elections respond to the development of the economy over the last 4 years (weighted according to the procedure outlined in Hibbs (1987)), then how do voters in congressional elections respond to the economy? Do they penalize a Congress for the economy of the Congress prior to it? Are they simply responding to their assessment of how the current president is handling the economy, or has handled the economy so far? If Bartels and Zaller used 6

⁸ Indeed, comparing the data of two scholars, the differences in their measures of presidential approval ranged from 0 to 31 percent, and disagreed by 6% on average.

specifications of the strength of the economy to predict presidential elections, how many does the intrepid congressional election forecaster need?

If a simple model can generate so many interesting problems that must be dealt with, what must be the fate of more complex models?⁹ Rather than discuss all of the models (and their various differences in specifications of variables), it will be more productive to simply discuss what other variables have been introduced into these models. There are essentially 4 variables that have also been included in various analyses. As Marra and Ostrom (1989) draw on many other models for theoretical inspiration, it makes sense to start with their added variables. Their first variable is meant to capture the effects of political events. In essence, it is an attempt to quantify the effect that a scandal or crisis or war might have on voters. The inclusion of such a variable is an admirable attempt to expand the qualitative approach. Problems of measurement obviously abound. These problems have been enough to discourage many researchers.¹⁰ Marra and Ostrom also attempt to control for the current political alignment by adjusting for the partisan distribution of the time, using National Election Studies data. This is also a laudable attempt to improve on previous models. Rather than dealing with the problems of past averages (see footnote 6), this variable is an attempt to capture the current partisan environment. What seems strange, then, is that the effect of this variable is so small. “A shift of slightly less than 5% in the partisan split will be translated into an additional seat.” (Marra and Ostrom, 559) The literature on congressional voters, on the other hand, continues to demonstrate the massive effect that partisanship has on vote choice. It seems strange that something that so strongly affects

⁹ Similar models to Tufte’s simple model certainly exist. For examples, see Erikson, 1990; Kramer, 1971; or Hibbs, 1982, 1987.

¹⁰ As a brief note, consider the evidence from 6 separate, yet similar, events and their radically different effects. Gary Hart’s infidelity likely cost him the Democratic nomination in 1984. Bill Clinton (and his popularity) survived infidelity and possible perjury. George W. Bush’s alleged DUI troubles didn’t appear to damage him, while Al Gore struggled mightily with an exaggeration about creating the Internet. A mysterious death or disappearance has either

voters should fail to matter much when aggregated. A third variable that Marra and Ostrom include is a measure of seats at risk. The logic here is quite simple; the party that has more seats that are in danger of being lost will likely lose more than the other party. As this is the central theme of this paper, we will leave discussion of this for another section.

A fourth variable that has been introduced by Lewis-Beck and Rice (1992) is the time in office of the incumbent president's party. Unfortunately, they offer no theoretical justification of this inclusion. Possibilities could include the semi-phenomenon of the decay of presidential popularity over time.¹¹ Lewis-Beck and Rice's work is important from another perspective, however. As mentioned earlier, Lewis-Beck and Rice's primary concern was the suitability of the model for forecasting. As such, they are intensely concerned with measures such as R^2 and the standard error of the estimate (SEE). The focus on R^2 , in particular, is troublesome. There are two reasons for this. First, the dependent variable is limited. Thus, the predictions generated by the estimation procedure, while generally not extendible outside the observed range of the data, are particularly hampered by the reality that predictions can quite possibly predict that the Republicans will have negative seats after an election. The second point is more substantial. R^2 simply tests how well the data fit to a linear assumption. With many of the variables in these models, there is substantial reason not to expect linearity of effects. For instance, the rate of RDI growth could reasonably be argued to not have a linear effect, particularly around small changes (and negative changes), which is actually what the typical values of RDI change are. The final point of criticism to be levied at those who would choose to compare these models in this fashion comes from the arguments made by the Bayesian modelers. This argument is nicely phrased by

ended a political career (Gary Condit) or simply prevented a senator from becoming president (Ted Kennedy). All of these context-specific effects make this researcher nervous about attempts to quantify 'political events.'

Bartels and Zaller (2001). Complementary to that point, however, is that regression is a very difficult tool to use correctly when one has so few cases. Micronumerosity¹² is a term that has been applied to this problem. Fortunately, micronumerosity does not lead to misestimation. But it does mean that alternative explanations may be rejected incorrectly. Unfortunately, we cannot simply collect more data to fix this problem.

There is one additional problem that besets most of the research out there, and that problem is mostly theoretical. As noted earlier, almost all of the models of national congressional elections assume that the elections are national events, when most scholarly research has shown that congressional elections are mostly local events. Jacobson's (2001) iterations of his model pay the most attention to the individual nature of congressional races. Building upon the work that has both demonstrated the 'incumbency advantage' (how much better incumbents fare in congressional elections) and that one of its main factors is not facing quality challengers, Jacobson's model includes a measure of the aggregate quality of the partisan challengers.¹³ This is a very important addition to the model. By incorporating a sense of the nature of the 435 races for Congress, Jacobson is paying attention to the nature of the enterprise—predicting an *aggregate* result. However, Jacobson loses a lot of the richness of his data when he comes to these totals. Who are these quality challengers' opponents? A former Republican state legislator would still be very unlikely to win an election in liberal San Francisco. While we believe that quality challengers are highly strategic,¹⁴ not all of them are.

¹¹ This decline is itself possibly explained by other factors. As this is a literature with which the author is not well-versed, only one reason will be offered here: the effect of vetoes on presidential popularity as demonstrated by Groseclose and McCarty (2001)

¹² For more on micronumerosity, see Gujarati 1995.

¹³ On the incumbency advantage, see Gelman & King(1990), Cox & Katz(1996), Jacobson(2001). On quality challengers, see Jacobson(2001), Cox & Katz(1996), Krasno(1994), Green & Krasno(1988).

¹⁴ See Jacobson(2001), Banks & Kiewiet (1989)

Furthermore, who will be the challengers is often not known until shortly before an election, giving Jacobson's model a much more explanatory than predictive function.

In sum, the vast majority of models err in modeling national congressional elections based on purely national variables. As national congressional election results are the aggregation of 435 elections, it only makes sense to model them as such. This is not to say that national conditions do not have strong effects on local congressional elections. No tests of the relative contribution of national versus local effects show that the national effects are nonexistent. Rather, the effects of national conditions on congressional election are filtered through the 435 contexts.

Theory

What is needed is a model that predicts the aggregate results as the consequence of individual elections. We cannot assume that all 435 seats are the same. In some seats, the party of the winner is a foregone conclusion. In fact, of the 12183 general elections for Congress since World War II, 1827 (15%) have featured one major partisan running unopposed.¹⁵ No matter what the state of the economy or unpopularity of the president, the outcome in those seats was a foregone conclusion. For the contested seats, we would certainly not expect all seats to equally be in danger of being lost by the incumbent party. Certainly, Mike Rogers (R-MI)—who won by 111 votes, much fewer than the Green Party candidate got—is a lot more scared of his chances of reelection than is Jose Serrano (D-NY), winner of his last election by a 92 point margin. Indeed, Serrano's 16th district (with its 1992 redistricted boundaries) has never given less than 91% of its congressional votes to the Democrat; Michigan's (for it is hardly Rogers') 8th district, on the

other hand, has ranged from 47 to 60% Democratic since the last time the boundaries were redrawn.

What local effects cause this (and greater) variations? A large variety of causes have been expounded upon in the rich literature on congressional elections. A few of the major causes that have been isolated include party affiliation, campaign funds, incumbency, quality of challengers, voter familiarity with candidates, personal contact with candidates, amount of time spent by candidates in the district, issue positions and particular votes.¹⁶ However, two of the most interesting are incumbency and the quality of challengers. Put simply, incumbents do better than non-incumbents, and high-quality challengers do better than poor challengers, *ceteris paribus*. These variables, however, are not simply random. Members in districts that are less safe are more likely to retire, as the value of staying in Congress is the same, but the cost (and risk) eventually may get to be too high. Quality challengers are more likely to wait until an incumbent retires to run, or if the incumbent is not apparently going to retire soon, run in a district where they have a better chance of winning.¹⁷ These two variables manage to capture a good deal of the variation in types of congressional elections. Quality candidates get better funded, spend enough time and money in the district to become familiar (or at least somewhat familiar) in the district, and generally know which positions are the right ones to take in the district. In a very real sense, challengers and incumbents are good measures of the safety of a district for the incumbent party.

¹⁵ Based on data provided by Gary Jacobson (personal communication).

¹⁶ Citations for the list of causes of vote choice would include an entire literature.

¹⁷ For the former situation, see Banks & Kiewiet. An example of district shopping happened in the run-up to the 2000 election in Orange County, California. Local Republicans, seeking to keep former MC Robert Dornan from running to fill an open seat, recruited Darrell Issa to run for the district. Issa, a former candidate for the Senate, could have run in a number of districts. While Dornan could have run in his old district, he also had a small house in the other district, allowing him to claim that he was a resident. Current Senators Hillary Rodham Clinton and Elizabeth Dole are other 'possible' examples of 'venue shopping.'

This gets us back to a common theme in the literature on national congressional elections. The models of Oppenheimer, Stimson and Waterman (OSW; 1986), Marra and Ostrom (1989), Lewis-Beck and Rice (1992), and Jacobson (2001) all include some measure of the number of seats at risk. As “seats at risk” is the focal variable of this analysis, let us discuss their measures in some detail. Following the work of OSW, this measure has generally been the number of seats the party in question¹⁸ has over some expected number. For OSW, that expected number is 254 seats for the Democrats (their average over the period that interests them). There is a problem in generating an independent variable from the dependent variable, however. In calculating the mean, we have removed one degree of freedom from the independent variable. In fact, the last point in their dataset should predict perfectly, as if it didn’t, 254 wouldn’t be the mean! There is another flaw in the measure, however (and one that might make the first flaw not problematic, but still leave the variable flawed). 254 is the mean for the whole period they studied, 1938 to 1984. They feel justified in using this period average because “although some secular realignments have no doubt occurred during this period, none has overturned the basic party system.”(234) Was 254 the norm in the late 1940s and 1950s, when the parties seemed equally matched?¹⁹ When are future researchers supposed to be able to determine that there was a big enough shift in the party system to stop adding cases to the model? Did Reagan’s election change the party system? The Republican takeover in 1994? The simple mean procedure leaves us with a variable that either predicts too accurately or is biased in different directions depending on trends.

¹⁸ Some models take as their dependent variable the party of the president, others use the democratic party and recode other variables. It makes no difference to the modeling procedure.

¹⁹ Before the reader thinks that Eisenhower is the cause of a temporary deviation from ‘normal’ behavior, recall the 80th Congress and the close race between Dewey and Truman in 1948.

Marra and Ostrom (1989) approach this problem from a slightly different angle. Akin to Tuftes adjustment of his dependent variable for the average of the last 8 elections, they create an independent variable that is the average seats for the last 8 elections. The primary advantage of this technique over OSW's is that it completely avoids the OSW variable's first problem. Since the election at hand is not being used to calculate the mean, we do not have the degrees of freedom problem. However, the second problem for OSW also has bite for them. An average produces the exact same prediction for an upward trend as for a downward trend (assuming both are centered on the same mean). But the intuitive political scientist knows that the 'expected' number of seats would be different for the two cases.²⁰

Lewis-Beck and Rice (1992) adopt the OSW measure for a different time period (their mean is 257); the earlier comments on that score shall suffice. Erikson (1990) and Jacobson (2001) adopt the same measure, namely, the number of seats held in the last Congress (Jacobson also uses the Marra & Ostrom specification for a different regression). The intuition behind this is clear; political trends are short term rather than long term (even the extensive party realignments of various periods took relatively little time to occur and stabilize at new levels). However, their measure is actually no different from those of Lewis-Beck and Rice or OSW. Their variable is a simple linear transformation (adding 254 compared to OSW, 257 compared to Lewis-Beck and Rice). Unless this variable is going to be interacted (which none of these models do), it makes no difference what the zero point is. However, none of these criticisms which have been made against these specifications individually is as important as the flaw that all of these measures have.

The main criticism to be levied against these models is that the distribution of seat vulnerability is not symmetrical across parties, elections or level of threat. Let's describe two

²⁰ See fn. 6 for more of this logic

similar situations about which we would have different expectations. In Situation A, the Democrats have 230 seats. However, of these 230 seats, only 5 are what any reasonable person would consider “unsafe.” In Situation B, the Democrats hold the same 230 seats, but 40 of them are “unsafe.” Suppose it’s going to be a mildly good year for the Republicans. In Situation B, there is a wide range of possible outcomes, depending on how the national forces interact with the local ones. In Situation A, majority committee staff should feel much safer about looking for a house out in Arlington, VA. The variables all expressed above assume that distribution of unsafe seats is perfectly correlated with the number of seats currently held. There is no reason not to test this assumption.

This brings us to a crucial decision point, however. We need to decide how to aggregate these individual race characteristics. There are 2 basic methodologies available. The first method treats national elections as our unit of analysis. This leaves us with a sample n of 27 elections (all the post-World War II elections for which we have data on the previous election, for that is how far back our more detailed information on candidate quality goes and is a common starting point for analysis).²¹ Thus, for each election, we would need a measure that captures the vulnerability of Republicans/Democrats to an electoral shift against them, including not only the raw advantage of one party over the other (which would be important if no major change in electoral landscape occurs) but also the degree of vulnerability of each party to shifts against them. A number of possibilities present themselves. The simplest is to classify each race according to the incumbent party, the presence or absence of an incumbent MC in the election, and whether or not we should consider the seat ‘unsafe’ or not.²² How can we classify seats as

²¹ See the methodological appendix on the availability and quality of data, particularly for years which must currently be left out of the data due to redistricting, reducing our n to 22.

²² Recall the logic that a combination of perceptions of national conditions and the vulnerability of the incumbent party (both due to incumbency and partisan balance in the district) are strong determinants of the decision of a

unsafe, especially given Mann's (1978) famous dictum that MCs are "unsafe at any margin?" Mann points out that the average previous margin of victory for losing incumbents has been growing.²³ Thus, it may be that margin of victory in the previous election needs to be adjusted for how much such a margin is worth at that time. There is also a very real possibility, however, that challengers and parties are not aware of this change in the system, and still operate under the old "rule of thumb" that any incumbent who won with less than 60% is vulnerable. If challenger decisions are the driving factor, then we would want to make our predictions based on their thinking. There are still other ways of determining that a seat is unsafe. Some districts are very consistent in their congressional elections (in the 1990s, the 11th district of Pennsylvania has cast between 66.5 and 68% of its major party votes for Democrats in each election); others are very unstable (Mississippi's 3rd district has gone from 81 to 26% Democratic over the same time). Thus, a possible measure of safety could be whether the district's previous margin is within some multiple of the district's average variation. A final possibility is that challengers could observe what the mean variation across districts was in the last two elections and figure that such interelection swings are possible and seek to take advantage of a vulnerable incumbent. While other operationalizations of this concept are certainly possible, these are fairly exhaustive. A summary of these possible measures of vulnerable members is presented in Table 1.

[INSERT TABLE 1 HERE]

quality challenger to run. While the presence of quality challengers could be used, such information is not available until much later, and the goal of this research is to produce sensible estimates a great deal of time in advance. On this logic, see Jacobson, 2001.

²³ Replication (and extension to later dates) of Mann's work on this subject was done by this author in formulating various model specifications. While not a smooth trend, the increasing average (or median or 3rd quartile, for that matter) margin that losing MCs had in the previous election is undeniable. Jacobson presents similar data on page 28 (Jacobson 2001)

Ideally, we could use the difference in the number of unsafe Democrats and Republicans as a variable in an OLS regression equation and simply substitute our new measure of seats-at-risk for the proxy variable that others have been using. However, that ignores the way in which this variable contributes to seat loss. To see why that is true, see Figure 1. This is akin to the way in which previous models have assumed congressional elections lead to congressional seats. Marginal members of the party at risk in an election are those we expect to lose. In years when the Democrats are likely to be picking up seats, what matters most is how many Republican seats are attainable. How many of those the Democrats pick up will depend largely on how good a year other factors lead to. Figure 2 is an example of the interactive effects of ‘seats-at-risk.’

[INSERT FIGURE 1 HERE]

Figure 1 is not a perfect representation of other models, but it approximates them. A shift in one direction²⁴ captures the vulnerable seats to that side. If we shift off the mean to the right, there would be more seats to be captured by a shift to the left at the second election. The difference between this simple model and those of Jacobson and Tufte is that the mean is allowed to shift between elections. In Figure 2, however, the expectation would be that a positive shift would lead to a similar situation as above, but with two important exceptions. The first is that the distribution varies from election to election. There’s little reason to impose assumptions on the distribution of vulnerable seats when we have data on the distribution.

²⁴ Figures 1 and 2 label shifts ‘positive’ or ‘negative.’ For which party is irrelevant.

Second, the effect of a shift of equal magnitude to either direction has different effects depending on the current distribution of vulnerable seats²⁵

[INSERT FIGURE 2 HERE]

The causal model being presented here proceeds as follows. National conditions such as presidential popularity and the state of the economy lead to some voters voting either for or against the party of the president. These votes, however, only have effect on the seat balance in Congress by affecting individual vulnerable legislators. Ideally, therefore, we would use a two-stage modeling procedure, as the effects of vulnerability are only realized as a result of national forces (as well as standard ‘random’ fluctuation within districts). However, two stage least squares (2SLS) is inappropriate to use in this case. In order to use 2SLS, you must have more regressors in the first stage than in the second. This poses a large problem, however, as we do not have any degrees of freedom to spare for these regressors. In short, the limited data do not allow us to utilize 2SLS. We are left, then, with a few alternatives. One will be explored within this paper: OLS models with interactions to approximate the desired 2SLS interactions. Another alternative is also being considered for future work, namely, an event count model (where the events being counted are partisan victories in a given election year).²⁶

²⁵ It is this second point that points out how Figure 1 is really a flawed representation of the models of prior research. Prior research assumes that the effects of the national variables are not conditional on current conditions (by including exposure, they have controlled for it). Thus, they are simultaneously assuming a uniform distribution for the effects of the national conditions and a ‘normal-like’ distribution for the effects of vulnerability.

²⁶ This thorny methodological mess has led this author to a better understanding of OSW’s (quote them where they talk about “other” specifications). Any suggestions that readers may have for better ways to model the theory described herein would be quite welcome. Better methodologists than this author have been consulted; their comments have been extremely helpful so far.

Data and Analysis

As a large portion of this project was the data collection and aggregation procedure, some simple descriptive statistics on the variables used are in order.²⁷ Table 2 contains summary statistics for the number of unsafe seats arrived at by the different models. A graphical representation of all the data is presented in Figure 3.

[INSERT TABLE 2 HERE]

[INSERT FIGURE 3 HERE]

What can be easily gleaned from Figure 3 is the remarkable similarity that all 5 measures display after 1968. No explanation for why the measures should become so highly correlated comes to mind. The great differences between these models and existing exposure models in the literature, however, are quite understandable. As has been argued earlier, the measures that have been used previously may be more correlated with election results than those presented here. However, these measures are an attempt to capture the aggregate individual race vulnerability of each party. That the measures diverge is an indication that assumptions about vulnerable seats made from numbers in Congress are not necessarily valid ones. Figure 4 confirms this, as do the correlations of these measures with those developed here. Oppenheimer, Stimson and Waterman's measure correlates negatively with these; the correlation coefficients range from

²⁷ In addition, interested readers may wish to consult the appendix for information on data sources, data collection, and unique problems posed by this data.

-0.086 to -0.316. The same is true of Jacobson's measure, but to a larger degree, as the coefficients range from -0.274 to -0.562.²⁸ One can guess from these numbers what the result of including these variables in regression would be, and those guesses are correct.

[INSERT FIGURE 4 HERE]

So far, we have determined that the widely varied measures yield dramatically different values for the variable of interest, "exposure." What, then, is the impact of using these different measures in a regression equation? In short, the new variables have nowhere near the predictive power of the old variables. No regression specification tested yielded better results than (or even results comparable to) those obtained by previous researchers.

The simplest comparison is between a model Jacobson (2001) presents as "an updated variant of Tufte's model" (144) and ones based on the versions of exposure presented here. The results for these regressions are presented in Table 3. It is obvious from this table that Jacobson's model easily outperforms its fellows here.²⁹ A discussion of the results will not be undertaken here, both for considerations of space and for the fact that the only model that shows any real promise is that of another scholar who has discussed the results of his own model at length. However, this test is rigged against the specifications presented here. Recall that one of the principal arguments raised against the simple specifications was that they assumed swings in either direction to have the same effects. The specifications of all the models in Table 3 continue this assumption. We need to test the new specifications interactively.

²⁸ The curious reader might be interested that the correlation between Jacobson's and OSW's measures is 0.251.

²⁹ Furthermore, the model presented here as Jacobson's is really a model that he uses as a baseline to present his model that includes challenger quality, which appears to outperform this simple model. In that model, Jacobson

uses a dummy variable for 1992 or later. The only way in which Jacobson's model was improved upon was by substituting a majority party dummy.

[INSERT TABLE 3 HERE]

As mentioned earlier, however, interaction in this case is quite difficult. A two-stage approach would predict the interaction terms in the first stage, but the additional variables that would have to be added would kill off the degrees of freedom to an intolerably low number. Note that we are already below the “rule-of-thumb” cutoff of 25 for the number of cases in our model. With the numbers problems this enterprise already entails, we need to find a method to ‘interact’ with the direction of effect without using up degrees of freedom. The approach used here is certainly econometrically flawed. However, it is the fairest test of these specifications of ‘exposure’ that has been devised.

Essentially, we can use a property of OLS regression to get some purchase on the issue. The OLS regression line, by definition, goes through the point of means (the means of all variables). If we assume that the other independent variable is held constant at its mean, then any scores on the independent variable that are below the mean will be correlated with scores on the dependent variable below the mean (or vice-versa if the affect of the independent variable on the dependent variable is negative). Thus, we can interact a mean adjusted variable with our variable of vulnerability.

This interaction is not a simple multiplicative one, however. We do not have a simple variable for ‘exposure’ in this case. The variable is actually two—how vulnerable is the party of the President, and how vulnerable is the opposition. Thus, the interaction takes this form:

a) If mean adjusted variable is negative:

Interaction term = mean adjusted variable * vulnerability of President’s
party

b) If mean adjusted variable is positive:

Interaction term = mean adjusted variable * vulnerability of opposition party

This complex interaction is easier to understand if we spell out our expected effects. When the President's party is in for a good year (either the economy is good or the President is popular), what matters is how vulnerable the opposition is, as it is their seats that we can expect to be in more danger. When the President is in trouble however, that trouble is only compounded by the difficulties that already faced his party. We interact the term because the greater the expected loss due to popularity or the economy, the more one's vulnerability matters. That is the logic. How do the models fare? Table 4 summarizes the results of the interaction when presidential approval is the interacted variable.

[INSERT TABLE 4 HERE]

What is one to make of these results? None of these represent substantial improvements over their 'mis-specified' versions in Table 3. The models for the interactions with the state of the economy are even worse. The standard errors of the estimates (SEEs) are never within 1 of that of Jacobson's simple model (4.46), meaning that the models are incredibly inaccurate. Most of the SEEs are so high as to represent almost no improvement over the simple mean of presidential seat loss. Simply put, all of the model specifications tested (and there are 4 additional specifications that have not been shown, so as not to bore the reader with tables that simply show a bad model) failed to demonstrate that the approach outlined here is useful at all.

Conclusion

What went wrong? At this point, two possible conclusions can be drawn. Modesty demands that the first possibility that be considered is that the various modeling procedures and specifications of the concept of differential vulnerability were incorrect or invalid in some way. Their lack of correlation with other intuitive measures of ‘exposure’ certainly indicates that this is possible. The data have been double-checked for accuracy and the calculations of the various variables (which are all aggregates of individual district data) have been verified as yielding the expected results for the independent variable. As a sub-category of this first possibility, the theory advanced here could be wrong conceptually. It seems intuitive to this author, but these results may point to the need for additional theorizing.

The second possible conclusion is that the models did exactly what they were supposed to do. The inspiration for this project was the observation that researchers have been testing a theory at the aggregate level when it operates at the level of individual districts. Perhaps the correlation that has been found in the aggregate research is only random. In statistical analysis, we accept hypotheses with some risk. These measures could be correlated with our dependent variable randomly. We cannot perform a new sampling procedure to rule this out; the election results of the last 50 years do not change no matter how many times we measure them.

At this point, both conclusions are possible. Theoretically, it makes sense that the aggregate results of 435 coincident elections should be modeled as groups of individual elections, not one election. However, the small data size, combined with admittedly imperfect realizations of the desired concept of individual district electoral safety, make it very possible

that the efforts to date have failed to fairly test either the existing theories or these versions of them.

Directions for Future Research

There are certainly other possible ways to conceive of electoral danger. Campaign finance, media, and other issues have been completely unexamined by this paper; challenger quality, while not totally ignored, has been dealt with by assuming a perfect relationship that doesn't exist. The examination of these issues is difficult in a multivariate context with the limited amount of data. That is not to say that work cannot be done to isolate the effects of these or other variables on national congressional elections. It is hoped that future work will shed some light on this perplexing issue of why national effects seem to be absent of local causes. This work is a failed attempt to examine the exposure thesis in a similar way to the examinations of the effect of the economy on voting. Hopefully, this author or others will be able to get some purchase on this interesting question.

Appendix

Data collection and auditing difficulty

As one of the prime focuses of this paper is on the vulnerability of a given congressional district in a given year, it was necessary to construct a series of data sets from multiple sources, making data auditing and error correction that much more difficult. This section will both highlight certain difficulties that posed a problem for research and describe the steps taken to overcome those difficulties and create a high quality data set. The purpose of this section is not to highlight work performed by the author, but rather to serve as a caution to those studying congressional elections to remember the local nature of not only the election itself, but of the reporting of the results and non-continuity of what often appears to be the same units.

The first problem encountered is in creating a time-series dataset. As the vulnerability variable was to be calculated based on previous electoral margin and incumbency status (see the Theory section for a theoretical explanation of why these two variables are good predictors of vulnerability), for any given election in a congressional district, data on the previous election in that district was required. However, this is exceedingly difficult to determine for years ending in 2 due to congressional reapportionment. The Constitution requires that congressional seats be reallocated after every census to reflect new populations. For some of these reallocations, states vary so much in the sizes of their delegations before and after the census that comparing the districts over that period is very difficult. This points to a closely related second problem.

Congressional redistricting due to the census reallocations can be dealt with rather easily, by simply leaving out redistricting years from the dataset entirely. That approach has been taken here. Less easy to deal with are the mid-decade redistrictings that have take place due to Supreme Court decisions (as in the 1963 and 1964 responses to *Baker v. Carr*, the late 1960s

responses to *Wesberry v. Sanders*, or the multiple redrawings of North Carolina's 12th district in the early 1990s) or state political decisions (such as Illinois' 1947 redistricting). If every year in which *any* kind of redistricting took place were discarded, there would remain only a few data points. For many of these, adjustments can be made to the data to account for the changes. Such changes are necessarily impressionistic in nature, as there is really no good way to predict how new constituents would respond to their new incumbents. For every mid-decade redistricting, maps of districts before and after the redistricting were compared.³⁰ Sources such as major metropolitan newspapers and *Congressional Quarterly* were consulted to get as much relevant information on the new districts political leanings. This data either consisted of generic statements such as 'leans Democratic' or estimates of a previous major election (presidential, congressional, or gubernatorial) in that district. However, we must remember that many congressional districts are not at all competitive; this is particularly true of many southern districts in the 1940s and 1950s. These districts make estimation of their prior congressional results difficult in general; impressions of primary sources and the author's impressions were used in these cases. Table A1 reports the total number of valid cases by year for 4 of the models discussed in this paper.

[INSERT TABLE A1 HERE]

Finally, there are holes in the datasets that are readily available. These omissions in no way reflect poorly on those who collected this data and generously provided such data to me.³¹

³⁰ Jonathan Katz was also kind enough to provide 'similarity scores' calculated by Cox and Katz (2002). These scores are being used to verify impressions of shifts as well as come up with new district political estimates.

³¹ Data on elections were obtained from ICPSR, study number 6311, Andrew Gelman and Gary King, principal investigators (1946-1992), and Gary C. Jacobson, personal communication. Data on incumbents obtained from Gary C. Jacobson, personal communication. I would like to thank all of these people for their kindness in providing the data and thoroughness with which the data had previously been audited. The two types of omissions in their data pose no problem at all for the projects for which the data was initially intended, estimating the advantage of

The omissions were limited to two types. The first type is unchallenged incumbents in states such as New York and California, where these incumbents ran as “Democrat/Republicans.” These incumbents were coded as belonging to the party that they affiliated with in Congress. The other type of omission was for multi-member systems, such as the 2 at-large seat elections used by Arizona and New Mexico in the 1940s. The previous election percentages were estimated by pairing the top vote-getter to the worst performing member of the other party, and the 2nd vote-getter to the best performing member of the other party (in no case were more than 2 Democrats or 2 Republicans running for these 2 seats). The front-runner only had to fear BOTH members of the other party getting more votes than he did, whereas the runner-up had to fear the best of those who they defeated.

incumbency and research on the nature of individual congressional elections (and summary measures of challenger quality), respectively.

Table 1 **Measures of Vulnerability**

<u>Model</u>	<u>Determination of “Safe” Members</u>
1	Won with more than 65% of two-party vote in previous election
2	Won with more than 60% of two-party vote in previous election
3	Won with more than 70% of two-party vote in previous election
4	Won by less than 1 standard deviation of district vote share in previous election
5	Won (in election t_1) by less than mean loser’s prior (in election t_2) margin of victory ³²

³² This measure is actually created from a smoothed trend function to extrapolate to the 2nd election years in a redistricting decade. The mean was generally above the median value, due to the inclusion of unchallenged incumbents. This was done because the median loser numbers seemed low for ‘rule of thumb’ predictions of incumbent safety (such as the 60% cutoff discussed by Mann and Jacobson).

Table 2 **Descriptive Statistics for Measures of Vulnerable Seats**

Model	Variable	Minimum	Maximum	Mean	Std. Deviation
Model 1	Pres. Party Vulnerable Seats	56	181	109.7	29.2
Model 1	Opp. Party Vulnerable Seats	69	168	108.9	22.9
Model 2	Pres. Party Vulnerable Seats	35	123	79.8	21.8
Model 2	Opp. Party Vulnerable Seats	45	118	76.3	19.7
Model 3	Pres. Party Vulnerable Seats	87	195	136.4	29.9
Model 3	Opp. Party Vulnerable Seats	104	201	136.3	23.5
Model 4	Pres. Party Vulnerable Seats	34	94	63.3	18.6
Model 4	Opp. Party Vulnerable Seats	43	104	69.3	15.0
Model 5	Pres. Party Vulnerable Seats	41	105	70.7	18.8
Model 5	Opp. Party Vulnerable Seats	28	106	68.5	17.0

Table 3 Simple Test of Exposure Variables

Variable	“Jacobson Model”	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	-16.683** (-2.980)	-12.373 (-1.618)	-12.425 (-1.525)	-12.966 (-1.882)	-11.039 (-1.227)	-12.128 (-1.441)
% change in RDI per capita	1.173* (2.637)	.718 (1.100)	.961 (1.423)	.603 (1.034)	1.238 (1.754)	1.154 (1.708)
Presidential Approval	.212* (2.140)	.169 (1.242)	.163 (1.123)	.184 (1.497)	.117 (.700)	.146 (.980)
“Exposure”	-.713*** (-4.895)	.075 (1.987)	.067 (1.157)	0.085** (3.007)	.034 (.432)	0.033 (.466)
Adjusted R ²	.585	.206	.099	.356	.042	.043
Durbin-Watson n ³³	2.12 22	2.367 22	2.271 22	2.194 22	2.169 22	2.170 22

* p<.05; ** p<.01; *** p<.001, two-tailed test

³³ To make the models more comparable, the sample years for the Jacobson model was restricted to the same years as those available to the other models. The effects on the Jacobson model regressions were minor.

Table 4 Interactive Model Results for Presidential Approval Interactions

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	81.935 (1.892)	72.743 (1.408)	91.883* (2.408)	10.476 (.254)	-10.582 (-.422)
RDI	.728 (1.152)	1.047 (1.648)	.626 (1.045)	1.122 (1.662)	1.164 (1.662)
Presidential Approval	-1.659 (-2.002)	-1.475 (-1.498)	-1.852* (-2.534)	-.289 (-.367)	.115 (.237)
Interaction of Mean Adjusted PA and 'Exposure'	.016* (2.212)	.020 (1.668)	.014* (2.782)	.061 (.568)	.0006 (0.079)
Adj. R ²	.239	.161	.323	.049	.032
N	22	22	22	22	22

* p<.05; ** p<.01; *** p<.001, two-tailed test

Table A1 Number of Congressional Districts Validly Coded

<u>Year</u>	<u>Models 1-4</u>
1948	435
1950	434
1952	0 ^a
1954	433
1956	434
1958	436 ^b
1960	437 ^b
1962	0 ^a
1964	414 ^c
1966	424 ^c
1968	422 ^c
1970	431 ^c
1972	0 ^a
1974	433
1976	435
1978	434
1980	435
1982	0 ^a
1984	435
1986	435
1988	435
1990	435
1992	0 ^a
1994	434
1996	433
1998	434
2000	433

- a. Census redistricting years. Future work will be done to come up with seat classifications for these years.
- b. Alaska and Hawaii's addition to the Union resulted in a temporary increase above the statutory total of 435 seats.
- c. References on the many 1960s redistrictings are available upon request.

Figure 1 **Depiction of Effects of Shift Assuming Symmetry of Vulnerability**

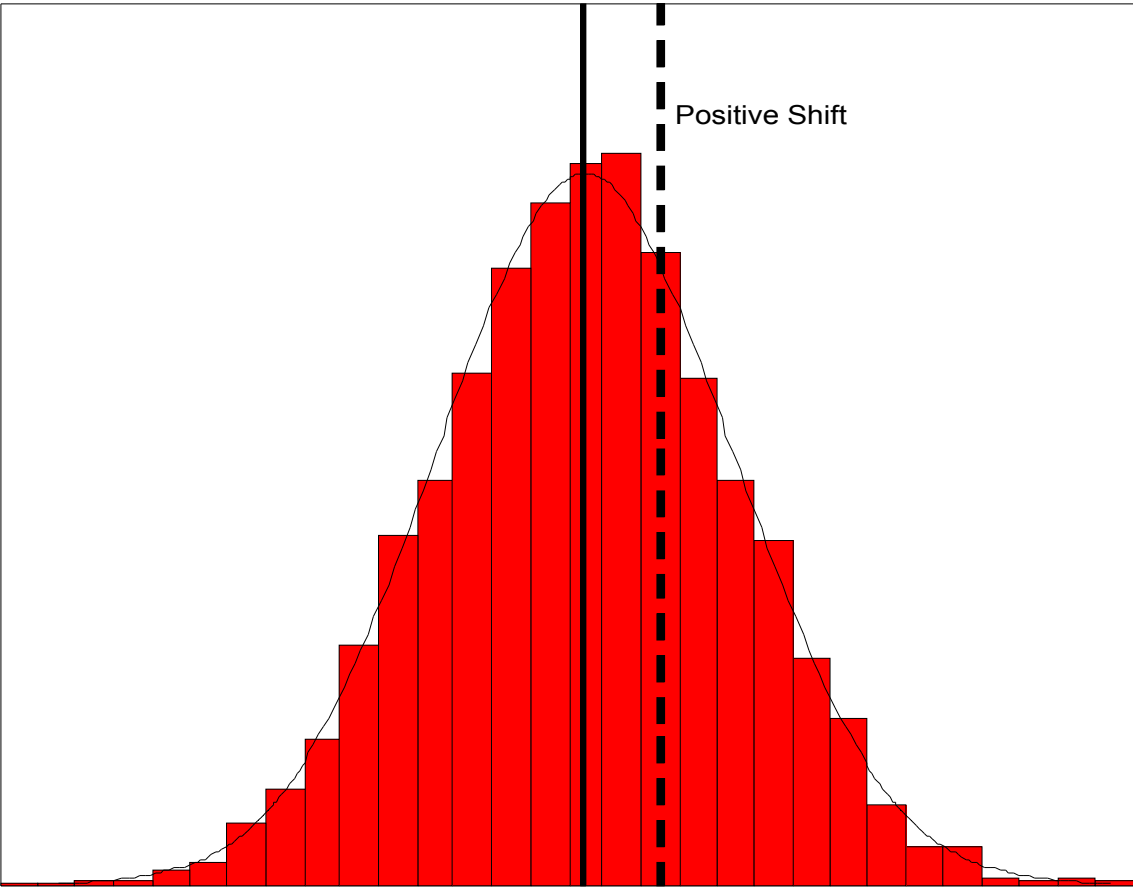


Figure 2 **Effect of Asymmetry on Expected Seat Loss/Gain**

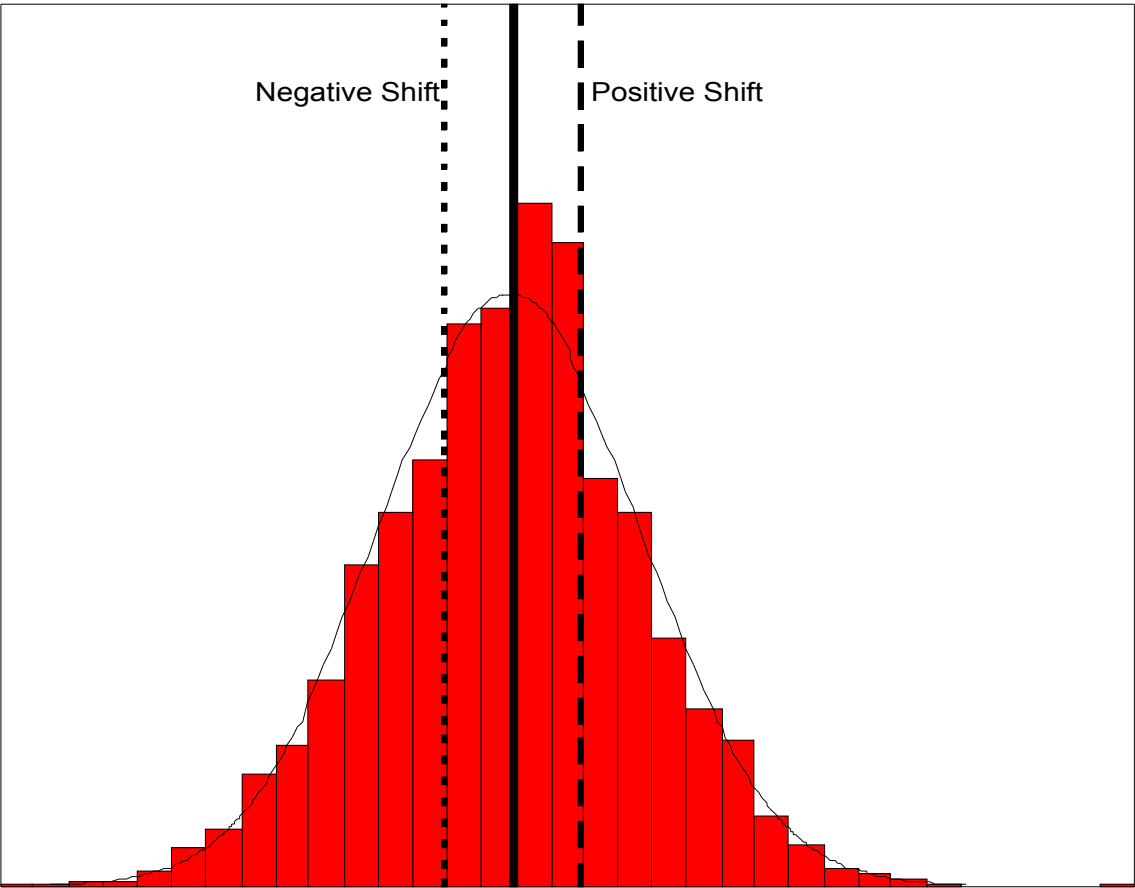
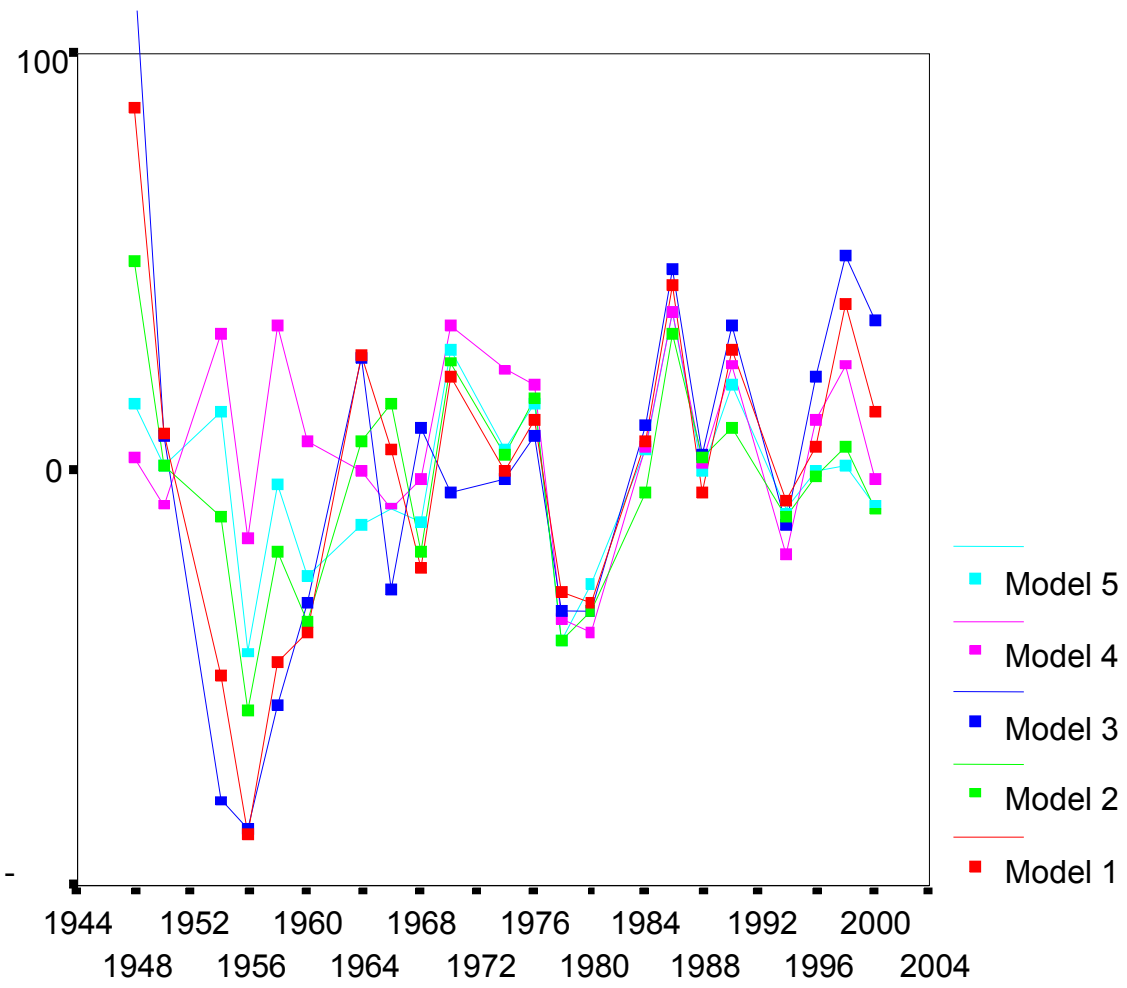
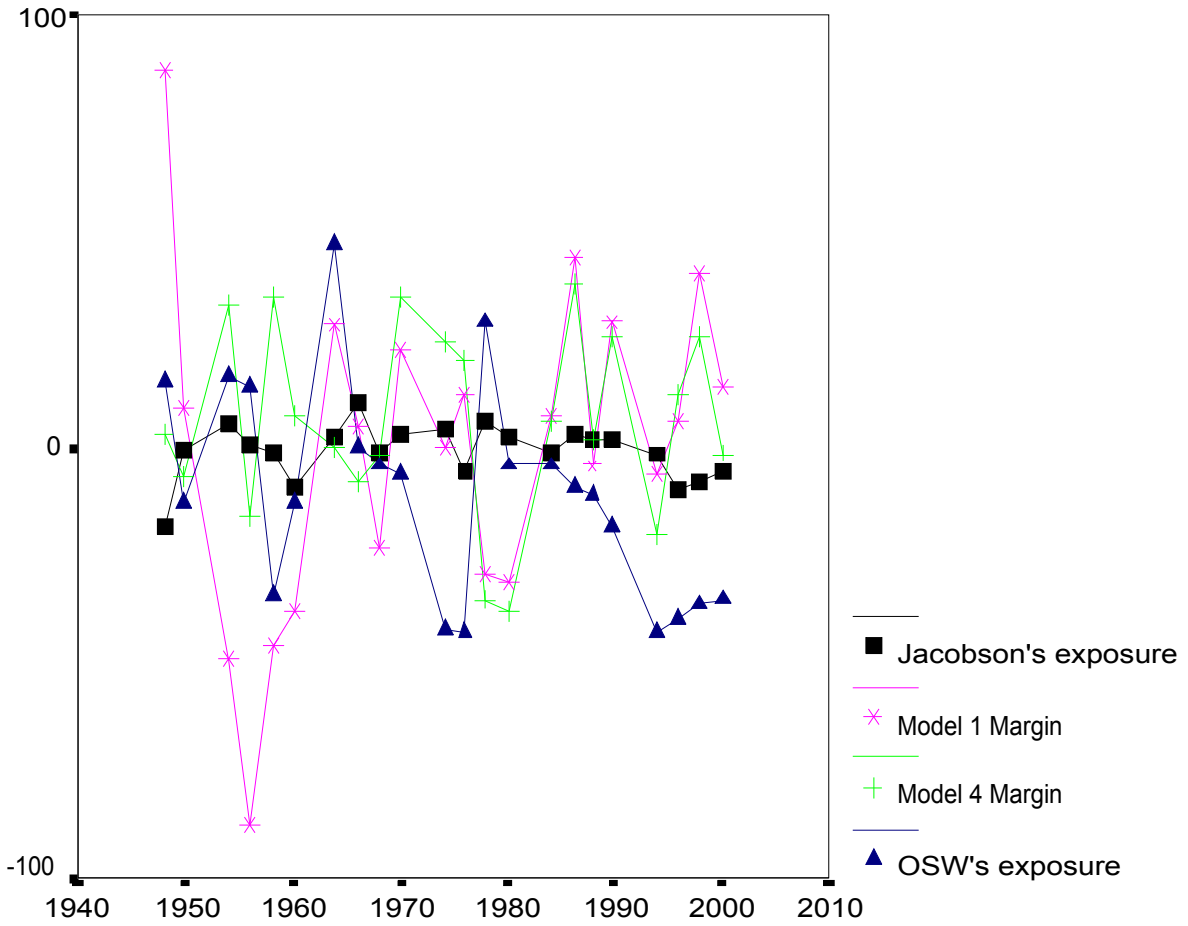


Figure 3 Margin of Vulnerable Seats



Break points exist at every year ending in 2.

Figure 4 Values of Various Models of “Exposure”



Break points exist at every year ending in 2 for Model 4 and Model 1 margins; Jacobson's and Oppenheimer, Stimson and Waterman's measures are valid at for every Congress.

References

- Banks, Jeffrey S. and D. Roderick Kiewiet, 1989. "Explaining Patterns of Candidate Competition in Congressional Elections." *American Journal of Political Science* 33:997-1015
- Bartels, Larry M., and John Zaller, 2001. "Presidential Vote Models: A Recount." *PS* 34:9-19
- Brady, David W., Robert D'Onofrio and Morris P. Fiorina, 2000. "Nationalization of Electoral Forces Revisited" in Brady, David W., James E. Cogan and Morris P. Fiorina, eds, *Continuity and Change in House Elections*. Stanford: Stanford University Press
- Cox, Gary W., and Jonathan Katz, 1996. "Why Did the Incumbency Advantage in U.S. House Elections Grow?" *American Journal of Political Science* 40:478-97
- _____, 2002. *Elbridge Gerry's Salamander*. Cambridge. Cambridge University Press
- Erikson, Robert S., 1990. "Economic Conditions and the Congressional Vote: A Review of the Macrolevel Evidence." *American Journal of Political Science* 34:373-99
- Erikson, Robert S., Joseph Bafumi, and Bret Wilson, 2001. "Was the 2000 Presidential Election Predictable?" *PS*, 34:815-19
- Erikson, Robert S., and Lee Sigelman, 1995. "A Review: Poll-Based Forecasts of Midterm Congressional Election Outcomes: Do the Pollsters Get It Right?" *Public Opinion Quarterly* 59:589-605
- Gelman, Andrew and Gary King, 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34:1142-64
- Groseclose, Tim and Nolan McCarty, 2001. "The Politics of Blame: Bargaining Before and Audience." *American Journal of Political Science* 45:100-19
- Gujarati, Damodar N., 1995. *Basic Econometrics*. New York: McGraw-Hill
- Hibbs, Douglas A., 1982. *The American Political Economy*. Cambridge: Harvard University Press
- Jacobson, Gary C., 2001. *The Politics of Congressional Elections*. New York: Addison Wesley Longman
- Krasno, Jonathan S., 1994. *Challengers, Competition and Reelection*. New Haven: Yale University Press

- Krasno, Jonathan S., and Donald Philip Green, 1988. "Preempting Quality Challengers in House Elections." *Journal of Politics* 50:920-936
- Lewis-Beck, Michael S., and Tom W. Rice, 1992. *Forecasting Elections*. Washington: CQ Press
- Mann, Thomas E., 1978 *Unsafe at Any Margin: Interpreting Congressional Elections*. Washington: AEI Press
- Marra, Robin F., and Charles W. Ostrom, Jr., 1989 "Explaining Seat Change in the U.S. House of Representatives, 1950-86." *American Journal of Political Science* 33:541-69
- Moore, David W., and Lydia Saad, 1997. "The Generic Ballot in Midterm Congressional Elections: Its Accuracy and Relationship to House Seats?" *Public Opinion Quarterly* 61:603-614
- Oppenheimer, Bruce I., James A. Stimson, and Richard W. Waterman, 1986. "Interpreting U.S. Congressional Elections: The Exposure Thesis." *Legislative Studies Quarterly* 11:227-47
- Pierseon, James E., 1975. "Presidential Popularity and Midterm Voting at Different Electoral Levels." *American Journal Of Political Science* 19:683-94
- Tufte, Edward R., 1975. "Determinants of the Outcomes of Midterm Congressional Elections." *American Political Science Review*, 69:812-26.