Measuring Stability and Change in Personal Culture Using Panel Data

Kevin Kiley

Stephen Vaisey

Duke University

Abstract: Models of population-wide cultural change tend to invoke one of two broad models of individual change. One approach theorizes that people actively update their beliefs and behaviors in the face of new information. The other argues that, following early socialization experiences, dispositions are stable. We formalize these two models, elaborate empirical implications of each, and derive a simple combined model for comparing them using panel data. We test this model on 183 attitude and behavior items from the 2006-14 rotating panels of the General Social Survey. Though the pattern of results is complex, it is somewhat more consistent with the settled dispositions model than the active updating model. Most observed change in the GSS appears to be short-term attitude change or measurement error rather than persisting changes. When persistent change occurs, it is somewhat more likely to occur in younger people than older people and more common for public behaviors and beliefs about high-profile issues than private attitudes. We argue that we need both models in our theory of cultural evolution but that we need more research on the circumstances under which each is more likely to apply.

INTRODUCTION

Culture is an important part of social life but cultures are continuously evolving. In 1972, for example, more than 40 percent of US adults supported a law outlawing interracial marriage. Three decades later, this opinion had become so uncommon that the question was removed from the US General Social Survey. How does this kind of cultural change happen?

Attempts to account for opinion changes in society have produced conflicting theories about the process of opinion formation and the ability of individuals to maintain consistent attitudes. Some models suggest that people lack the cognitive tools to maintain consistent beliefs on social and political issues. As a result, they construct responses on the fly in the interview setting, drawing on ideas from opinion leaders and changing their attitudes as elite discourse changes (Zaller 1992; Converse 1964; Perrin and McFarland 2011). By contrast, cohort replacement theories posit that people do hold opinions and are unwilling to alter them in the

face of societal change, and thus public opinion changes only with generational turnover (Mannheim 1952; Ryder 1965). Another set of models claim that people hold "a number of real, stable, and sensible opinions about public policy" and change their opinions in response to new information (Page and Shapiro 1992: p. xi; Achen 1992). More recently, sociologists have suggested that people attempt to align malleable peripheral beliefs with relatively fixed "core" beliefs using social cues (Boutyline and Vaisey 2017; Lakoff 2002; Goldberg and Stein 2018).

Despite their diversity, models of aggregate attitude and behavior change often implicitly invoke one of two broad models of individual change with strong connections to theories in cultural sociology. The first of these models is an *active updating* model that emphasizes the role of changing discourses, environments, and interactions on attitude formation. This model is rooted in a broadly pragmatist approach to action, which claims that contemporary social environments and problems provoke people to adapt their views and make new meaning (e.g., Joas 1996, Gross 2009, Swidler 2001). The second is a *settled dispositions* model, which emphasizes the continuing influence of durable dispositions acquired early in life. This model has affinities with the Bourdieusian tradition, which de-emphasizes (though does not ignore) the current environment in favor of the "past conditions of production" (e.g., Bourdieu 1990, Vaisey and Lizardo 2016). These two models represent different approaches to understanding how people come to hold diverse forms of *personal culture*, or the declarative and non-declarative attitudes, worldviews, values, dispositions, and associations that manifest at the individual level (Lizardo 2017).

In this paper, we make these two models of personal culture change explicit, deduce some of their empirical implications, and derive a statistical model for estimating the prevalence of active updating using panel data. In contrast to previous models that tend to assume one of

these data-generating processes to measure stability and reliability of estimates over time, our approach separates persisting change from non-persisting change to estimate if there is evidence of individuals in the population making persistent changes in attitudes and behaviors. We then apply this method to 183 items from the 2006-2014 General Social Surveys (GSS). By classifying the pattern of change in personal culture, we can clarify when different accounts of aggregate change are more likely to apply. Because of data limitations, we cannot speak to all types of cultural objects (e.g., music styles, baby names). We do, however, investigate a wide variety of opinions, including views on politics, free speech, race, and gender roles, and practices including socializing at bars or attending church, that are important in contemporary U.S. society.

Our analysis yields several results. First, we find that the majority of what appears to be individual-level change in attitude or practice probably reflects short-term (i.e., non-persisting) change or measurement error rather than actual persistent change. Simply put, there is little evidence that large numbers of U.S. adults changed their beliefs or practices in lasting ways over this period of study. Second, settled opinions vary in how consistently individuals report the same answer. Consistent with theories that argue that people lack clear opinions, some survey items appear to elicit inconsistent or random responses. At the same time, people are consistent on both high- and low-profile public policy items, suggesting a greater degree of "real" attitudes than these theories suggest. Third, the persistent change that we *do* see in the data is somewhat more concentrated among younger respondents. On several items it appears that younger adults are still in the process of acquiring dispositions and habits they will take into later life. Fourth, we find that changes in social behavior (e.g., church attendance, political party membership, socializing) are more likely to persist than changes in private attitudes (e.g., political ideology),

and people are more likely to report these attitudes and behaviors consistently. This suggests that interactional and institutional mechanisms may provide stronger support for lasting change than pressures for intrapsychic consistency.

Our findings offer broad support to theories claiming that cultural change comes through generational turnover rather than persuasion and social influence. However, there is a pattern of exceptions and caveats that can help us understand how institutions and events shape the process of cultural change and that challenge the idea that change in all attitudes follows a similar trajectory over the life course. The pattern of results also supports models of attitude change that put ideological identification at the center of a network of political beliefs and suggest that individuals are more likely to make lasting changes in their partisan identification than to their general political beliefs. We argue that there is a place for both the active updating and settled dispositions models in our theory of cultural evolution but that we need more research on the circumstances under which each is more likely to apply.

THEORETICAL BACKGROUND

Belief Formation in Cultural Sociology

How does cultural change happen at the individual level? Let us make the question concrete by imagining a person who answers the same question each year for several years. The question could be anything, but assume it is this GSS question: "Please tell me whether you strongly agree, agree, disagree, or strongly disagree [with this statement]: 'a working mother can establish just as warm and secure a relationship with her children as a mother who does not work.'"

How does the respondent formulate a response to that question each time, year after year? To make things as explicit as possible, we can write the data-generating process formally.

Although this presentation may make it seem like we are assuming rationality or conscious deliberation, this way of writing the models makes no particular cognitive assumptions. We will explain this in greater detail below. For now, consider the following two simple models:

$$y_{it} = y_{it-1} + \nu_{it} \qquad (1)$$

$$y_{it} = U_i + \nu_{it} \qquad (2)$$

These models may seem similar at first glance, but they have different implications for the pattern of individuals' responses we would observe over time. Equation (1) represents an *active updating* model and Equation (2) represents a *settled dispositions* model, as we explain below. Figure 1 shows these models in graphical form, which helps highlight their differences. In the next two sections, we consider each of these models and briefly discuss their connections to sociological theories.

[Figure 1 about here]

Active Updating Model

Equation (1) represents the active updating model. The respondent forms her answer by starting with what she said last time (y_{it-1}) and then incorporating any new considerations (ν_{it}). There is no need to remember responses from earlier time points (e.g., y_{it-2}) because this information gets folded into the updated response each time. Formally, Equation (1) is a Markov model, where future states depend entirely on the current state. This formal property is often assumed to underlie the data generation process in studies of change in and reliability of repeated survey measures (Alwin 2007; Krosnick and Alwin 1989). In this framework, we use the term "updating" to refer to the change to a new baseline for whatever reason, not necessarily requiring the incorporation of new or better information.

More informally, this AUM posits a person who is updating her views in the face of social experience. There are formal Bayesian ways of modeling updating, and this model underlies theories of rational updating in the face of new information (Achen 1992; Bartels and Jackman 2014; Page and Shapiro 1992), but we need not rely on any assumptions of rationality, optimality, or conscious thought for this basic process to apply. Following Gross's (2009, 367) pragmatist account, we could instead regard this model as consistent with an "active and creative relation to the world" that "lead[s] actors to see themselves in new ways, to value different kinds of goods, and to become attached to problem solutions that they could not have imagined previously."

Active updating lies at the heart of most theories that suggest social environments influence individual attitudes and behaviors. When sociologists posit that adults adopt the beliefs and behaviors of their friends (Christakis and Fowler 2009; Centola 2018; DellaPosta, Shi, and Macy 2015), or that individuals adopt behaviors that they view as consistent with their cultural dispositions (Goldberg and Stein 2018), they invoke this model. When social scientists argue that people adapt their partisan affiliations to conform to their social groupings, change their ideological commitments to conform to their partisan identifications (Green, Palmquist, and Schickler 2002), change their partisan affiliations to match their ideological commitments (Baldassarri and Gelman 2008; Abramowitz and Saunders 2006), or respond to exposure to diverse opinions by becoming more extreme in their viewpoints (Bail et al. 2018), the claim is rarely that these changes are temporary shocks that will revert to a stable baseline over time. Instead, the claim is that the changing social environment changes dispositions, which will then remain stable until the environment changes again.

Regardless of whether the exact process is rational, heuristic, discursive, embodied, or any combination of these, the key notion is that the actor continues to be influenced by the (social) environment in ways that might lead to long-term, persistent change in beliefs, practices, and identities (Gross 2009; DeGloma 2014). As the person encounters new considerations throughout her life (e.g., adding working mothers to one's social network, seeing changing media representations of working mothers), she can continue to revise her views.

The AUM makes no assumption about the distribution of ν_{it}. Specifically, it does not assume that ν_{it} has an expected value of 0, either for any time t or for any person i. This leaves open the possibility of a population-wide shift in responses as many people react to the same changes in the environment.

Several theories of cultural change at the societal level require evidence of active updating at the individual level, such as arguments that the individuals in the population are becoming more polarized on political issues (DiMaggio, Evans, and Bryson 1996; Iyengar and Westwood 2015), arguments that individuals adjust their political preferences based on the performance of the governing party (Page and Shapiro 1992; Bartels and Jackman 2014), arguments that new ideas diffuse across the population through person-to-person contact (Centola 2018), and arguments that changes in elite discourse drive change in mass opinion (Zaller 1992). In the absence of evidence that people change their minds in persisting ways, these theories cannot be valid at the aggregate level, and we must seek other explanations for aggregate changes.

Settled Dispositions Model

Equation (2) represents the settled dispositions model. Here each respondent begins the study period with a set tendency to respond to the question in a particular way (U_i). Temporary

considerations, like current events, can play a role in what response she gives at each time (part of the ν_{it}), but these considerations have no lasting impact beyond time t. As the right panel of Figure 1 makes clear, there is no mechanism by which a particular consideration can "propagate up" into the settled disposition and change the baseline for future responses. Such considerations are thus temporary influences only. At its core, this model reflects the Bourdieusian model of action that emphasizes "the past conditions of production" (Bourdieu 1990, 66ff). In other words, updating happened, but it happened in the past, prior to the time of the study. In this sense, U_i reflects the "habitus."

The SDM underlies theories that suggest that individuals are stable in their dispositions throughout life or, if they do change, tend to revert back to a relatively stable baseline in short time, including "control" theories such as Identity Control Theory and Affect Control Theory (Robinson 2007; Smith-Lovin and Heise 1988). The notion that individuals are stable in their dispositions is commonly associated with cohort replacement theories of attitude change at the population level. These models posit that people hold relatively stable opinions, that few people change their attitudes as they age, and that most change in public opinion must come from older people dying and being replaced in the population with young people who hold different opinions (Mannheim 1952; Ryder 1965). Vaisey and Lizardo (2016), looking at a range of opinion items in the GSS, suggest that population-wide cultural change most likely comes in the form of cohort replacement.

On its face, the settled dispositions model might seem to imply that individuals hold strong, consistent opinions, but this need not be the case. The settled dispositions model is also consistent with theories that suggest that people lack stable opinions and are unable to report them as such in a survey context (Perrin and McFarland 2011; Bourdieu 1979). These theories

tend to focus on questions of public policy, since the act of opinion creation requires more information than individuals typically have (Zaller 1992; Converse 1964). Such accounts imply that peoples' responses at any particular time are a deviation from a baseline, with this deviation often assumed to be random and commonly referred to as a form of "measurement error" (Page and Shapiro 1992; Converse 1964). As long as these deviations are random, there is no reason to perceive these changes as a form of updating, even if wave-to-wave differences are quite large.

Like the AUM, the SDM does not require ν_{it} to be 0 in expectation for the population at time t, but it does assume that ν_{it} has an expected value of 0 within an individual over time. That is, current considerations can move a person from their baseline temporarily, but there is a tendency to "bounce back" to that baseline over time. The SDM thus allows for population-wide shifts in beliefs, practices, or identities at a particular time (temporary period effects), but it assumes that within people these shifts will be erased over time as people return to their baselines.

Like with the active updating model, there are theories of attitude development and change at societal level that would require evidence of the settled dispositions model at the individual level. For example, if people's attitudes change in a way that reflects updating, then it would be wrong to consider their deviations from their average to be "measurement error," and theories that suggest that changes in belief are not real (Page and Shapiro 1992) would be unsupported. Similarly, an item would have to be relatively stable for most adults for aggregate cultural change to come primarily through cohort replacement (Ryder 1965).

The two general models thus differ fundamentally on their emphasis on the character of personal change. In the AUM, changes tend to persist because they propagate into future responses, shifting the baseline over time. In the SDM, changes tend to revert because each

person returns to their baseline. We can therefore summarize the models' implications for individual change as follows: the AUM predicts *persisting change* whereas the SDM predicts *non-persisting change*.

Consistency in Response

Neither model makes any assumptions about the magnitude of the variance of ν_{it}, meaning that both models allow for the possibility of either high or low levels of change in responses from wave to wave. This means that we must consider separately the overall *amount* of change in responses from wave to wave and the *persistence* of this change. Knowing that there is more wave-to-wave change in one item than another tells us little about how these two items relate to underlying theories of attitude development. However, knowing that two attitudes have the same rate of wave-to-wave change, but that one follows an active updating pattern with little random fluctuation, while the other follows a settled disposition pattern with a high amount of random fluctuation allows us to make more precise inferences about the applicability of different theories to these items. As such, we designate a second term, *consistency*, to reflect this second dimension of change. *Consistency* here simply refers to the degree that the attitude development process departs from pure randomness.

Toward Theoretical Synthesis

We do not suggest that all attitudes and behaviors captured by survey questions follow a settled dispositions model or that all display active updating. Instead, we suggest that by classifying the pattern of change at the individual level, we can use that as a tool for adjudicating debates of change at the aggregate level.

The clearest theoretical contrast is between models that assume beliefs are settled during adulthood, such as cohort replacement theories, which would produce a pattern of low active

updating, and those that posit that individuals update beliefs throughout their lives, which would produce an active updating pattern. This disagreement thus forms the primary structure of our analysis. However, it is possible that some items display active updating while others do not. There are at least three ways to understand how both models could be present in the population without contradiction: age-based differences, core-periphery models, and public-private differences.

Age-based differences: The concept of a "cohort effect" is that the environment influences the baseline response tendency when a person is young and then stabilizes for the remainder of one's life (Elder 1974; Bartels and Jackman 2014). To put it differently, cohort formation requires an early period of updating followed by a later period characterized by either absolute stability (where the variance of ν_{it} is low) or by temporary, non-persisting changes that disappear as people revert to baseline.

Items that reflect a cohort replacement pattern should demonstrate low active updating at the individual level for most adults but high active updating for young people, as this is the period most commonly found to be susceptible to attitude change (Glenn 1974; Inglehart and Baker 2000; Visser and Krosnick 1998; Danigelis, Hardy, and Cutler 2007).

Core-periphery models: A number of theories of attitude development posit that people hold "core" beliefs, which they use to adjudicate peripheral attitudes (Boutyline and Vaisey 2017; Goldberg and Stein 2018). This work has predominantly focused on political beliefs and adjudicating whether people use ideological identity, policy positions, moral beliefs, or partisanship to form their positions on other issues (Lakoff 2002; Green, Palmquist, and Schickler 2002; Kinder and Kalmoe 2017). In their analysis of cross-sectional data, Boutyline and Vaisey (2017) find support for ideological identification – assessment of oneself as a

conservative or a liberal, not as a Republican or Democrat – as the "core" belief in the political space. Similarly, Baldassari and Gelman (2008) find increased alignment between partisan identification and policy positions but little increase in the correlation between specific policy positions over time. They interpret this finding as indicating that "voters are splitting along party lines according to the issues that are most salient to them, while they do not bother to adjust their (weak) preferences on the remaining issues" (Baldassarri and Gelman 2008: pp. 442).

The picture that emerges from these previous works is that ideological identity should be a core, if poorly reported, disposition. Changes in ideological identity should be non-persisting, occurring because people either do not fully understand these terms and report their views with error (Converse 1964) or because people bounce around a stable mean, meaning this item should show low updating and medium to low consistency. On the other hand, we should see evidence that individuals adjust their partisan identification to match their ideological identification as they come to find that one party better represents their core positions (Baldassarri and Gelman 2008). In other words, liberal Republicans should become liberal Democrats, not conservative Republicans. This means that partisan identification should display high levels of updating. Again, this does not mean that many people change, only that changes tend to persist.

Alongside ideological identification, moral issues and general political sentiments might serve as core organizing principles, but what these beliefs are will not be consistent across people (Baldassarri and Gelman 2008). For example, some people might hold strong, consistent views on abortion while others care about the amount of government intervention in the economy. Because of this heterogeneity, these beliefs should have a moderate level of consistency, since those people who do not value them should report them inconsistently. There should be no

evidence of updating in these beliefs, as previous work suggests that people do not bring these attitudes in line with other beliefs, even if they are discordant.

Peripheral beliefs that "grow" out of core beliefs, such as preferences for government spending on various priorities and specific policy positions, should show some evidence of updating, as individuals attempt to align these with their core beliefs. However, because of the cognitive effort required to map peripheral beliefs onto core beliefs, we generally expect the former to be inconsistently reported (Zaller 1992; Converse 1964). However, there are conditions that might decrease that cognitive effort and therefore increase the amount of updating and consistency with which peripheral issues are reported, to which we now turn.

Public vs. Private Culture: A final dimension that likely influences the degree of active updating and consistency of attitude and behavior reports is the publicness of the item, or whether it is in some way externalized in public symbols, discourses, and institutions. The act of opinion construction often taps what Lizardo calls "personal declarative culture," the "explicit, symbolically mediated culture" such as language that people use to reason, evaluate, judge, and categorize objects (2017: p. 91). This kind of knowledge is contradictory and flexible, meaning a person can reach and justify a range of different opinions in response to the same prompt, depending on the circumstances. However, when public culture provides a strong "scaffold" through clear signals of how identities and social locations should influence opinions, it becomes easier for people to maintain consistency (Lizardo and Strand 2010).

Theories that argue that individuals do not hold consistent opinions also posit that issues that receive significant media attention can be reported more consistently than other issues (Zaller 1992; Converse 1964). This means a small number of high-profile issues might demonstrate a higher rate of consistency, but we do not expect high rates of active updating for

most of these, since signals are consistent over time. However, we do highlight one issue that saw significant change in elite opinion, which should have lead to active updating in the general population. In the time frame we observe (2006-2014), there was a major shift in elite opinion on gay rights. Leaders of the Democratic Party (including President Barack Obama) came out in favor of legalizing gay marriage, and a number of Republican Party leaders also began to express support for gay marriage. For this issue, we expect to find evidence of active updating throughout the population, not just among young people.

Items that tap some public dimension of behavior and attitudes are also likely to show higher levels of active updating because social mechanisms, such as publicly joining a group, can help maintain changes in ways that intrapsychic forces cannot. Any item where the burden of maintaining consistency is externalized and not only "in a person's head" might demonstrate both higher rates of active updating and greater consistency.

As we noted previously, we do not expect that all items will align with just one model. If we find a diverse array of patterns, with items displaying a range of active updating and consistency, then the above theories provide a window through which to interpret these distinctions. Our goal here is neither to declare victory for one of the theoretical perspectives nor to simply say that all "matter." Rather, our objective is to improve sociological models of cultural evolution by more precisely specifying when and where different types of processes are at work. We believe achieving a better understanding of these processes will be relevant for many sub-fields of sociology as well as for other social science fields that study changes in beliefs and behaviors.

RESEARCH QUESTIONS AND EXPECTATIONS

With these considerations in mind, we ask the following questions:

First, *to what extent are patterns of personal cultural change generally better described by an active updating model or a settled dispositions model?* Previous work on cultural change using cross-sectional data has suggested that cohort effects are generally more important than period effects in explaining broad cultural change (Vaisey and Lizardo 2016). This implies that, in the repeated measures data on adults we will use here, we should find that the settled dispositions model performs better on most items because cohort formation should be (mostly) complete.

Second, *is there evidence that younger respondents are doing more active updating than older respondents?* In a sample of adults, the possibility exists that cohort formation may be complete for most beliefs and behaviors before the study period (i.e., before age 18). However, if some cohort formation is still occurring among younger respondents, we should see that evidence consistent with the AUM is disproportionately located among younger respondents.

Third, *are there systematic differences in item content between questions that exhibit different levels of active updating and consistency?* The preceding sections have made some predictions based on existing literature, but we cannot enumerate predictions for all sorts of beliefs and behaviors. As we note, our approach is to use these theories as a lens through which to interpret the overall pattern of results.

ANALYTIC STRATEGY

To investigate these ideas empirically, we examine 183 survey items from the 2006-2014 General Social Survey panels in search of evidence in favor of an active updating model. This period of the GSS contains three different three-wave panels, each of which surveys a sample of adults three times over a four-year period (e.g., 2008-2010-2012). Three waves of data is the minimum amount needed to compare the predictions of the active updating model and the settled

dispositions model. We discuss item selection below. While a broader range of years would be preferable, the panel component of the GSS began in 2006 and was discontinued in 2014, so these years represent the full range of what we could analyze using the GSS.

Statistical Model

BASIC MODELS

Our main goal is to estimate separately the amount of active updating and the amount of non-persisting change in responses over time. We first consider measuring the amount of active updating. Our two different models make different predictions in the three-wave panel context. The AUM makes the following prediction for wave 3:

$$E(y_{i3}) = y_{i2} \qquad (3)$$

That is, the AUM predicts that a respondent's most recent response is the best available predictor of her next response, and that wave 1 carries no additional information in predicting wave 3 once we control for wave 2. If change is persisting, then our best guess is that a person's response will be close to what they said last time, and previous responses will provide no additional predictive power.

The SDM makes the following prediction:

$$E(y_{i3}) = U_i \qquad (4)$$

Since the best estimate of U_i is the mean of the respondent's two previous answers, we can rewrite the SDM prediction as:

$$E(y_{i3}) = \frac{y_{i2} + y_{i1}}{2} \qquad (5)$$

That is, the SDM predicts that the average response of previous waves is the best predictor of the next response. If change is non-persisting, then taking the average of the last two responses will be our best guess about a person's underlying position.

COMBINED MODEL

Both of the models above include y_{i2} as a predictor of y_{i3}, but only the settled dispositions model includes y_{i1} as a predictor. If the SDM is correct, y_{i1} should be just as predictive as y_{i2} because both are (on average) equally informative about the respondent's stable disposition. Therefore, to test for evidence of active updating we use a model that evaluates whether y_{i2} carries any additional predictive power over y_{i1}. If the two previous estimates are equally predictive, then we can be relatively confident that the data we observe came from a settled dispositions model. However, if y_{i2} is a better predictor of y_{i3} than y_{i1} is, this provides evidence that some respondents engage in active updating. We use the following non-linear model to estimate the relative influence of y_{i1} and y_{i2}:

$$E(y_{i3}) = \alpha + \phi\beta y_{i2} + (1-\phi)\beta y_{i1} \qquad (6)$$

Rather than generate separate coefficient estimates for y_{i2} and y_{i1}, this model generates two parameter estimates of interest: β, which captures how well any combination of previous waves predicts a person's response at wave 3, and ϕ, the relative proportion of wave 3 explained by wave 2 compared to wave 1. If the Settled Dispositions Model is the preferred data-generating process for an item, then both y_{i2} and y_{i1} should be equally predictive of y_{i3}, and ϕ will equal .5, meaning the best estimate of wave 3 is a function of the mean of previous waves, consistent with Equation (5). If the active updating model is present in at least some respondents and wave 1 provides no additional predictive power when we control for wave 2, then ϕ will increase toward 1 to converge with Equation (3) in certain circumstances.

Our estimates of β provide a measure of the consistency of individuals' responses, contingent on the degree of active updating. We can think of this parameter as analogous to an R^2 measure in a traditional linear model, capturing the total "predictiveness" of the model. If

individuals pick a random response at each wave, the best predictor for a person at wave 3 will be the sample average, and $\beta = 0$. If there is little random fluctuation between waves, once the amount of active updating is accounted for, β will approach 1.

COMPARISON TO OTHER APPROACHES

Our model is not the first to measure stability and change in panel data, but existing models make assumptions that eliminate the distinction between data-generating processes we seek to test. Hout and Hastings (2016) use a hierarchical model to measure reliability in GSS responses. This model assumes that there is no change in the underlying latent item other than wave-specific period effects (akin to our settled dispositions model), so the design precludes the possibility of quantifying the level of active updating in an item over time, assuming that this change is just measurement error. These authors also test a structural-equation model designed by Alwin (2007) and Heise (1969) that assumes the process that generates the data is the Markovian active updating process we outline earlier. While this approach gets closer to our model by generating a parameter for stability and reliability, it would require us to make an assumption that the amount change is consistent across waves. This approach also combines persisting and non-persisting change into two similar but distinct kinds of change: "structural" and "non-structural," both of which can be persisting and non-persisting. This distinction, while important for some theoretical questions, is not our focus.

A number of other approaches seek to understand the consistency of latent beliefs by combining and scaling responses to questions that represent the same latent concept (Ansolabehere, Rodden, and Snyder 2008), assuming that wave-to-wave changes in responses tend to represent measurement errors around a "true" latent belief. This raises the distinction between the *stability of a belief* and the *stability of a survey question response*. Since we at times invoke both these models, we include in our analysis several composite scales of related items. If

wave-to-wave changes in survey responses are non-persisting measurement errors, then scales should have higher consistency than the measures they comprise, but we should see no difference in their levels of active updating.

Another possibility is that participation in the survey itself produces change or stability, referred to as panel conditioning bias (Warren and Halpern-Manners 2012; Oh, Yeatman, and Trinitapoli 2019). While Warren and Halpern-Manners (2012) outline several forms of panel conditioning, we can group these into two broad patterns. One set of patterns suggests that people's responses become more consistent over time as participation in the survey forces them to crystallize their beliefs, seek out new information that helps them form beliefs, realize their beliefs are out of sync with the general population, or learn to "game" the survey to get through it faster. If this were taking place in the GSS, it would result in a pattern of high active updating and high consistency, as respondents would change between waves 1 and 2, and wave 2 would become a better predictor of wave 3.

We do not view this as a problem for our theoretical models. If people change their attitudes or behavior as a result of participating in a survey, they are conforming to the theoretical active updating model, being open to change throughout their life course, and the source of that change is irrelevant. This might lead to a higher estimate of active updating than we might observe in a population that did not take the survey, which would hinder our ability to extrapolate our findings, but would still provide evidence that people update beliefs over time.

A second form of panel conditioning posits that people exhibit low levels of updating because of commitment bias, or an attempt to maintain consistency in their responses over time, even if they actually change. In this scenario, individuals respond to a question at wave 1 and give the same response in subsequent waves, even if their true beliefs or positions change. This

would be problematic for our study, as it would under-estimate the amount of real change in the population. However, if this is the case, we should observe no active updating in responses and high levels of consistency, as it would be illogical for individuals to report random changes if they were attempting to maintain consistency.

Finally, a number of approaches exist for evaluating theoretical process of belief formation and change for the population, such as examining the association between theoretically related values (Baldassarri and Gelman 2008; Boutyline and Vaisey 2017), or by looking at changes in the distribution of responses over time (DiMaggio, Evans, and Bryson 1996). These tools are well designed to address the questions they set out to answer. However, since our theoretical questions focus on the process of belief change within individuals, these do not speak to our core concerns.

LIMITATIONS OF THE METHOD

Three challenges limit our ability to evaluate the presence of settled dispositions and active updating models using our approach and therefore limit the conclusions we can reach. The first is that our model is designed to allocate variance explained to each of the prior waves rather than to assign probabilities to each data-generating process. Because of this, a few individuals making large persisting changes can inflate the ϕ estimate even if most individuals make small non-persisting changes.

The second is measurement error, which is a form of non-persisting change. For some researchers, measurement error represents the inconsistency that results from constructing responses anew each wave (Zaller 1992; Converse 1964), and in that case it should not be considered "error" so much as an indicator of that process at work, since there is no "true" item to measure. For other theories, measurement error reflects individuals' inability to accurately

report their response. It is also possible that measurement error reflects errors of selection and interpretation, such as misunderstanding the question or incorrect coding.

Because measurement error looks the same as non-persisting real change (and because the latter is sometimes interpreted as the former), ϕ estimates will be biased toward .5, since responses with error would be random departures from baseline. There is evidence that many of the items explored in our analysis are measured with significant error (Alwin 2007; Hout and Hastings 2016). On the other hand, previous studies of reliability tend to conflate measurement error and non-persisting real change in attitudes, meaning that while we might have good estimates for the combination of these two processes, we cannot separate them. Because of measurement error, it is unlikely that ϕ and β will reach 1 for any item, even if the underlying process is fully based on active updating.

The third challenge of our analysis is that we focus on predicting wave 3. If individuals have a high likelihood of changing between waves 2 and 3, our ability to predict responses at wave 3 will be limited and β will be low. Our model relies on the assumption that "persisting" change is relatively rare and that individuals who change between waves 1 and 2 do not also make persisting changes in the opposite direction between waves 2 and 3. If the rate of active updating is so high that individuals make changes between each wave, then the model becomes indistinguishable from the settled dispositions model with high measurement error, and it may not be reasonable to consider this sort of change "persistent."

In addition to these three challenges, there are two forms of change that our model is not well designed to account for. The first is a uni-directional shock to the population. Since our model includes an intercept, changes that shift all responses toward one end of the scale are absorbed into that term and not accounted for in our ϕ estimates. The second is change in the

variance of responses. If all individuals shift outward or inward toward the population mean but maintain their relative position in the overall distribution, this change will be absorbed into β but not enter into ϕ.

Despite these limitations, the model is capable of detecting the presence of persisting change even in the presence of high levels of measurement error. Because of this, it is best to think of our approach as seeking any evidence in favor of active updating, rather than allocating probabilities to each model. We can only detect whether there is any evidence of persisting belief changes, and therefore whether there is any evidence that active updating is taking place in the population.

Analysis Steps

Our analysis proceeds in three steps to answer our three research questions. First, we evaluate the overall evidence in support of the active updating model. To do this, we compare for each item the Bayesian Information Criteria (BIC) of a model estimated using Equation (6) with a free estimate of ϕ to a model that constrains $\phi = .5$. We calculate the posterior probability that the model with the free parameter fits the data better. If the model with the constraint is preferred, then we conclude that both wave 1 and wave 2 are equally good predictors of wave 3, meaning there is no evidence that respondents are actively updating on that item.[1]

Second, for variables that show at least some evidence of active updating ($\phi > 0.5$), we ask whether the persistent change is concentrated among younger respondents. To test this, we re-estimate our original model and allow ϕ to have different values above and below a given age

[1] All data used in this analysis is publicly available through the General Social Survey at the National Opinion Research Center at http://gss.norc.org/. R code for how we cleaned the GSS panels, all analyses performed in this paper, and all figures created in this paper can be found at http://github.com/krkiley/panel_change.

cutoff. Rather than test a single age cutoff, we again use BIC comparisons to evaluate whether including the dummy variable improves the model fit using a cutoff of every age between 20 and 45. We test a range of cutoffs to ensure robustness of the overall pattern to specific ages.

Finally we consider whether there are any meaningful patterns in the relative distribution of evidence for active updating across variables as suggested by existing theories. Although, as we discussed above, previous work gives some indications about what we might expect, the approach here will necessarily be inductive.

Item Selection

To test our model on as broad a range of items as possible, we sought measures of attitudes, beliefs, self-assessments, self-perceptions, and social behaviors that were asked in three waves of the GSS panels. We excluded from our analysis questions that focused on demographic characteristics (marital status, household size, region, gender, race, ethnicity), work activity (employment status, income, hours worked, size of workplace), objective socioeconomic status (years of education and highest degree, home ownership), and evaluations of a respondent by the interviewer. We follow Hout and Hastings (2016) and group questions into 15 categories based on subject material. Questions in the same category tend to be asked in the same block during the survey and have the same structure, such as questions about confidence in institutions, questions about government spending, and questions about social life.

We also follow Hout and Hastings (2016) in recreating common scales about gender roles, access to abortion, and social trust. This includes a six-question scale of support for abortion and a seven-question scale which includes the question asking about abortion under any circumstances ("abany"). We use Smith's (1997) scale of "misanthropy" by combining questions about how helpful, fair, and trustworthy people are. We use four questions to create a scale of

gender role attitudes (Cotter, Hermsen, and Vanneman 2011). Like Hout and Hastings, we combine civil liberties items into six scales about the freedom of atheists, communists, militarists, racists, and, in the 2010-14 panel, Muslim clergy. We combine four parallel questions about how frequently individuals socialize to create a "social life" scale. We combine four questions about support for suicide under different circumstances. We also created a scale of support for police use of violence against criminal suspects by combining five binary questions about the conditions under which individuals support police use of violence.

In total, we test the model on 183 GSS items, including the composite scales.[2] For each question, we use all cases for which the respondent gave responses in all three waves. Models are estimated using weights that account for the GSS's sampling design as well as non-response adjustment.

RESULTS

Our model estimates two parameters of interest for each GSS item: β, our measure of consistency, captures how well any combination of previous waves predicts a person's response at wave 3. High values of β indicate that individuals are relatively consistent in their responses, once we control for the amount of active updating. Our measure of active updating, ϕ, captures the relative proportion of wave 3 variance predicted by wave 2. If responses are generated through a true settled dispositions model, then ϕ will be .5 (i.e., both wave 1 and wave 2 are equally good predictors of wave 3). As the evidence of active updating increases, ϕ will increase toward 1. Both ϕ and β equaling 1 would indicate that all individuals who changed between

[2] To ensure that our estimates of ϕ are not simply artifacts of response scale construction, we estimate the model on coarsened versions of items, generated by collapsing responses to questions with more than three response options into scales of two or three response options. These results are reported in Appendix B.

waves 1 and 2 persisted in their change, that an item was measured with no measurement error, and there was no additional change between waves 2 and 3.

Evidence for Active Updating

To evaluate the evidence in favor of the active updating model, we compare for all 183 items the Bayesian Information Criteria (BIC) of a model with a free estimate of ϕ to a model that constrains $\phi = .5$. If the model with the constraint is preferred, then there is no evidence that respondents engage in an active updating process with respect to that item.

Figure 2 plots the distribution of ϕ estimates for the 183 items evaluated in this analysis and the posterior probabilities that the model without the constraint fits the data better, generated by comparing the BIC from models with and without the constraint.

[Figure 2 about here.]

On the left side of the figure, we see that the majority of ϕ estimates fall between .5 and .6, meaning that wave 2 is only a slightly better predictor than wave 1 for most items. This suggests that if active updating is happening in these responses, it is relatively infrequent or small compared to temporary change and measurement error.

To provide a concrete example, consider the GSS question that asks respondents whether they think it should be possible for a woman to receive a legal abortion if she became pregnant as a result of rape, to which individuals can respond either "yes" or "no." This produces a ϕ estimate of 0.62, above the 75th percentile of all ϕ estimates. Of the 2259 people who responded to the question in three waves, 257 changed between waves 1 and 2. Under the settled disposition model, these responses would reflect either measurement error or a temporary shift at either wave 1 or wave 2, and we would expect that about 50 percent or 129 individuals would maintain the same response into wave 3. Only 147 of the 257 people (57 percent) who changed between

waves 1 and 2 maintained the same response at wave 3. Therefore, we only have evidence that about 18 individuals (less than one percent of the sample) showed evidence of persisting change.

Considered this way, the majority of GSS items demonstrate persisting change at a rate of less than one percent of the total sample, and none shows evidence of persisting change greater than 5 percent, with confidence in the leadership of the executive branch of the federal government, confidence in banks and the financial system, and the respondent's belief about whether they will be able to find a good job topping the list. In other words, even for items that show strong evidence of some active updating, the overall amount of attitude change in the population is likely small.

The right side of Figure 2 shows that although the majority of items prefer the free parameter, 75 items (about 40 percent of the total) prefer the constraint, meaning these items show no evidence of active updating over this period. That is, although respondents might give different answers to these items in any particular wave because of measurement error or a transient change of opinion, they tend to revert to their previous position. This group includes many items about abortion, civil liberties, confidence in institutions, and views on race and gender.

We will discuss in more detail below how different items perform. To answer our first question, however, we need only focus on the overall distribution. Forty percent of items show no evidence for active updating, and among items that do show some evidence of persistent change, very few come close to approaching 1. This means that for almost all items, measurement error or non-persistent change tend to be much more common than persistent change. We can only be really confident in detecting substantial amounts of persistent change (greater than 2 percent of the population) among a small minority of items, perhaps 1 in 5. This

means that most of the "change" that shows up in the GSS panels reflects some combination of measurement error or non-persistent change.

Age Heterogeneity

Our second research question asks whether there is evidence that younger respondents update their views more than older respondents. Although it's impossible to determine what proportion of people are following each data-generating process, it is possible to compare the age distribution of evidence for updating in each item.

Of the 108 items that showed any evidence for active updating in the last section, 22 showed differential effects of age for more than 50 percent of cutoff ages we tested, meaning the majority did not. Figure 3 plots the estimates of these 22 items for individuals above 30 and individuals equal to or below 30 to get a sense of the magnitude of difference between older and younger individuals on these items.

[Figure 3 about here.]

The majority of items that show evidence for age concentration show that active updating is more prevalent among younger respondents than older respondents. These items include views on affirmative action, women in the workforce, and politics; several civil liberties items; general views of whether people can be trusted; and views on whether doctors should let terminal patients die. These items tend to be in subject areas where a large proportion showed no evidence of active updating, which suggests an overall trend of these views being formed earlier in life (i.e., prior to becoming eligible for the GSS at 18) and remaining relatively stable over time.

For some items, such as whether individuals can be trusted, political views, whether physicians should allow terminal patients to die, and whether companies should make special efforts to hire and promote women to address past discrimination, all evidence of active updating

disappears for people over 30. This suggests that these items follow an "impressionable years" pattern, where early adulthood is a time when these opinions are still malleable, but beliefs quickly harden into "durable dispositions" (Vaisey and Lizardo 2016; Krosnick and Alwin 1989; Alwin and Krosnick 1991). For other items, such as how important people believe it is for children to be popular and views on how much the government should spend on health care, there is still evidence of active updating in older individuals even though it is substantially less than for younger individuals. This is consistent with the "increasing persistence" hypothesis, where attitude change gradually becomes less likely as individuals age (Glenn 1974; Inglehart and Baker 2000).

Eight items show a negative effect of being below the age cutoff on the ϕ value, meaning that younger people showed *less* evidence of active updating than older individuals. These items include how often individuals were active in religious activities, views on suicide in the case of bankruptcy, and views on whether aging parents should live with their children. Some of these items might be things people are not forced to consider until later in life and as a result do not form clear opinions on while young. This pattern where older individuals change their attitudes and behaviors at higher rates than younger individuals is somewhat unanticipated in the attitude change literature (Visser and Krosnick 1998; Danigelis, Hardy, and Cutler 2007), and suggests greater heterogeneity in the relationship between age and attitude change than previously theorized.

The remaining 86 items (just under half of all items explored here) show evidence for some active updating but do not show consistent evidence for age heterogeneity, suggesting a more complicated relationship between age and attitude change than previously theorized.

However, this does not mean these show strong evidence of an active updating model. These items may simply be susceptible to updating for a small proportion of the population.

Item heterogeneity

While 40 percent of items show no evidence of active updating, and those that do show evidence tend to show only weak support for active updating, it is difficult with just these findings to draw any broad conclusions about how these results speak to theories of attitude development and change. Here we bring in our second dimension of attitude change, consistency in responses, to clarify the overall pattern.

Figure 4 plots the ϕ and β estimates for items shaded by whether they preferred the $\phi =$ .5 constraint or not and label a few items that stand out. Items tend to prefer the $\phi = 0.5$ constraint for a combination of two reasons: because wave 1 and wave 2 have equal predictive power (ϕ is close to .5) or because the measure is so unpredictable (β is low) that neither wave 1 nor wave 2 has much predictive power, making any observed active updating close to meaningless.

[Figure 4 about here.]

Items showing evidence of active updating tend to have ϕ estimates greater than .55, and most have β estimates greater than .6. A small group of variables, including confidence in the leadership of the executive branch of the federal government have low β estimates, meaning that prediction at wave 3 is difficult, but have large ϕ estimates, meaning that wave 2 is still a better predictor than wave 1.

Which items show the strongest evidence for active updating? There is no way we can discuss all 183 items in detail without the discussion becoming tedious. Although we include ϕ estimates for all items in Appendix A, Figure 5 summarizes the distributions of ϕ by the content

of the question. We constrain items that showed no evidence of active updating to $\phi = .5$. In addition to showing the median and interquartile range of each distribution, the figure also highlights the item in each group that shows the greatest degree of evidence for active updating.

[Figure 5 about here.]

There is a lot to process, even in this summary figure. The main takeaway is that, even for items that show some evidence for active updating, the values of ϕ are still quite low. Only two groups of items have median ϕ values above 0.6: public spending and religious activity. In general, then, it is accurate to say that most of the "change" measured by the GSS is not persistent but some combination of measurement error and short-term fluctuations.[3]

Consistent with the findings of Vaisey and Lizardo (2016), we see that more than half of items in the gender, family, race, sex, civil liberties, and confidence in institutions groups show no evidence for updating (see Appendix B). These categories also contain several items that show evidence of active updating in younger cohorts, suggesting that these items became settled by the time respondents entered the GSS sample. Views on these issues are likely shaped by early socialization experiences and mostly settled by the time a respondent reaches adulthood. This means that, for most of these items, social change must occur through cohort succession rather than through individual change.

---

[3] While we cannot control for measurement error in our analysis, we can take steps to mitigate its impact. Appendix B presents results comparing items with more than three scale points to coarsened versions of these question with either two or three scale points. As noted previously, low measurement error might be a reasonable assumption for some items. Previous studies using different approaches to measuring the reliability of survey reports suggest that some items captured in our study, such as whether a person owns a gun, are measured with a high degree of reliability (> 0.9) (Hout and Hastings 2016). For other items, such as confidence in the leadership of major companies ($\phi = .58$), reliability might be as low as 0.5. There is very little correlation ($\rho = .165$) between our ϕ estimates and reliability estimates.

There are several exceptions to this general pattern, however, even in categories with otherwise low ϕ values. Items with the largest values generally have one or more attributes in common. We consider these attributes to give some general impressions of the pattern.

Some of the high-ϕ items rely on *external mechanisms* that help maintain them. If a person starts going to church or starts socializing with friends at a bar, she builds social networks that make this behavior more likely to continue. This is clear when contrasted with how often individuals socialize with friends, relatives, or neighbors, which are more nebulous questions that display less active updating. Switching political parties (which involves changing public registration) is a more persistent change than changing political ideology (which can happen privately in the mind). Owning a gun has a high ϕ value because a new physical object either enters or leaves the person's possession.

Other high-ϕ items have a *changing referent*. That is, although the item wording is the same, the object to which the question refers may change between survey waves. The most obvious example of this is the item about confidence in the executive branch of the federal government (which has the highest ϕ value of all items in the analysis). The president changed between the 2008 and 2010 waves of the GSS, meaning that the question no longer referred to the same administration. If we generate estimates for this item for each of the three panels (2006-10, 2008-12, and 2010-14) for the "confidence in the leadership of the executive branch of the federal government" item, it is only for the middle panel (2008-12), where the president changed between waves 1 and 2, that shows significant evidence of persistent change ($\phi = .95, \beta = .51$). In the 2006-10 panel, waves 1 and 2 have almost no predictive power ($\beta = .10$). For the 2010-14 panel, which takes place entirely during the Obama administration, ϕ moves much closer to .5 and predictive power increases ($\phi = .57, \beta = .65$).

Likewise, all public spending items refer to whether we're spending "too much, too little, or the right amount" on different areas. The change of administration and changing federal spending policies likely affected these items. The same applies to most of the questions about subjective SES, where the respondent is asked questions about her personal financial or work situation (which changed for many Americans during the time of the study due to the Great Recession). If the environment is changing we should see exactly this sort of pattern.

Perhaps the most striking pattern in our findings pertains to questions about *gay rights*. There are 6 items that ask about some aspect of gay rights, and all show evidence for active updating. Questions about civil liberties for gays and about gay marriage are the highest in their categories. The huge public and political salience of this issue throughout the study period likely made this issue one where more people than usual were open to revising their views. This pattern is consistent with Zaller's (1992) argument that highly salient issues where elite opinion shifts can lead to large changes in public opinion.

A few items lack external mechanisms, represent forms of private culture, and did not achieve salience during this period, but still display active updating, including some views on abortion and the morality of different forms of sex. We address these items in greater detail in the discussion section.

Political Beliefs

Because theories of public opinion formation and the development of political beliefs form the bulk of the theoretical tension that frames this analysis, and because many of these beliefs display evidence of active updating, we examine these beliefs in greater detail than other items. Figure 6 plots the ϕ and β estimates for questions about political identity, the role of

government, and specific policies. We break these items up into categories and remove a few items for ease of viewing.

[Figure 6 about here.]

There are several notable features of the figure. First, political party affiliation is a clear outlier, with greater consistency and greater active updating than other items. In contrast, questions asking about specific public spending priorities tend to have low consistency compared to other items. Political views – including ideological identification (Named "Political views" in the figure) and general views of the role of government ("Gov't do more or less", "Gov't reduce inequality") – display higher consistency than specific policy questions but weak evidence of active updating. It is important to restate that for political views, all evidence of active updating disappears by the time individuals reach 30, and individuals become more consistent in reporting their ideological identity as they age.

While people tend to be inconsistent in their responses to spending priorities, views on other policy questions with both high and low salience are reported quite consistently. Views on abortion, especially when considered as a scale, are consistently reported. As with all other scales, aggregating the composite items increases consistency but does not affect the estimate of the amount of active updating, validating our method's ability to separate persisting and non-persisting change. As discussed previously, views on gay marriage display high levels of active updating. In contrast, the item asking about support for legalizing marijuana use is reported with consistency but weak active updating. This is notable, as public opinion and policies regarding both have shifted considerably in recent years. Previous work suggests that change in both could be driven by the same underlying process (Schnabel and Sevell 2017). Our findings suggest that,

at least in recent years, many individuals have changed their views on gay marriage, but changing views on marijuana have been driven primarily by cohort replacement.

DISCUSSION AND CONCLUSIONS

This study was motivated by a theoretical contrast in cultural sociology between the settled dispositions model (SDM), which emphasizes the power of the past, and the active updating model (AUM), which emphasizes contemporary meaning making. These two models of individual change are implicit in analyses of belief change at the society level, but they are rarely compared empirically. We asked three general questions in light of this distinction and developed an empirical approach to adjudicating (albeit imperfectly) between these models. We now revisit our research questions to summarize what we have learned.

First, *to what extent are patterns of cultural change generally better described by an active updating model or a settled dispositions model?* In general terms, consistent with previous research (e.g., Vaisey and Lizardo 2016), we see a greater degree of evidence in support of the settled dispositions model. Around 40 percent of all items show no evidence for updating. For the items that do show evidence of active updating, the overall rate of persisting change in the population is likely low. For most items measured by the GSS, less than one percent of the population appears to make any persisting change in their views in a two-year period, and most changes appear to be short term or random deviations. What that means in practice is that knowing what a person said two years ago provides almost no better prediction of their current views than knowing what they said four years ago. Unfortunately, because of measurement error, we cannot be sure exactly how much updating there is. But the average level of updating for most views appears to be low.

Second, *is there evidence that younger respondents are doing more active updating than older respondents?* For a limited subset of items, there seems to be evidence that younger respondents are updating their views whereas older respondents are not. This is consistent with a "cohortization" model that views young respondents as susceptible to updating shocks and older respondents as relatively insensitive to such shocks (see e.g., Bartels and Jackman 2014). Because the youngest respondents in the GSS are 18 years old, we may lack the ability to detect updating that occurs even earlier, so this is probably an underestimate. This may be the case for the roughly 70 items that show no evidence of persistent change over time. Overall, however, there is significant heterogeneity in the relationship between age and attitude and behavior change.

Third, *are there systematic differences in item content between questions that are better described by each model?* With 183 items tapping very different kinds of opinions and demonstrating a range of active updating and consistency, there is no one overall pattern of attitude change in the population. We find patterns of responses that provide a range of support for diverse theories of attitude formation, which we discuss below. At the same time, our results are consistent with some general expectations about updating. Questions with public (or otherwise changing) referents and questions tapping high-salience topics over the study period (such as gay rights) showed the most evidence for active updating. Items assumed to be "peripheral" beliefs, such as specific policy questions, also showed evidence for updating and low consistency. Most items about gender, family, race, and institutions showed the least evidence for updating, suggesting most people's views on those topics are settled by the time they turn 18.

There are a handful of items that demonstrate active updating and enough consistency to merit further consideration, but are not easily explained by theoretical mechanisms outlined so far. These include questions about abortion in the case of rape or poverty, the morality of premarital and teen sex, the ideal number of children in a family, and whether children should obey adults or think for themselves. These items did not achieve high salience during the time of the study and there were not, to the best of our knowledge, clear changes in elite opinion on these issues. This suggests additional mechanisms that can drive persisting change at the individual level, such as exposure to these ideas through mass media or by directly interacting with people in these groups. In the case of questions about children and family structures, it might be the experience of having children that leads to changes in these attitudes. It is important to note that even though these items demonstrate active updating, it is a small proportion of the population that are changing their views during the study window.

The overall pattern of results, while diverse and at times hard to reconcile, does shed light on several theoretical debates. We now consider these implications.

Implications for Cultural Sociology

In the domain of cultural theory, our findings support the view that a great deal of cultural change happens slowly through the mechanism of cohort succession. Most beliefs about gender roles, sexual morality, and abortion appear to be settled by early adulthood. The settled nature of these beliefs is often coupled with a high rate of consistency, suggesting that individuals truly hold these beliefs or at least have sufficient external support to consistently report them over time. In contrast, many views about race were so inconsistent that it would be difficult to call them either settled or updating. Even those items that did display strong evidence

for active updating (e.g. “whites rich or poor” and “whites work hard”) were still hard to predict wave after wave.

Our results suggest that one reason attitudes are largely stable is because most issues simply do not reach the level of salience necessary to shift opinions. In contrast to other beliefs, the pattern of findings for gay rights show that a high degree of public salience and social movements can accelerate change by encouraging people to update their views. By definition, salience is a limited resource, meaning only a few beliefs and behaviors could change at this rate during any given period. The baseline process of attitude change appears to be more consistent with a model that shows that people do not really change; rather, they die and are replaced by cohorts with different views. This general model is more consistent with a Bourdieusian theory that emphasizes the “conditions of past production” rather than processes of active meaning construction with little long-term memory.

While the dominant pattern is stability during adulthood, the results suggest greater heterogeneity in the relationship between age and attitude change than what is emphasized in existing theories. The most prominent and well supported theories of attitude change suggest a peak of susceptibility to attitude change early in adulthood and either a rapid or gradual decline in attitude change with age (Alwin and Krosnick 1991; Visser and Krosnick 1998). An additional view, the “life stages” hypothesis, suggests a similar pattern with a surge in attitude change late in life (Visser and Krosnick 1998). While this age-related decline appears to be true for many political attitudes, the pattern is not nearly as consistent for other kinds of attitudes. Some items, such as views on the Bible, suggest equal openness to attitude change as individuals age. Others, such as views on most abortion questions, suggest that early adult socialization is so strong that lifetime opinions are settled by the time most people reach even 18 years old.

For some beliefs, such as whether aging parents should live with their adult children and whether divorce laws are too lenient, persistent change becomes more common with age, a pattern that is not accounted for in any major hypothesis about the relationship between age and attitude change. This pattern suggests that the salience of an issue can matter at the individual level, as well as the societal level. Rather than supporting a single theory linking age to attitude change, our results call for more work linking attitude content to social factors that encourage openness to change at different ages. Rather than assuming that "attitudes" in general are more or less likely to change at particular ages, we should explore the relationship between age and stability for a range of attitudes. Doing so will expand our understanding of the institutional and developmental factors that that give rise to stable or variable attitudes (Howe and Krosnick 2017). Recent work that has attempted to provide social explanations for (mostly political) attitude stability in middle age is an important step in that direction (Eaton et al. 2009; Visser and Mirabile 2004).

Implications for Political Sociology

A major takeaway of our analysis is that ideological identity (identity as a liberal or conservative, and the extremity of this identification) was in all practical terms stable for individuals over 30. Respondents might express different positions from wave to wave, but in guessing what a person will say in the future, we are better off guessing the mean of their previous responses than their most recent response. While our sample does not cover a large enough window of time or the life course to say for certain whether this represents a regular pattern (perhaps there is some period-specific reason that younger individuals in our data changed while older individuals did not), the pattern is consistent with theories and previous findings that political dispositions become settled by age 30. With the exception of some low-

profile government spending questions, most policy questions show greater evidence of active updating than this question of ideological identification, though the overall level of active updating is still limited. Partisan identification, in contrast, showed the highest degree of active updating of all political questions, as well as some of the highest consistency, and this updating was active across all age ranges.

These results are consistent with theories that posit that, at least in the current era, ideological identification (view of oneself as liberal or conservative), rather than some other belief such as partisan identification, moral views, religion, or particular policy positions, is the "central" political belief (Boutyline and Vaisey 2017; Converse 1964).

It is important not to overstate the role of this belief in forming other beliefs. While our findings suggest that individuals appear to bring their partisan identification in line with their ideological identification, there is limited evidence of adjusting other beliefs. This pattern is most consistent with Baldassari and Gelman's (2008) model of "partisans without constraint," which suggests that individuals hold a few strong beliefs and align their partisan identification with these, rather than adopting beliefs as a function of their partisan identification.

Since there is no evidence of ideological change for the majority of the sample, and only very weak evidence of changes in specific policy positions, even high profile ones, the pattern of results is not consistent with the popular conception of political polarization in which individuals become more extreme in their views over time. This is notable given the time frame of our study, which covered the Obama administration, a time that is commonly assumed to observe a conservative shift for Republicans and a liberal shift for Democrats. Similarly, our findings are not consistent with the idea that individuals adopt a partisan affiliation based on their social groupings and subsequently adjust their ideological commitments to conform to that (Green,

Palmquist, and Schickler 2002). For this to be true, ideological identification and other political beliefs would have to demonstrate at least as much updating as partisan identification, which is not true in our data.

The finding that ideological identification is comparatively stable should not be taken to imply that most people in the population have the kind of tightly knit belief structures that political scientists typically call "ideologies." Outside of a handful of high-profile items such as partisan identity, abortion, and gay marriage, individuals appear to lack clear opinions on most specific policy questions. The low degree of consistency in beliefs is consistent with the idea that people are "ideologically innocent" (Kinder and Kalmoe 2017). What we can say is that individuals over 30 in American during the period we studied did not make lasting changes in whether and how strongly they thought of themselves as "liberal" or "conservative."

The overall picture that emerges from evaluating the active updating and stability of political items in the GSS is one in which the majority of respondents, especially those over 30 years old, hold a general political identity and a few clear views on issues like abortion, attempt to align their partisan identification to their views, and respond to elite opinion change when it provides clear signals. It is not a picture of a rapidly polarizing society or one wholly ignorant of public debates.

Methodological Implications

Methodologically, our results highlight the challenges of evaluating population-wide attitude change using short-term panel studies. The evidence strongly suggests that most of what might be interpreted as "change" in the GSS panels is some combination of measurement error or non-persistent change. It does not matter whether measurement error or short-term change is the predominant driver behind this pattern; what matters is that substantive change is too rare in a

sample of adults to measure accurately on the vast majority of items. This strongly argues against using two-wave panels to measure attitude change, which do not allow researchers to separate persisting from non-persisting changes.

The fact that persistent change is practically nonexistent for many items bolsters the case for using repeated survey responses to measure the reliability of survey items (Alwin 2007; Hout and Hastings 2016), since it is often a valid assumption that the underlying view is unchanging. At the same time, our results call for greater focus on methodological tools that can separate short-term attitude change from measurement error. While we generally assume that lasting changes in attitudes are more likely to influence behavior, this is not necessarily true. Short-term attitude changes might be meaningful in shaping short-term behaviors, but identifying it is difficult.

We said at the outset that patterns of change might be the result of panel conditioning, or that the process of participating in the survey leads to more active updating or stability than might be expected in the absence of survey participation. One could view our results through this prism and claim that items that exhibit high active updating and high consistency (such as views on gay marriage or partisan identification) do so because of panel conditioning bias, or that items that exhibit low active updating and high consistency (such as views on abortion or the legalization of marijuana) do so because of commitment bias, but it becomes difficult to explain why these biases operate for specific questions and not others.

We believe the overall pattern of results we observe is more consistent with other theoretical models of belief change than those outlined by panel conditioning. We see too much inconsistency in responses for commitment bias to be a major explanatory factor. Items that we have theoretical reasons to suspect might succumb to the updating form of panel conditioning

bias, such as questions where "respondents' initial attitudes are less crystallized" (Warren and Halpern-Manners 2012: p. 499), questions that "increase respondents' knowledge of the behavior and/or their motivation to engage in it" (Warren and Halpern-Manners 2012: p. 500), or questions that "induce respondents to provide socially non-normative or stigmatized responses" (Warren and Halpern-Manners 2012: p. 501), tend to show low active updating and low consistency.

We cannot (and would not want to) rule out the possibility that panel conditioning is taking place in the GSS. We believe it is worthwhile to explore these same GSS panels for evidence of panel conditioning. However, we do not believe that panel conditioning bias is the principal driver of the overall pattern of change and consistency we observe.

Because many attitudes, including views on abortion, race, gender roles, social trust, and institutional confidence, have mostly stabilized by the time individuals enter the GSS, our results also call for greater emphasis on surveying the attitudes of adolescents and children to understand how these attitudes are formed. Panel studies tracing the political socialization of adolescents are rare but could be highly fruitful. In a similar vein, it does not seem worthwhile to ask certain GSS questions repeatedly. Questions about racial stereotypes, which show almost no consistency from wave to wave but have been asked every wave since 1996, strike us as particularly problematic. Those that are repeated should be specifically targeted to topics that are believed to be changing broadly (e.g., politics, gay rights).

Our results ultimately suggest that real, persistent attitude change is an uncommon phenomenon among adults. Understanding the social origins of individuals' attitudes requires greater focus on the "conditions of past production"—childhood and adolescence—that give rise to persistent beliefs in adulthood.

APPENDIX A: ϕ VALUES FOR ALL VARIABLES

Figures 7 through 10 plot ϕ estimates for all items included in the analysis, grouped by subject material.

[Figures 7 through 10 about here.]

APPENDIX B: COARSENED VARIABLE ESTIMATES

Estiates of ϕ for coarsened items are presented in Figures 11 through 13. If scales had symmetrical scales with no clear midpoint (e.g., strongly agree/agree/disagree/strongly disagree), we coarsened those scales to two points (agree/disagree). If scales included a clear midpoint, we included that and coarsened responses on either side. For items on large scales such as hours of TV watched, we split responses into greater than the median or less than or equal to the median. If changes between ends of each scale are more persistent than changes within ends of these scales, then ϕ estimates of coarsened item should be higher than estimates for items with more response options.

[Figures 11 through 13 about here.]

Generally speaking, most items with no evidence of active updating continue to have no evidence of active updating when coarsened. For items that showed evidence for active updating, the coarsened estimates tended to be very similar to the uncoarsened estimates, suggesting that persisting and non-persisting changes happen about as often between ends of these scales as they do within ends of these scales.

There are a couple notable departures from this general pattern. Several political questions – general political views, views on whether the government should reduce inequality, and views on whether the government should help blacks – show decreased evidence of persisting change when coarsened. This suggests that changes around the midpoint of the scale

tend to be measurement error, a finding that is consistent with previous work suggesting that individuals without settled political views tend to choose points around the middle of the response scale (Converse 1964).

REFERENCES

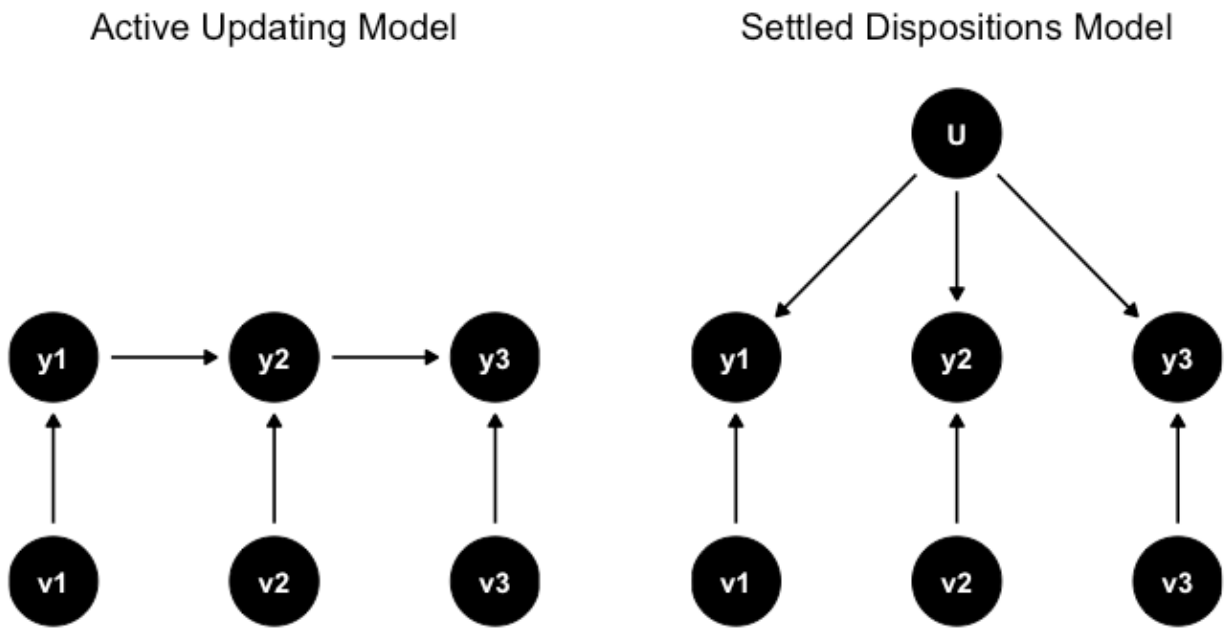

Figure 1: Graphical representation of active updating and settled dispositions models over time.

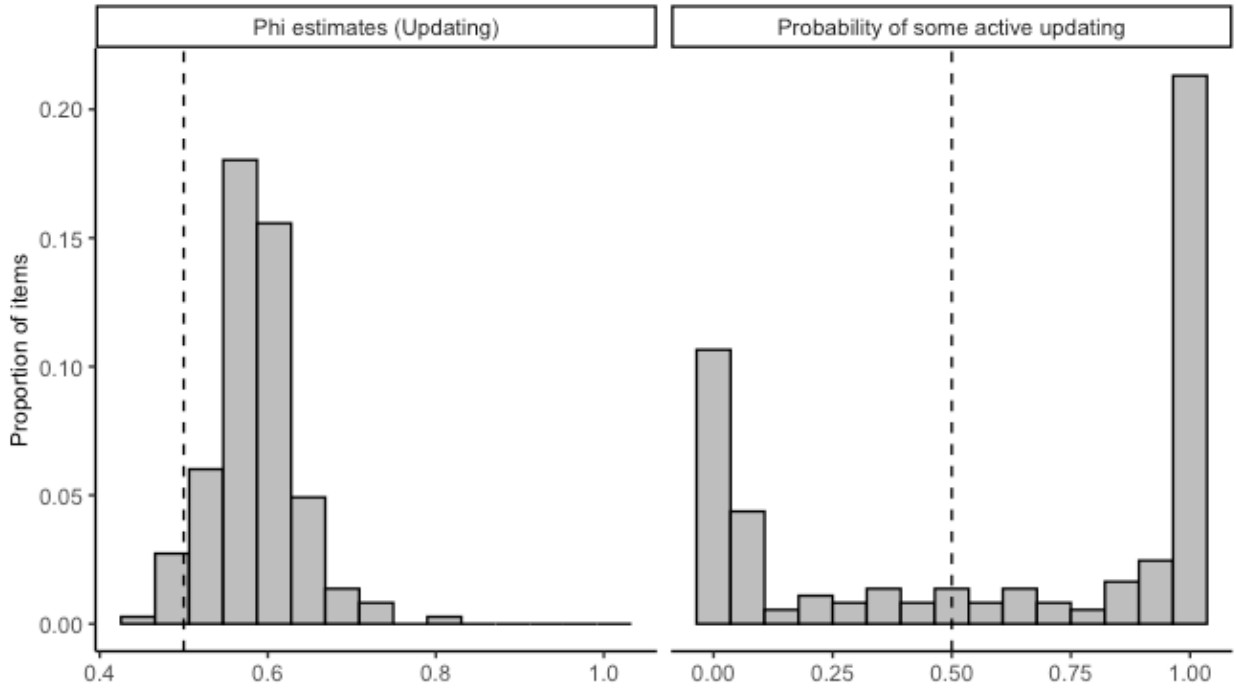

Figure 2: Distribution of phi estimates and probabilities that items show evidence of active updating.

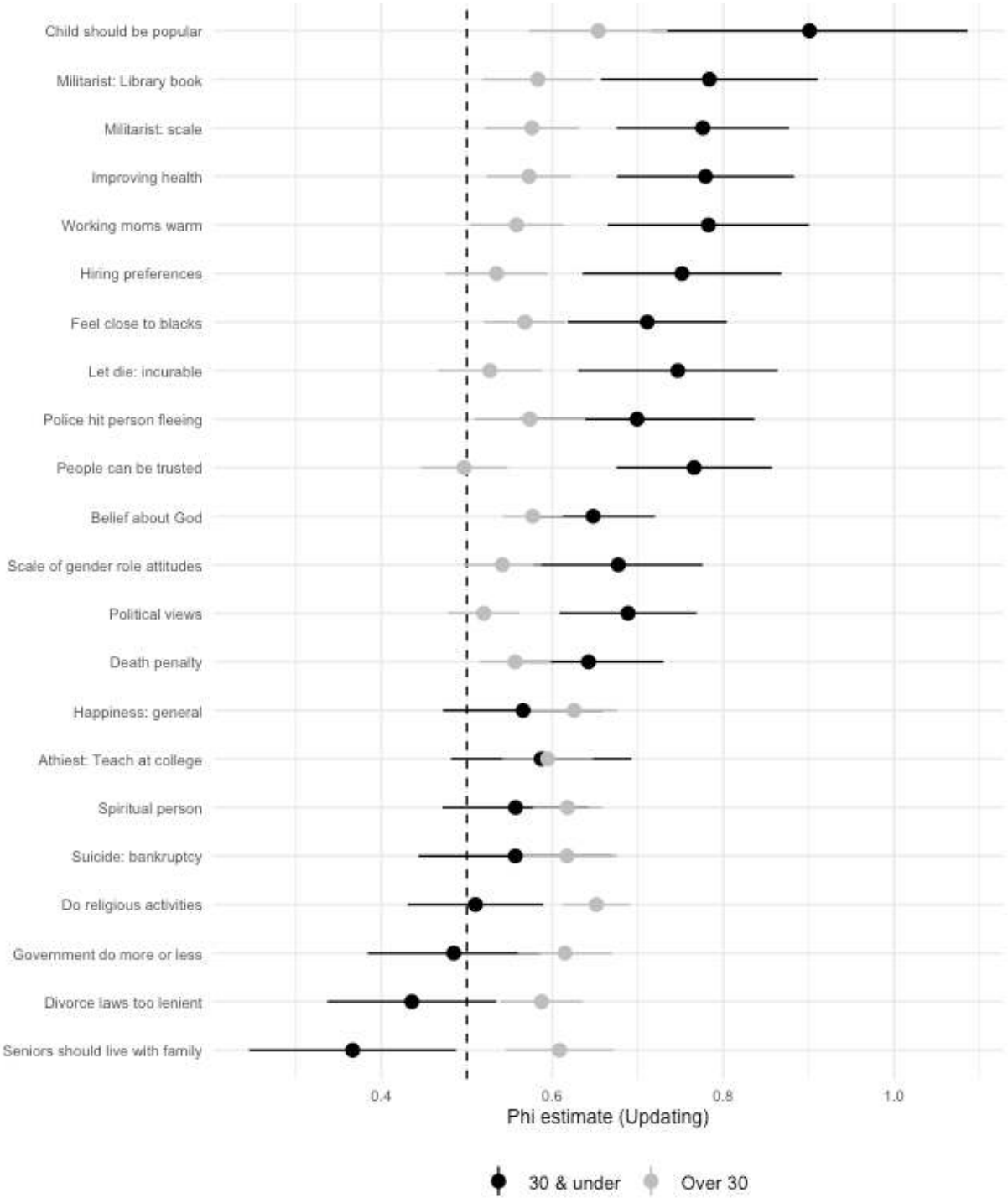

Figure 3: Comparison of phi estimates for individuals over and equal to or less than 30 years old, with 95 percent confidence intervals.

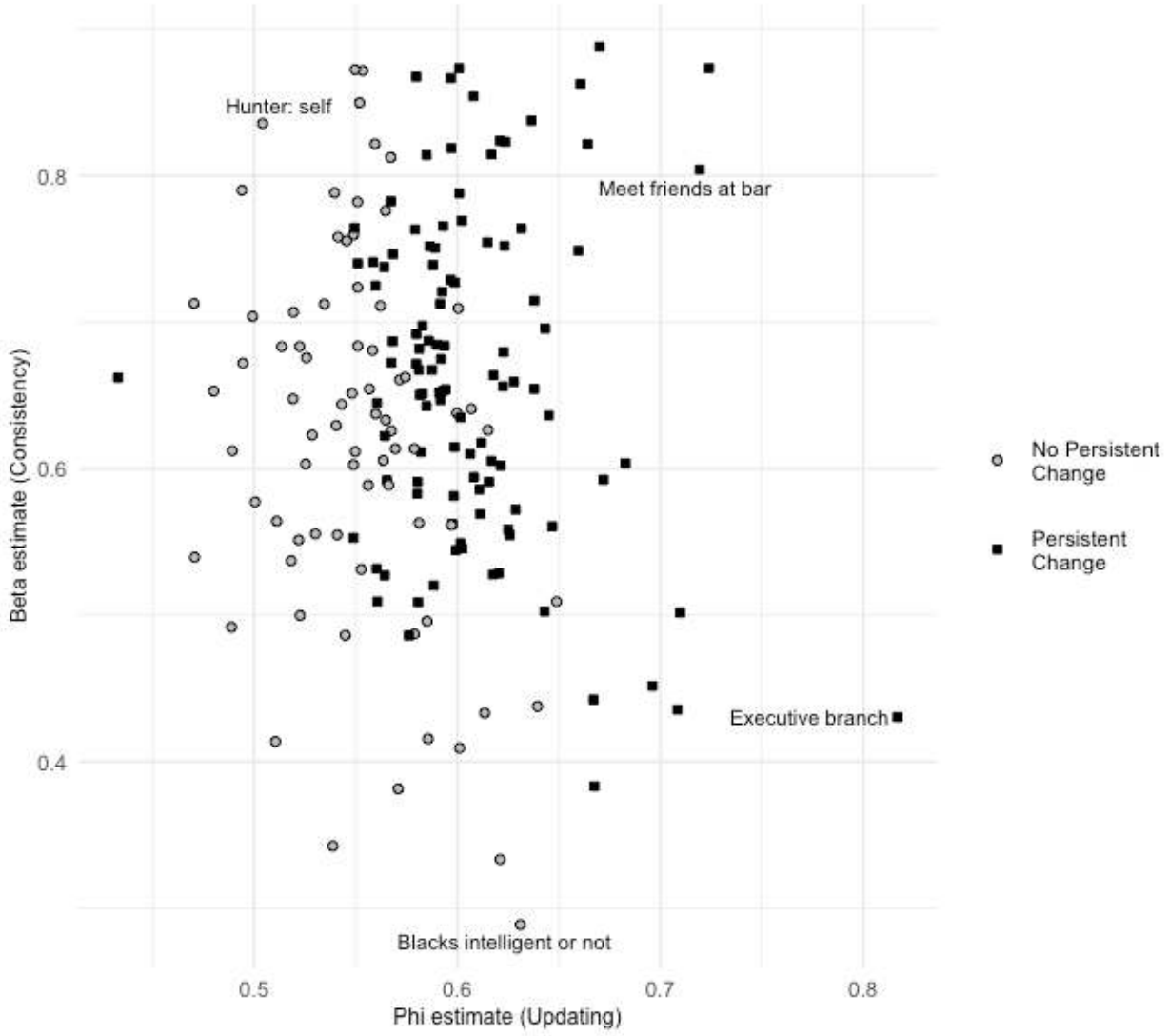

Figure 4: Distribution of beta and phi estimates for GSS items, by whether model prefers phi = .5 constraint.

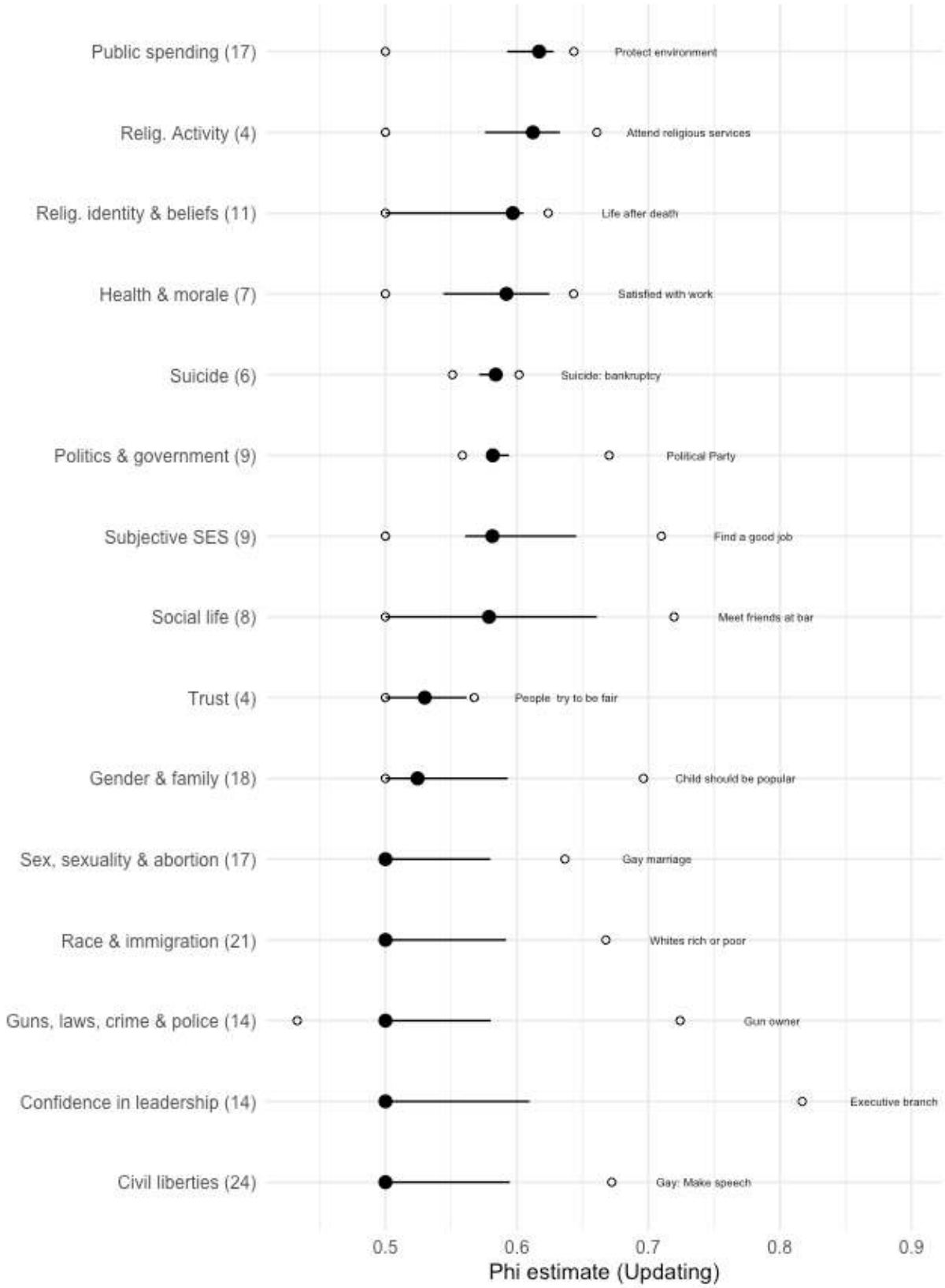
Public spending (17)
Protect environment
Relig. Activity (4)
Attend religious services
Relig. identity & beliefs (11)
Life after death
Health & morale (7)
Satisfied with work
Suicide (6)
Suicide: bankruptcy
Politics & government (9)
Political Party
Subjective SES (9)
Find a good job
Social life (8)
Meet friends at bar
Trust (4)
People try to be fair
Gender & family (18)
Child should be popular
Sex, sexuality & abortion (17)
Gay marriage
Race & immigration (21)
Whites rich or poor
Guns, laws, crime & police (14)
Gun owner
Confidence in leadership (14)
Executive branch
Civil liberties (24)
Gay: Make speech
0.5
0.6
0.7
0.8
0.9
Phi estimate (Updating)

Figure 5: Summary of phi estimates for all items, by topical group. Numbers in parentheses indicate the number of items in each topical group.

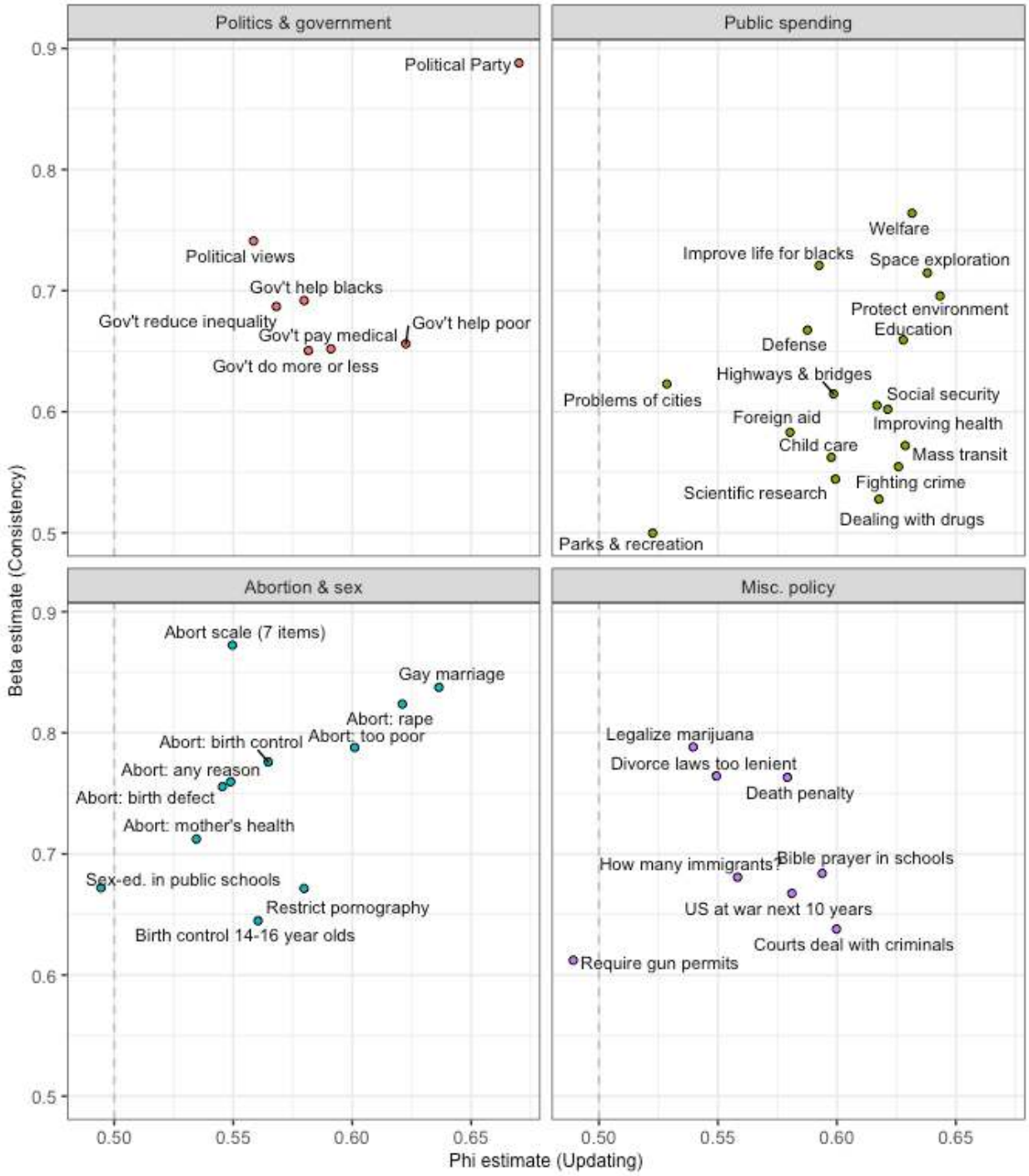

Figure 6: Phi (updating) and beta (consistency) estimates for 43 political attitudes and identities.

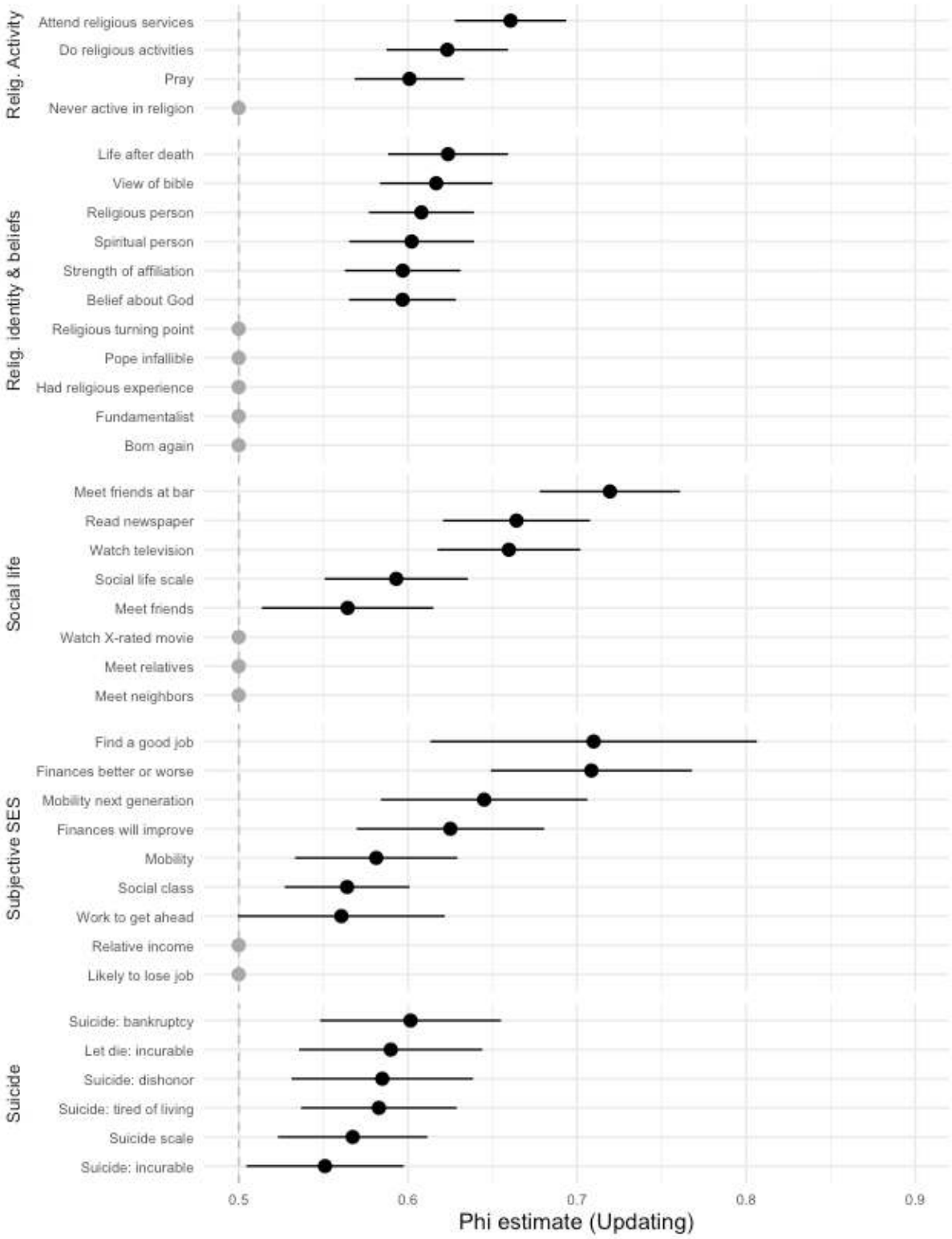

Figure 7: Phi estimates for items about religious activity and beliefs, social life, subjective SES, and suicide.

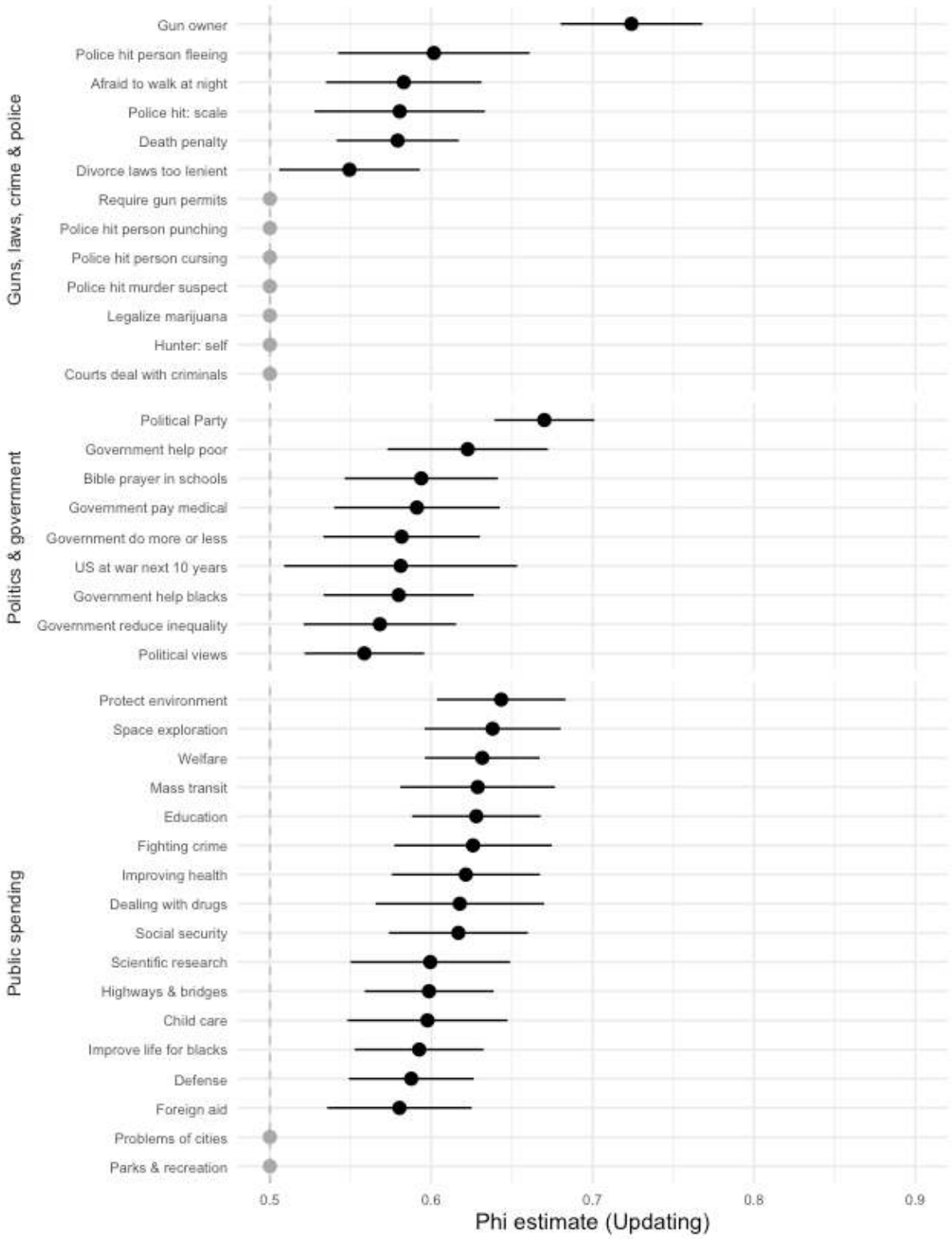

Figure 8: Phi estimates for items about guns, law, crime and policing; politics and government; and public spending. The item 'Police can hit citizens', which has phi = .43, has been removed for ease of viewing

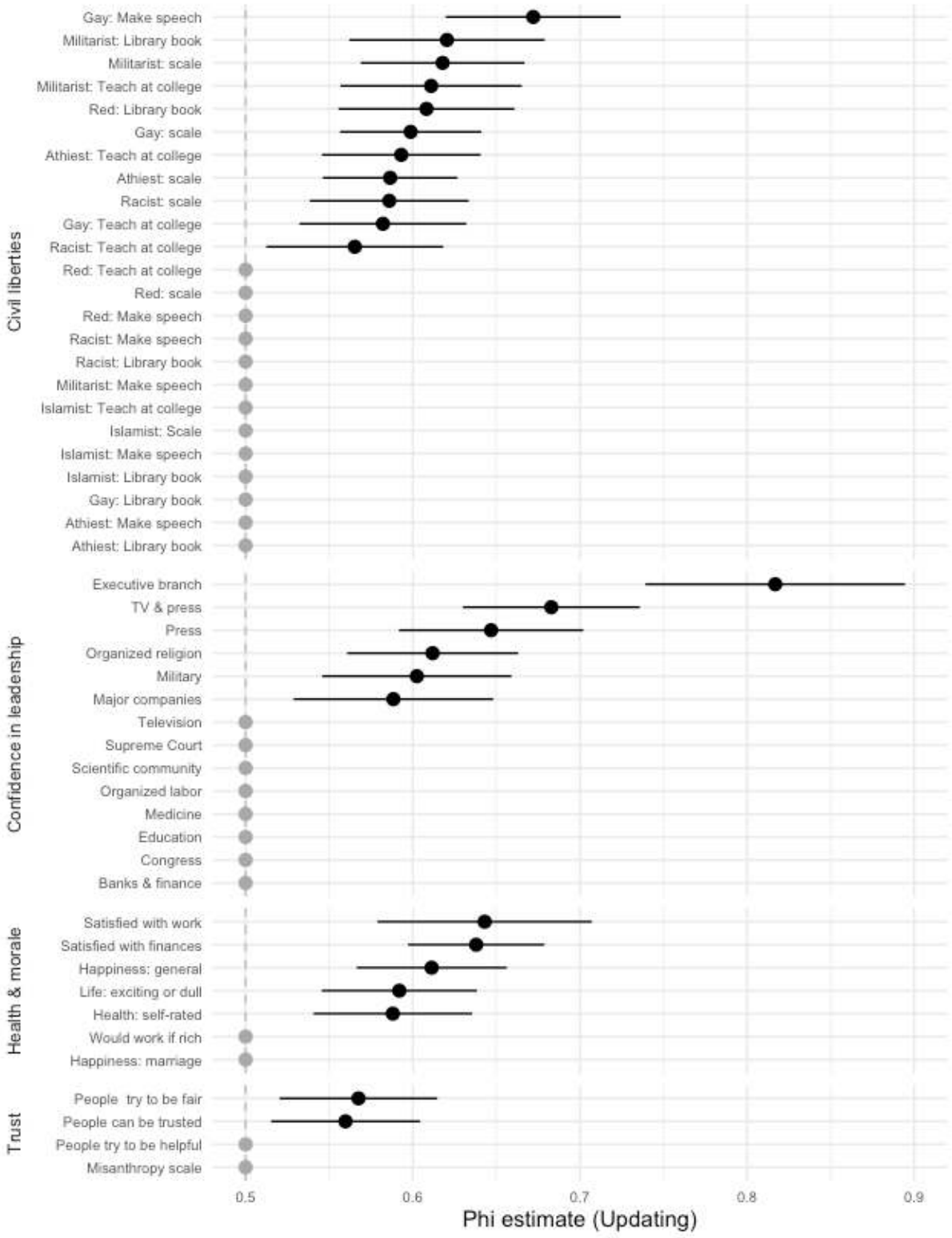

Figure 9: Phi estimates for items about civil liberties, confidence in leadership, health, morale, and social trust.

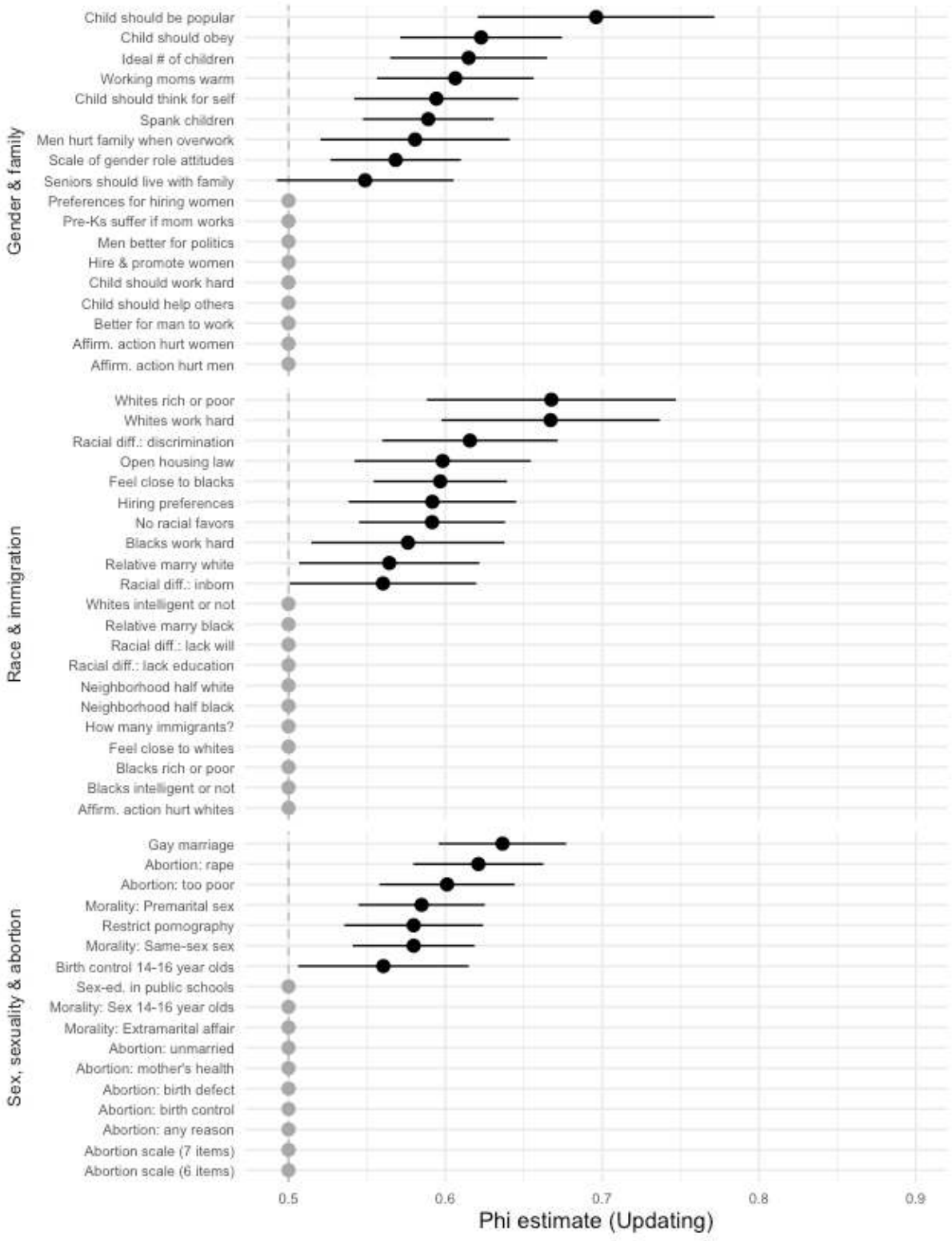

Figure 10: Phi estimates for items about race, gender, sex, sexuality, and abortion.

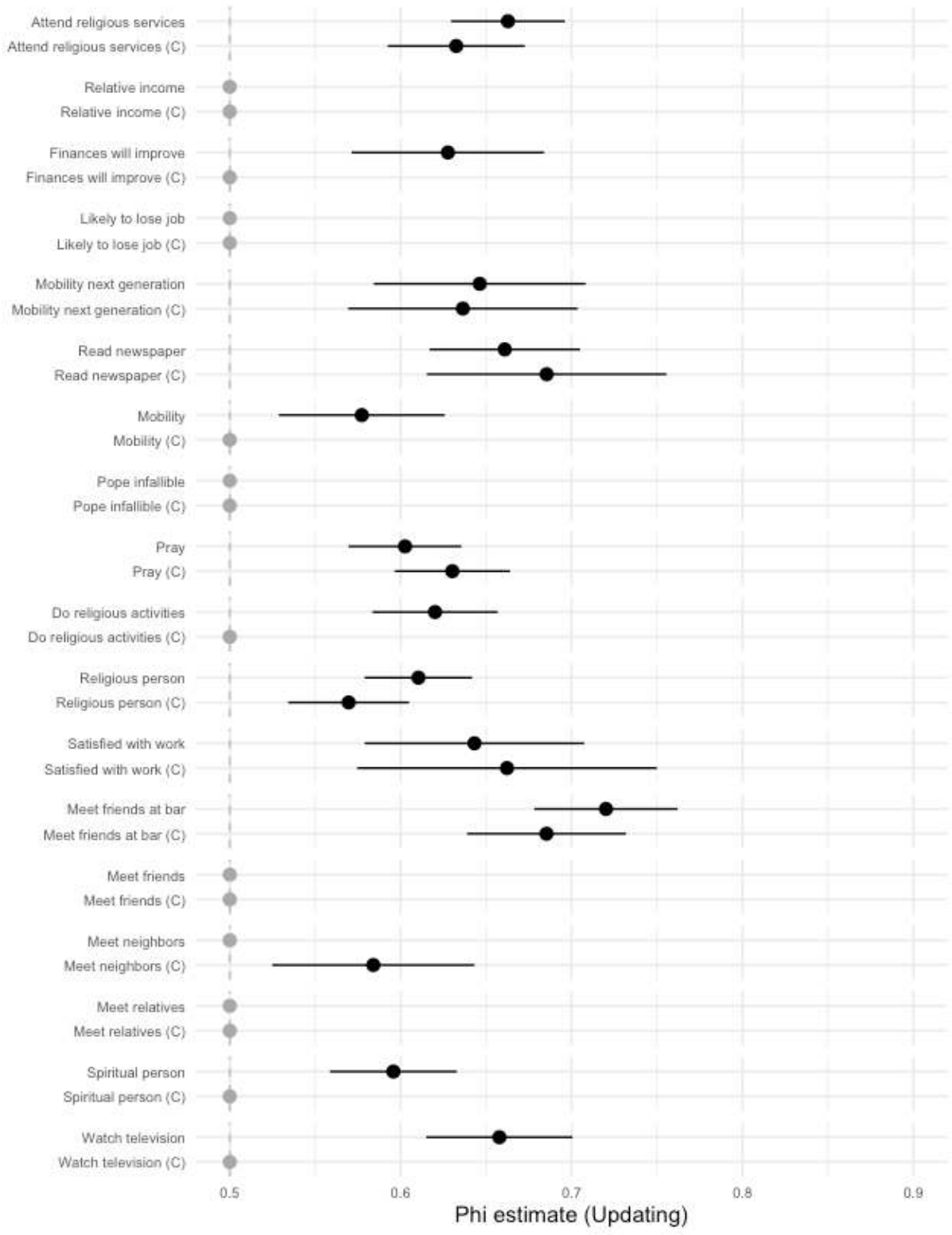

Figure 11: Phi estimates for coarsened questions about religion, subjective socioeconomic status, and morale.

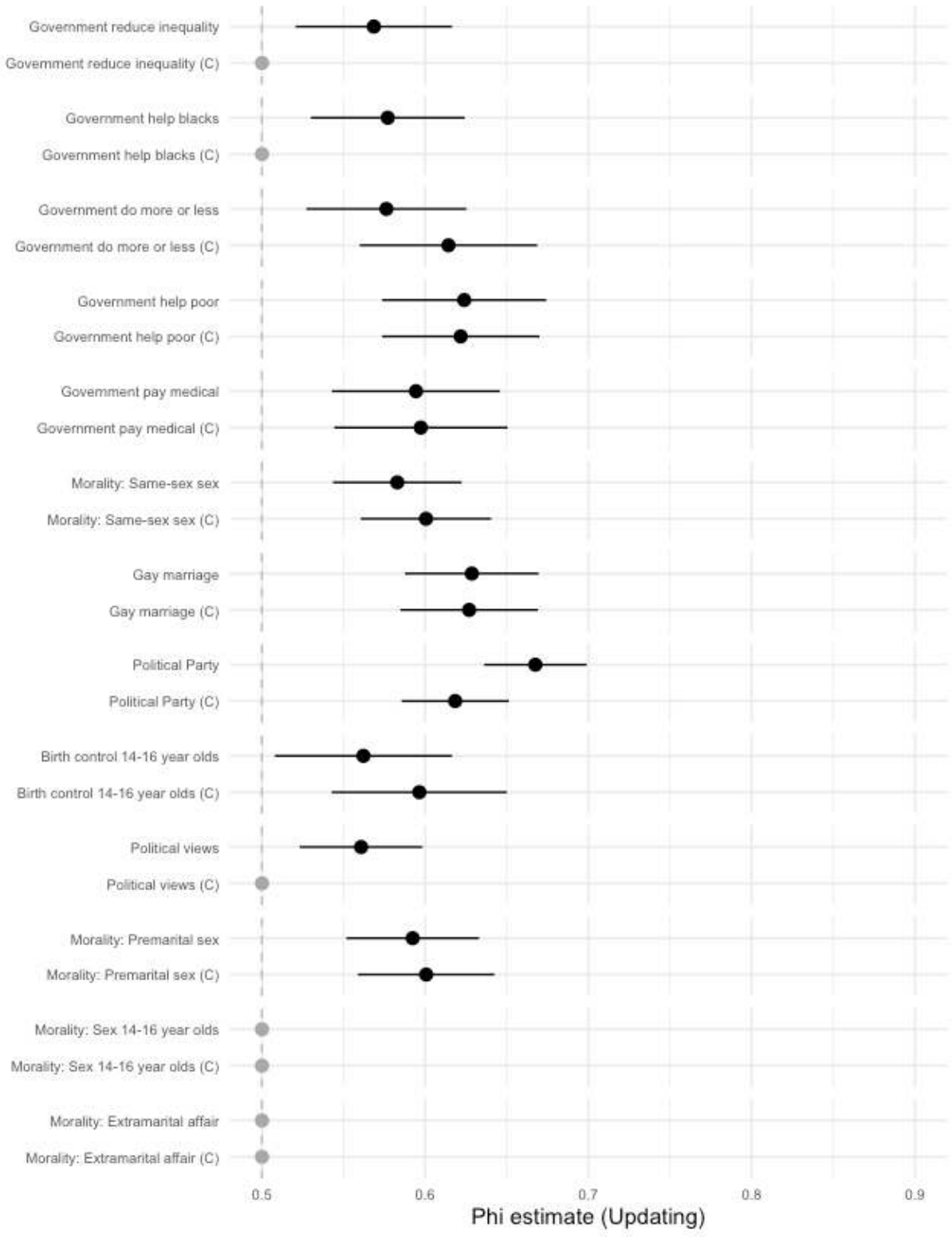

Figure 12: Phi estimates for coarsened questions about politics, government, sex and sexuality.

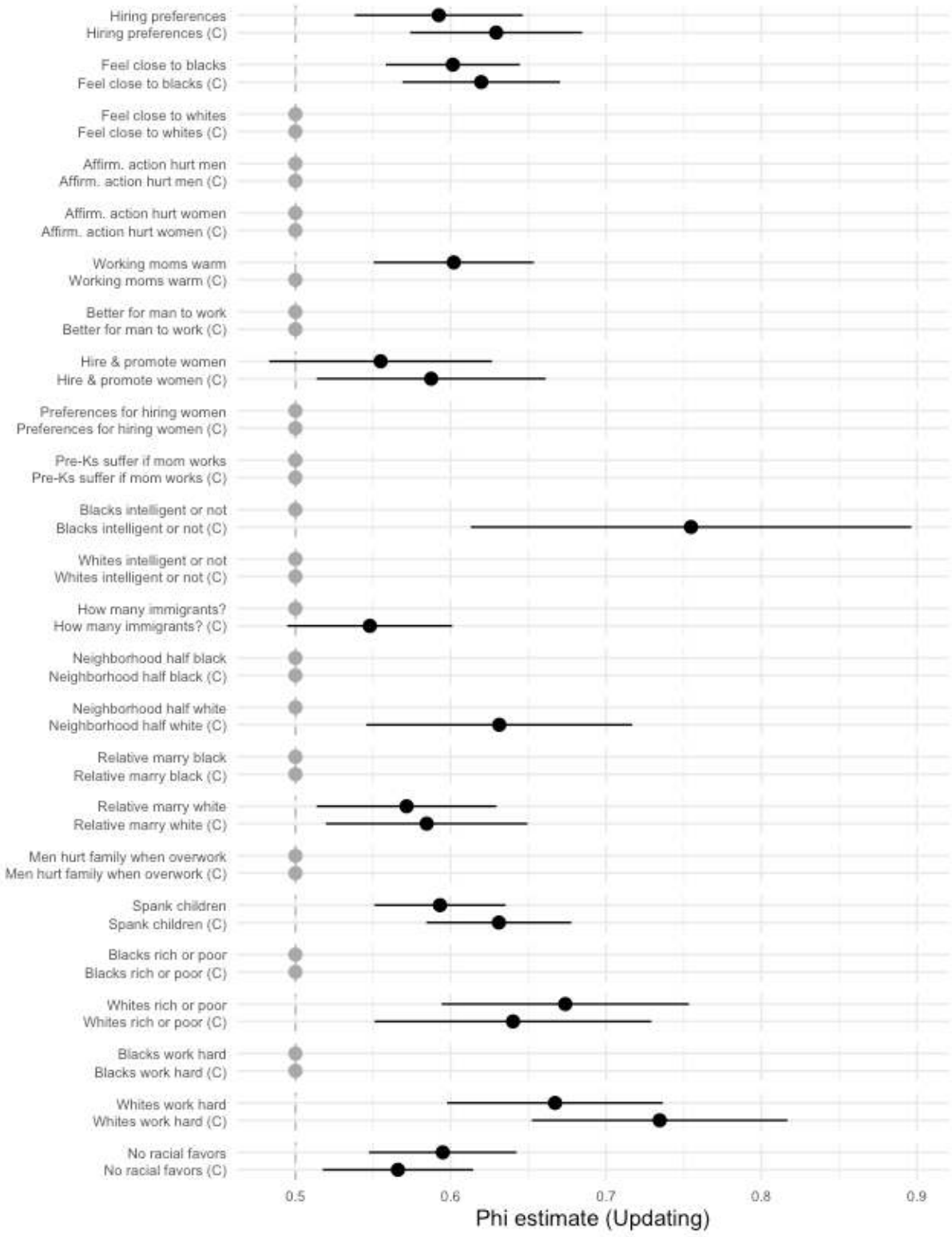

Figure 13: Phi estimates for coarsened questions about race and gender.

Abramowitz, Alan I., and Kyle L. Saunders. 2006. “Exploring the Bases of Partisanship in the American Electorate: Social Identity Vs. Ideology.” *Political Research Quarterly* 59 (2): 175–87.

Achen, Christopher H. 1992. “Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research.” *Political Behavior* 14 (3): 195–211.

Alwin, Duane F. 2007. *Margins of Error: A Study of Reliability in Survey Measurement*. New York: John Wiley & Sons.

Alwin, Duane F., and Jon A. Krosnick. 1991. “Aging, Cohorts, and Stability of Sociopolitical Orientations over the Life Span.” *American Journal of Sociology* 97 (1): 169–95.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder Jr. 2008. “The Strength of Issues: Using Multiple Measures to Gauge Preferennce Stability, Ideological Constraint, and Issue Voting.” *American Political Science Review* 102 (2): 215–32.

Bail, Christopher A., Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. “Exposure to Opposing Views on Social Media Can Increase Political Polarization.” *Proceedings of the National Academy of Sciences* 115 (37): 9216–21.

Baldassarri, Delia, and Andrew Gelman. 2008. “Partisans Without Constraint: Political Polarization and Trends in American Public Opinion.” *American Journal of Sociology* 114 (2): 408–46.

Bartels, Larry M., and Simon Jackman. 2014. “A Generational Model of Political Learning.” *Electoral Studies* 33: 7–18.

Bourdieu, Pierre. 1979. “Public Opinion Does Not Exist.” In *Communication and Class Struggle*, edited by A. Mattelart and S. Siegelaub, 124–30. Bagnolet.

———. 1990. *The Logic of Practice*. Stanford, California: Stanford University Press.

Boutyline, Andrei, and Stephen Vaisey. 2017. “Belief Network Analysis: A Relational Approach to Understanding the Structure of Attitudes.” *American Journal of Sociology* 122 (5): 1371–1447.

Centola, Damon. 2018. *How Behavior Spreads: The Science of Complex Contagion*. Princeton, New Jersey: Princeton University Press.

Christakis, Nicholas A., and James H. Fowler. 2009. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. New York: Little, Brown; Company.

Converse, Philip. 1964. “The Nature of Belief Systems in Mass Publics.” In *Ideology and Discontent, International Yearbook of Political Behavior Research*, edited by David Apter. Vol. 5. Free Press.

Cotter, David A., Joan M. Hermsen, and Reeve Vanneman. 2011. ""The End of the Gender Revolution? Gender Role Attitudes from 1977 to 2008." *American Journal of Sociology* 117: 259–89.

Danigelis, Nicholas L., Melissa Hardy, and Stephen Cutler. 2007. "Population Aging, Intracohort Aging, and Sociopolitical Attitudes." *American Sociological Review* 72: 812–30.

DeGloma, Thomas. 2014. *Seeing the Light: The Social Logic of Personal Discovery*. Chicago: University of Chicago Press.

DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why Do Liberals Drink Lattes?" *American Journal of Sociology* 120 (5): 1473–1511.

DiMaggio, Paul, John Evans, and Bethany Bryson. 1996. "Have American's Social Attitudes Become More Polarized." *American Journal of Sociology* 102: 690–755.

Eaton, Asia A., Penny S. Visser, Jon A. Krosnick, and Sowmya Anand. 2009. "Social Power and Attitude Strength over the Life Course." *Personality and Social Psychology Bulletin* 35 (12): 1646–60.

Elder, Glen. 1974. *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press.

Glenn, Norval D. 1974. "Aging and Conservatism." *Annals of the American Academy of Political and Social Science* 415: 176–86.

Goldberg, Amir, and Sarah K. Stein. 2018. "Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation." *American Sociological Review* 83 (5): 897–932.

Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.

Gross, Neil. 2009. "A Pragmatist Theory of Social Mechanisms." *American Sociological Review* 74: 358–79.

Heise, David R. 1969. "Separating Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34: 93–191.

Hout, Michael, and Orestes P. Hastings. 2016. "Reliability of the Core Items in the General Social Survey: Estimates from the Three-Wave Panels, 2006–2014." *Sociological Science* 3: 971–1002.

Howe, Lauren C., and John A. Krosnick. 2017. "Attitude Strength." *Annual Review of Psychology* 68: 327–51.

Inglehart, Ronald, and Wayne E. Baker. 2000. "Modernization, Cultural Change, and the Persistence of Traditional Values." *American Sociological Review* 65: 19–51.

Iyengar, Shanto, and Sean J. Westwood. 2015. “Fear and Loathing Across Party Lines: New Evidence on Group Polarization.” *American Journal of Political Science* 59 (3): 690–707.

Kinder, Donald R., and Nathan P. Kalmoe. 2017. *Neither Liberal nor Conservative: Ideological Innocence in the American Public*. Chicago: University of Chicago Press.

Krosnick, John A., and Duane F. Alwin. 1989. “Aging and Susceptibility to Attitude Change.” *Journal of Personality and Social Psychology* 57 (3): 416–25.

Lakoff, George. 2002. *Moral Politics: How Liberals and Conservatives Think*. Chicago: University of Chicago Press.

Lizardo, Omar. 2017. “Improving Cultural Analysis: Considering Personal Culture in Its Declarative and Nondeclarative Modes.” *American Sociological Review* 82 (1): 88–115.

Lizardo, Omar, and Michael Strand. 2010. “Skills, Toolkits, Contexts and Institutions: Clarifying the Relationship Between Different Approaches to Cognition in Cultural Sociology.” *Poetics* 38 (2): 205–28.

Mannheim, Karl. 1952. “The Problem of Generations.” In *Essays on the Sociology of Knowledge*, 276–322. New York: Oxford University Press.

Oh, Jeong Hyun, Sara Yeatman, and Jenny Trinitapoli. 2019. “Data Collection as Disruption: Insights from a Longitudinal Study of Young Adulthood.” *American Sociological Review* 84 (4): 634–63.

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

Perrin, Andrew J., and Katherine McFarland. 2011. “Social Theory and Public Opinion.” *Annual Review of Sociology* 37: 87–107.

Robinson, Dawn T. 2007. “Control Theories in Sociology.” *Annual Review of Sociology* 33: 157–74.

Ryder, Norman B. 1965. “The Cohort as a Concept in the Study of Social Change.” *American Sociological Review* 30 (6): 843–61.

Schnabel, Landon, and Eric Sevell. 2017. “Should Mary and Jane Be Legal? Americans’ Attitudes Toward Marijuana and Same-Sex Marriage.” *Public Opinion Quarterly* 81 (1): 157–72.

Smith, Tom W. 1997. “Factors Relating to Misanthropy in Contemporary American Society.” *Social Science Research* 26: 179–96.

Smith-Lovin, Lynn, and David R. Heise. 1988. *Analyzing Social Interaction: Advances in Affect Control Theory*. New York: Gordon; Breach Science Publishers.

Vaisey, Stephen, and Omar Lizardo. 2016. “Cultural Fragmentation or Acquired Dispositions? A New Approach to Accounting for Patterns of Cultural Change.” *Socius* 2: 1–15.

Visser, Penny S., and Jon A. Krosnick. 1998. “Development of Attitude Strength over the Life Cycle: Surge and Decline.” *Journal of Personality and Social Psychology* 75 (6): 1389–1410.

Visser, Penny S., and Robert R. Mirabile. 2004. “Attitudes in the Social Context: The Impact of Social Network Composition on Individual-Level Attitude Strength.” *Journal of Personal and Social Psychology* 87 (6): 779–95.

Warren, John Robert, and Andrew Halpern-Manners. 2012. “Panel Conditioning in Longitudinal Social Science Surveys.” *Sociological Methods and Research* 41 (4): 491–534.

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