

**Is Public Opinion Stable?
Resolving the Micro-Macro Disconnect in Studies of Public Opinion**

by

James N. Druckman
druckman@northwestern.edu
Department of Political Science
Northwestern University

Thomas J. Leeper
leeper@u.northwestern.edu
Department of Political Science
Northwestern University

March 22, 2012

Abstract

Public opinion matters, both as a central element of democratic theory and as a substantive foundation for political representation. The origins and nature of public opinion have long attracted the attention of social science. Yet, a number of questions remain and among the more perplexing asks whether – and under what conditions – public opinion is stable. We argue that an answer to this debate depends in large part on whether one looks at aggregations of individual opinions (i.e., macro public opinion) or at the individual opinions themselves (i.e., micro public opinion). In this paper we explore the macro/micro divide and offer a framework for when opinions are likely to be stable or volatile that reflects both the content of the political environment and the nature of individuals’ opinions. With reference to research on public opinion dynamics surrounding the USA Patriot Act, we discuss the implications of opinion stability for interpreting public opinion and for understanding the normative implications of public preferences.

Public opinion matters. In theory, it serves as the foundation from which democratic governmental action should spring (e.g., Dahl 1971). In practice, elected officials tend to respond to public opinion (e.g., Shapiro 2011); moreover, given its role in determining election outcomes, it is not surprising that politicians invest massive resources into tracking and attempting to influence opinions. The origins and nature of public opinion has attracted the attention of scholars since the emergence of the modern social sciences. Yet, a number of questions remain and among the more perplexing: *is public opinion stable?* If the public's preferences exhibit stability, it suggests that sentiments expressed at one point in time will largely sustain and thus may reflect clearly held beliefs. On the other hand, if instability dominates, it *could* suggest less stock should be placed into the meaning of the public's preferences at a given point in time.¹

It turns out that whether one concludes stability rather than instability to a significant extent depends on whether one looks to macro trends in aggregated opinions (e.g., the percentage of the public that support increased defense spending) or micro level individual opinions (e.g., an individual's specific preference for defense spending). Consider the conclusions from two highly influential books from 1992, the first focused on macro opinion and the second on micro:

- “Our data reveal a remarkable degree of stability in America's collective policy preferences” (Page and Shapiro 1992: 45).
- “Opinion statements vary randomly across repeated interviews of the same people; entirely trivial changes in questionnaire constructions... can easily produce [large] shifts in aggregate opinion” (Zaller 1992: 28).

¹ We emphasize that it *could* suggest less meaning of public opinion because another possibility is that instability stems from systematic and thoughtful opinion changes in response to meaningful events.

These conclusions are not time-bound, as similar conclusions can be found in very recent macro (e.g., Erikson et al. 2002: 235, Jones and Baumgartner 2005, Wood and Vedlitz 2007: 553, Baumgartner et al. 2009: 175-178, Soroka and Wlezien 2010: 69) and micro work (e.g., Achen and Bartels 2004, Hibbs 2008, Chong and Druckman 2010: 665).

In this paper, we explore the sources of the micro-instability and macro-stability divide. We begin in the next section with a general discussion of micro versus macro studies, via extended example of public opinion surrounding the USA Patriot Act. We then offer a framework for understanding when opinions should be stable or volatile. Next we build further by identifying three sources of the micro-macro disconnect that we believe explain why the type of data employed yield such distinct conclusions. We end the paper by discussing the implications of our argument for both understanding public opinion and for interpreting what (in)stability implies from a normative perspective and why this matters for those who report on and read about public opinion in the news. One of our more intriguing conclusions is that stability – often presumed to indicate “higher quality” opinions – may bring with it some undesirable features.

Micro-Macro Perspectives

The divide between micro and macro perspectives in the social sciences is certainly nothing new, as discussion can be found in works by prominent scholars James, Laswell, Lewin, and Schelling (see Eulau 1996 for historical discussion; also see Schelling 1978). In his aptly titled autobiography – *Micro-Macro Dilemmas in Political Science* – Heinz Eulau (1996) explains, “The fancy terms ‘micro’ and ‘macro’ have come to mean large and small or individual and aggregate or part and whole... Once micro and macro had been attached to persons or groups... It was only a small step to insist on ‘bridging’ the micro-macro gap” (37-38). While this

gap pervades a large range of topics, it is most interesting – for our purposes – when it comes to public opinion and communication.

It is important to be clear what we mean by micro and macro public opinion data. For micro data, the unit of analysis is an individual (e.g., survey respondent). Typically, the researcher’s interest lies in knowing what opinion(s) that person holds, why, and with what effects. For example, one may be interested in knowing whether an individual respondent opposes or supports the Patriot Act, which is a piece of legislation enacted by the U.S. Congress and signed by President George W. Bush shortly after the September 11th, 2001 terrorist attacks that increases the powers of law enforcement agencies to monitor communications, records, and financial transactions in an effort to identify terror threats.² It is instructive to understand why the individual holds an opinion – does it reflect deeply held values, knowledge about the issue, social experiences, and/or media coverage? – *and* whether the opinion shapes subsequent behavior (e.g., is the individual willing to sign a petition in support of the Act?). Much of this work employs surveys that measure an individual’s support for the Act, asking, for example:

- The Patriot Act was enacted in the weeks after September 11, 2001, to strengthen law enforcement powers and technology. What do you think—do you oppose or support the Patriot Act?

Researchers then correlate answers to this opinion measure (typically measured on a 7 point scale ranging from 1 = oppose strongly to 7 = support strongly) with other variables such as demographic features (e.g., gender, income), partisan attributes (e.g., partisanship), experiences (e.g., media exposure), values (e.g., importance of law and order), etc. Very early

² The Act contains a number of other elements such as redefining terrorism so as to include domestic incidents. The actual name of the Act is the “USA PATRIOT Act” which stands for: “Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism.” The Patriot Act is a good issue on which to focus in so far as it resembles many other issues in being periodically salient, touching on both economic and social dimensions.

survey research reported responsive instability, meaning individuals' opinions measured at one point in time changed at a later point in time (e.g., Converse 1964).

More recent work has built on this by employing experiments that randomly assign respondents to different types of questions (for discussion, see Druckman et al. 2011). For example, some respondents randomly receive the following (civil liberties) version of the Patriot Act question:

- The Patriot Act was enacted in the weeks after September 11, 2001, to strengthen law enforcement powers and technology. Under the Patriot Act, the government has access to citizens' confidential information from telephone and e-mail communications. As a result, it has sparked numerous controversies and been criticized for weakening the protection of citizens' civil liberties. What do you think—do you oppose or support the Patriot Act?

Others receive a distinct (terrorism) version that asks:

- The Patriot Act was enacted in the weeks after September 11, 2001, to strengthen law enforcement powers and technology. Under the Patriot Act, the government has more resources for counterterrorism, surveillance, border protection, and other security policies. As a result, it enables security to identify terrorist plots on American soil and to prevent attacks before they occur. What do you think—do you oppose or support the Patriot Act?

Much of the work that takes this (experimental) approach finds that respondents' opinions, on average, widely differ depending on which version of the question they receive. Many conclude that this suggests opinions are not grounded and quite malleable based on whatever rhetoric is most recently heard by respondents (e.g., Zaller 1992, Bartels 2003). In many ways, these conclusions offered an explanation for responsible instability by showing that instability stems, at least in part, from alternative rhetoric found in discourse or in the survey questions.

Other relevant work has tracked individuals' opinions over time by asking the same respondents the same question (e.g., about the Patriot Act) several weeks apart. The modal

finding here is that opinions change and any effects (e.g., from a certain type of question at one point in time) quickly decay (e.g., Chong and Druckman 2010). For example, when individuals receive the terrorism version of the Patriot Act question, they likely become more supportive of the Act. Yet, for the modal individual, that support quickly dissipates and, in fact, may flip if the individual later receives the civil liberty frame. Chong and Druckman (2010) state, “when competing messages are separated by days or weeks, most individuals give disproportionate weight to the most recent communication because previous effects decay over time” (663).

Whether all this instability suggests that citizens’ opinions are baseless and of little value is a topic of debate – for example, it could be that the changes reflect reasonable movements rather than ineptitude (for discussion, see Druckman 2001, Sniderman and Theriault 2004). Regardless, for us the most important point is that on many issues, when studied at the micro-level, individuals’ political attitudes appear unstable.³ Such dynamics led Best and McDermott (2007) to conclude that “...reported opinions on ... the USA Patriot Act – vary greatly due to simple variations in question wording, content, and response options” (1).

This view of public opinion as fickle is somewhat puzzling because it appears contradictory to the conclusion of macro level studies. For macro studies, the unit of analysis is not the individual per se but rather it is a given issue or a given point in time with the focus most often being on the overall percentage of individuals who support or oppose a perspective – such as the percentage that support the Patriot Act or the frequency of each response at a given point in time.⁴ Much macro level work studies whether government policies respond to aggregate trends in opinions (e.g., does the government increase Patriot Act spending when support

³ As is typical, we treat the terms “opinion” and “attitude” as interchangeable.

⁴ The measure could be the same as the previously presented individual level measure where support is construed as any score above 4. Alternatively, a measure could report percentages for each of the 7 response options (e.g., percentage who registered a 6) or use a distinct set of response options (e.g., “support,” “not sure,” “oppose”).

increases, over time?), and conversely, does public opinion react to governmental actions (e.g., does support wane once spending increases?) or other events (e.g., effect on support of a terrorist threat) (e.g., Page and Shapiro 1992, Erikson et al. 2002, Soroka and Wlezien 2010).⁵ As intimated, studies of macro opinions towards the Patriot Act report tremendous stability, contradicting the micro findings: a 2011 report from Pew states, “Public views of the Patriot Act, whose renewal is being debated by Congress, have changed little since the Bush administration.” What this means, for example, is that the level of support for the Act at one point in time is near equivalent to support at a later time. These findings of micro instability and macro stability are not unique to the Patriot Act but rather extend across countless issues and times (see, e.g., Bartels 2003). Mortensen (2006) explains, “Studies convincingly demonstrate that aggregated voter opinions are rather sticky... [yet there are] random fluctuations at the individual level” (18).

This contradiction emerges even though macro opinion is simply the aggregation of micro attitudes: macro support for the Patriot Act comes from simply counting the number of individual respondents that expressed support. What explains this striking micro/macro instability/stability inconsistency?⁶ Unraveling the ostensible micro-macro inconsistency is more than a pedantic exercise (see Jackson and Kollman 2011: 507). Politicians often turn to aggregate opinion for guidance (e.g., Erickson et al. 2002) and media outlets typically report on aggregate trends (e.g., Traugott and Lavrakas 2008). To interpret these trends and understand how one may go about altering them, it is essential to contemplate their micro-foundations – do these trends reflect reasoned judgments or are they of less substantive meaning?

⁵ There is a related debate on whether political actors respond to issue specific opinions (e.g., specific trends regarding the Patriot Act) or more generalized ideological trends (e.g., liberalism versus conservatism) (Druckman and Jacobs 2006).

⁶ The micro-macro inconsistency is a prototypical example of an ecological inference problem where stable trends at the macro level belie the underlying volatility at the micro level.

What Generates Stability?

To be clear, we define stability as when opinions sustain or do not change when measured at two or more points in time. We posit two factors as being critical for *instability*. The first factor is a weak attitude. Attitudes can range from nonexistent (i.e., a “non-attitude”) to weak to extremely strong (Fazio 2007). For example, an individual may be asked for her opinion on a policy about which she has never heard (e.g., regulation of vending machines) or be asked about an issue on which he is highly committed to his position (e.g., abortion). As attitudes become stronger, they also exhibit greater stability – indeed, by some definitions, a strong attitude is (tautologically) one that persists and resists change (Krosnick and Smith 1994, Miller and Peterson 2004, Visser et al. 2006). Thus, change occurs mostly when attitudes are weak.

Importantly, attitude strength is a multi-dimensional concept. The strength of a given attitude depends on the nature of the attitude (e.g., more extreme opinions tend to be stronger), the attitude’s structure (e.g., more accessible attitudes tend to be stronger), and the process by which one forms their attitudes (e.g., attitudes based on more elaborative thinking tend to be stronger, as are attitudes formed in an “online” fashion; e.g., Petty and Krosnick 1995, Bizer et al. 2004: 215-216). Attitudes also tend to be stronger when they are deemed personally important (e.g., Krosnick 1988) or are viewed as more certain (e.g., Visser et al. 2006). Finally, attitude strength grows when individuals think about their attitudes or have attitude relevant experiences (e.g., Krosnick and Smith 1994, Visser et al. 2006, Glasman and Albarracín 2006: 78), including being exposed repeatedly to the same information such as due to continuous media coverage (e.g., Chong and Druckman 2011). Attitude strength lies on a continuum from weak to strong; however, in what follows, we simplify for presentational purposes by focusing on either strong or weak attitudes.

Second, for an attitude to change there typically must be a stimulus that induces the change; this might include an ostensibly persuasive argument (even if not consciously processed), a world event, a novel experience, and/or re-thinking of a viewpoint. That said, most micro studies attend to stimuli contained in communications such as the above experimental example. These are meant to mimic the types of rhetoric found outside of the study context (communications that may influence macro trends). Macro movements and hence instability could be driven by other factors such as world events and experiences. Since we seek to explain the micro instability and macro stability (rather than vice versa), we limit our following discussion to communications. We also attend to stimuli that are potentially persuasive – that is, information that has sufficient credibility to induce change at least under some conditions.

[Table 1 About Here]

Our attitude strength X stimulus framework maps into four model scenarios that appear in Table 1. We suspect that, all else constant, stability at both the micro and macro level should occur in three of the four situations. In the first two cases – where there are no stimuli, we expect stability because there are no experiences, such as encountering information that would stimulate a reconsideration of an attitude. We expect instability when opinions are weak and there is a stimulus (assuming the stimulus is sufficiently credible to induce change). As explained, weak attitudes are relatively open to change and thus a stimulus may induce such modifications (assuming the stimulus pushes the opinion in a direction counter to the prior opinion).

Perhaps the most interesting case is when an individual possesses a strong opinion and encounters a potentially persuasive stimulus (that counters the present opinion, such as a terrorism argument to an individual who opposes the Patriot Act). When this occurs, we expect, all else constant, that the individual will reject the stimulus and cling to the extant opinion. This

occurs because individuals with strong attitudes tend to engage in motivated reasoning whereby they seek out information that confirms priors (i.e., a confirmation bias), view evidence consistent with prior opinions as stronger (i.e., a prior attitude effect), and spend more time counter-arguing and dismissing evidence inconsistent with prior opinions, regardless of objective accuracy (i.e., a disconfirmation bias) (e.g., Lodge and Taber 2000, Kunda 2001, Bartels 2002, Redlawsk 2002, Rudolph 2006, Gerber and Huber 2009, 2010, Goren et al. 2009, Druckman et al. 2010, Druckman and Bolsen 2011, Lavine et al. n.d.).

When individuals possess strong attitudes, they are likely to “come inescapably to mind, whether consciously recognized or not, and for better or worse these feelings guide subsequent thought” (Lodge and Taber 2008: 33). For example, when people receive new information about George W. Bush, those with strong feelings interpret that information in light of their existing opinions about Bush. Thus, a pro-Bush voter might interpret information suggesting that Bush misled voters about the Iraq war as either false or as evidence of strong leadership in a time of crisis, rather than an accurate indication of incompetence or deception. Such voters maintain their support of Bush and may even become even more supportive. Alternatively, an individual strongly opposed to the Patriot Act will reject arguments about its utility for combating terrorism, even if the argument is otherwise objectively sound. This leads to the fairly ironic dynamic that those with less developed, weaker attitudes “are processing information more ‘objectively’ than those with [stronger] attitudes” (Houston and Fazio 1989: 64).⁷

The Micro-Macro Disconnect

⁷ Motivated reasoning is less likely to occur if the individuals are motivated to process information in an accurate way. Taber and Lodge (2006) argue, however, that in most political contexts, this is unlikely (although see Druckman n.d.).

As explained, micro level public opinion work tends to suggest instability while macro level work suggests stability. With our strength X stimulus framework in mind, we can now turn to explaining three possible sources of the inconsistent micro-macro findings.

Measurement Error

Measurement error can generate instability on individual survey responses that, when randomly distributed in the sample, cancel out at the macro level. Measurement error occurs when a survey response departs from its “true value” – for example, on a 7-point Patriot Act support scale, from strongly opposed to strongly supported, it could be that respondent’s true attitude should land him/her around 5.5. If the survey were administered twice, the respondent may once report a “5” and the other time a “6.”

Measurement error can stem from characteristics of the respondent (e.g., he/she is not paying attention or did not understand the question), the interviewer (e.g., misreading the question including the response options), the questionnaire (e.g., the order in which questions are asked), or other factors such the context of data collection. At the micro level, measurement features can cause a respondent to offer different answers at distinct points in time, leading to ostensible instability over-time. Yet, at the macro level, random measurement error cancels out since the randomness means roughly the same number of respondents who move in one direction (e.g., 40% offer a lower level of support at time 1 than time 2) will move in the other (e.g., 40% offer a higher level support at time 1 than time 2).⁸ Thus, stability exists in aggregation (e.g., even though 40% increased support between times 1 and 2 and 40% decreased support, the

⁸ There also is the possibility of nonrandom measurement error such as a question written a biased fashion (e.g., “Most support the Patriot Act since it helps prevent terrorism. What do you think...?”) This type of measurement error would only generate micro instability if it was corrected or changed in over-time surveys.

averages at each time are the same) (see Page and Shapiro 1992, Mortensen 2006: 18; also see Bartels 1996).

Measurement error can generate micro instability in any of our four Table 1 scenarios. At first glance, it would seem that such error is less likely among individuals with strong opinions since they tend to cling to those attitudes. Yet, since measurement error is not about substantive changes, it means susceptibility is not contingent on attitude strength. Krosnick and Abelson (1992) state the “relatively simple hypothesis that these effects [i.e., responsive instability] are greater in the case of weaker attitudes has clearly been disconfirmed” (193; see also Visser et al. 2006 for a more recent discussion).

Ansolabehere et al. (2008) offer compelling evidence that once corrections for measurement error are put in place (e.g., using multiple measures and taking averages), the result is micro level (and macro level) stability. They explain “the low correlations of individuals’ issue preferences over time... are easily reconciled with a model in which there is a high degree of measurement error and a high degree of stability in preferences” (216).

Sample Inconsistencies

Most discussions about the sample focus on the selection of respondents in the survey (e.g., are the respondents representative of the target population?). Yet, sampling also includes the selection of issues and times. Researchers aim to draw inferences about opinions on the universe of issues across time, however, they have no choice but to focus on select issues at particular times. We suspect that, at least some of the micro-macro discrepancy, can be traced to distinct foci in the issues examined and the timing of the studies. Aggregate studies almost always rely on publicly available survey data from credible polling organizations (e.g., Gallup, American National Election Study); consequently, these studies focus on public opinion toward

the issues that were asked in these surveys. This turns out to be a very small and likely non-random sample of the possible universe of issues (e.g., all issues the government addresses over a term). Burstein (2003) explains that “the entire set of issues studied may be so small that it is unrepresentative of the set of all issues and an inadequate basis for generalization... what should be emphasized is how our capacity to generalize is limited by the narrowness of the range of issues studied” (31, 36). He also states that “it’s no secret that public opinion data don’t exist for most policies legislatures consider” (68).

Importantly, the issues that tend to be included in public surveys are those that are more salient – it is sensible that survey organizations prefer to gauge issues salient to the public. Druckman and Jacobs (2006: 470) explain that there is a pressure “to collect policy opinion data on issues seen as important by the public.” In his survey of extant work, Burstein (2003) shows that these issues include social welfare, taxes, defense, etc: issues that have the potential to directly affect citizens.⁹

What this means is that macro-studies may be biased toward issues on which citizens possess stronger opinions since the issues are more likely to be of personal importance (a key dimension of attitudes strength; see Visser et al. 2006).¹⁰ Also these issues are more likely to be covered in the media, thereby providing citizens with more repeated exposures to the issue which, as mentioned, enhance attitude strength (e.g., Chong and Druckman 2011). In Figure 1, we chart the number of questions asked regarding the Patriot Act (by all survey organizations contained in the iPoll database) along with media coverage of the Patriot Act (as captured by

⁹ Another point is that publicly available surveys tend to focus on policy generalities rather than specific topics – such as questions about support for welfare rather a particular welfare provision. It may be that attitudes toward such general areas are more stable.

¹⁰ While we focus on the issue and time, we also note that another possible reason for the micro-macro gap is the nature of the samples used (e.g., more student samples in many micro studies). Recent work though that relies on representative sample survey experiments leads us to put less emphasis on this possibility.

non-editorial mentions of “Patriot Act” in Section A of *The New York Times*). The number of survey questions in the field (gray bars) peaks when media coverage increases. Survey questions are not asked consistently across the period and none were in the field during initial October, 2001 authorization and few were asked between the July, 2005 and May, 2011 reauthorizations. Effectively, polls that are responsive to media coverage select upon opinions that are strong and salient; this nonrandom selection of times for assessing public opinion problematizes the assessment of stability. Indeed, as mentioned, access to information tends to generate stronger attitudes which in turn lead to stability.

[Figure 1 About Here]

In short, the strong opinions on issues that are polled during times of increased activity lead to stability. This sharply contrasts with the foci of many micro-level studies that typically choose issues for the exact opposite reason. These studies search for issues on which prior opinions are weak since that may allow for change (the focus of many of these studies) and/or issues which have been absent from recent media coverage. Chong and Druckman (2010: 667) echo many other micro studies in stating that they selected issues because “opinions on these issues are liable to change, which allows us to test hypotheses [about opinion change].” Studies also opt to select issues that “that receive scant attention outside of the experiment itself (e.g., de Vreese 2004, Chong and Druckman 2010)” (Druckman et al. n.d: 3). Examples from the micro-studies that demonstrate volatility include attitudes about a particular ballot proposition (Albertson and Lawrence 2009), an election involving a new candidate about whom individuals have scant prior opinions (e.g., Gerber et al. 2011), regulation of hog farms (Tewksbury et al. 2000), opinions about urban sprawl with respondents not directly affected (e.g., Chong and Druckman 2007), or attitudes toward abstract and impersonal subjects such as people’s trust in

institutions (e.g., de Vreese 2004, Mutz and Reeves 2005; for discussion, see McGraw and Dolan 2007).

In sum, varying levels of stability in macro and micro-level studies of opinion may stem, in part, from differences in the issues explored and the timing of that exploration (also see Wood and Vedlitz 2007). The root of the disconnect stems from incomparable samples of issues and times.

Empirical Example

To see how opinion strength can generate distinct patterns of stability, consider Chong and Druckman (2010)'s survey experiment. Their December, 2009 study involved a nationally representative sample of about 1,300 individuals and focused on opinions about the Patriot Act. Their specific dependent measure was the same as that presented above where respondents reported their support for the Patriot Act a 7-point scale with higher scores indicating increased support. They measured opinions at two points in time (t1 and t2), separated by about 10 days.

There are two critical features of their study. First, the study employed versions of the aforementioned terrorism (“Pro”) and civil liberties (“Con”) frames. (The frames, however, were presented as a series of statements rather than in question wording as in the example above.) Respondents received different mixes of these frames at t1 and t2. Second, Chong and Druckman randomly assigned respondents to conditions that induced them to form strong opinions at t1, or induced them to form weaker opinions at t2.¹¹ We will not go into the details of the specific opinion strength manipulations, but suffice it to say that Chong and Druckman offer evidence

¹¹ They also included conditions with no attitude strength inducement but we do not discuss those results here.

that their inducements (which are commonly used in psychology) did in fact generate stronger or weaker t1 opinions about the Patriot Act.¹²

[Figures 2 and 3 About Here]

Figure 2 reports the average opinions at t1 and t2 for the weak attitude conditions, for various frame combinations.¹³ Figure 2a includes conditions that did not include a frame at t2 while Figure 2b is from conditions with a t2 frame. There clearly is substantial over-time volatility with opinions at t1 reflecting the direction of whatever frame the respondents received but then either moving toward the control group at t2 (i.e., the t1 “No,” t2 “No” condition) when no t2 frame is offered or flipping to reflect the direction of the t2 frame when a t2 frame is offered. There is no stability whatsoever.

Figure 3, which contains analogous results but this time for those induced to form strong opinions, presents an entirely different portrait. Here we see tremendous stability when no t2 frame is offered (Figure 3a). Moreover, Figure 3b shows similar stability even in the presence of a contrary t2 frame; individuals with strong attitudes reject it and cling to their t1 opinion (which was affected by the t1 frame). This latter dynamic reflects motivated reasoning whereby respondents counter-argue and reject contrary evidence.¹⁴

¹² They did this by inducing on-line or memory-based processing (see Chong and Druckman 2010 for details).

¹³ We recognize that it may be ironic that we are using aggregated averages to make a case about micro opinion, which we have argued differ from macro opinion because it is not aggregated. Nonetheless, we take this approach for presentational clarity, noting that if we instead looked at individual opinion change across periods, the story would be the same. What is critical here are the conditions that generate different types of micro level opinions – and these conditions accentuate common differences in micro and macro samples given when measurement typically takes place (as explained).

¹⁴ Direct evidence on this comes from a question that asked strong attitude individuals to rate the “effectiveness” of the t1 and t2 statements in terms “providing information and/or making an argument about the Patriot Act,” on a 7-point scale with higher scores indicating increased effectiveness. Participants who received the Con frame at t1 reported an average effectiveness score of 4.84 (1.65; 83) whereas those who received the same frame at t2 (after having received the Pro frame at t1) registered a significantly lower 4.40 (1.44; 83) average ($t_{164} = 1.83$, $p \leq .05$ for a one-tailed test). Similarly, the average time 1 Pro rating is 5.05 (1.54; 83), but only 4.29 (1.59; 83) at t2 ($t_{164} = 3.13$, $p \leq .01$ for a one-tailed test). In short, those induced to form strong attitude at t1 downgraded the t2 frames that contradicted their t1 priors.

These results have been replicated with various other issues including attitudes about urban sprawl, a state funded casino, new scientific technologies, and health care (see, e.g., Druckman 2010, Druckman and Bolsen 2011, Druckman et al. 2011, Druckman and Leeper n.d.). The implication is that if macro studies do indeed focus on issues at times when individuals develop strong attitudes, stability is to be expected, whereas instability would be the norm for micro studies to the extent they focus on less developed issues.

While the results of Chong and Druckman's experiment reveal a source of the macro-micro disconnect, it ironically cannot explain the discrepancy in the case of the Patriot Act given it focuses on one issue during one time period. Moreover, there are undoubtedly issues on which most possess weak opinions that nonetheless lead to differing macro and micro dynamics (putting measurement error aside). We suspect these issues as well as the aforementioned Patriot Act inconsistency stem from the final possible cause which we now discuss.

Ecological Validity of the Rhetorical Environment

One possible reason why micro instability on a given issue in a certain time period would exhibit macro stability is that the instability cancels out. For example, consider the weak attitude conditions in the Patriot Act experiment. In that case, proportional numbers of individuals were exposed to the Pro and Con frames at each point in time. Thus there was considerable movement but since the numbers were largely equivalent (due to the assignment to conditions), the consequence is a cancelling out. Indeed, if we merge all the weak attitude scenarios, it would appear as if there was aggregate macro stability as the overall t1 mean is 4.40 (std. dev. = 1.79; N = 575) and the t2 mean is 4.38 (std. dev. = 1.70; N = 575).

This suggests that one possibility is that stability stems from a macro environment that includes a broad array of contrasting information. Such environments would differ from micro

studies that often expose individuals to information pushing them in a single direction (e.g., just a Pro frame). When it comes to the Patriot Act, Chong and Druckman (2011: 257) report that the civil liberties and terrorism frames appeared with nearly identical frequencies in the *New York Times*, from 2001 through 2005. When this occurs, the competing frames often cancel out, leading to no effect on opinions.¹⁵ The results in Figures 2 and 3 support this contention. Notice for both the strong and weak attitude conditions, when individuals receive Pro *and* Con frames at t1, their opinions are unmoved relative to the control and consequently sustain until t2.¹⁶

The larger point here is that micro work may be lacking in ecological validity, which refers to the extent to which studies approximate situations in “real-life” situations. If most stimuli in the world and thus in macro studies involve competing information streams, but micro-studies explore asymmetric information, the disconnect may simply reflect a lack of ecological validity in micro studies (particularly experimental studies on which we have focused). Chong and Druckman (2011) make this exact point upon their discovery that across many issues, media coverage incorporates competing information: “Because news stories typically contain *more* than one or two effective frames, readers rarely encounter a scenario – common in experimental studies – in which they are restricted to a single monolithic frame of the issue. Thus, framing effects that occur outside of controlled experimental settings are not well understood” (255; also see Sniderman and Theriault 2004: 141, Kinder 2007).

The implication is that stability is the norm, due to competing communications, and that micro studies overstate instability due to scant attention to competition. This raises the question

¹⁵ Once challenge to this, however, is if individuals select information consistent with their prior opinions and thus segment such that those with pro (con) prior opinions only view pro (con) information. Indeed, this is a manifestation of motivated reasoning. In this scenario those, we would see micro-stability rather than instability (see Druckman et al. 2011).

¹⁶ Alternatively, stimuli in the world may be less likely to change in macro contexts (for example, see Baumgartner et al. 2009: 175–178).

of how these competing communications work. On the one hand, Sniderman and Theriault (2004) suggest that “political debate, being exposed to opposing sides, tightens the linkages of mass belief systems and increases the constraint between basic principles and specific issue choices” (158). In other words, individuals exposed to competing messages largely ignore them and fall back on their well formed values. On the other hand, Zaller (1996) suggests that “the mass media routinely carry competing political messages [and] each message... has its effects, but the effects tend to mutually canceling in ways that produce the illusion of modest impact” (20). That is, citizens do not rely on well-formed, reasoned values but rather sway back and forth reflecting the messages.¹⁷

What Does This All Mean?

We began by asking whether public opinion is stable. Our answer may be less than satisfying insofar as it amounts to: “it depends.” More important than this bottom line, however, is our identification of *when* we may expect stability. We predict opinions will be stable on issues and at times where individuals possess strong opinions or, putting measurement error aside, when there is a lack of persuasive stimuli in the environment. We argued that micro level studies significantly over-state the malleability of the mass public due to the focus on issues on which individuals possess weak attitudes. On the flip side, macro studies likely over-state the extent of stability due to relying on publicly available data that over-represent issues that receive substantial media coverage and on which individuals possess strong opinions.

Our argument offers a fairly clear blueprint for steps that can be taken to vitiate the micro-macro gap:

¹⁷ Based on an explicit testing of these alternative accounts, Chong and Druckman (2007) report “The participants in our experiments were open to argumentation on both sides of the issue and did not merely revert to standing positions” (651).

- All possible efforts should be put forth to reduce measurement error in surveys.

While some approaches to doing so – such as using multiple items as Ansolabehere et al. (2008) suggest – come with costs (e.g., the cost of survey time, demand effects), there also are more straightforward steps that can be taken to minimize error (for general discussion, see Krosnick and Presser 2010).

- Studies should consciously assess the representativeness of the issues and times on which they focus.

A first step along these lines is to carry out a more systematic appraisal of the exact issues and times that have been the focus in micro and macro study. Then, going forward, studies should attempt incorporate multiple issues (that are likely to having varying distributions of attitude strength), or at least recognize, the consequences of not doing so. While most instruments include a vast array of respondent demographic questions, they rarely incorporate attitude strength questions that could provide insight into expected stability or instability.¹⁸ Krosnick and Abelson (1992) made a plea, twenty years ago, for the regular inclusion of attitude strength measures in public opinion surveys, but thus far, it has gone largely unheeded.

- More attention should be paid to issues of ecological validity.

The intellectual evolution of many political communication studies led to an overemphasis on documenting the possibility of effects (for discussion, see Druckman et al. 2009). This is no longer a critical goal and scholars should invest more in identifying the nature of the rhetorical environment that surrounds an issue. They then should seek to theorize and emulate the effects of that environment. We recognize this brings with it a host of challenges, not

¹⁸ Page and Shapiro (1992: 285-320) explore potential volatility differences among various demographic subgroups (based, for example, on age, income, education, region, race). They conclude that “the bulk of the evidence indicates that different groups do not tend to *change* their preferences very often in very different ways” (318; italics in original). Our point, however, is that much more relevant subgroup differences may be found if one focuses on attitudes strength instead of conventional demographics.

only in terms of requiring more intensive content analyses, but also it introduces the likelihood of fewer statistically significant findings, which then face a publication bias. This, in turn, brings with it a larger concern about the publication process and the biases that result from a narrow focus on p-values (for discussion, see Gerber et al. 2010).

To observers of politics, our argument suggests that they - and all of us - should be cautious about inferring much at all from survey evidence that mass opinions either have changed or remained stable on a given issue. Politicians frequently legitimize their stances with reference to public opinion, particularly when majorities are on their side or opinions seem to be shifting their way. Whether a bare majority or a few-percentage-point shift is meaningful requires an understanding of survey practice - question wording, sampling, and so on - but also some sense of why opinions might behave the way they do. To fully grasp the latter, reports of mass opinion need to be contextualized with information about the environment from which opinions were measured and some sense of the strength of those opinions. Unfortunately, present reporting rarely mentions either. Observers and reporters should aim to present richer narratives to make sense of public opinion.

A final point concerns the normative implications of our argument. Strong opinions and stability are often seen as signs of an engaged and thought-out citizenry, and thus, coveted attributes. Attitude strength promotes constraint (Federico and Schneider 2007) and engagement (e.g., Leeper 2012). Yet, strong attitudes also lead to motivated reasoning that can cause individuals to resist consideration of relevant alternative perspectives. At the extreme, such individuals can be close-mindedly dogmatic, which might be as problematic as extremely labile preferences. Theorists should not presume that well developed and thought-out opinions always trump fleeting ones when it comes to opinion “quality” (see Druckman n.d.).

As we mentioned, micro-macro gaps pervade the social sciences and we have focused on just one example. In so doing, however, we affirm Eulau's (1996) concluding hope that the field of communication and public opinion is an area where the gap could be bridged: "the new 'discipline' of Communication represents the fulfillment of the dream for... Interdisciplinary behavioral science that can address the gap" (359).

References

- Achen, Christopher H., and Larry M. Bartels. 2004. "Musical Chairs: Pocketbook Voting and the Limits of Democratic Accountability." Presented at the Annual Meeting of the American Political Science Association, Chicago, IL.
- Albertson, Bethany L., and Adria Lawrence. 2009. "After the Credits Roll: The Long-Term Effects of Educational Television on Public Knowledge and Attitudes." *American Politics Research* 37(2): 275-300.
- Ansolabehere, Stephen D., Jonathan Rodden, and James M. Jr. Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102(2): 215-232.
- Bartels, Larry M. 1996. "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40(1): 194-230.
- Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24(2): 117-150.
- Bartels, Larry M. 2003. "Democracy With Attitudes." In Michael Bruce MacKuen, and George Rabinowitz, eds., *Electoral Democracy*. Ann Arbor: The University of Michigan Press.
- Baumgartner, Frank R., Jeffrey M. Berry, Marie Hojnacki, David C. Kimball, Beth L. Leech. 2009. *Lobbying and Policy Change: Who Wins, Who Loses, and Why*. Chicago: University Of Chicago Press.
- Best, Samuel J., and Monika L. McDermott. 2007. "Measuring Opinions vs. Non-opinions." *The Forum* 5.
- Bizer, George Y., Jon A. Krosnick, Allyson L. Holbrook, S. Christian Wheeler, Derek D. Rucker, and Richard E. Petty. 2004. "The Impact of Personality on Cognitive, Behavioral, and Affective Political Processes." *Journal of Personality* 72: 995-1027.
- Burstein, Paul. 2003. "The Impact of Public Opinion on Public Policy: A Review and an Agenda." *Political Research Quarterly* 56: 29-40.
- Chong, Dennis, and James N. Druckman. 2007. "Framing Public Opinion in Competitive Democracies." *American Political Science Review* 101(4): 637-655.
- Chong, Dennis, and James N. Druckman. 2010. "Dynamic Public Opinion: Communication Effects over Time." *American Political Science Review* 104(4): 663-680.

- Chong, Dennis, and James N. Druckman. 2011. "Public-Elite Interactions." In Robert Y. Shapiro, and Larry R. Jacobs, eds., *The Oxford Handbook of the American Public Opinion and the Media*, New York: Oxford University Press.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David Apter. New York: Free Press.
- Dahl, Robert A. 1971. *Polyarchy*. Chicago: University of Chicago Press.
- de Vreese, Claes H. 2004. "The Effects of Strategic News on Political Cynicism, Issue Evaluations, and Policy Support: A Two-Wave Experiment." *Mass Communication and Society* 7(2): 191-214.
- Druckman, James N. 2001. "On the Limits of Framing Effects: Who Can Frame?" *Journal of Politics* 63: 1041-1066.
- Druckman, James N. 2010. "Competing Frames in a Political Campaign." In Brian F. Schaffner and Patrick J. Sellers, eds., *Winning with Words*. New York: Routledge.
- Druckman, James N. N.d. "The Politics of Motivation," *Critical Review*. Forthcoming.
- Druckman, James N., and Toby Bolsen. 2011. "Framing, Motivated Reasoning, and Opinions About Emergent Technologies." *Journal of Communication* 61(4): 659-688.
- Druckman, James N., Jordan Fein, and Thomas J. Leeper. 2011. "Framing and Biased Information Search." Presented at the Annual Meeting of the American Political Science Association Annual Meeting, Seattle, WA.
- Druckman, James N., Cari Lynn Hennessy, Kristi St. Charles, and Jonathan Weber. 2010. "Competing Rhetoric Over Time." *The Journal of Politics* 72: 136-148.
- Druckman, James N., and Lawrence R. Jacobs. 2006. "Lumpers and Splitters: The Public Opinion Information That Politicians Collect and Use." *Public Opinion Quarterly* 70(4): 453-476.
- Druckman, James N., James H. Kuklinski, and Lee Sigelman. 2009. "The Unmet Potential of Interdisciplinary Research." *Political Behavior* 31: 485-510.
- Druckman, James N., and Thomas J. Leeper. N.d. "Learning More from Political Communication Experiments: The Importance of Pretreatment Effects." *American Journal of Political Science*. Forthcoming.
- Erikson, Robert S., Michael MacKuen, and James A. Stimson. 2002. *The Macro Polity*. New York: Cambridge University Press.

- Eulau, Heinz. 1996. *Micro-Macro Dilemmas in Political Science: Personal Pathways Through Complexity*. Norman, OK: University of Oklahoma Press.
- Fazio, Russell H. 2007. "Attitudes as Object-Evaluation Associations of Varying Strength." *Social Cognition* 25(5): 603-637.
- Federico, Christopher M., and M. C. Schneider. 2007. "Political Expertise and the Use of Ideology: Moderating Effects of Evaluative Motivation." *Public Opinion Quarterly* 71(2): 221-252.
- Gerber, Alan S., James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2011. "How Large and Long-Lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Large Scale Randomized Experiment." *American Political Science Review* 105(1): 135-150.
- Gerber, Alan S., and Gregory A. Huber. 2009. "Partisanship and Economic Behavior: Do Partisan Differences in Economic Forecasts Predict Real Economic Behavior?" *American Political Science Review* 103: 407-426.
- Gerber, Alan S., and Gregory A. Huber. 2010. "Partisanship, Political Control, and Economic Assessments." *American Journal of Political Science* 54: 153-173.
- Gerber, Alan S., Neil Malhotra, Conor M. Dowling, and David Doherty. 2010. "Publication Bias in Two Political Behavior Literatures." *American Politics Research* 38: 591-613.
- Glasman, L. R., and Albarracín, Dolores. 2006. "Forming Attitudes that Predict Future Behavior: A Meta-analysis of the Attitude-Behavior Relation." *Psychological Bulletin* 132: 778-822.
- Goren, Paul, Christopher M Federico, and Miki Caul Kittilson. 2009. "Source Cues, Partisan Identities, and Political Value Expression." *American Journal of Political Science* 53(4): 805-820.
- Hibbs, Jr., Douglas A. 2008. "Implications of the 'Bread and Peace' Model for the 2008 Presidential Election." *Public Choice* 137: 1-10.
- Houston, David A. and Russell H. Fazio. 1989. "Biased Processing As A Function Of Attitude Accessibility: Making Objective Judgments Subjectively." *Social Cognition* 7(1): 51-66.
- Jackson, John E., and Ken Kollman. 2011. "Connecting Micro- and Macropartisanship." *Political Analysis* 19: 503-518.
- Jones, Bryan D., and Frank R. Baumgartner. 2005. "A Model of Choice for Public Policy." *Journal of Public Administration Research and Theory* 15: 325-51.
- Kinder, Donald R. 2007. "Curmudgeonly Advice." *Journal of Communication* 57(1): 155-162.

- Krosnick, Jon A., and Robert P. Abelson. 1992. "The Case for Measuring Attitude Strength in Surveys." In Judith M. Tanur, ed., *Questions About Questions: Inquires into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation.
- Krosnick, Jon A., and Stanley Presser. 2010. "Question and Questionnaire Design." In Peter V. Marsden, and James D. Wright, eds., *Handbook of Survey Research 2nd Edition*. Bingley, United Kingdom: Emerald.
- Krosnick, Jon A., and Wendy A. Smith. 1994. "Attitude Strength." In V. S. Ramachandran, ed., *Encyclopedia of Human Behavior*. San Diego, CA: Academic Press.
- Kunda, Ziva. 1999. *Social Cognition: Making Sense of People*. Boston: The MIT Press.
- Lavine, Howard, Christopher Johnston, and Marco Steenbergen. N.d. *The Ambivalent Partisan: How Critical Loyalty Promotes Democracy*. Oxford University Press.
- Leeper, Thomas J. 2012. "Information Choice and Opinion Change Over Time." Presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL.
- Lodge, Milton, and Charles S. Taber. 2000. "Three Steps Toward a Theory of Motivated Political Reasoning." In Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin, ed.s., *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*. New York: Cambridge University Press.
- Lodge, Milton, and Charles S. Taber. 2008. *The Rationalizing Voter: Unconscious Thought in Political Information Processing*. Stony Brook University: Unpublished Manuscript.
- McGraw, Kathleen M., and Thomas Dolan. 2007. "Personifying the State: Consequences for Attitude Formation." *Political Psychology* 28: 299-328.
- Miller, Joanne M., and David A. M. Peterson. 2004. "Theoretical and Empirical Implications of Attitude Strength." *Journal of Politics* 66(3): 847-867.
- Mortensen, Peter B. 2006. *The Impact of Public Opinion Public Policy*. Ph.D. Dissertation, Aarhus University.
- Mutz, Diana C., and Byron Reeves. 2005. "The New Videomalaise: Effects of Televised Incivility on Political Trust." *American Political Science* 99(1): 1-15.
- Page, Benjamin I., and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago, IL: University Of Chicago Press.
- Petty, Richard E., and Jon A. Krosnick, eds. 1995. *Attitude Strength: Antecedents and Consequences*. New York: Psychology Press.

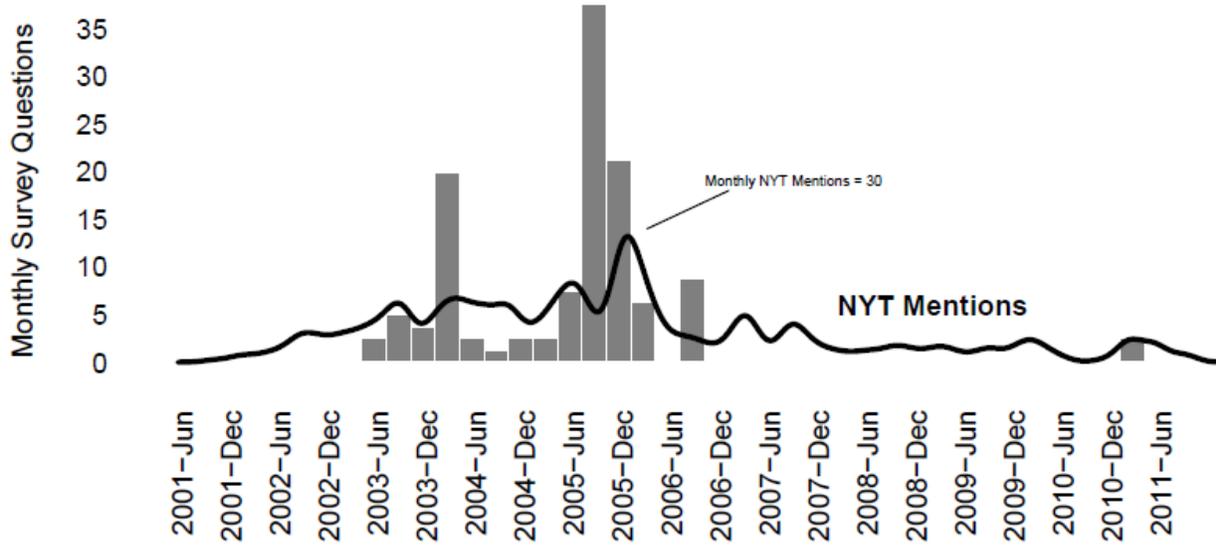
- Pew Research Center. 2011. "Public Remains Divided Over the Patriot Act." Available at <http://pewresearch.org/pubs/1893/poll-patriot-act-renewal>.
- Redlawsk, David P. 2002. "Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making." *Journal of Politics* 64(4): 1021–1044.
- Rudolph, Thomas J. 2006. "Triangulating Political Responsibility: The Motivated Formation of Responsibility Judgments." *Political Psychology* 27(1): 99–122.
- Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: Norton.
- Shapiro, Robert Y. 2011. "Public Opinion and Democracy." *Public Opinion Quarterly*. Forthcoming.
- Sniderman, Paul M., and Sean M. Theriault. 2004. "The Structure of Political Argument and the Logic of Issue Framing." In William E. Saris and Paul M. Sniderman, eds., *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*. Princeton, NJ: Princeton University Press.
- Soroka, Stuart N., and Christopher Wlezien. 2009. *Degrees of Democracy*. New York: Cambridge University Press.
- Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50(3): 755–769.
- Tewksbury, David, Jennifer Jones, Matthew W. Peske, Ashlea Raymond, and William Vig. 2000. "The Interaction of News and Advocate Frames." *Journalism and Mass Communication Quarterly* 77(4): 804-829.
- Traugott, Michael W., and Paul J. Lavrakas. 2008. *The Voter's Guide to Election Polls*. Lanham, MD: Rowman & Littlefield Publishers.
- Visser, Penny S., George Y. Bizer, and Jon A. Krosnick. 2006. "Exploring the Latent Structure of Strength-Related Attitude Attributes." *Advances in Experimental Social Psychology* 38(6): 1-67.
- Wood, B. Dan, and Arnold Vedlitz. 2007. "Definition, Information Processing, the Politics and of Global Warming." *American Journal of Political Science* 51(3): 552-568.
- Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zaller, John. 1996. "The Myth of Massive Media Impact Revived: New Support for a Discredited Idea" In Diana Mutz, Richard Brody, and Paul Sniderman, eds., *Political Persuasion and Attitude Change*. Ann Arbor: University of Michigan Press.

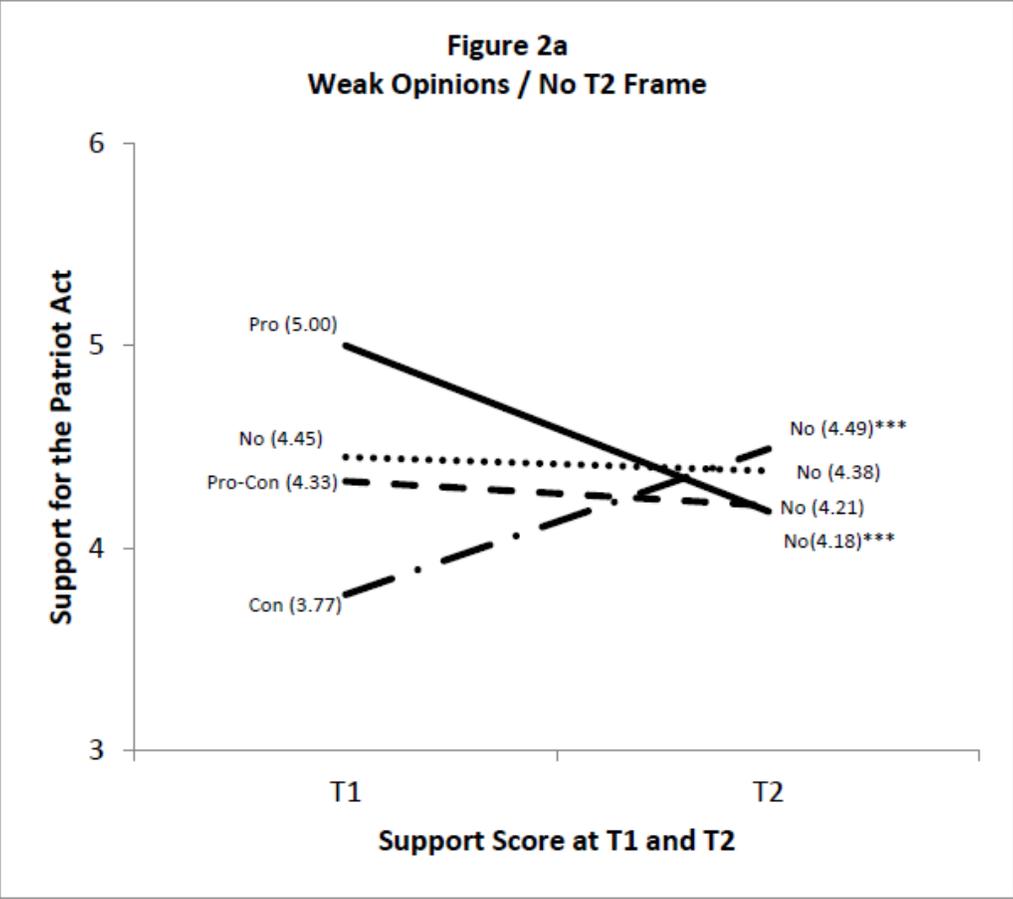
Table 1. Conditions for Opinion Stability

Attitude Strength	Stimuli	Stability?
Weak	No Stimuli	Yes
Strong	No Stimuli	Yes
Weak	Stimuli	No
Strong	Stimuli	Yes

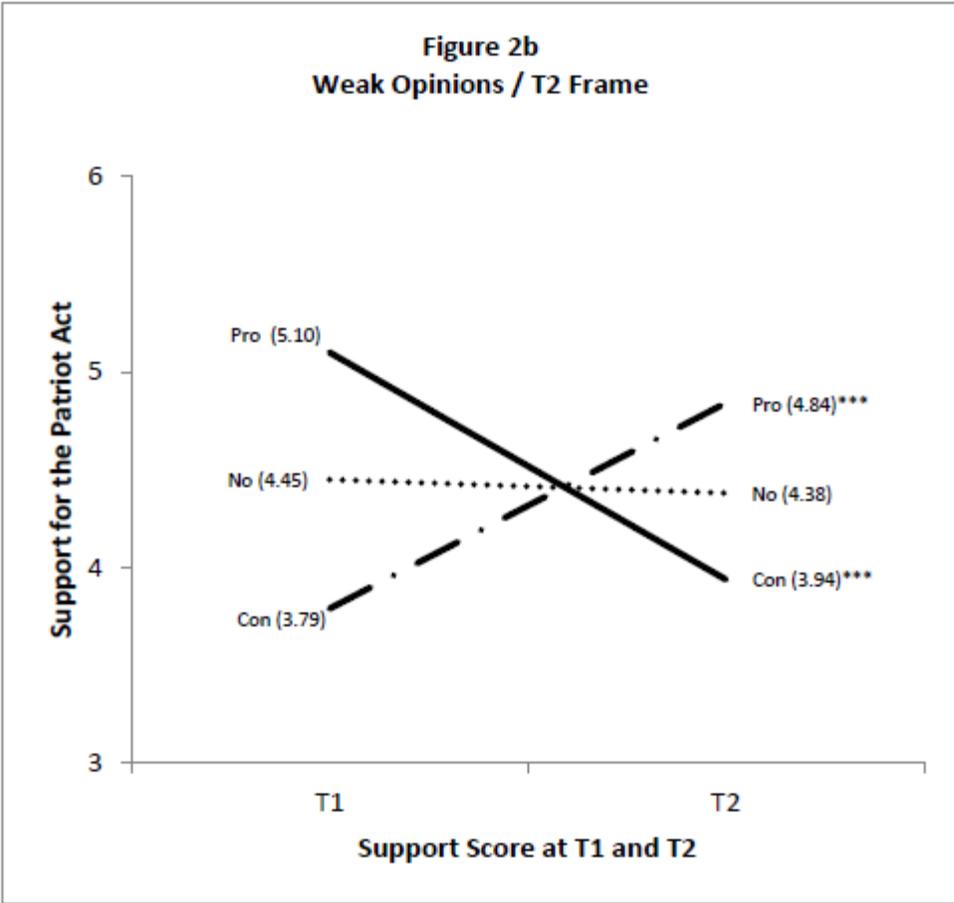
Figure 1. Patriot Act Survey Questions and Mentions in *The New York Times*



Note: Histogram is number survey questions fielded each month (grouped into three-month intervals for clarity of presentation). Dark black line is a kernel-smoothed density plot of non-editorial mentions of “patriot act” in Section A of *The New York Times* over the same period.

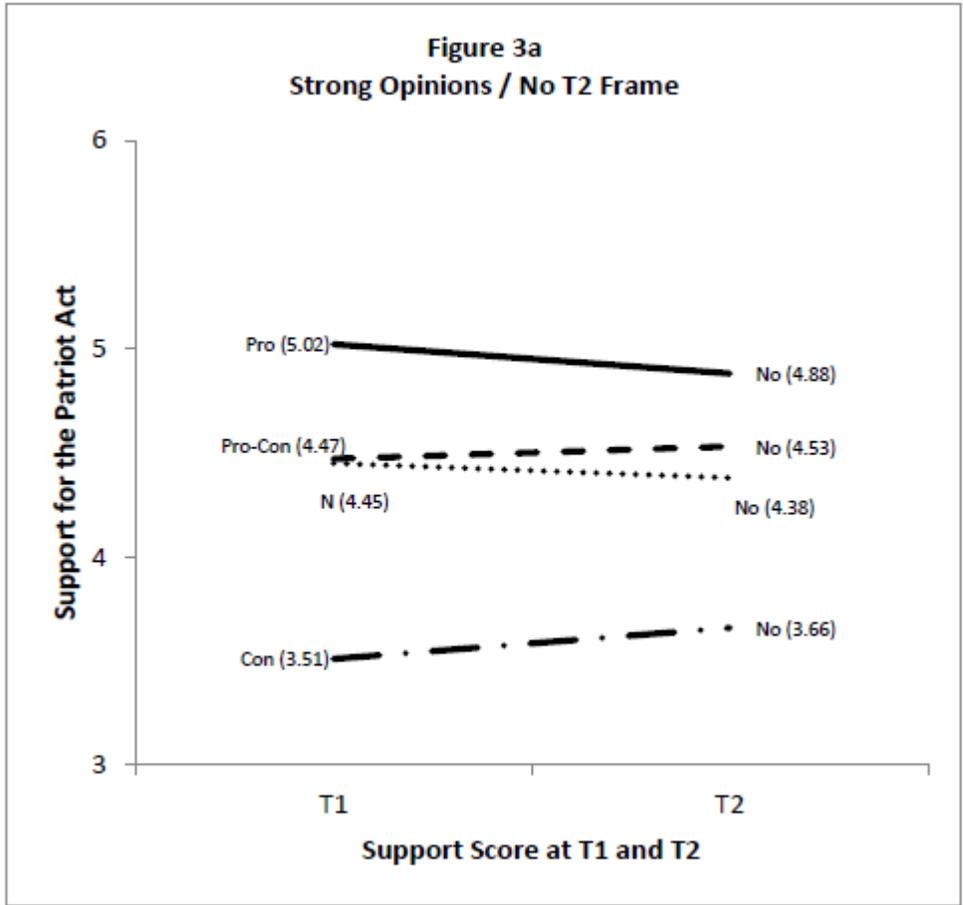


Source: Chong and Druckman (2010) , used here with permission of the authors.
 *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$ for one-tailed tests (for changes between t1 and t2).

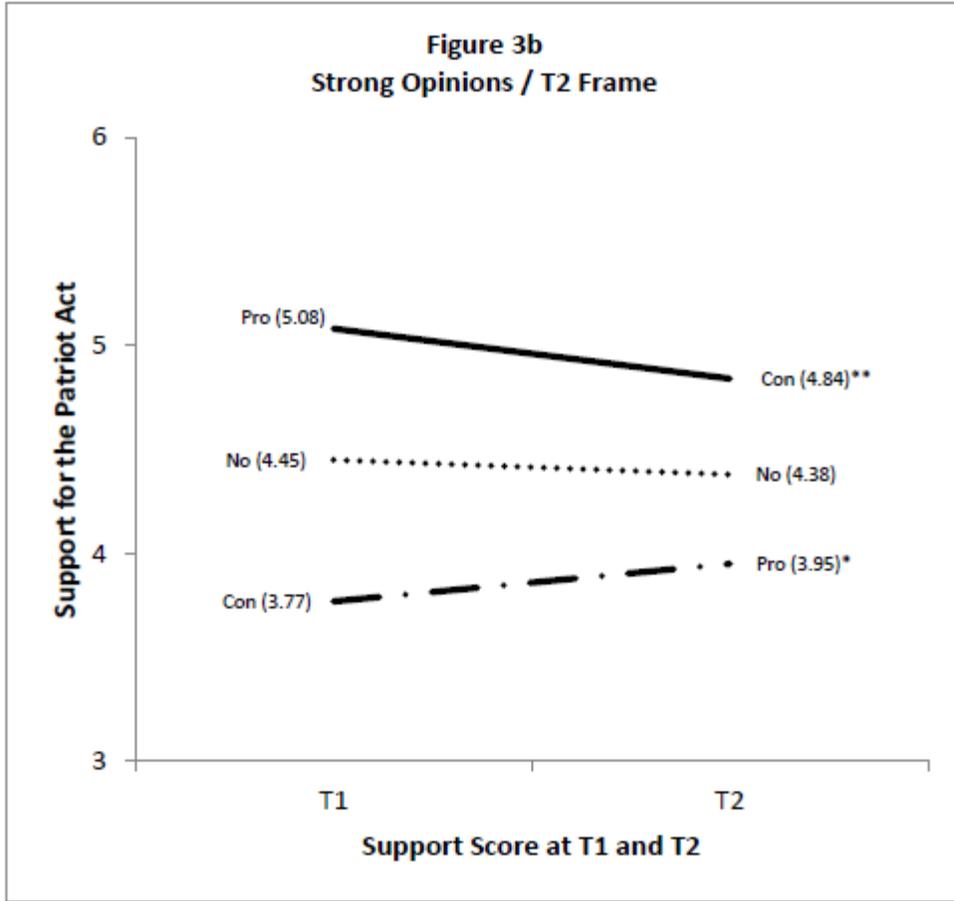


Source: Chong and Druckman (2010) , used here with permission of the authors.

*** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$ for one-tailed tests (for changes between t1 and t2).



Source: Chong and Druckman (2010) , used here with permission of the authors.
 *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$ for one-tailed tests (for changes between t1 and t2).



Source: Chong and Druckman (2010), used here with permission of the authors.
 *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$ for one-tailed tests (for changes between t1 and t2).