Democratic responsiveness concerns the degree to which government policies match public preferences. Studies of responsiveness typically use national surveys to characterize public opinion, but whether or not poll questions overlap with the policy agenda is unknown. The first of two empirical studies here with hundreds of issues on the national agenda in the United States from 1947 to 2000 reveals that public opinion is mostly unrelated to policy outcomes. The picture appears to be more ominous—i.e., opinion and policy are negatively related—on highly salient issues that attract media attention. A second analysis revisiting published work confirms that responsiveness patterns look different depending upon whether opinion-policy connections are evaluated based upon issues from the national legislative agenda as opposed to starting with available survey data and then looking for policy developments. Thus, conclusions about democratic responsiveness depend upon the issues that are examined, and many times opinion surveys do not include questions about tangible public policy options. In that sense, future changes in democratic responsiveness might go undetected because scholars often lack data on what goes into the denominator of democracy.
Democracies are often judged by the degree to which leaders respond to public preferences. Scholars have shown that representation of public opinion varies across issues (Miller and Stokes 1963; Monroe 1998; Wlezien 1996; 2004) and salient topics (Page and Shapiro 1983). Opinion-policy responsiveness also takes different forms across the American states (Erikson, McIver, and Wright 1993; Lax and Phillips 2012; Pacheco 2013), institutions (Erikson, MacKuen, and Stimson 2002; Kuklinski 1978), and cross-nationally (Brooks and Manza 2007; Soroka and Wlezien 2010). However, and quite provocatively, recent studies indicate that some opinions matter more than others and there also appear to be longitudinal trends pointing to a decline in responsiveness. That is, public policy tends to favor business leaders (Jacobs and Page 2005) or affluent individuals (Gilens 2005; Bartels 2008; cf. Soroka and Wlezien 2010) over the mass public (also see Enns and Wlezien 2011). Likewise, levels of responsiveness may have been declining in recent eras (Jacobs and Shapiro 2000; Monroe 1998).

While it is hard to overstate the normative importance of studies like these, virtually all of them share the same liability; they depend upon survey data to characterize public policy preferences. That is, many researchers collect available public opinion data and then look to policy outcomes rather than starting with a comprehensive set of national issues before assessing what the public wants. So, while theorists stress the importance of representation (e.g., Dahl 1956; Pitkin 1967) and studies like those mentioned above provide an empirical scorecard, the score nearly always depends on data availability—i.e., whether opinion data exists and on what topics. With few exceptions (see Burstein 2014), this limitation has not been the subject of sustained scholarly inquiry. Instead, responsiveness studies have been criticized for causal ambiguity (Page 1994) or for a bias toward the status quo (Gilens 2005), but a more subtle and potentially insidious problem is that traditional measures of democratic responsiveness obscure
the degree to which poll questions match the national issue agenda. Stated another way, survey research is the primary tool used to measure public opinion (e.g., Asher 2011; Herbst 1993; Glynn et al. 2004). Yet if the relationship between opinion and policy is important, then we should consider how patterns in polling affect perceptions of democratic responsiveness.

In contrast to many previous studies of opinion-policy responsiveness, this study begins with a measure of the national policy agenda and then incorporates public opinion data to the extent that it exists. Statistical analyses of hundreds of issues in the United States reveal a less optimistic impression of democracy in action than has typically been reported in the past. On issues from the national legislative agenda that were captured in public polls, opinion appears to be unrelated to policy or perhaps even negatively related—especially on salient issues—once efforts are undertaken to assess responsiveness on the entire agenda. All of this is to say that measurement decisions affect conclusions about opinion-policy linkages.

**Measuring Democratic Responsiveness**

Scholars often measure opinion-policy responsiveness in one of three ways. One approach is dyadic, where the behavior of elected representatives varies with public opinion (e.g., Bartels 1991; Hill and Hurley 1999; Karol 2007; Miller and Stokes 1963). A second tactic assesses whether changes in the public preferences are associated with changes in public policy (Page and Shapiro 1983) or the degree to which changes in global measures of public mood play out across various governmental institutions (Erikson, MacKuen, and Stimson 2002). A third and related variant considers the relationship between majority opinion and policy at any given point in time (e.g., Monroe 1979; 1998). Viewed statically (i.e., not over time), policy tends to correspond with majority opinion most of the time, especially on salient issues (see Shapiro 2011; Manza and Cook 2002a; 2002b; Burstein 2003; or Kuklinski and Segura 1995 for
reviews). All three approaches have virtues, but the last two address “…the responsiveness of the political system as a whole” (Page and Shapiro 1983, 176).¹

One normatively unsettling finding is that responsiveness seems to have declined during the late-20th century. For example, two important studies by Monroe (1979, 1998) show that government policies in the 1980 to 1991 period were less consistent with the preferences of a majority of the American public than during the 1960-1979 period; consistency declined from 63 percent in the first period to 55 percent in the second. Similarly, Jacobs and Shapiro (1997) found that congruency of opinion and policy change on welfare, crime, Social Security, and health fell from 67 percent during the Reagan Administration (1984-1987) to 40 percent during the Bush administration (1988-1991) before bottoming at 37 percent during the early years of the Clinton administration (1992-1994). Most recently, Gilens (2005, 784) found low levels of responsiveness in the 1981 to 2002 period; in a sample of more than 1,700 policy questions, only 35% of the policy changes in his study took place despite majority support for change in roughly 59% of the survey questions.²

These trends are unsettling. Citizens appear to be getting less from politicians than they were decades ago. The implication is that democracy in America does not work as well as it once did, especially not for the poor (Gilens 2005; Bartels 2008; cf. Soroka and Wlezien 2010). Among the causes that Jacobs and Shapiro (2000) cite is the rise of “crafted talk” where politicians package their proposals to resonate with members of the public so that they appear responsive without actually being so. Feeding these developments are party polarization,

¹ According to Jacobs and Shapiro (2000), responsiveness occurs when, “…the public’s substantive preferences point government officials in specific policy directions…” (p. 302; also see Page and Shapiro 1983). It is similar to representation, which according to Pitkin (1967) is, “…acting…in a manner responsive [to the represented]…” (pgs. 209-10). Representation and responsiveness occur when governmental actions reflect public opinion.

² Page and Shapiro (1983) also report declining responsiveness over time, from 67% congruence during 1935-45 to 54% during the 1960s before rebounding somewhat in the 1970s.
individualization in Congress, incumbency, interest groups, and divisive interbranch relations. As Jacobs and Shapiro (2000) conclude in their book, *Politicians Don’t Pander*, “Our main point is that the influence of public opinion on government policy is less than it has been in the past…” (p. xvi) and they point to “The general decline in responsiveness since the 1970s…” (p. 5). Others agree that responsiveness stagnated in the late-20th century (Mooney and Lee 2000).

However, some critics dispute this, arguing instead that pandering by politicians is widespread (Simon 2006). For example, in his review, Burstein (2003) reports that overtime comparisons within the same study “find more evidence of increase than decline…” (p. 36). Likewise, analysts who employ aggregate-level time-series approaches tend to paint more optimistic pictures (e.g., Erikson, MacKuen, and Stimson 2002; Hobolt and Klemmensen 2005; Wlezien 2004; Soroka and Wlezien 2010). Thus, responsiveness may or may not be declining, but in either case it is possible that the findings could be plagued by subtle measurement biases.

**Calculating Democratic Responsiveness**

Democratic responsiveness has been measured as the number of times public policy is (or moves in a manner) consistent with public opinion majorities. In other words, responsiveness represents the number of instances in which policies are consistent with public preferences over the number of issues for which there is public opinion data on policy preferences. Monroe (1998) adopts this basic formulation. He looks for majority sentiment and asks, is policy consistent with majority sentiment? It if is, then that opinion-policy pair is added to the numerator. All available instances of policy preferences measured by polls appear in the denominator. A variant of this basic calculation appears in Page and Shapiro’s (1983) classic study. They look for sizable—six percentage point—changes in public opinion. The number of times policy shifted in a congruent
manner goes in the numerator while the total number of available instances appears in the
denominator (also see Weissberg 1978).

What this means is that studies of responsiveness often depend upon the availability of
polling data. In other words, the reliance on public opinion surveys means that scholars can only
perform this type of calculation in areas where data exist. Page (2002, 332) acknowledges as
much when he states, “…one fundamental type of sampling bias subtly, and almost inescapably,
affects nearly all studies of opinion-policy links. We can study the impact of public opinion only
to the extent that public opinion is measured.” Similarly, in a review of the democratic
responsiveness literature, Burstein (2003, 38) writes, “Studies of the impact of policy on opinion
always begin with public opinion—that is, with issues for which public opinion data are
available. But such data are available for only a small fraction of all issues, those controversial
enough to warrant attention from survey organization.” These concerns have surfaced in other
works as well (e.g., Jones and Jenkins-Smith 2009, 54, Lax and Phillips 2012, 150, Manza and
Brooks 2012, 105; Shapiro 2011, 986), although not with accompanying analytical corrections.

An exception is Burstein (2014). He randomly samples legislative proposals introduced
before Congress in an attempt to provide what he hopes will be an unbiased assessment of
democratic responsiveness. Specifically, Burstein (2014) starts with 5,977 public bills introduced
in the 101st Congress during 1989-1990 and then chooses 60 to study, with 50 selected at random
from the entire set and another 10 selected randomly from bills reported out of committee.
Burstein finds no statistical link between opinion and policy (2014, 59). Opinion and policy were
consistent for 18 times, which was nearly half of the time that public opinion data existed (i.e.,
36 times out of 60). However, and consistent with the motivation underlying this examination,

---
3 Gilens (2012; also 2005) adopts a variant by modeling how the intensity of public preferences for policy change
(i.e., 55% versus 90%) are associated with actual policy change. Analyses reported later revisit his formulation.
public opinion measures were unavailable for nearly two dozen of the proposals. This meant Burstein’s estimate of consistency dropped to 31% of the time once the entire agenda was taken into consideration (i.e., 18 of 58=.31 or 31%; Burstein omitted two of the original sixty due to conflicting opinion indicators).

To recap, then, researchers often examine public policy outcomes once they identify a set of issues where public preferences have been assessed in surveys. This could lead to misleading inferences. While pollsters ask questions on a variety of topics (e.g., Stevens 2002; Shaw 2000; 2003; Shaw and Mysiewicz 2004), they are under no obligation to field a representative set of questions. Indeed, constraints (e.g., finances, time pressures, or the desire to keep surveys short) all conspire to make some questions more likely to appear than others. As Burstein (2014, 45) writes, “Public opinion polls focus on issues important to the public—the very issues on which the public is most likely to hold elected officials accountable, and on which, therefore, democratic governments are mostly likely to do what the public wants.” Later Burstein (2014, 56) speculates that, “…previous research that included only issues on which opinion had been measured (however indirectly) probably ignored at least 40 percent of the proposals that Congress considers.” The problem is that while this could mean that, “…current estimates of the impact of opinion on policy are too high…” (2014, p. 70) due to such biases, Burstein examines only a tiny fraction of congressional policy proposals (i.e., 60 issues out of nearly 6,000 bills introduced in a single congressional session). And yet even then it is instructive that Burstein could only locate specific issue preference data for a handful of policy proposals. So, even though he started with policy first before moving to the survey data, by in large Burstein found that the polling did not correspond to the issue agenda.
A New Denominator

In calculating democratic responsiveness, scholars are frequently limited to publicly available poll data. However, lawmakers might act on legislative proposals not covered by poll questions just as they may fail to act on issues of national importance that escape the attention of survey researchers. A problem of this nature plagued studies of legislative gridlock for years; accurately characterizing the amount of legislation that fails to pass as a result of factors like divided government (the numerator) depends on how one specifies the legislative agenda (the denominator). In a landmark study, David Mayhew (1991) concluded that divided government did not affect legislative output. Years later, however, Sarah Binder (2003; 1999) argued that gridlock reduced output once the scope of the national policy agenda was considered.  

Denominator revisions have altered conclusions in other studies as well. For instance, recharacterizing the denominator of electoral participation—from the voting age population to eligible voters—changed perceptions of turnout patterns in America (McDonald and Popkin 2001). Voter participation is higher if calculated with the voting eligible population as opposed to all individuals old enough to vote, which includes many ineligible individuals. Likewise, scholars examining roll call voting in legislatures discovered via recasting the denominator that studies of party cohesion over-estimated the degree of party unity in legislatures because the sample was artificially limited to recorded votes rather than voice votes (Carrubba et al. 2006).

A related problem looms in the background of democratic responsiveness calculations as well. Polling takes place on a wide variety of issues, but some topics are more popular than others. If citizens are not asked about matters that leaders are considering, then determining

---

4 Binder (2003) uses New York Times editorials to determine the policy agenda and then uses this in the denominator of a gridlock calculation with the number of policies that failed to pass in the numerator. Interestingly, a lagged public mood variable does not affect gridlock (Ch. 4 and App. C), which suggests that public preferences do not influence public policy output once other factors are considered. For a review of other divided government works, see Binder (2003, especially footnotes 5 and 6 on pages 161-2).
whether opinion plays a role becomes much harder. The basic intuition is that responsiveness is a function of elite behavior (i.e., policy outcomes), polling, and the policy agenda. Whether coverage-bias exists depends, to some degree, on whether conclusions differ after accounting for the entire policy agenda. Figure 1 shows, via diagrams, the essence of the potential problem. Public opinion on policy might only tangentially overlap with the national policy agenda as shown with the oval in Figure 1 labelled “minimal overlap.” In other instances, it might “somewhat overlap” (diagonal stripes) as Burstein suspects, with pollsters only partially covering what legislators are considering. Alternatively, it could be the case that polling organizations ask about the national agenda often, so that the there is “substantial overlap” of polling and the policy agenda as shown in the oval with dark blue shading.

Revisiting the denominator in this manner has the potential to influence judgments of democratic responsiveness. As Brehm (1993) writes, “There is hardly an aspect of American political life untouched by polling and survey research” (p.3). Yet the polling agenda probably does not match the national policy agenda perfectly. If government leaders spend a good deal of their time on issues that are ignored by pollsters, then it might seem like responsiveness is declining even though it is not; in other words, the public might get what it wants on topics in polls, but those topics might not reflect the legislative agenda.5 Thus, responsiveness scores might be lower (or higher) if poll questions fail to correspond to the national policy agenda.

Burstein’s core intuition—that the study of democratic responsiveness may be flawed due to sampling bias—has merits; in such studies, it is important to define a set of issues being

---

5 Charting the numerator is hard too, especially when pollsters ask about a policy area in general without specifying a particular piece of legislation (e.g., questions about welfare spending might pertain to food stamps or Temporary Assistance for Needy Families [TANF], but what about the earned income tax credit?). As Page and Shapiro (1983, 176) note, “…some opinion items are so ambiguous that they are not easily matched with specific policies.”
considered and then assess public opinion. Regrettably, though, Burstein’s study was perilously small and the methodology he employed has shortcomings as well. In particular, Burstein is conscious of sampling coverage bias as revealed in his decision to use random sample selection, but sampling pertains to how it is done as well as how many are selected. Larger sample are more precise (i.e., less sampling error). Few survey researchers would be comfortable drawing inferences about the population from a sample of 50 to 60 units. At the very least, margin of error with a sample of this size would be so large as to make any conclusions extremely tentative.\(^6\) Burstein’s sample shrinks further because public opinion data does not exist on some of the legislative proposals he investigated. Thus, it is not surprising that Burstein failed to find a significant relationship between opinion and policy in such a small sample, one that lacks statistical power.

Similarly, the proposals Burstein employs are created by lawmakers. That is, he only studies (or samples from) what has been introduced, but agenda control is important (Carrubba et al. 2006; Krehbiel 1991). Legislative leaders determine what is or is not going to be considered. Therefore, by studying bills that have been introduced—rather than the broader set of issues that lawmakers could be acting on—Burstein’s sample might not be representative in the sense that it depends upon the strategic actions of politicians, many of whom are already invoking public opinion strategically in an attempt make it look like they are responsive (Cook, Barabas, and Page 2002; also see Jacobs and Shapiro 2000). In that sense, Binder’s (2003) characterization of the national legislative agenda, while not perfect, casts a wider net in the sense that issues do not need to be formally advanced by legislators to be worthy of consideration.

\(^6\) Burstein (2014, page 32, footnote 5) defends his sample as being large enough and he points to other studies which only consider a few dozen issues (e.g., Soroka and Wlezien 2010 consider fewer than 35 issues).
Empirical Analyses

To what degree does the polling agenda match the public policy agenda? Furthermore, how do polling patterns influence our perceptions of democratic responsiveness? To address these questions, ideally we would have data on every survey question that has ever been asked and, in an ideal world, polls would overlap with the policy agenda to a considerable degree. Unfortunately, such data do not exist. However, it is possible to see how reality deviates from the ideal based upon the survey questions that have been methodically archived at the Roper Center for Public Opinion Research at the University of Connecticut. The Roper Center is “the world's largest archive of survey data” with more than 600,000 questions. Importantly, the iPoll database shows opinion frequencies for individual questions as well as the name of the organization conducting the study, field interview dates, sample sizes, and other methodological details.

The analyses reported here make use of the iPoll archive in two ways. The first study below starts with Binder’s data as an indicator of the national policy agenda and then attempts to locate poll questions on those topics. In doing so, we will be able to learn how survey questions map onto the national policy agenda as well as the degree to which opinion is associated with policy outcomes. A second empirical study adopts a somewhat different approach by revisiting the main analyses from Martin Gilens (2012) in his book titled Affluence and Influence to see if the conclusions change once responsiveness is recalculated for issues on (or off) the national agenda. The analyses are divided into two studies with details on the methods appearing in each section below.
Study 1: Democratic Responsiveness on the National Policy Agenda

Patterns in polling on policies could alter perceptions of democratic responsiveness, especially if survey questions do not correspond to the policy agenda. To investigate this possibility, opinion-policy connections are evaluated based upon the national issue agenda. In particular, responsiveness scores were calculated using the entire set of Sarah Binder’s (2003) data on the U.S. policy agenda from the 80th (covering the years 1947 to 1948) through the 106th (1999 to 2000) congressional sessions. As noted earlier, Binder’s measure is based upon unsigned editorials appearing in a prominent national newspaper, the New York Times, during each two-year congressional session. The issues mentioned at any given point were diverse (e.g., price controls in the post-war period of the late-1940s, statehood for Hawaii in the 1950s, desegregation in the 1960s, ratification of nuclear arms treaties in the 1980s, welfare reform in the 1990s) and the number of issues on the agenda varied too. For example, the Congress with the most issues was the 99th (1985-1986) with 141 distinct issues while the 86th (1959-1960) had the least issues with 62 legislative concerns. Across the 27 sessions, the average was 104 issues per congressional session.

In the aggregate, there were 2,818 issues on the national policy agenda from 1947 through 2000. They are, however, disconnected from public opinion. To match opinion to policies, two coders reviewed more than 20,000 polls from the Roper collection to determine whether polling data existed on the national legislative issue agenda. Poll questions selected for inclusion were all from nationally representative surveys (drawn from more than 70 organizations—with most coming from well-known firms like Gallup, Princeton Survey

---

7 Although some congressional sessions spanned three calendar years (e.g., January 3, 1957 to January 3, 1959 for the 85th), they are reported here in two year intervals given that most work takes place in the first and second years.

8 The coders achieved a high Krippendorf alpha reliability score of .82 on a random subsample of cases. The issues were also sorted randomly to ensure that one coder did not code certain time periods or issues.
Research Associates and Louis Harris & Associates, major news media, and academic survey researchers from the University of Michigan and the National Opinion Research Center). The principle requirements were that questions (A) correspond to the issues on the agenda and (B) capture what the public prefers on an issue. Once a question was deemed relevant, a database was constructed with all available questions for every issue. As with the legislative agenda, there was considerable variation. Of the nearly 3,000 issues on the agenda in the second half of the 20th century, 658 had at least one question containing public preferences on those same issues.

These numbers mask considerable heterogeneity because for some Congresses, as few as 6 percent of the issues on the agenda were queried in polls (4 out of 62 issues in the 86th Congress) while other times more than a third of the issues on the agenda appeared in survey questions (the 104th Congress with polling on 39 of the 107 issues). Yet, on average only 24 percent of the issues on the agenda were queried in Roper archived questions, although Figure 2 shows the variations in the size of the agenda as well as the proportion with poll coverage. Thus, the polling agenda does not necessarily align with the policy agenda; indeed, nearly three-quarters of issues on the policy agenda are not covered in national polls. Scenarios akin to the “minimal” to only “some overlap” depicted earlier (figure 1) seem to describe reality, providing even less coverage than Burstein (2014) suspected would exist.

Nevertheless, survey researchers were asking questions. On the 658 (out of 2,818) issues from the national agenda that were the subject of at least one poll question, some issues featured

---

9 Any question wording that could be used to characterize public preferences was included (e.g., “favoring,” “supporting,” “approving” of a policy) as long as the question was asked of the entire sample, not just those who passed a prior filter item. Also, the time frame searched for each congressional session was the entire Congress (i.e., the official start and ending dates) as well as the weeks before the official start of the session dating back to the election on the premise that once a new Congress is elected, the public might be queried about issues ahead of their formal work period. As it turned out, most of the questions deemed relevant were during the time period that Congress was in session.
dozens of questions (e.g., limits to FBI powers in the mid-1970s as well as aid to the Nicaraguan rebels or tax reform in the 1980s). In fact, more than sixty percent of the items with poll coverage featured more than one poll (410 of the 658, or 62 percent), with an average of nearly 4 questions (mean=3.7) on the issues that garnered any attention from pollsters. Overall, then, most of the policy agenda could have been covered with a more equitable distribution of polling (i.e., if the 2,448 questions were spread across the 2,818 issues instead of focusing on just a quarter).

Notwithstanding the prevalence of polling, it could be that responsiveness calculations are unaffected by the lack of coverage. This would be the case if the polling coverage, even if small, was representative. While that could have been what happened, it is possible to predict which issues receive coverage as well as the number of questions, so it is not a random sample.\footnote{These analyses are akin to “failed” randomization checks in an experiment; specifically, a probit model predicting polling coverage (0=no questions vs. 1=at least 1 poll question) as well as a Poisson model for the count of poll questions on an issue both reject the null hypothesis of no significant predictors (likelihood ratio tests against the null are $p < .001$).}

In analyses (not shown), issue salience, as captured in Binder’s (2003) measure of the number of editorials in the New York Times, is a statistically significant ($p < .01$, two-tailed) and positive predictor of polling. Certain congressional sessions—especially those occurring after the 1970s—were also more likely to have received attention from pollsters. Institutional configurations, such as Democratic control of the Presidency, House, or Senate were also significant predictors of polling counts, positively so in the case of the first two. Issue dummies for the most numerous categories (Crime/Legal/Civil Rights, Environmental/Energy/Science, Economic/Fiscal/Tax, ForeignPolicy/Trade, Defense/Terrorism/Intelligence) were all associated with more polling coverage, except for the topic of governmental administration/election procedures which was negative relative. Net of these factors, legislative success (i.e., issue...
passage or not) was negative but unrelated to the existence of polling, although it was close \( p < .14 \), two-tailed) and this helps set the stage for the next set of analyses.

In particular, Binder (2003) characterizes the national policy agenda as well as the legislative success on those same polices (i.e., congressional passage and presidential approval or an override of his veto). Of the 2,818 issues on the agenda from 1947 to 2000, just less than half of them were counted as legislative successes \( (n=1,350 \text{ or } 47.9\%) \). The proportions are roughly similar for the subset of issues with polling coverage; just under half \( (46\% \text{ or } 305 \text{ out of } 658) \) were instances of policy passage. However, just because the pollsters ask about policies that have a roughly equal chance of succeeding does not mean that public will prevails. Figure 3 demonstrates that the public often fails to secure representation.

**INSERT FIGURE 3 HERE**

Figure 3 shows the proportion of democratic responsiveness on the vertical axis (instances when a majority of the public prefers change and gets it as well as when majorities prefer no change and policy does not change). The horizontal axis shows biannual congressional sessions from the 80th in 1947-1948 until the 106th in 1999-2000. There are two series depicted, one for responsiveness overall in black and another for responsiveness on the subset of issues with polling coverage. In either case, both are low. The average of the series in black is .11 with a range of 0 to .22. The average for the light line is .47 with a wider range of 0 to .78. The figure highlights particular congressional sessions with high responsiveness (e.g., 1953-54, 1967-68, 1981-82, and 1995-96) as well as some low points for certain sessions relative to what took place just before or after (the 1974-75 period as well as 1959-60 with no responsiveness but also with only four agenda items that have poll coverage). For both series, they end below where they started, but there is so much year to year volatility that it is hard to say definitively that
responsiveness is on the decline. It is perhaps easier to say that responsiveness appears to be low, especially overall (in black); it rises above .20 just twice.

A more comprehensive analysis of the relationship between public opinion and policy appears in Table 1. The table shows coefficients for models where the dependent variable is legislative passage on an issue from the national policy agenda. The main predictors are public opinion as measured by the percentage of the public in favor of the issue in question, issue salience as captured in Binder’s data with the number of editorials, and finally an interaction of these two. The first three columns report models (one with only public opinion, one for public opinion and its interaction with salience, and a final one that includes the interaction along with several control variables) for the “original” data—that is, the 658 issues with at least one poll question. The next set of three models in Table 1 under “missing opinion data imputed” attempt to recover the missing public opinion predictors using the variables discussed earlier that predict polling coverage. These analyses make use of multiple imputation techniques (King et al. 2001) to generate five estimates of what public opinion was for any given topic that lacks coverage and then combines the estimates in single model. As it turns out, however, the estimates are similar across the columns.

INSERT TABLE 1 HERE

Public opinion, on its own, is negatively related to policy passage in the first column (coeff.=-.0042) but the standard error (.0031) is too large to attain statistical significance. Adding salience and its interaction does quite a bit; salience is positive (and represents instances with high NYT coverage but low public opinion) but the interaction of salience and opinion is negative, suggesting that strong public demand is less likely to result in policy passage when the issue is one that is featured regularly in editorials. This is true even with control variables for
congressional session, issue dummies, the survey organization conducting the poll, the research assistant researching poll coverage, and coder confidence.\textsuperscript{11}

The models in the first three columns of Table 1 analyze 658 issues that received attention in public opinion polls, but sometimes there are several poll questions in Roper for any given issue on the agenda, which means the total number of observations in the “original” data is 2,448.\textsuperscript{12} The number of observations increases for the last three columns of Table 1 because public opinion has been multiply imputed for the issues which lacked coverage in surveys (using techniques in King et al. 2001). In these analyses, the number of issues is 2,818 and the total number of observations in the model is 4,622 due to the repeated polling on issues. Yet, the conclusions are similar. The coefficients are roughly the same size and sign, but the significance improves due to the added power. Opinion on its own seems negatively relative to policy passage, but an interaction with salience reveals that this is true on the issues that are featured often in the press (perhaps because they are issues which should be enacted but fail).\textsuperscript{13}

The best way to illustrate the effects is by calculating predicted probabilities of legislative passage for various combinations of opinion and salience. In other words, do issues on the agenda pass if public opinion is in supportive, especially once salience is taken into account? Table 2 shows that opinion and salience are indeed important predictors of legislative success, but in a manner opposite what might be preferred normatively. The first row puts the average probability of passage at .45 (with a 95% confidence interval from .43 to .47) for issues with

\footnotesize
\begin{itemize}
  \item \textsuperscript{11} Most of the time the coders expressed high confidence in their ability to match public opinion to the policy agenda items. Occasionally, though, there was some ambiguity in the way the topic was described. In these instances, comprising 24% of the policy agenda, they were often only somewhat unsure (616 of 2818, or 22%) and rarely very unsure (45 out of 2818, or 1.6%). The terms in the model for coder confidence attempt to control for this possible measurement error. The terms are insignificant and the findings look similar if unsure cases are dropped.
  \item \textsuperscript{12} The unit of analysis for these analyses is an opinion-policy pair. Since issues could have more than one poll, the models employ clustered standard errors to account for the lack of independence.
  \item \textsuperscript{13} They include a diverse array of topics like the war on drugs, steel import quotas, tuition tax breaks, grazing permits, high speed rail, domestic violence, repeal of the Glass-Steagall banking regulations, endangered species protection, online copyright infringement, flight caps at airports, forest firefighting, and bankruptcy laws.
\end{itemize}
average public support (roughly 54%) and average salience (~5 editorials).\textsuperscript{14} Six other scenarios are shown in Table 2. The first set varies opinion from low to high (i.e., two standard deviations below/above the mean) to show that moving opinion dramatically in support of a bill has a modest negative effect on the probability of passage of -.06 (from .48 on issues with low support but average salience to .42 with high levels of public support, with the difference in bold typeface), albeit an insignificant one (se=.04, interval from -14 to .03). On issues with low salience (only 1 editorial, the lowest of the sample), the probability of passage on issues as we move from low to high opinion rises negligibly (.04) in a statistically insignificant fashion (se=.05). The most dramatic scenario appears in the last set of predicted values for Table 2. On issue with high salience (i.e., 20 editorials), passage is very likely when opinion is unsupportive (pr=.79, se=.03), but it drops precipitously to a less than even chance of passage (pr=.41, se=.04) for high salience issues when the public is highly supportive. This 39 percentage point drop in the likelihood of passage is statistically significant (95% interval from -.51 to -.26), and it is a shocking finding. Public opinion appears to be unrelated to policy passage except on highly visible issues, in which case it is negatively related.

\textbf{INSERT TABLE 2 HERE}

It is possible to analyze data differently, such as whether a majority of the public prefers change and policy changes as well as instances when the public prefers no change and policy remains the same (i.e., akin to Burstein 2014 or Monroe 1998). Those analyses are similar in that responsiveness looks different when evaluated relative to the policy agenda. Opinion is unlikely to produce legislative action overall, but especially for the subset of issues that are the subject of

\textsuperscript{14} The predicted probabilities in Table 2 are based upon simulations (King, Whittenberg, and Tomz 2000) from the model shown in column 2 of Table 1 for only cases where public opinion exists and without control variables. However, the substantive patterns are the same in simulations based with models including controls (column 3 of Table 1) or with models based upon the imputed data (columns 5 or 6).
repeated editorials. Also, while most of the policy agenda lacks polling coverage, there are still hundreds of cases across more than fifty years (i.e., many more than Burstein 2014, even without data imputation). Yet, the picture becomes even clearer, from a statistical standpoint, with attempts to recover what the public may have preferred on more than three-quarters of the agenda uncovered by polls. In either case, the normative conclusion is not especially encouraging. Democratic responsiveness in the late-20th century is low with opinion mostly unrelated to policy or even negatively so on issues that draw media attention.

**Study 2: Reanalysis of Gilens (2012)**

The analyses above start with the policy agenda and then move to polling coverage. Yet, it is possible, to revisit a past work on responsiveness with the more modest goal of seeing if the findings are sensitive to whether the issues are on the national agenda. To accomplish this, it was necessary to find a published opinion-policy responsiveness study that (A) dealt with specific policy issues, (B) covered a fairly long time period in sufficient depth, and (C) made the data publicly available for reanalysis. One prominent work that met all three criteria was Gilens (2012) who studied the relationship between public opinion and public policy over many of the same years as studied earlier. While his criteria for inclusion of issues was more limited (i.e., he began by identifying survey questions using the word “oppose” while in the previous study many different styles of policy support questions were included) it is an example of starting with polling first (as opposed to the agenda). Thus it is possible to look within his analyses for issues that happen to be on the policy agenda versus those that were not.15

---

15 The data for Gilens (2012) is posted on the Russell Sage Foundation website, although code to reproduce the findings was not publicly available. In personal communication, Gilens graciously provided help on data coding decisions that permitted a successful replication of the main set of empirical findings (Table 3.1 in Gilens 2012).
Gilens (2012) is a complex study, diving into numerous aspects of responsiveness (e.g., which income groups get represented, policy topic variations, interest group influence, over time patterns, electoral concerns). While it not possible to revisit every analysis here, his core analysis in the first empirical chapter relates public opinion to policy responsiveness. The first column of Table 3 below perfectly reproduces the results shown in Gilens (2012, page 76, column 1, Table 3.1). In particular, his methodology uses a public opinion logit coefficient value to predict policy change on issues from 1981 to 2002. The entries in lower rows of column 1 of Table 3 replicate the other results reported in Gilens that show the predicted probability of policy responsiveness if 20% favored the policy (which is estimated at .19) versus predictions if 80% favored the policy (estimated to be .43). This 24 percentage point gain (i.e., .19 to .43) is then recast as a ratio of 2.2 based upon high opinion over low (i.e., .43/.19=2.2). The higher the ratio, the more responsive policy is to the preferences of the upper income group relative to the poor.

Columns 2 through 4 of Table 3 show the same analyses for a subset of Gilens’ data. His analysis (Gilens 2012, Chapter 3) spanned 1981 to 2002. The models in the columns beyond the first omit the last two years because Binder’s agenda data do not extend that far. Setting aside these two years, however, does not change the substantive story. Public opinion is related to policy change overall (coeff.=.31, p < .01), albeit more strongly for those in the 90th income percentile than for members at the bottom 10th percentile of income—as evidenced by the higher relative ratio in the predicted probabilities at low and high levels of opinion support. The last three columns of Table 3 include a dummy variable for whether the issue was on the national policy agenda (810 of the 1,520 issues) as well as an interaction between opinion and being on the policy agenda. The patterns show that Gilens’ analyses are unaffected by whether or not the
issue was on the agenda or not. Specifically, the interaction term representing issues on the agenda is positive but insignificant, although the relationship is less pronounced for the poor.

While the significant relationship between opinion and policy—even for issues off the policy agenda—might seem inconsistent with the patterns reported earlier, several points deserve comment. First, Gilens began with polling data and then moved to policy while the analyses reported earlier with Binder’s agenda data began with policy. This means that all of Gilens’ data were considered but then a portion was deemed on or off the agenda; in reality, there were still many issues off the agenda. Second, there was a subtle difference in the procedures for coding policy. In particular, Gilens looks for policy change for a period of four years beyond the survey question while Binder evaluated legislative success for issues “…enacted into law by the end of the Congress” (p. 38). To addresses both of these limitations, and to better equate the two styles of research, a final set of analyses (not shown) revisits the analyses presented in Study 1 here, but this time for a subset in which is possible to match Gilens’ opinion and outcome data to issues on the national policy agenda. For this smaller set of issues on the agenda with overlapping dates (i.e., for 181 of the 1,150), Gilens’ opinion measure is a significant predictor of policy only when using his outcome measure. In other words, public opinion for the whole sample and for the 90th percentile of income is positively and significantly related to policy ($p < .05$ for both), albeit not significantly so for the 10th percentile ($p < .23$). If the outcome measure is switched to Binder’s measure, then the patterns disappear; opinion is unrelated to policy.\(^\text{16}\)

Moveover, salience plays a role as it did earlier. Even with Gilens’ policy outcome measure, an interaction of opinion and salience produces the same patterns identified earlier (i.e., as in columns 2 or 3 and 5 or 6 of Table 1). Public opinion on low salience issues (the

\(^{16}\) Null results like these and others reported earlier are important to combat potential biases in favor of publishing statistically significant findings (e.g., Gerber and Malhotra 2008; Gerber et al. 2010).
constitutive term of the interaction) is positive and significant, but just as it was earlier the interaction of opinion and salience interaction is negative and statistically significant. Thus, we see confirmation of the basic intuition guiding the analyses. Perceptions of democratic responsiveness are heavily influence by the denominator employed. Starting with opinion data and then moving to policy tends to show legislative success if opinion is highly supportive or maintenance of the status quo if the public is unsupportive. The picture looks different—and less optimistic normatively—when responsiveness is recalculated as the relationship between opinion-policy on issues from the national agenda.

Conclusions

How well does democracy work in the United States? Answers vary, but the general trend seems to be movement in the wrong direction from the standpoint of democratic theory. However, democratic responsiveness calculations depend on polling data. The irony is that even though polls seem to be ubiquitous, relatively few questions pertain to policies on the national agenda. In other words, while democratic responsiveness appears to be low on observable issues, but it is hard to tell for the vast majority of the national policy agenda.

Survey experts periodically re-evaluate survey methods (e.g., Smith 1987; Keeter 2012; Keeter et al. 2006; Marsden and Wright 2010) as well as the evidence on the opinion-policy linkages (Manza and Cook 2002a; 2002b; Burstein 2003). Distilling democracy to a single number makes overtime comparisons easier, but doing so risks glossing over some of the underlying trends. As with legislative gridlock (Binder 1999) or voter turnout (McDonald and Popkin 2001), how the denominator is calculated influences judgments of democratic vitality.

---

17 Specifically, for overall opinion, the interaction is -.09 with a standard error of .04, which is \( p < .04 \), two-tailed. It is slightly less significant for the 10th income percentile \( p < .10 \), but it is highly significant \( p < .02 \) when using opinion from members of the 90th income percentile.
There, of course, are limitations to this study. Chief among them is the reliance on data from the Roper collection. The iPoll database is the world’s largest repository of public opinion questions, but it does not include all polls ever asked. For example, the Roper data do not incorporate the private polls that politicians conduct (e.g., Druckman and Jacobs 2006). Another limitation concerns the operationalization of the national policy agenda, which is based upon editorials from one major newspaper. One could envision other ways of operationalizing the issue agenda that might influence the results. Yet, irrespective of what measures are used, scholars investigating opinion-policy connections in the future should begin by considering which issues are on the national agenda.


Carrubba, Clifford J., Matthew Gabel, Lacey Murrah, Ryan Clough, Elizabeth Montgomery, and Rebecca Schambach. 2006. “Off the Record: Unrecorded Legislative Votes, Selection
Bias, and Roll-Call Vote Analysis.” *British Journal of Political Science* 36 (Oct.): 691-704.


Wlezien, Christopher. 1996. “Dynamics of Representation: The Case of U.S. Spending on


<table>
<thead>
<tr>
<th></th>
<th>Original Data</th>
<th>Missing Opinion Data Imputed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Opinion</td>
<td>-0.0042</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
<td>(0.0037)</td>
</tr>
<tr>
<td>Issue Salience</td>
<td>--</td>
<td>0.0694</td>
</tr>
<tr>
<td></td>
<td>(0.0207)</td>
<td>(0.0189)</td>
</tr>
<tr>
<td>Public Opinion X Salience</td>
<td>--</td>
<td>-0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Control Variables</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Constant</td>
<td>0.1746</td>
<td>-0.3778</td>
</tr>
<tr>
<td></td>
<td>(0.2027)</td>
<td>(0.2380)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2,448</td>
<td>2,448</td>
</tr>
<tr>
<td>Number of Issues</td>
<td>658</td>
<td>658</td>
</tr>
</tbody>
</table>

**Note:** The table displays probit coefficients with clustered robust standard errors in parentheses. The dependent variable is policy output with value of 1 for instances of legislative success on an issue from the national policy agenda and zero otherwise (Source: Binder 2003). Issue salience is a count of the number of editorials on the issue from the *New York Times* (Source: Binder 2003). The models with "control variables" include dummy variables for congressional session, issue, survey organization, coder identity, and coder confidence. The models with "missing data imputed" employ multiple imputation techniques (e.g., King et al. 2001) to recover the missing public opinion responses on issues from the national policy agenda without survey coverage. **p < .01; * p < .05 (two-tailed).**
Table 2. Predicted Probability of Legislative Success in Various Opinion and Salience Scenarios

<table>
<thead>
<tr>
<th></th>
<th>Probability of Passage (S.E.)</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability</td>
<td>Low</td>
</tr>
<tr>
<td>Average Public Opinion Support, Average Salience</td>
<td>.45 (.01)</td>
<td>.43</td>
</tr>
<tr>
<td>Low Public Opinion Support, Average Salience</td>
<td>.48 (.02)</td>
<td>.43</td>
</tr>
<tr>
<td>High Public Opinion Support, Average Salience</td>
<td>.42 (.02)</td>
<td>.38</td>
</tr>
<tr>
<td><strong>Difference (High to Low)</strong></td>
<td><strong>-.06 (.04)</strong></td>
<td><strong>-.14</strong></td>
</tr>
<tr>
<td>Low Public Opinion Support, Low Salience</td>
<td>.39 (.03)</td>
<td>.33</td>
</tr>
<tr>
<td>High Public Opinion Support, Low Salience</td>
<td>.43 (.03)</td>
<td>.38</td>
</tr>
<tr>
<td><strong>Difference (High to Low)</strong></td>
<td><strong>.04 (.05)</strong></td>
<td><strong>-.06</strong></td>
</tr>
<tr>
<td>Low Public Opinion Support, High Salience</td>
<td>.79 (.03)</td>
<td>.73</td>
</tr>
<tr>
<td>High Public Opinion Support, High Salience</td>
<td>.41 (.04)</td>
<td>.34</td>
</tr>
<tr>
<td><strong>Difference (High to Low)</strong></td>
<td><strong>-.39 (.06)</strong></td>
<td><strong>-.51</strong></td>
</tr>
</tbody>
</table>

Note: The table shows predicted probabilities via simulation (King, Tomz, and Whittenberg 2000) based upon the model estimates in column 2 of Table 2 for the specification using all available instances of agenda items with poll coverage (n=2,448). Average Public Opinion Support is 54.8% in favor of the issue with high and low points set at two standard deviations above and below the mean (low=15.8, high=93.8). Average salience is 5.1 based upon the number of editorials in the New York Times for valid cases while the high and low values of two standard deviations above and below the mean is 1 to 20.3).
Table 3. Replication of Gilens (2012) and Reanalysis Based Upon Issues On/Off the National Policy Agenda

<table>
<thead>
<tr>
<th></th>
<th>Gilens (2012) Replication</th>
<th>Gilens On/Off Agenda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Opinion Logit Coefficient (Standard Error)</td>
<td>.41 ** (.05)</td>
<td>.31 ** (.06)</td>
</tr>
<tr>
<td>On National Policy Agenda (Standard Error)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Public Opinion X On Agenda (Standard Error)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Intercept</td>
<td>-.85</td>
<td>-.86</td>
</tr>
<tr>
<td>Probability if 20% Favor</td>
<td>.19</td>
<td>.22</td>
</tr>
<tr>
<td>Probability if 80% Favor</td>
<td>.43</td>
<td>.39</td>
</tr>
<tr>
<td>Relative Ratio (80%/20%)</td>
<td>2.21</td>
<td>1.82</td>
</tr>
<tr>
<td>N</td>
<td>1,779</td>
<td>1,520</td>
</tr>
<tr>
<td>-2*Log-Likelihood</td>
<td>2,198</td>
<td>1,877</td>
</tr>
<tr>
<td>Likelihood Ratio X2</td>
<td>60.16</td>
<td>28.28</td>
</tr>
<tr>
<td>Significance</td>
<td>p &lt; .01</td>
<td>p &lt; .01</td>
</tr>
</tbody>
</table>

Note: As reported in Gilens (2012, 76), cases consist of survey questions about proposed policy changes asked between 1981 and 2002. The dependent variable is policy outcome coded 1 if the proposed policy change took place within four years of the survey date and 0 if it did not. Predictors are the logits of the percentage of respondents favoring the proposed policy change (column 1) or the predicted percentage of respondents at a given income percentile favoring the proposed policy change (columns 3, 4, 6 and 7). Columns 2-7 analyze a subset of Gilens for the years of 1981 to 2000 that overlap with Binder's (2003) data on the national policy agenda. The subset for "10th Inc." refers to the 1981 to 2000 time period for poor respondents at the 10th percentile of income. The "90th Inc." is the same but for wealth respondents at the 90th percentile. ** p < .01; * p < .05 (two-tailed).
Figure 1. Three Possible Scenarios Depicting How Survey Questions Might Overlap with the National Policy Agenda

Note: The oval in light blue shading in the middle depicts the “National Policy Agenda,” which represents all issues being considered for governmental action at a given moment in time. The oval in light yellow shading marked “Minimal Overlap” shows a hypothetical situation in which the polling agenda rarely corresponds to the national policy agenda. The gray oval with diagonal strips shows “Some Overlap” with the policy agenda, but survey questions also cover many other issues not being considered by lawmakers. The oval in dark blue illustrates a scenario of “Substantial Overlap” when poll questions often cover the policy agenda.
Figure 2. Size of National Policy Agenda and Proportion with Polling Data, Congressional Sessions from 1947-1948 (80th Congress) to 1999-2000 (106th Congress)
Figure 3. Democratic Responsiveness from 1947 to 2000 in U.S. Congressional Sessions from the 80th to the 106th

Note: The overall responsiveness series (black line) uses the entire national policy agenda for each congressional session in the calculation of opinion and policy correspondence (i.e., policy changes when a majority support change or does not change when majorities oppose). A second measure, depicted in the light (orange) line, shows agreement between policy and opinion for only the issues from the national policy agenda with polling data. In both cases, the national policy agenda data come from Binder (2003) and polling data come from the Roper Center for Public Opinion Research.