Tracking the Development of Media Frames within and across Policy Issues

Amber E. Boydstun, University of California, Davis*
Dallas Card, Carnegie Mellon University
Justin H. Gross, University of North Carolina, Chapel Hill
Philip Resnik, University of Maryland, College Park
Noah A. Smith, Carnegie Mellon University

August 19, 2014

Abstract

Framing is a central concept in political communication and a powerful political tool. Thus, it is hugely important to understand: a) what frames are used to define specific issues, b) what general patterns are evidenced by the evolution of frames over time, and c) how frames diffuse, or spread, across policy areas and venues. These tasks also pose a serious challenge, thanks to the volume of text data, the dynamic nature of language, and the variance in applicable frames across issues (e.g., the 'innocence' frame of the death penalty debate is not used in discussing smoking bans). We describe a project that advances framing research methodology in two ways. First, we present a unified coding scheme for content analysis across issues, whereby issue-specific frames (e.g., innocence) are nested within high-level dimensions (or frame *types*) that cross-cut issues (e.g., fairness). We call this the "Policy Frames Codebook" with an eye toward doing for frame categorization across issues what the Policy Agendas Codebook has done for issue categorization across agendas. Second, we validate the policy frames coding scheme by applying it to news coverage of three issues—smoking, immigration, and same-sex marriage—in the United States over a twenty-two year period. This pilot dataset is the first cut from a larger data collection effort that will eventually span news coverage of five policy issues over several decades. Using this data, our long-term aim is to identify and assess empirical patterns in which frames tend to get selected across policy debates in the United States, how frames within policy debates tend to evolve, and the conditions under which frames spread from one issue to the next and/or across policy venues (e.g., between states, or from the media to Congress). Toward this aim, we employ strategies heavily informed by existing work in natural language processing, but tailored to the specific needs and professional sensibilities of framing and policy scholars.

1 Introduction

Framing—portraying a policy issue from one perspective to the necessary exclusion of alternative perspectives—is a central concept in political communication (see Schaffner and Sellers, 2010, Introduction for a nice overview, and the remainder of the edited volume for several illustrations). It is widely accepted that framing can have a significant influence on public attitudes toward important policy issues (e.g., Chong and Druckman, 2007; Nelson et al., 1997) and on the application of policy directly (e.g., Baumgartner et al., 2008). Understanding, for a given issue, what frames are used by politicians, the media, and the voting public to define that issue—and how these frames evolve

*Corresponding author: aboydstun@gmail.com

and diffuse—is therefore a crucial task for advancing our understanding of politics. It is also an enormous challenge, due to the dynamic and creative nature of language and the growing volume of data in which frames appear and develop over time. As engagement by citizens in the political discourse broadens via the widespread adoption of blogging, commenting, and other social media, scientific study of the political world requires reliable analysis of how issues are framed, ideally in real time. Yet the process by which a political scientist or communication scholar identifies the catalogue of frames in a political discourse about a particular issue (**frame discovery**) is complex and labor-intensive; so is the secondary process of coding instances of framing in text (**frame analysis**) in order to reveal patterns in frame usage.

Moreover, the very definition of framing has been notoriously slippery. The most widely employed definition among current researchers in political communication is provided by Robert Entman: "Framing essentially involves *selection* and *salience*. To frame is to *select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation* for the item described" (Entman, 1993, p. 52, emphasis by the author). Beyond the challenge of gaining consensus on a conceptual definition, the matter of operationalizing the notion of issue framing presents its own set of difficulties. As Matthes and Kohring point out in a recent article aimed at improving the reliability and validity of content analytic measures of media frames, "a frame is a quite abstract variable that is hard to identify and hard to code in content analysis" (Matthes and Kohring, 2008, p. 258). As a result, measurement is challenging; in the words of another researcher, "it is extremely difficult to neutralize the impact of the researcher in framing research" (Van Gorp, 2005, p. 503, cited in Matthes and Kohring, 2008).

Despite the important nuances of different conceptual definitions of framing, all these definitions treat language as central, making the tools of natural language processing especially important. Whether the aim is to study frames defined as memes, dimensions of debate, or specific arguments for or against something, framing according to all of these characterizations is amenable to identification and analysis through signals in language use, ranging from simple lexical clues to word clusters to choices of syntactic structure. As we will discuss below, natural language processing promises to help us meet some of the challenges inherent in operationalizing and measuring frames as the volume and diversity of relevant data continues to increase.

Additionally, beyond looking at frames associated with specific issues, scholars and citizens alike face the challenge of being able to trace how frames are deployed across multiple policy debates. We can examine how the death penalty is framed in terms of "an eye for an eye" vs. "cruel and unusual punishment," for example, but understanding the mechanisms and effects of framing as broader phenomena requires better methods that allow us to trace overarching tropes connected with party or ideological position *across* a variety of issue debates. This would include, for example,using divergent metaphors of the nation as a family, such as "strict father" and "nurturant parent" family types (Lakoff, 2004); invoking a value that is widely embraced, yet whose meaning is deeply contested (e.g. "freedom"), across a spectrum of settings (Lakoff, 2006); or framing various policies in terms of potential losses or potential gains (Kahneman and Tversky, 1979). Here, too, language is key to identifying and analyzing frames that cross-cut issues, and again, therefore, natural language processing is a key research tool.

Whether we are looking at politicians' communications, at traditional news media, or particularly at social media, the availability of online data creates an unprecedented opportunity to track framing in near-real-time, and to understand it as an evolutionary process across time, issues, and communication venues (e.g., different media formats, different states, different political institutions). Thus, understanding framing as a general phenomenon requires large-scale text data analysis well beyond what has been accomplished by expert manual annotation alone.

Our three-year interdisciplinary project, funded by the U.S. National Science Foundation,[1] takes first steps toward data-driven, expert-informed, computational modeling of framing that augments the ability of experts to discover frames and analyze their use in textual discourse. The project's goals are to:

1. Develop algorithms for automated frame **analysis**, leveraging high-level domain knowledge and issue-specific knowledge from expert political scientists, annotated examples, and unannotated text and contextual metadata linked to the text. Concretely, we are developing algorithms to produce similar results to human coders who achieve high inter-coder reliability at identifying frames used in text passages from contemporary political discourse, relating to a range of issues.

2. Integrate insights from political science experts with statistical analysis of political discourse in order to enable frame **discovery**. Concretely, we are developing tools that speed up the work of human experts and will reveal frames that might be missed by the subjective "naked eye."

3. Apply both of these methods to better understand the **evolutionary** process behind frame development, dynamics, and diffusion for specific issues in current American politics, across multiple years and multiple traditional and social media streams.

This paper describes some of our key research activities during the project's first two years. We focus primarily on the initial development of the Policy Frames Codebook (§3), presenting an initial dataset (§4), and offering preliminary evidence from computational models designed to automatically characterize framing in text (§5).

2 Substantive Questions: Frame Selection, Evolution, and Diffusion

Political scientists, and politicians, too, have learned that issue framing can drive how policy debates unfold and how citizens and government respond. Yet we know relatively little about framing as a general phenomenon; that is, about its empirical regularities. Past studies are invaluable for understanding framing in specific issues like the death penalty, but rarely can they unlock findings that generalize across issues. Past work does, however, suggest hypotheses about general traits of framing across issues in three linked concepts: 1) *frame selection*: certain frames tend to be used more than others, contingent on the institutional venue and political/economic context (e.g., during a recession, economic frames should be pervasive); 2) *frame dynamics*: frames evolve in important ways within issue debates over time; 3) *frame diffusion*: frames spread, contagion like, across venues (e.g., for state-level initiatives like same-sex marriage, frames should spread across states in a cascade-like fashion). While we do not tackle direct answers to these questions in this preliminary paper, we discuss them here as motivation for our work.

Consider the current national and state-by-state debates about same-sex marriage. As political scientists, we have instincts about the many (and complex) factors that will shape public opinion and public policy in this shifting debate. Moreover, past issue-based studies suggest that framing may play an important role. However, we have very little general understanding of framing to bring to bear on the questions of: a) Which frames will be used in the debate (frame selection)? b) How will framing of the debate change over time within a given venue, such as national newspaper coverage (frame dynamics)? c) How will framing of the debate spread across states or across venues, such as from national newspapers to local media, or vice versa (frame diffusion)? Note that in many

[1] `http://www.ark.cs.cmu.edu/Compuframes`

other realms of political science, we do have such general knowledge. For instance, in the literature on voting, we have a wealth of empirical evidence to understand the general traits of which types of voters will vote which way, how vote preferences change over time, and the effects of things like candidate valence on vote choice. The framing literature, by contrast, is rich but much more specialized. A large-scale empirical advance is required for us to make a large-scale advance in our knowledge of how (and when) framing works.

Our central hypothesis, grounded in past research, is that issue framing exhibits many empirical regularities across policy debates. In broad strokes, agenda setting and issue framing studies suggest we should be able to take at least some of the framing traits uncovered in experimental and single-issue studies and apply them to studies of framing of multiple issues in the wild. Below, we discuss key hypotheses derived from the literature. For each, a test of the hypothesis will mark significant advances for the study of framing, regardless of whether the null hypothesis is supported or rejected. Of course, many other hypotheses could be imagined, and the framing data we are gathering will allow testing of those hypotheses, too.

2.1 Hypotheses about Frame Selection

Work on agenda setting shows how a given venue (*New York Times*, the budget committee, etc.) will systematically pay more attention to some issues than others, and that these selection patterns are shaped by institutional incentives (Baumgartner and Jones, 2010). For example, journalistic incentives should lead to a higher use of sensational frames and political frames in news coverage (e.g., Gans, 2004; Graber and Dunaway, 2014). And the incentives of both media and Congress suggest that both venues should select more loss frames than gain frames (Grimmer, 2013, Soroka, 2006). But because context matters as well as institutional incentives, we should expect to see some frames used across all venues under certain conditions. For example, we should expect to see political frames increase across the board during election periods (e.g., Iyengar et al., 2004) and a pervasive use of economic frames during periods of recession (Neuman et al., 1992). Thus, in broad terms, we expect that frames are selected as a function of the interaction between institutional incentives and context.

2.2 Hypotheses about Frame Dynamics

Thinking about the dynamics of frame selection within an issue, work on the crisis framing cycle suggests we should see a common shift from loss frames to gain frames in media coverage of a crisis as time from the crisis event elapses (Glazier and Boydstun, 2012). This dynamic trend may hold for non-crisis issues, as well; news coverage might tend to shift from loss to gain frames over time after an issue breaks in the news, driven by a human tendency, perhaps endowed by natural selection, to prioritize loss-related information first and foremost (McDermott et al., 2008).

More generally, work in nonparameterized text analysis suggests that political debate about issues indeed does not maintain a static set of dimensions over time but, rather, evolves (e.g., Quinn et al., 2010). Supporting this notion is the aforementioned work on framing in the case of the death penalty (Baumgartner et al., 2008). That study found evidence that during key periods when attention to the death penalty surged around one dominant frame, in fact that frame was reinforced by the heightened activation of several related frames in the debate, manifested by an increase in the diversity of attention across frames (rather than a decrease, as would be produced by a true frame dominance). In other words, what may look like a single frame monopolizing the discourse may in fact be a bandwagoning of frames, producing the attention cascade. Thus, we expect that frames within an issue debate evolve over time, and do so in systematic, bandwagoning ways.

2.3 Hypotheses about Frame Diffusion

Studies of tipping points, policy diffusion and related dynamic processes suggest that framing should exhibit cascading properties over time and between venues (e.g., Bikhchandani et al., 1992; Granovetter, 1978). For example, frames can spread between government and media in the case of foreign policy making (Entman, 2003). Similarly, work on social contagion of policy across state lines suggests that frames might spread, geographically, from one state to another (Pacheco, 2012). This hypothesis is especially salient for current state-level discussions of issues like same-sex marriage and legalizing marijuana. Additionally, the idea of contagion suggests that some frames might spread from one issue debate to another. For instance, the recent economic crisis might have increased the likelihood of economic frames being adopted across issues in a snowball fashion, with one possible result being increased traction of marijuana on state agendas when framed in terms of potential state revenue. Thus, we expect frame diffusion between issues and venues to change not gradually but rather in cascade fashion.

3 Measuring Framing: The Policy Frames Codebook

It is perhaps fitting that issue-framing, also known as issue-definition, has been defined in many different ways. These varied definitions have helped produce a burgeoning framing literature, but they leave us without the ability to examine patterns in framing both within and across issues over time. Here, we describe the benefits of framing schemas that cross-cut policy issues, and we introduce just such a schema: the Policy Frames Codebook. Just as the Policy Agendas Codebook[2] provides a system for categorizing topical cues across policy agendas, our Policy Frames Codebook provides a system for categorizing framing cues across policy issues.

As a key outcome of our project intended for use in the wider community, this codebook is a carefully validated resource (more on this below) that is useful in both human and automated content analysis. For those who wish to use it for conventional hand-coded content analysis, it will provide a common framework for cross-project comparison and replication, while remaining general enough to allow project-specific code development based on idiosyncrasies of individual issues and research questions about these issues. Should such researchers then wish to scale up to analysis of a larger corpus than can be efficiently handled by a small team of human coders, automated content analysis will be an option, without having to start from scratch, as the codebook is designed and validation exercises conducted with scalability in mind. In this section we briefly motivate our approach and describe the codebook's present state of development.

3.1 Issue-General and Issue-Specific Approaches to Framing

Framing research has already benefitted from certain well-established schemas that generalize across issues. Iyengar (1991), for instance, identifies episodic frames (focused on specific incidents or cases) as distinct from thematic frames (focused on larger trends or context). Such general schemas facilitate invaluable insights into high-level patterns of political communication and, most importantly, their influence on public attitudes. For example, people who consume stories about poverty that are framed episodically by focusing on unemployed individuals are more likely to blame poverty on individual failings. People who consume thematic poverty stories, focused on national unemployment rates, are more likely to blame poverty on the government or other forces beyond an individual's control (Iyengar, 1991). However, existing generalized frame schemas do not unpack the topical content of frames as second-level agenda items (in the sense of McCombs, 2002), offering very little

[2] `http://www.policyagendas.org`

information about *how* the nature of a given policy debate shifts from one substantive dimension of the issue to another.

Other frame schemas are issue specific. For example, Baumgartner and colleagues trace framing in the case of capital punishment using an extensive codebook of frames specific to that issue (the death penalty does/does not deter crime, the death penalty system is/is not subject to error, etc.; Baumgartner et al., 2008). These issue-specific schemas are wonderfully detailed, but they do not allow us to examine patterns—and test hypotheses—across issues. In the death penalty study, for instance, the authors find suggestive evidence that conceptually-linked frames can "piggyback" on one another. The rise of the "innocence" frame in the mid-1990's, for instance, was accompanied by a rise in frames related to evidence, due process, classism, and racism. These related frames likely gained attention on the coattails of the innocence frame, but then in turn helped fuel that frame's momentum and attention to the death penalty overall. From this single case, it appears that the appearance of one frame may increase the likelihood of substantively linked frames being used (Baumgartner et al., 2008). However, without a coding schema that cross-cuts issues, we have no way of testing this hypothesis or others.

3.2 The Policy Frames Codebook

Our Policy Frames Codebook is intended to provide the best of both worlds: a general system for categorizing frame cues across policy issues designed so that it can also be specialized in issue-specific ways.[3] The codebook contains fourteen categories of frame "dimensions" (plus an "other" category) that are intended to be applicable to any policy issue (abortion, immigration, foreign aid, etc.) and in any communication context (news stories, Twitter, party manifestos, legislative debates, etc.). The dimensions are listed below, followed by discussion of how we created and validated them.

1. **Economic frames**: The costs, benefits, or monetary/financial implications of the issue (to an individual, family, community, or to the economy as a whole).

2. **Capacity and resources frames**: The lack of or availability of physical, geographical, spatial, human, and financial resources, or the capacity of existing systems and resources to implement or carry out policy goals.

3. **Morality frames**: Any perspective—or policy objective or action (including proposed action)—that is compelled by religious doctrine or interpretation, duty, honor, righteousness or any other sense of ethics or social responsibility.

4. **Fairness and equality frames**: Equality or inequality with which laws, punishment, rewards, and resources are applied or distributed among individuals or groups. Also the balance between the rights or interests of one individual or group compared to another individual or group.

5. **Constitutionality and jurisprudence frames**: The constraints imposed on or freedoms granted to individuals, government, and corporations via the Constitution, Bill of Rights and other amendments, or judicial interpretation. This deals specifically with the authority of

[3]We use the term "frame cues" to acknowledge that in many cases the signals our codebook helps identify are not explicit memes, like "death panels" or "lipstick on a pig", that some scholars equate with frames, but rather second-level issue cues that may or may not be hooked to explicit memes. Thus, one benefit of our approach will be to see whether we can connect the ground-level frame cues seen across media to the explicit frame memes iconic of ideologically polarized debate. We might ask, for example, whether the prevalence of a type of frame cue in second-level agenda-setting form tends to be an antecedent or a consequence of polarized use of frame memes.

government to regulate, and the authority of individuals/corporations to act independently of government.

6. **Policy prescription and evaluation**: Particular policies proposed for addressing an identified problem, and figuring out if certain policies will work, or if existing policies are effective.

7. **Law and order, crime and justice frames**: Specific policies in practice and their enforcement, incentives, and implications. Includes stories about enforcement and interpretation of laws by individuals and law enforcement, breaking laws, loopholes, fines, sentencing and punishment. Increases or reductions in crime.

8. **Security and defense frames**: Security, threats to security, and protection of one's person, family, in-group, nation, etc. Generally an action or a call to action that can be taken to protect the welfare of a person, group, nation sometimes from a not yet manifested threat.

9. **Health and safety frames**: Healthcare access and effectiveness, illness, disease, sanitation, obesity, mental health effects, prevention of or perpetuation of gun violence, infrastructure and building safety.

10. **Quality of life frames**: The effects of a policy on individuals' wealth, mobility, access to resources, happiness, social structures, ease of day-to-day routines, quality of community life, etc.

11. **Cultural identity frames**: The social norms, trends, values and customs constituting culture(s), as they relate to a specific policy issue.

12. **Public opinion frames**: References to general social attitudes, polling and demographic information, as well as implied or actual consequences of diverging from or "getting ahead of" public opinion or polls.

13. **Political frames**: Any political considerations surrounding an issue. Issue actions or efforts or stances that are political, such as partisan filibusters, lobbyist involvement, bipartisan efforts, deal-making and vote trading, appealing to one's base, mentions of political maneuvering. Explicit statements that a policy issue is good or bad for a particular political party.

14. **External regulation and reputation frames**: The United States' external relations with another nation; the external relations of one state with another; or relations between groups. This includes trade agreements and outcomes, comparisons of policy outcomes or desired policy outcomes.

15. **Other frames**: Any frames that do not fit into the above categories.

Researchers may choose to employ only these categories as listed here, or they could also nest issue-specific frames (or arguments) within each category. For example, in the case of capital punishment, the "innocence" frame would be a frame specific to that issue but categorized under the dimension of "fairness and equality." Figure 1 and Figure 2 illustrate how morality-based issue-specific frames within immigration and same-sex marriage would be nested within the generalizable morality frame dimension. In this way, scholars can apply the Policy Frames Codebook to new content analysis projects or take existing datasets that employed issue-specific frames and categorize those frames into the dimensions provided here.

We developed these categories through several iterative rounds of inductive and deductive testing. We began by brainstorming—amongst our team and several colleagues—categories that we

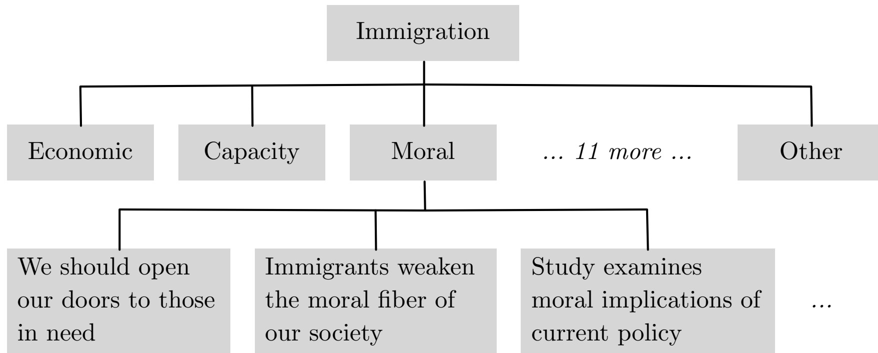


Figure 1: Illustration of hierarchical policy frames coding scheme: Immigration.

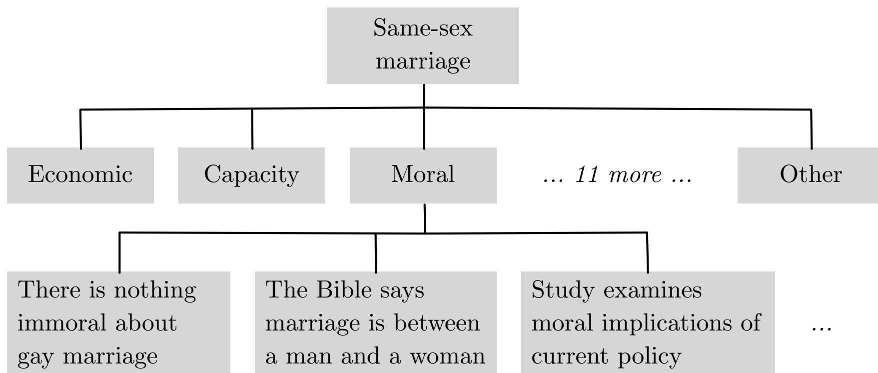


Figure 2: Illustration of hierarchical policy frames coding scheme: Same-sex marraige.

imagined would cross-cut most, if not all, policy issues, while also examining a random sampling of newspaper stories and blog posts to see which frames appeared and how we might categorize them. Then we tried applying our preliminary list of frame categories to a random sample of front-page newspaper stories, TV news stories, blog posts, and tweets covering a wide range of issues, first, to see if all (or at least most) of our frame categories could be applied and, second, to look for any frame cues not captured by our categories. Again, we revised our categorization scheme accordingly. Next, we shopped our list around, sending it to additional colleagues and presenting it at multiple conferences, both nationally and internationally (i.e., at the 20th International Conference of Europeanists), again revising our schema based on this feedback. Finally, we did another round of test coding. Throughout this testing process, we developed and revised not only our list of categories but also a codebook that defines and gives examples for each category.

Scholars can apply these categories in whatever way suits their research aims. However, we advocate coding each piece of communication (e.g., newspaper story, blog post, Congressional bill) according to the *primary* frame category used, as well as the presence of any additional frames employed. For example, a news story focused on the economic impacts of immigration but with additional discussion about the challenges of immigrants' quality of life and cultural assimilation would receive three frame dimension codes—economic, quality of life, and cultural identity—but the economic dimension would be marked as primary. In hand-coding documents for frames that appear throughout the text (i.e., as opposed to the primary frame for each document), we work at a granular level, letting coders select specific passages (paragraphs, sentences, phrases) that evoke particular frame dimensions. Specifically, we instruct coders to identify frame dimensions cued to readers in each document and then to highlight the selection of text that gives us evidence of those frame cues.

One way of validating the general codebook we have developed is to see how often it fails to capture frames that appear. In other words, what percentage of the time must we resort to the "other" category at the end of our list of frame dimensions? Looking at our manually-coded data, 1990–2012, we see that in fact nearly all stories were codeable using our fourteen substantive dimensions, with 1.5% or fewer stories receiving "other" as the primary code for any policy issue (immigration, smoking, and same-sex marriage). As a much more conservative estimate, we can look at whether *any* coder *ever* assigned "other" to *any* passage in the document; according to this extremely cautious criterion, the "other" category appears as a non-primary choice for some piece of language in 6%, 6.1%, and 9.8% of the stories on immigration, smoking, and same-sex marriage, respectively. At this stage, therefore, we view primary-frame coding as reasonably comprehensive and reliable, and we are conducting further analysis of passage-level coding to determine whether further codebook revisions are needed.

3.3 Coding for Tone

In addition to tracking the frame cues in each text, we track the tone, or position cue, of each text and encourage other scholars to do the same. We differentiate among *positive*, *negative*, and *neutral* tones, where the precise definition varies according to issue being studied and the partition according to these designations will depend on the researcher's operationalization choices. For example, in pilot testing our immigration codebook on newspaper articles, we define these tones from the perspective of immigrants and their advocates; we might have instead defined them from the perspective of supporters of greater restrictionism, likely (but not necessarily) with equivalent results.

- **Positive tone**: Immigration and immigrants' rights are portrayed in a positive light or from a generally sympathetic point of view, so that immigrant advocates and supporters of less restrictive immigration laws would be pleased to see the news article.
- **Negative tone**: Immigration and immigrants' rights are portrayed in a negative light or in a non-sympathetic manner, so that immigrant advocates and supporters of less restrictive immigration laws would be disappointed or upset to see the news article.
- **Neutral tone**: Immigration and immigrants' rights are portrayed using both positive and negative tones that balance each other out, *or* the news article does not appear to discuss the issue either positively or negatively.

One can imagine other partitions we might have drawn on the space of tones. For example, one could define the perspective to be that of undocumented immigrants, so that an article drawing attention to the positive image of authorized immigrants in contrast with the undocumented would be negative in tone; in the coding above, it would be labeled "neutral" due to the ambiguity. Furthermore, some researchers may wish to study only *explicit* framing of an issue by clear advocates for a position—we expect examples to abound within editorials and op-ed columns, blog posts, and opinion commentary on cable news or talk radio, but they may also be found in "straight news" stories via the quotes of activists, politicians, and interest group members. In our definition of the three tone types above, we allow for detection of *implicit* frames as well. The coder is simply asked to put herself in the position of an individual directly affected by the issue at hand and must essentially decide whether the article would be appealing or distressing. In this case, the aspects of the issue receiving attention from a journalist may themselves rub certain readers the wrong way despite not overtly taking sides in a conflict. (Within recent computational linguistics literature, Recasens et al. (2013) draw a related distinction between *framing bias*, which involves explicitly subjective words or phrases linked with a particular point of view, and *epistemological bias*, which involves implicit assumptions and presuppositions in ostensibly neutral writing.)

Within the structure we propose, many options are left to the judgment of the researcher, but adopting this structure ensures that such judgment will be made explicit and defended within the context of one's research program.

4 Pilot Data

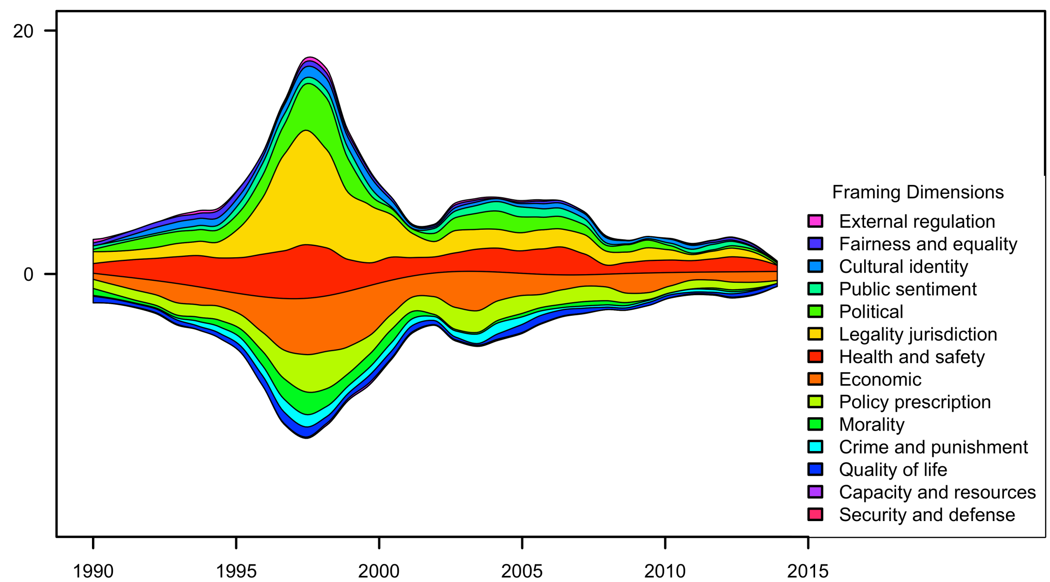
To test the validity of the Policy Frames Codebook, including our plan to track tone of coverage, we apply our approach to a random sample of news stories from national U.S. newspapers on three policy issues—smoking, immigration, and same-sex marriage—between 1990 and 2012 using the LexisNexis archives. We used iterative testing to minimize false positives and false negatives in our archive searches, arriving at keyword strings that did a relatively good job of identifying relevant stories for each issue. We retrieved this initial population of potentially relevant stories from thirteen national U.S. newspapers.[4]

As of this writing, 3,478 immigration articles, 3,341 smoking articles, and 2,683 same-sex marriage articles have been coded by trained coders, after discarding lingering irrelevant stories. Another set of 730 immigration articles have also been coded for tone in the same manner. Subsets of these have been coded by multiple coders (409 for immigration framing, 457 for tobacco, 2,167 for same-sex marriage, and 612 for immigration tone). All of these were annotated for primary frame or primary tone, as well as the presence of framing or tone cues throughout the article. We focus on these primary coding decisions here. In the case of agreement among a majority of coders, the agreed-upon primary code is retained. Going forward, for those articles where there was disagreement on the primary frame, a more experienced "master" coder will adjudicate. In applying the Policy Frames coding system across policy issues, we experienced no fewer but also no more challenges than in any other manual coding project, suggesting the tractability of our approach.

To explore the results of this pilot data, we provide below a series of figures. Our goals in this section are twofold. First, we establish the validity of our unified framing coding scheme by seeing whether our results across two issues pass muster with our understanding of events in the history of each policy debate. Second, we take a first look at the patterns in framing across issues and in the patterns of tone coverage of immigration over time (although in the latter case our N is too small to support interpretation).

We begin with Figure 3 (top portion), which shows as a stream graph the total count of our sampled smoking stories over time (at the month level), broken out by the primary frame cue each story contained. We see a surge in attention to smoking in the late 1990s, which appears to coincide with an increased discussion of legal frame cues, likely surrounding the surge in court cases related to lung cancer during that period. Figure 3 (bottom portion) plots the total number of stories coded as containing cues for each framing dimension, whether primary or not; note the difference in scale of the counts (y-axis), though the proportions are similar. In addition to the highly intuitive correspondence between the coded data and the known historical trend, the value of our coding approach is supported by the interesting variation in frame dimensions: the dimension most readily associated with smoking—health—does not come close to dominating the discussion.

[4] *Atlanta Journal-Constitution*, *Daily News (New York)*, *Denver Post*, *New York Times*, *Palm Beach Post*, *Philadelphia Inquirer*, *San Jose Mercury News*, *St. Louis Post-Dispatch*, *St. Paul Pioneer Press*, *Tampa Bay Times*, *The Herald-Sun (Durham)*, *USA Today*, and *Washington Post.*



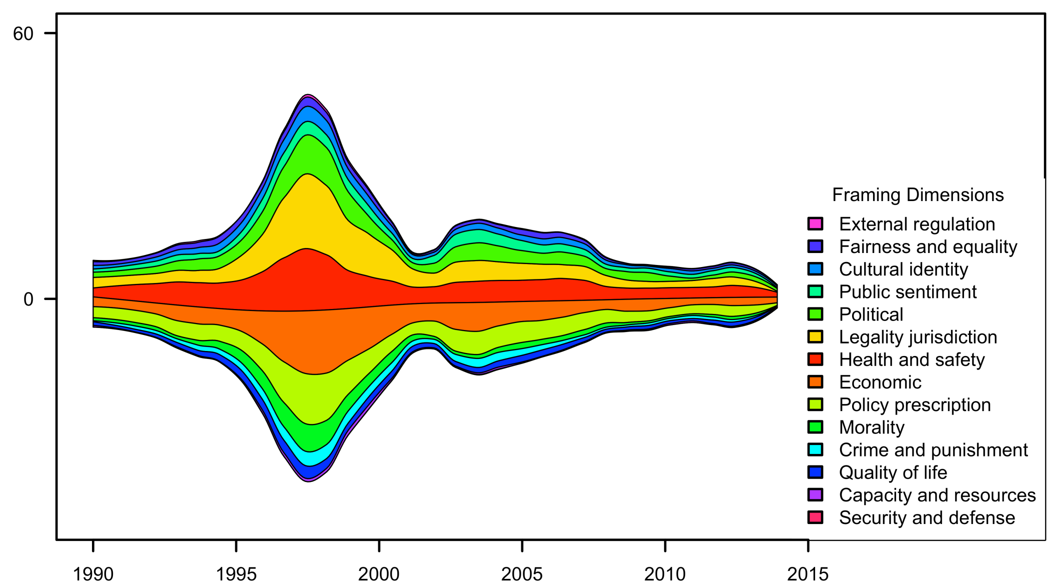
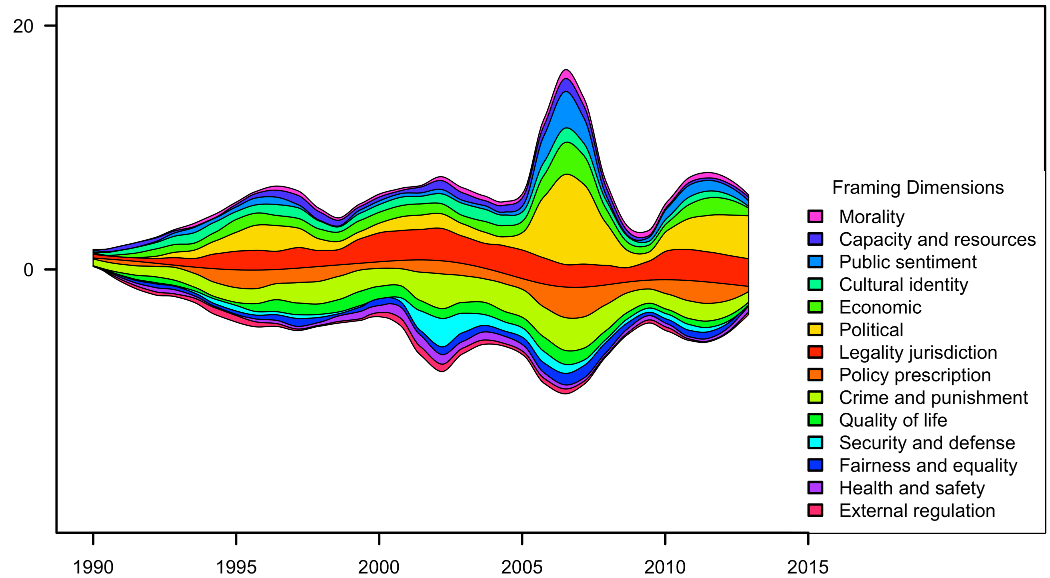


Figure 3: Frames used in news coverage of **smoking**, 1990–2012. The graph above shows the total count of stories in our sample over time (at the month level), partitioned by the *primary frame* selected by coders. The graph below shows the total count of stories containing each framing dimension (not necessarily as a primary frame).



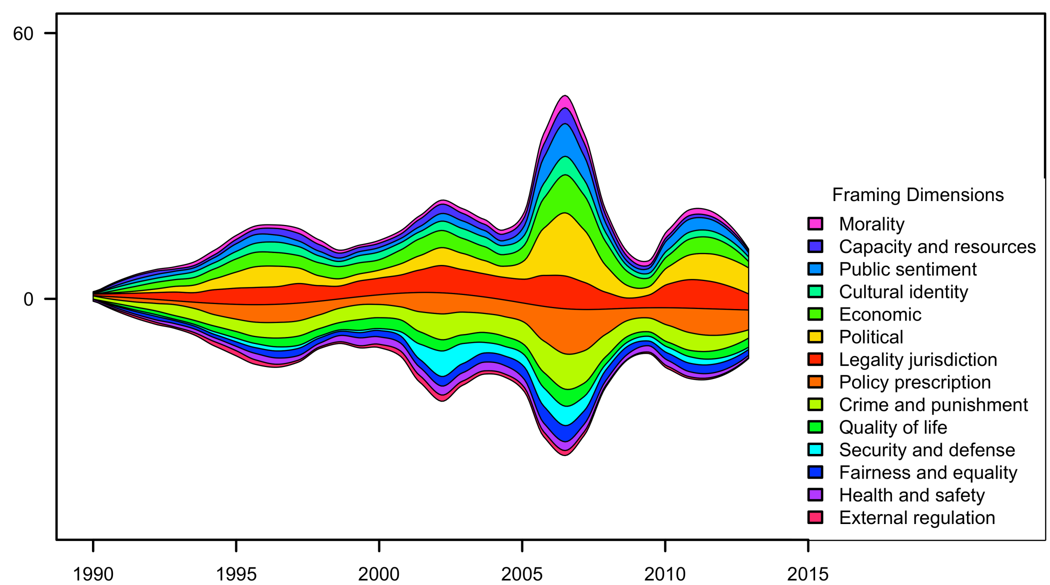
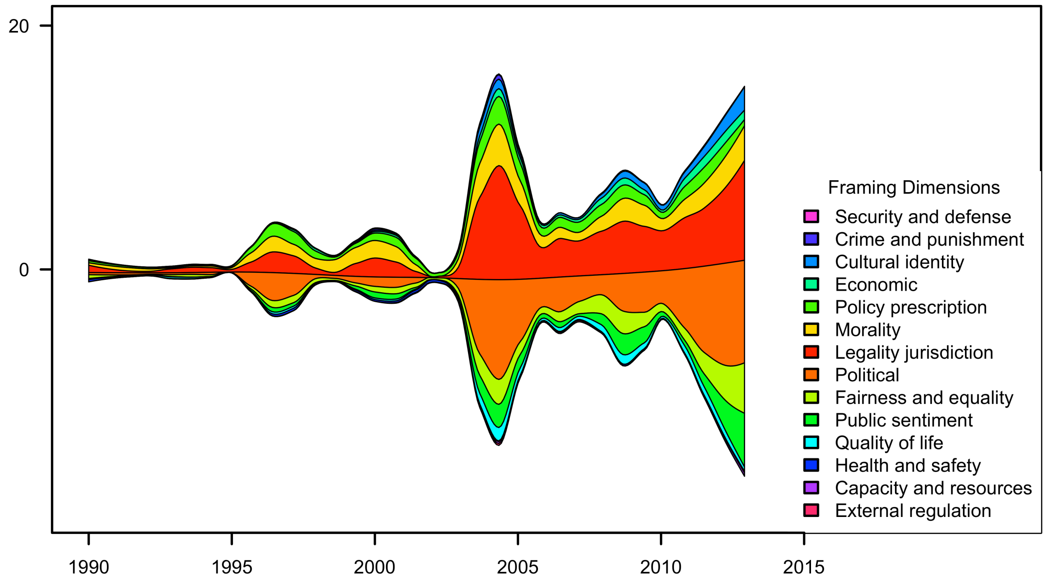


Figure 4: Frames used in news coverage of **immigration**, 1990–2012. The graph above shows the total count of stories in our sample over time (at the month level), partitioned by the *primary frame* selected by coders. The graph below shows the total count of stories containing each framing dimension (not necessarily as a primary frame).



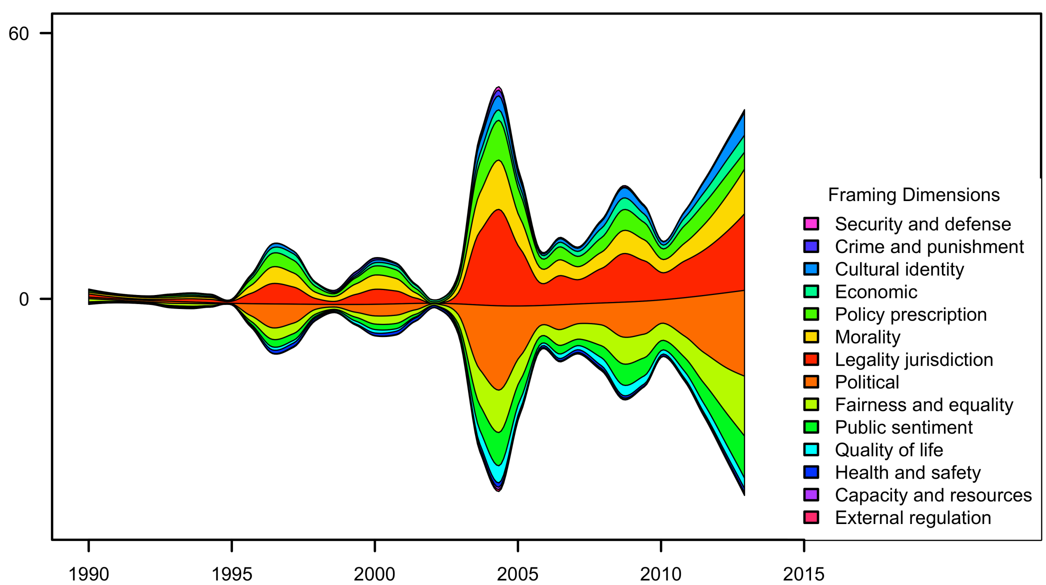


Figure 5: Frames used in news coverage of **same-sex marriage**, 1990–2012. The graph above shows the total count of stories in our sample over time (at the month level), partitioned by the *primary frame* selected by coders. The graph below shows the total count of stories containing each framing dimension (not necessarily as a primary frame).

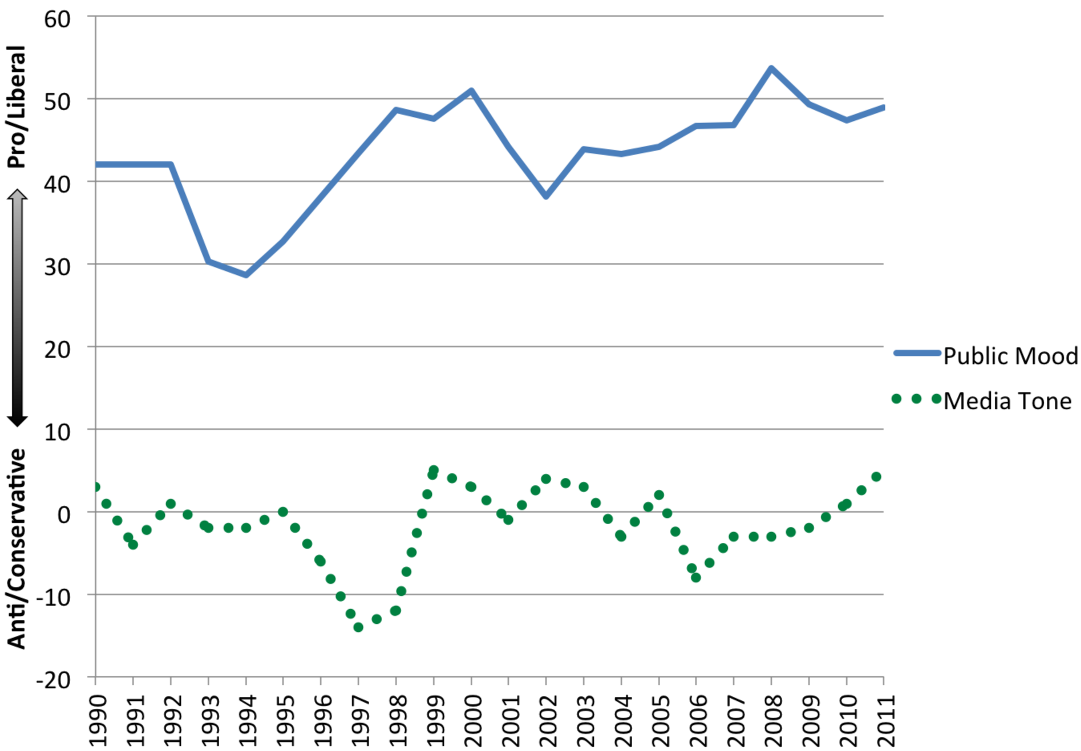


Figure 6: Tone of news coverage of immigration (number of pro-immigration stories minus number of anti-immigration stories) vs. public mood on immigration, 1990–2011.

We can parallel this initial look at frames used in the smoking and tobacco debate with those used in the immigration and same-sex marriage debates. Figures 4 and 5 show stream graphs generated similarly (based on primary frames and totals) for these two issues, respectively. For immigration, we see security frames increasing in 2001 (presumably after the September 11 attacks) and political frames increasing around the 2006 midterm elections. Again, we see interesting variance in frames used, well beyond the security and fairness frames that tend to be most associated with this issue. In the case of same-sex marriage, the legal and political frames are strongest, and volume fluctuates in tandem with the election cycle since 1996.

We can also take a very preliminary look at the small sample of immigration stories we coded for tone, where "media tone" is calulated as the total number of pro-immigration stories minus the total number of anti-immigration stories in each year. Eventually, we will compare the tone of stories to the frames employed, testing for a hypothesized link between the two. Additionally, we plan to analyze the effects of media framing and tone on public opinion. Figure 6 shows a sneak peak at one such comparison, mapping U.S. public attitudes about immigration against our measure of media tone about immigration. The public mood data is taken from Stimson's compilation of public opinion marginals, and tracks the relative position of respondents between more conservative attitudes toward immigration (e.g., tighter border control and increased deportations) and more liberal attitudes (e.g., more generous asylum policies and better resources for undocumented immigrants). These data are much too limited to allow inferences, but the apparent correlation is intriguing, as is the observation that, to the extent the visible pattern holds up to more rigorous analysis, changes in public mood precede changes in media tone rather than the reverse.

Issue	Train Docs.	Test Docs.	Vocabulary
Smoking	2,453	817	10,313
Immigration	2,466	821	10,594
Same-sex marriage	1,797	598	7,825

Table 1: Text classification data.

5 Automated Frame Analysis

As discussed in Section 1, one key goal in this project is to dramatically scale up the coding process conventionally employed in political science research, in order to facilitate analysis of larger and more diverse datasets. There are many ways to introduce automation into manual coding, ranging from an ideal of accurate fully automated analysis to a variety of semi-automated techniques that include a combination of automatic processing and human coding or review.[5] Here we describe initial efforts to automate frame analysis in documents, and then present estimates of frame use in a larger corpus.

5.1 Frame Analysis as Text Classification

For each of each of three issues (smoking, immigration, and same-sex marriage) and fourteen framing dimensions (plus "other" and "irrelevant" to the issue), we constructed a binary text classifier. Each classifier predicts whether a particular frame dimension cue is present, given the text of a document. Note that this is somewhat different from predicting the *primary* frame of a document, although it represents a concrete first step in that direction.

Each of the 3×16 classifers is a logistic regression classifier trained with L_1 regularization. The regularizer strength was tuned on a held-out sample separate from the test set. As features, each classifier uses binary indicators for the presence of each word in the (issue-specific) vocabulary.[6] Table 1 shows the size of each issue's training set, test set, and vocabulary. During training, when a document was annotated by more than one annotator, the label was taken from a single annotator selected uniformly at random.

Using a test sample distinct from the training set used to estimate the parameters of the logistic regression model, we estimate the performance of each classifier.

We report AUC, which stands for "area under the receiver-operating characteristic curve." This statistic is used to characterize a classifier's performance at all possible tradeoffs between false positives and false negatives. It can be intuitively understood as (an estimate of) the probability that a classifier, given a randomly chosen positive instance and a randomly chosen negative instance, will rank the positive one higher. (For logistic regression classifiers, this means that the classifier assigns a greater log-odds score to the positive instance than to the negative one.) Note that a random classifier, or one that always assigns the same answer, will achieve an AUC of 0.5. Any reasonable classifier should score above this value, and our plots are scaled accordingly.

AUC performance is presented in Figures 7, 8, and 9, for smoking, immigration, and same-sex marriage, respectively. In each graph, the frame dimensions are ordered by decreasing prevalence in the test set, with the test-set frequency of each, averaged across annotators, shown in parentheses.

[5] Within healthcare, natural language processing has begun to play an analogous role in medical coding and clinical decision support; see, e.g., Demner-Fushman et al. (2009); Resnik et al. (2006).

[6] NLTK's "Punkt" was used to tokenize. Certain patterns including phone numbers, dates, URLs, numbers, etc. were collapsed manually, and terms occurring in fewer than three training-set articles were excluded or occurring in a standard stoplist. Preliminary experiments with bigrams, preselected words, and annotator variables were inconclusive, so we report only this simple model here.

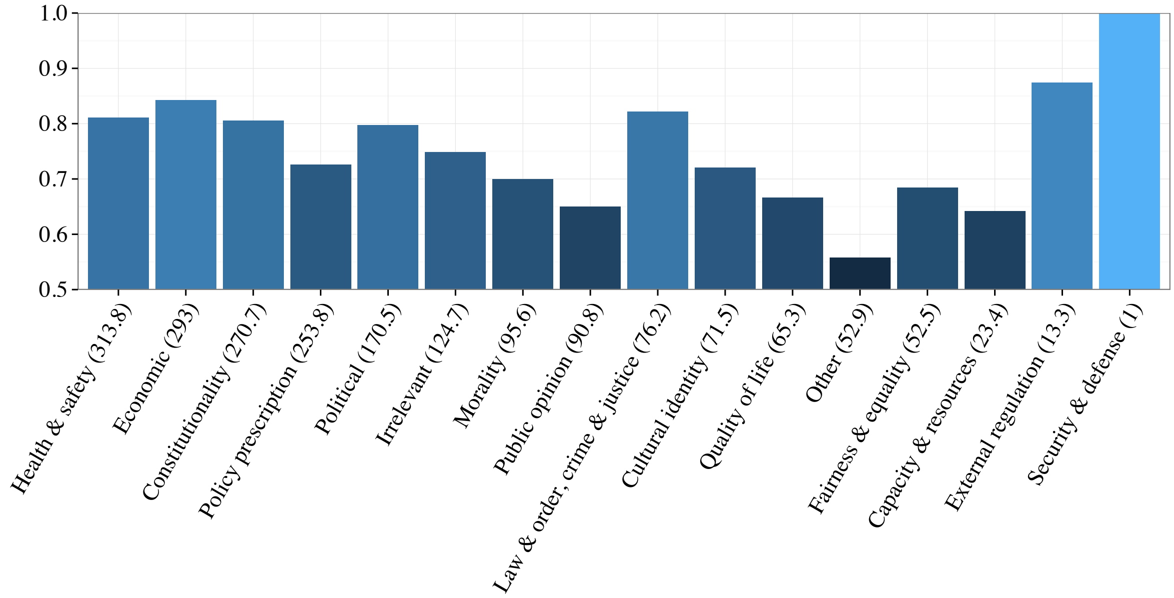


Figure 7: AUC of single-dimension classifiers, for **smoking**. Dimensions are ordered by frequency in the test set (shown in parentheses; counts represent averaging across annotators). We include a classifier for "irrelevant," which signifies that the document is not relevant to this issue.

Our classifiers all show better-than-chance performance, though some framing dimensions are clearly more easily detected using this simple bag-of-words approach than others. Though it is not evident from the graphs (which summarize classifier performance at all false positive/negative tradeoffs), when applied to the test set, our classifiers generally work better on the more frequent framing dimensions.[7]

For the three most prevalent framing dimensions for each issue, we show in Table 2 the twenty features (words) with the greatest impact on positive classification decisions in the held-out data. For a given feature, impact is estimated as the product of its learned coefficient (or weight) and the feature's average value in the test set (Yano et al., 2012). This downplays strongly-weighted (but rare) terms, instead favoring positively-weighted terms that show up frequently out-of-sample. Subjectively, we find these high-impact terms to have good face validity.

We believe that improved methods for detecting framing in text can significantly improve over the performance of these classifiers. However, we believe that, at least for the more frequently occurring frames, these classifiers are sufficiently accurate to apply to a larger dataset and take a preliminary look at frame dynamics and diffusion effects.

5.2 Estimates in a Larger Corpus

To explore patterns of framing in a larger corpus of articles, we applied these binary classifiers to a larger set of uncoded documents in the smoking, immigration, and same-sex marriage corpora from LexisNexis. We then grouped these articles by month of publication to obtain, for each frame, an estimate of the proportion of documents in each month that exhibit that frame. Although this approach provides the best available prediction of framing in each individual article, taking

[7] The true positive rate and true negative rate correlate with test-set frequency of the target class at $\rho = 0.705$ and $\rho = -0.895$, respectively.

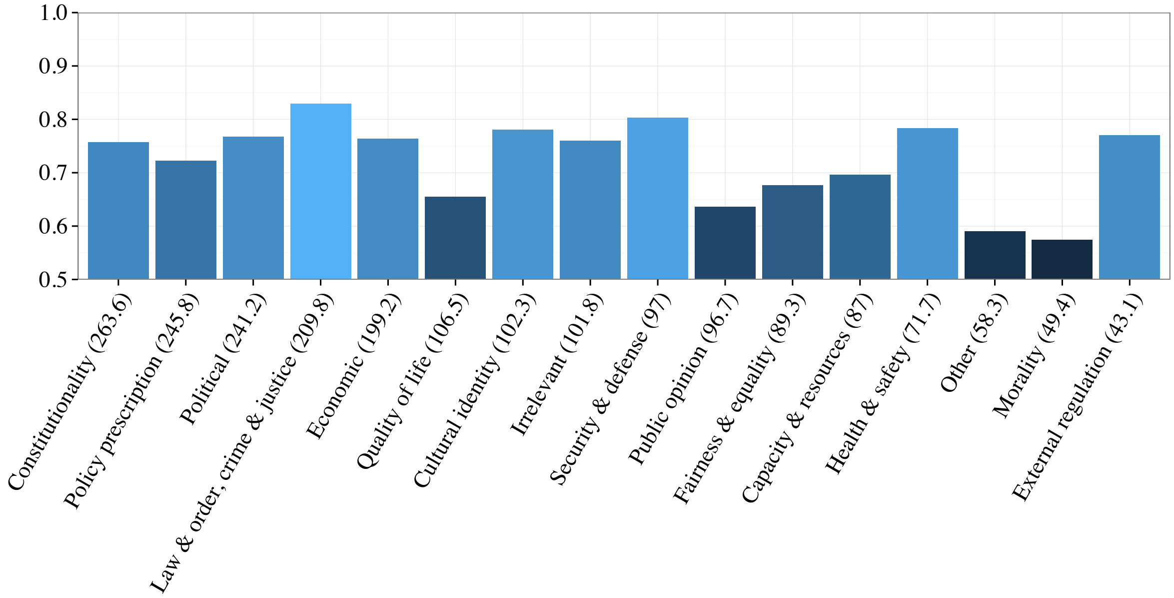


Figure 8: AUC of single-dimension classifiers, for **immigration**. Dimensions are ordered by frequency in the test set (shown in parentheses; counts represent averaging across annotators). We include a classifier for "irrelevant," which signifies that the document is not relevant to this issue.

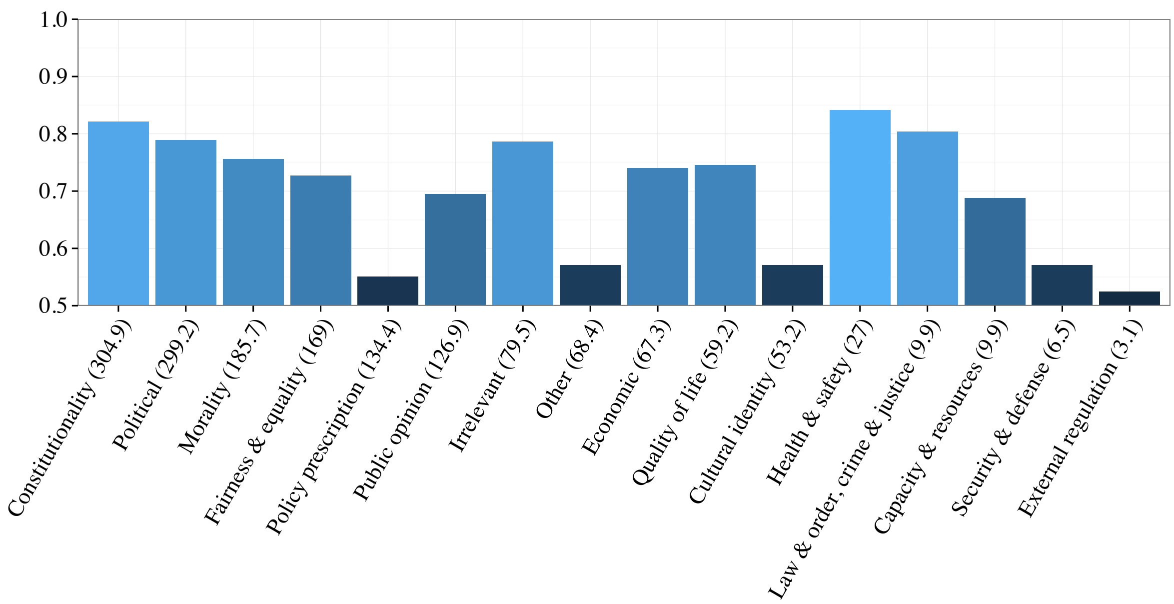


Figure 9: AUC of single-dimension classifiers, for **same-sex marriage**. Dimensions are ordered by frequency in the test set (shown in parentheses; counts represent averaging across annotators). We include a classifier for "irrelevant," which signifies that the document is not relevant to this issue.

Issue	Dimension	Words
Tobacco	Health and safety	health smoking # cigarettes cancer smokers smoke lung secondhand study nicotine evidence addiction quit death RJ become young price smoked
	Economic	billion $ tobacco million market money tax cost sales costs health business settlement pack revenue industry taxes companies farmers increase
	Constitutionality	smoking tobacco court law settlement ban legal lung lawsuit health pay administration public jury seeking $ anti-smoking attorney makers yesterday
Immigration	Constitutionality	law legal states court state came illegal immigration citizenship US case hearing rights week judge monday come government visas announced
	Policy prescription	law state policy year new officials two country program government immigration border director illegal residents would service nation reform proposal
	Political	immigration states president year congress bill public support federal today administration illegal legal make could bush want political expected making
Same-sex marriage	Constitutionality	court state marriage law legal constitution supreme marriages constitutional states san could amendment massachusetts marry decision couples family judge laws
	Political	marriage said senate bill amendment support campaign president gov political democratic mayor republican debate would republicans constitution measure vote obama
	Morality	marriage same-sex religious church woman traditional conservative one man rev issue family added passed members called presidential ban measure come

Table 2: Highest-impact words in classifiers.

the aggregate of these predictions in this manner will not necessarily give unbiased estimates of the proportions of interest, since each classifier may itself be biased. Following Hopkins and King (2010), we therefore apply a correction based on the performance on our classifiers on the held-out test set.

In particular, by comparing the predictions of our classifiers to the true, manually-coded labels of the test set (§5.1), we can estimate the sensitivity and specificity of each classifier. These represent the probabilities that each the classifier will be correct, given that the frame of interest is actually present or absent in a document, respectively. A statistically consistent estimate of the actual proportion of documents that exhibit each frame in the unlabeled corpus is then given by

$$\hat{p} = \frac{p - (1 - \widehat{\text{specificity}})}{\widehat{\text{sensitivity}} - (1 - \widehat{\text{specificity}})}, \tag{1}$$

where p is the uncorrected positive rate estimated in the unlabeled corpus. Because this correction depends upon estimating sensitivity and specificity from the test set, it can be undefined in the cases where there are very few test articles coded with a particular frame. In order to avoid meaningless proportions, we therefore threshold the resulting true proportions to be between 0 and 1. By multiplying the resulting true proportions by the number of articles published in each month, we obtain estimates of the number of articles in each month that would be coded with each frame.

Note that the analysis here is based not on the use of *primary* frames but rather on the number of news stories guessed (by people or programs) to use a framing dimension at some point in the story; a story that uses k dimensions counts k times. Note also that within each issue we built a set of binary classifiers, one for each dimension, not a multi-class classifier, based on a preference not to force competition among frames. Hence the analyses here are commensurate with the stream graphs shown in the lower panels of Figures 3, 4, and 5.

Each of Figures 10, 11, and 12 shows two stream graphs, corresponding to the tobacco, immigration, and same-sex marriage issues, respectively. In each case, the graph above shows the estimated frequencies in the unlabeled pool, and the graph below is a rescaling of the estimates from labeled data, shown in the lower panels in Figures 3, 4, and 5 (respectively), for direct comparison.

Classifying a larger set of unlabeled data not only provides more data to look at, but also allows us to observe interesting findings we might have missed looking only at the labeled data. For example, in the case of smoking, the unlabeled data graph in Figure 10 (top) suggests that two frame dimensions not immediately associated with the debate—fairness/equality and morality—are in fact important parts of the political discourse surrounding this issue. Additionally, using the unlabeled data brings to the fore the prominent role that the political frame dimension has across time within all three issues. This fact offers some face validity for our method, showing that by applying our general frame codebook and computational methods to analyzing framing, we uncover both the differences in types of frames used between these very different issues but also some commonalities. Politics is, quite unsurprisingly, a common staple of political discourse. Looking more closely at Figures 10, 11, and 12 provides preliminary support for our hypothesis about framing dynamics and also allows us to consider additional hypotheses we might derive, in true deductive/inductive iterative fashion. For example, we hypothesized that surges in attention to an issue are not usually produced by a single dominant frame but, rather, by multiple frames "bandwagoning" together, producing the attention cascade. In each of the three issue areas examined, we see at least one major surge in attention: around 1997 for smoking, around 2006 for immigration, and around 2004 for same-sex marriage. For smoking and immigration, these surges in attention map onto the surges we saw in Figures 3 and 4, respectively—as they should, since the total use of primary frame dimensions, captured through manual coding (shown in Figures 3 and 4) should correlate directly with the total use of frame dimensions throughout news stories as a whole, captured through Figures 10 and 11. In each of the three issues shown here, we see suggestive yet pronounced evidence that surges in attention are not driven by a single dominant frame. Rather, in each case we see multiple frames activated in the policy debate at the time of the surge in attention.

This preliminary evidence also holds implications for additional hypotheses we might derive about frame diffusion. If frames tend to bandwagon during surges in attention to an issue within a single venue type (here, newspaper coverage), it supports the notion that what is really at work is not a powerful "reframing"—to which some venues might be more amenable than others—but rather a heightened activation of political discourse across many facets of the issue, thereby increasing the chances that a different venue (e.g., state legislatures) will pick up on one of the frames activated in the debate. In other words, to the extent that framing diffuses across policy venues, it may be the case that the mechanism is not at the frame level—where the frame itself is the agent of contagion—but rather at the level of the debate itself—where the level of *activity* in the debate is what helps it, and the frames at play, to spread across venues.

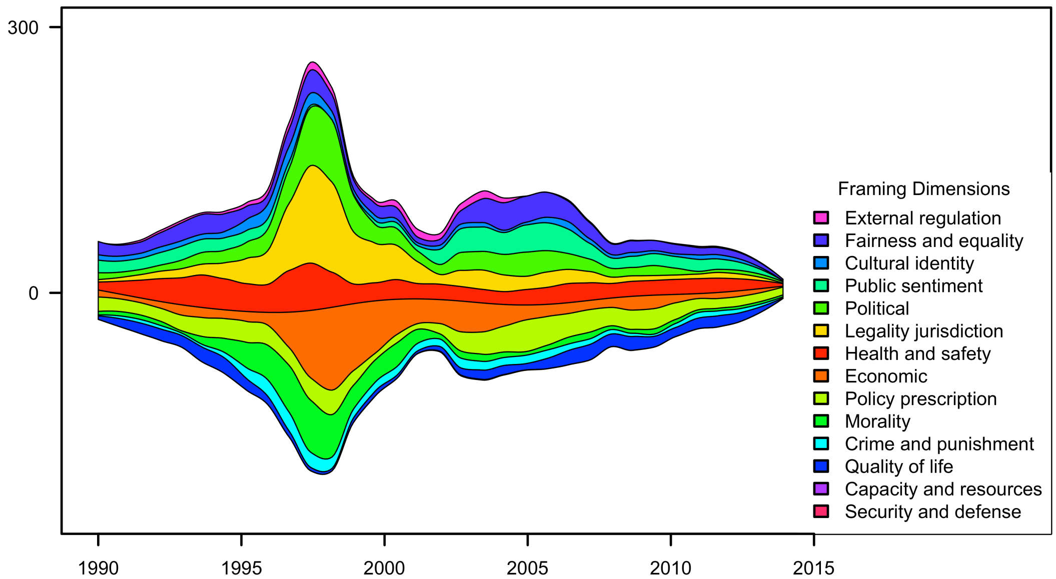
Figures 10, 11, and 12 suggest other hypotheses as well. For example, the cases of tobacco and immigration suggest that surges in attention to an issue tend to reverberate, with aftershocks of sorts following a major punctuation. Interestingly, in both cases these reverberations contain a notable amount of attention to the "public sentiment" frame dimension. Additionally, in the case of immigration, we see that the increase in attention to security and defense frames that began in 2001 held through the next major punctuation in 2006, but then died off. This evidence is anecdotal, of course, but falls in line with general punctuated equilibrium theory, suggesting that we should see the distribution of frames within an issue debate follow a structural-break pattern of punctuation, whereby a frame can maintain prevalence in a debate for an extended period but then be uprooted by a new surge in discussion that focuses on different frames.

6 Conclusion

We have presented an overview of a project whose goal is to bring together the empirical study of framing with scalable, state of the art computational modeling of text. A key research activity during our first two years has been the initial design and development of the Policy Frames Codebook, a theoretically motivated resource that provides both a general, cross-issues inventory of frame categories and the instantiation of those categories for a set of issues highly relevant in U.S. policy studies. Based on analysis of manual coding across three quite different issues, our inventory of framing dimensions demonstrates cross-issue applicability and comprehensiveness, and our initial efforts using manual and automated coding of framing dimensions provide suggestive evidence for the potential of our approach, both as the basis for traditional manual-coding efforts and as a foundation for large-scale automated analyses not possible using manual effort alone. We are optimistic that the Policy Frames Codebook, and tools derived from our associated computational efforts in frame analysis and frame discovery, will make it possible for the broader community of political scientists and communication scholars to join us in approaching the phenomenon of framing using a combination of theoretical and large-scale data driven methods.

Acknowledgments

The authors gratefully acknowledge project members and collaborators who contributed to this work: Brice Acree, Aaron Bonner, Jordan Boyd-Graber, Matthew Lesenyie, Viet-An Nguyen, Yanchuan Sim, and Kristina Victor. This research was supported in part by NSF IIS collaborative research funding under grants 1211201, 1211266, 1211277, and 1211153. Any opinions, findings,



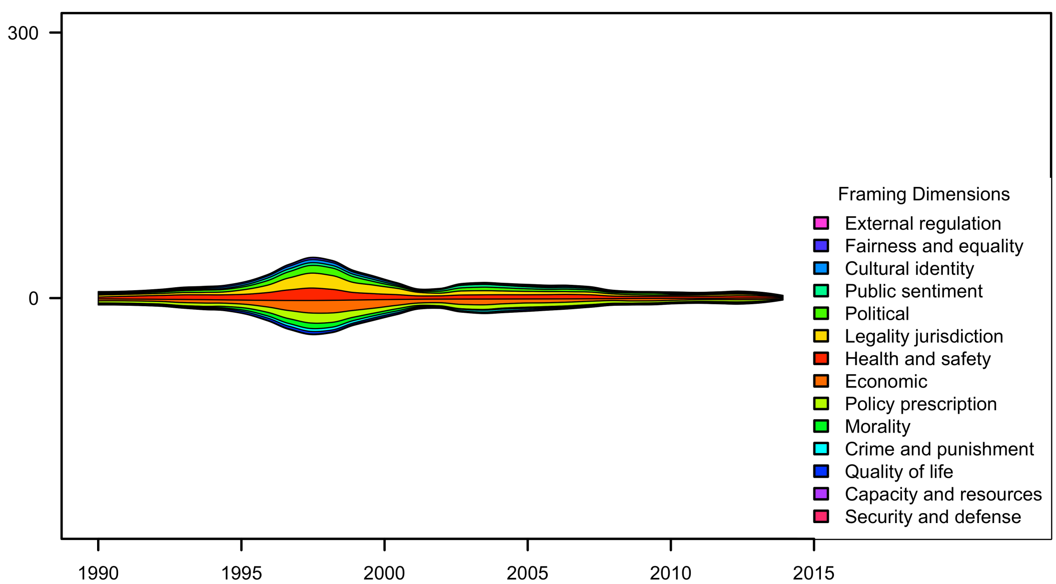
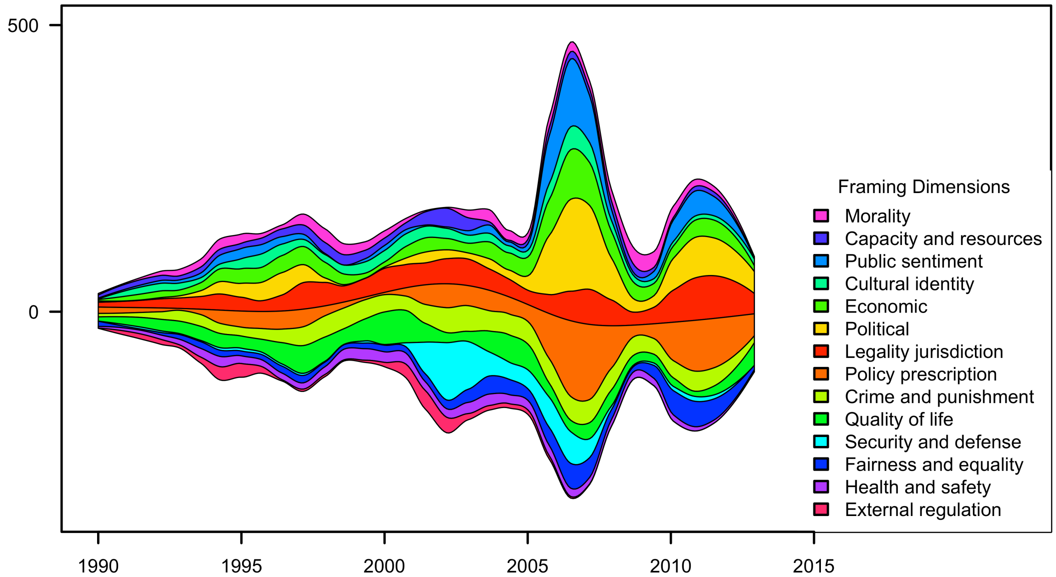


Figure 10: Frames used in news coverage of **smoking**, 1990–2012. The graph above shows the estimated counts of stories in our *unannotated* sample over time (at the month level), for each framing dimension. The lower graph shows the counts from our manually coded sample, on the same scale.



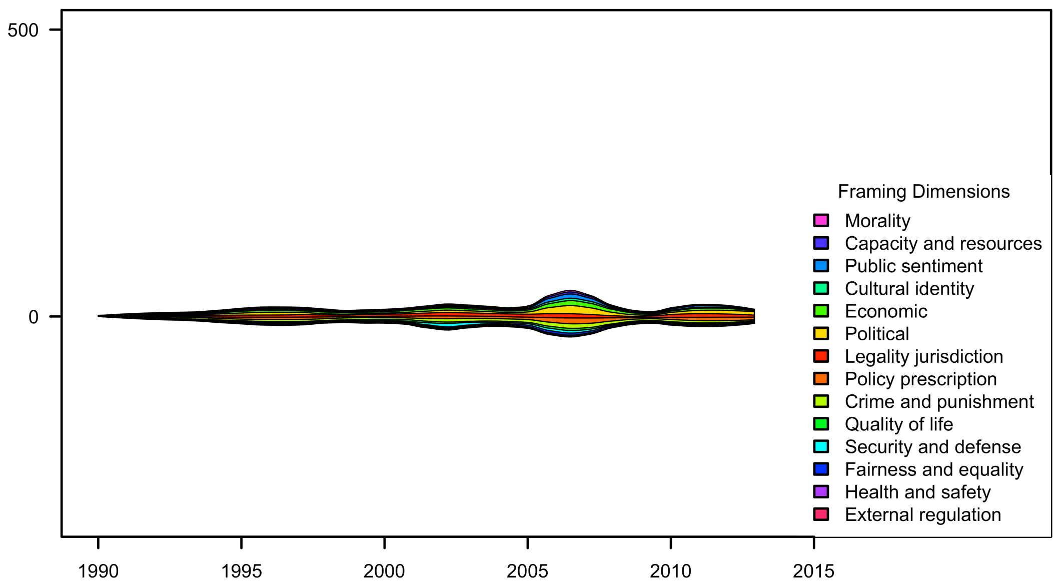


Figure 11: Frames used in news coverage of **immigration**, 1990–2012. The graph above shows the estimated counts of stories in our *unannotated* sample over time (at the month level), for each framing dimension. The lower graph shows the counts from our manually coded sample, on the same scale.

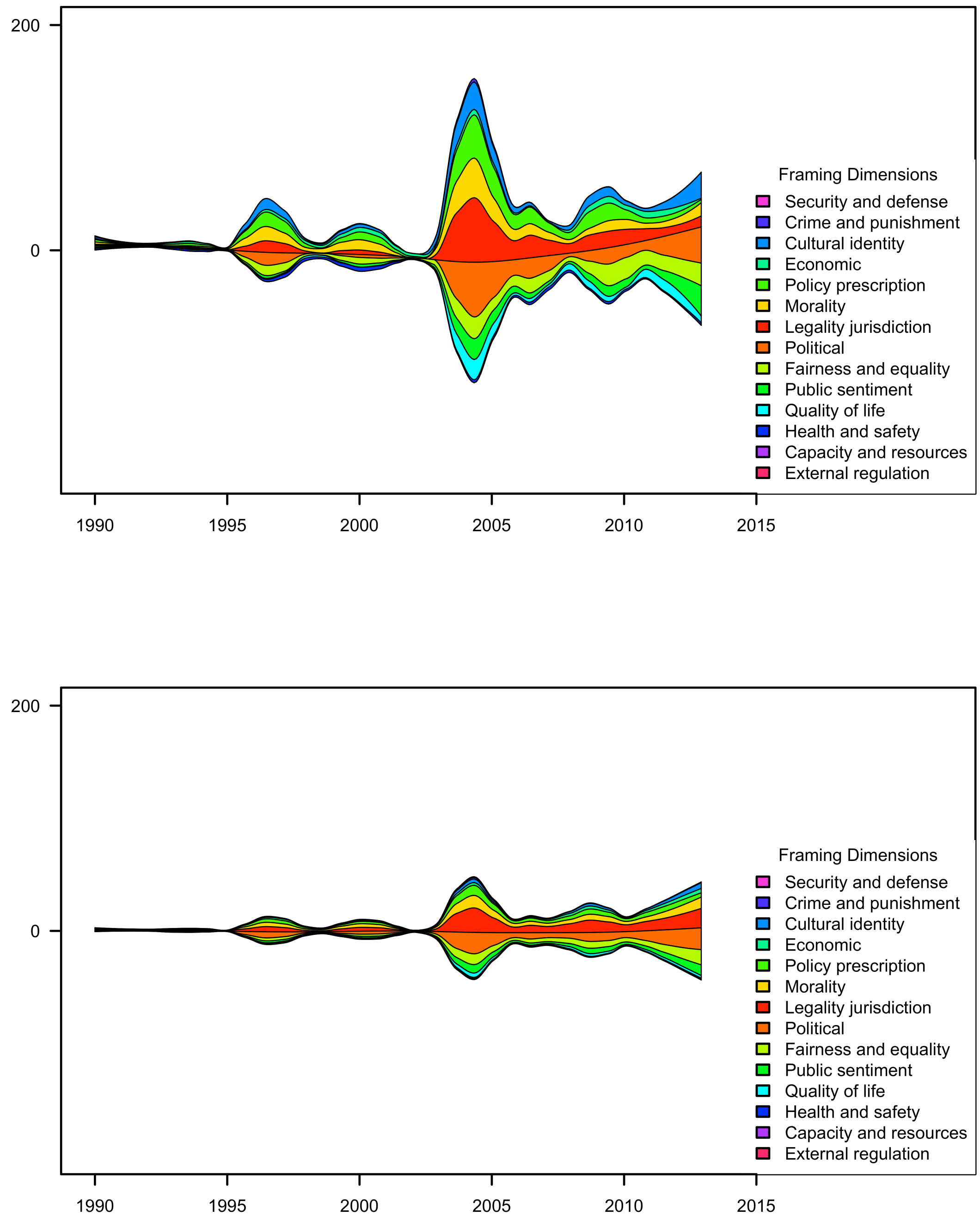


Figure 12: Frames used in news coverage of **same-sex marriage**, 1990–2012. The graph above shows the estimated counts of stories in our *unannotated* sample over time (at the month level), for each framing dimension. The lower graph shows the counts from our manually coded sample, on the same scale.

conclusions, or recommendations expressed here are those of the authors and do not necessarily reflect the view of the sponsor.

References

Frank R. Baumgartner and Bryan D. Jones. *Agendas and Instability in American Politics*. University of Chicago Press, 2010.

Frank R. Baumgartner, Suzanna L. De Boef, and Amber E. Boydstun. *The Decline of the Death Penalty and the Discovery of Innocence*. Cambridge University Press, 2008.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):9921026, 1992.

Dennis Chong and James N. Druckman. Framing theory. *American Political Science Review*, 10: 103–126, 2007.

Dina Demner-Fushman, Wendy W. Chapman, and Clement J. McDonald. What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics*, 42(5): 760–772, 2009.

Robert Entman. Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4):51–58, 1993.

Robert M. Entman. Cascading activation: Contesting the white house's frame after 9/11. *Political Communication*, 20:415–532, 2003.

Herbert J. Gans. *Deciding Whats News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Northwestern University Press, 2004.

Rebecca A. Glazier and Amber E. Boydstun. The president, the press, and the war: A tale of two framing agendas. *Political Communication*, 2012.

Doris A. Graber and Johanna Dunaway. *Mass Media and American Politics*. CQ Press, 2014.

Mark Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83: 1420–1443, 1978.

Justin Grimmer. *Representational Style in Congress: What Legislators Say and Why It Matters*. Cambridge University Press, 2013.

Daniel J. Hopkins and Gary King. A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, 54(1):229–247, 2010. ISSN 1540-5907. doi: 10.1111/j.1540-5907.2009.00428.x. URL `http://dx.doi.org/10.1111/j.1540-5907.2009.00428.x`.

Shanto Iyengar. *Is Anyone Responsible? How Television Frames Political Issues*. University of Chicago Press, 1991.

Shanto Iyengar, Helmut Norpoth, and Kyu S. Hanh. Consumer demand for election news: The horserace sells. *The Journal of Politics*, 66:157–175, 2004.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

George Lakoff. *Don't Think of an Elephant! Know Your Values and Frame the Debate—The Essential Guide for Progressives*. Chelsea Green Publishing, White River Junction, NY, 2004.

George Lakoff. *Whose freedom?: the battle over America's most important idea*. Macmillan, 2006.

J. Matthes and M. Kohring. The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, 58(2):258–279, 2008.

Maxwell McCombs. The agenda-setting role of the mass media in the shaping of public opinion. *North*, 2009(05-12):21, 2002.

Rose McDermott, James H. Fowler, and Oleg Smirnov. On the evolutionary origin of prospect theory preferences. *The Journal of Politics*, 70:335–350, 2008.

Thomas E. Nelson, Rosalee A. Clawson, and Zoe M. Oxley. Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review*, 91(3):567–583, 1997.

W. Russell Neuman, Marion R. Just, and Ann N. Crigler. *Common Knowledge. News and the Construction of Political Meaning.* University of Chicago Press, 1992.

Julianna Pacheco. The social contagion model: Exploring the role of public opinion on the diffusion of antismoking legislation across the american states. *The Journal of Politics*, 74(1):187–202, 2012.

Kevin M. Quinn, Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science*, 54:209–228, 2010.

Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. Linguistic models for analyzing and detecting biased language. In *Proceedings of ACL*, 2013.

Philip Resnik, Michael Niv, Michael Nossal, Gregory Schnitzer, J. Stoner, A. Kapit, and R. Toren. Using intrinsic and extrinsic metrics to evaluate accuracy and facilitation in computer-assisted coding. In *Perspectives in Health Information Management: Computer Assisted Coding Conference Proceedings*, 2006.

Brian F. Schaffner and Patrick J. Sellers. Introduction. In Brian F. Schaffner and Patrick J. Sellers, editors, *Winning with Words: The Origins and Impact of Political Framing*, volume 1, page Chapter 1. Routledge, 2010.

Stuart N. Soroka. Good news and bad news: Asymmetric responses to economic information. *The Journal of Politics*, 68:372–385, 2006.

B. Van Gorp. Where is the frame?. victims and intruders in the belgian press coverage of the asylum issue. *European Journal of Communication*, 20(4):484–507, 2005.

Tae Yano, Noah A. Smith, and John D. Wilkerson. Textual predictors of bill survival in Congressional committees. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics*, pages 793–802, June 2012.