

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 model theorizes that people actively their beliefs in the face of new information, while the other argues that, following early socialization experiences, individuals' dispositions are stable. We formalize these two models, elaborate empirical implications of each, and derive a simple model for comparing their prevalence using panel data. We test this model on more than 184 attitude and behavior items from the 2006-14 rotating panels in the General Social Survey. The pattern of results is more consistent with the settled dispositions model than the active updating model, and most observed change in the GSS appears to be short-term attitude change or measurement error rather than persisting changes in belief. When persistent change occurs, it is somewhat more likely to occur in young people. Evidence for active updating is more common among public behaviors than private attitudes and for high-profile issues. We argue that there is a place for both models in our theory of cultural evolution but that we need more evidence on the circumstances under which each is more likely to apply.

1 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?

Sociological analyses of cultural change tend to focus on population-wide changes in attitudes, beliefs, or practices that, while rarely made explicit, invoke one of two broad models of individual cultural change. The first is an *active updating* model that emphasizes the role of changing discourses, environments, and interactions. In this model, people continuously revise their views and practices based on the people, discourse, and events around them. The second is a *settled dispositions* model, which emphasizes the continuing influence of durable dispositions acquired early in life. In this model, adults' views and practices are relatively robust to changes in the social environment.

Both accounts have affinities with traditions in sociological theory. The active updating model is most compatible with a broadly pragmatist approach to social 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 settled dispositions model has affinities with the Bourdieusian tradition that 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).

In this paper, we make these two core models explicit, deduce some of their empirical implications, and derive a simple strategy for comparing their prevalence using panel data. We then apply this method to 184 items from the 2006-2014 General Social Surveys (GSS). 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 society.

Our analysis yields several results. First, we find that the majority of what appears to be individual-level change in belief or practice probably reflects short-term (non-persisting) change or measurement error rather

than actual persistent change. Simply put, there is little evidence that U.S. adults are changing their beliefs or practices in lasting ways over this period of study. Second, the persistent change that we *do* see in the data is somewhat more concentrated among younger respondents. That is, on several items at least, it appears that younger adults are still in the process of acquiring dispositions and habits they will take into later life. Third, we find that changes in social behavior (e.g., church attendance, political party membership, socializing) are more likely to be persistent than changes in private attitudes (e.g., political ideology). This suggests that interactional and institutional mechanisms may provide stronger support for lasting change than pressures for intrapsychic consistency.

These findings are generally more consistent with the settled dispositions model than with the active updating model. However, there is a pattern of exceptions and caveats that can help us understand how institutions and events shape the process of cultural change. We argue that there is a place for both models in our theory of cultural evolution but that we need more evidence on the circumstances under which each is more likely to apply.

In addition to their theoretical significance, our findings also have practical methodological implications. We argue that short-duration panel studies of adults will be of limited utility for studies of cultural change unless they have a clear focus on the sorts of things that are likely to change in persistent ways over the study period. Otherwise, researchers will waste resources measuring “change” that turns out to be noise. In general terms, then, we suggest that social scientists would do well to conduct more panel studies of children and adolescents while using cross-sectional designs to study adults.

2 Theoretical Background

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 at all. We will explain this in greater detail below. For now, consider the following two models:

$$y_{it} = y_{it-1} + \nu_{it} \tag{1}$$

$$y_{it} = U_i + \nu_{it} \tag{2}$$

These models may seem similar at first glance, but they have very different implications for the pattern of individuals’ responses 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 theory. In the section after that, we discuss the conditions under which each model is more likely to apply.

2.1 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 even 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

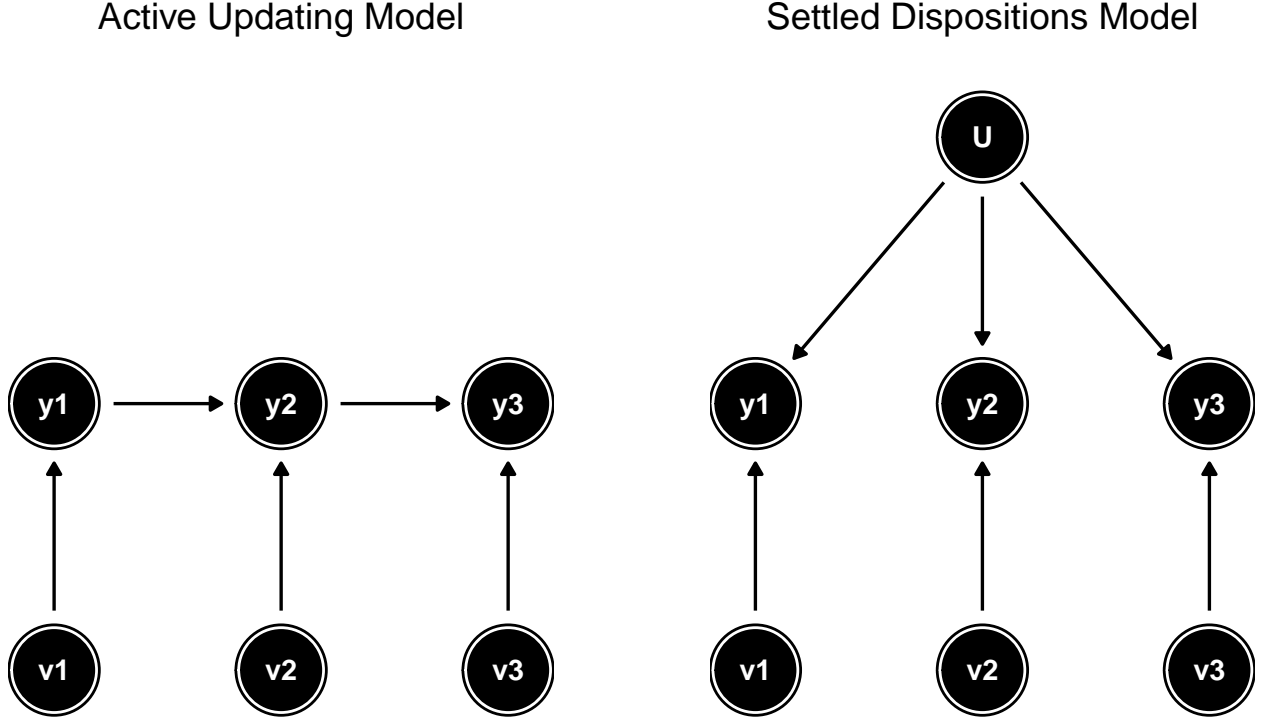


Figure 1: Different Models of Response Over Time

current state. This formal property, and this active updating model generally, is often assumed to underlie data generation in studies of change in and reliability of repeated survey measures (Alwin 2007; J. A. Krosnick and Alwin 1989; Hout and Hastings 2016).

More informally, this AUM posits an agent that is updating her views in the face of social experience. There are formal Bayesian ways of modeling updating, 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.”

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.

2.2 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 temporary influences only.

The SDM is most consistent with 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.” We will say more about this below.

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 individuals these shifts will be erased over time as individuals return to their baselines.

Metaphorically, one can think of the active updating model as a “traveler” and the settled dispositions model as a “visitor.” That is, in the active updating model, the traveler wakes up in the morning, walks in a direction for a while, goes to sleep, and then wakes up the next day to pick up the journey exactly where she left off. In the settled dispositions model, the visitor also travels in a particular direction each day but goes home at the end of the day and wakes up in the same bed each morning.

The 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. In the SDM, changes tend to revert because each person returns to their existing 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*. We consider the implications of these predictions below.

2.3 Combining the Models

Although our presentation above highlights differences between these models, they are not necessarily contradictory. There are at least two ways to understand how both models could apply without contradiction.

2.3.1 Age and Cohort Formation

The easiest way to reconcile the AUM and SDM is to posit that they operate sequentially over a person’s life. A range of sociological theories provide conflicting hypotheses and evidence for how susceptible individuals are to attitude and behavior change at different life stages. Most of these hypotheses suggest that early adulthood is a time of heightened attitude change. Two hypotheses – the “increasing persistence” hypothesis (Glenn 1974; Inglehart and Baker 2000) and the “impressionable years” hypothesis (Visser and Krosnick 1998) – suggest that susceptibility to attitude change is highest early in life and decreases (gradually or quickly, respectively) as individuals age. The “life stages” hypothesis (Visser and Krosnick 1998; Danigelis, Hardy, and Cutler 2007) posits that both early (<40) and late (60+) life are times of more attitude change but that mid-life is characterized by limited attitude change.

To continue the metaphor above, one can imagine travelers wandering around for a while (active updating) before setting up their permanent home base (settled disposition). This is, in fact, the logic that underlies the notion of a “cohort effect.” The idea of a cohort effect is that the contemporary environment influences the baseline response tendency when a person is young and then stabilizes and stays “baked in” for the remainder of one’s life (Elder 1974; Bartels and Jackman 2014). Any notion of a settled disposition (including the Bourdieusian one) requires an earlier “formative period” where something like the active updating model was operative (Vaisey and Lizardo 2016). To put it slightly differently, cohort formation requires an early period of persistent change 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.

This discussion of cohorts brings us close to the vast literature on age-period-cohort decomposition models, which is focused on statistical techniques to partition variance among sources of social change in repeated cross-sectional data (Yang and Land 2006). Although this work is important, it is focused on modeling

changes in population aggregates whereas we are concerned with modeling cultural change from the point of view of the same actor moving through time.

The two are related and have implications for each other, of course. Again, the idea of a cohort requires that some people (i.e., younger people) are influenced by the environment to make persistent change in ways that others (i.e., older people) are not. This implies a transition, perhaps a gradual one, from active updating to a settled disposition. A further connection between the two is that the world implied by the SDM can only change in the aggregate due to cohort replacement. That is, if persisting change is extremely rare among adults, the only way for persisting change to occur in the population is for older cohorts to die and be replaced with younger cohorts with (perhaps) different baselines. For our purposes, the most important idea here is that “cohortization” implies that the AUM might describe the data-generating process better among younger respondents whereas the SDM might describe it better among older respondents.¹

2.3.2 Different Items, Different Processes

Another way to reconcile the AUM and the SDM is to consider that they apply to different sorts of questions a researcher might ask a respondent. We can illustrate this idea with (non-cultural) examples we understand well and then consider how less obvious cultural processes might be similar or different.

First, if we repeatedly ask a sample of women “how many children have you ever given birth to?”, we know the data is generated by a process that looks like equation 1. Leaving measurement error aside, the response is a simple combination of how many children the respondent had last time we asked (y_{it-1}) and whether or not she has had any since (ν_{it}).

Second, if we ask repeatedly ask a sample of adults “how many full siblings did you have at age 16?”, we know the true answer cannot change and therefore equation 2 is the best model. This is because U_i refers to a fixed quantity and cannot change over time. Responses could still vary due to measurement error (e.g., misunderstanding the question) meaning that the variance of ν_{it} need not be 0. But the underlying reality is, by definition, a settled matter.

Finally, there are responses that have no stable component but are entirely functions of local circumstances. If we repeatedly ask a sample of people in the same city “did it rain yesterday?”, their responses would not be described better by Equation 1 or Equation 2. In this case, everything that matters is in the contemporary circumstances (ν_{it}) and thus the terms that distinguish Equation 1 from Equation 2 (y_{it-1} or U_i) are ignorable. These types of items are not important for adjudicating our research questions, but we still need to be aware of this possibility.

What we want to determine, then, is how much a particular item resembles these different processes. Consider repeated requests for (dis)agreement with the prompt “a working mother can establish just as warm and secure a relationship with her children as a mother who does not work.” Does the aggregate pattern of responses look more like an active updating process where a person might be changing her mind? Or does it look more like asking a person to report repeatedly on a settled matter?

When it comes to beliefs and behaviors, which items will demonstrate persistence likely depends on a number of factors. Vaisey and Lizardo (2016) found evidence that beliefs about gender, race, and other “core” opinions are more driven by cohort effects and are therefore less likely to be actively updated by adults. Using repeated cross-sectional data, Danigelis, Hardy, and Cutler (2007) find that “intracohort” change (a good indicator that some cohort members are changing their minds) is larger for young individuals (18 to 39 years old) than older individuals (over 60 years old) for attitudes about economic gender equality and structural reasons for racial inequality, while other attitudes pertaining to civil liberties, political gender equality, and privacy tended to show similar rates of intracohort change at both young and old ages during the time frame they studied.

¹We do not have space for a full discussion here, but we regard these approaches as complementary. Repeated cross-section approaches must make strong assumptions about individual processes but are able to use many years of data. A panel approach is better able to capture micro-level processes but at the cost of using only a few years of data; this requires making strong assumptions about the generalizability of the process across historical time.

Previous studies have also found that political attitudes are more volatile during early and late adulthood but stabilize during middle age (J. A. Krosnick and Alwin 1989). While volatility of responses increased as individuals aged, this is mostly due to increases in measurement error, rather than persisting change. This generally suggests that political views will show decreasing evidence of persistent change as age increases.

“Public culture,” like participating in external activities, may show similar rates of persistence over time because social mechanisms (like joining a group) can help maintain changes in ways that intrapsychic forces cannot. Finally, items that capture issues that are highly salient in the media during the study period may be more likely to show active updating because of heightened exposure. Sears and Valentino (1997) find that during political campaigns adolescents tended to adopt new beliefs about issues that were salient during the campaign but not those issues peripheral to it. Given these different possibility, we do not begin this study with clear predictions on all issues. Because of the large change on gay marriage and gay rights during this period, however, we do expect to see active updating on related items.

2.4 Toward Theoretical Synthesis

The active updating model and the settled dispositions model are different and they do have affinities with different theoretical traditions in sociology. Indeed, the difference in emphasis between environmental cues, creative meaning-making, and public discourse on the one hand and settled dispositions, “habitus,” and cohort replacement on the other has been a basic theoretical tension in cultural sociology for more than a decade (Vaisey 2009; Lizardo 2017; Vaisey and Lizardo 2016).

One model may, in fact, be a better starting place than the other. Previous work using cross-sectional data suggests that settled dispositions are hugely important (Vaisey and Lizardo 2016). Despite these findings, the evidence is still mixed (Danigelis, Hardy, and Cutler 2007). Therefore our goal here is neither to declare victory for one of the theoretical perspective nor to simply say that both “matter.” Rather, our objective is to improve sociological models of cultural evolution by more precisely specifying when and where different types processes are at work. We believe achieving a better understanding of these processes will be relevant for many subfields of sociology as well as for other social science fields that study changes in beliefs and behaviors.

3 Research Questions and Expectations

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

First, *are patterns of cultural change generally better described by the active updating model or the 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?* As we just mentioned, 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 are better described by each model?* Items differ from each other in many ways and it would be challenging to come up with clear predictions for all possible types of items. In general, however, we expect (following Vaisey and Lizardo) that core beliefs about gender, family, race, and trust should be more consistent with settled dispositions because their formation should be already be complete for adults. Items that require public commitments such as

changing political parties, socializing, or religious attendance and affiliation should be open to persistent change because there are social mechanisms (not just cognitive ones) to *maintain* the change.

4 Analytic Strategy

To investigate these ideas empirically, we examine 184 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.

4.1 Statistical Model

4.1.1 Basic Models

Our two different models make different predictions in the three-wave panel context. The AUM makes the following prediction for wave 3:

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

That is, the AUM predicts that a respondent's most recent response is the best available predictor of her next response. If change is persisting (i.e., if our metaphorical traveler wakes up in the same place she went to sleep in last night), then our best guess is that she will be close to where we saw her last, and previous responses will provide no additional predictive power.

The SDM makes the following prediction:

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

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

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

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 (i.e., if our metaphorical visitor eventually always returns home), then taking the average of the last two places we saw her will be our best guess about the location of her home base.

4.1.2 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, we begin to observe evidence that at least some individuals in the population engage in active updating. We use the following non-linear model to estimate the relative influence of y_{i1} and y_{i2} :

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

Rather than estimate separate β parameters for y_{i2} and y_{i1} , this model generates two parameter estimates of interest for each item in our analysis: 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). In the active updating model is present in at least some respondents, then wave 2 will be more predictive than wave 1, and ϕ will increase to converge with Equation (3) in certain circumstances.

4.1.3 Limitations of the Method

Three challenges limit our ability to evaluate the relative presence of settled dispositions and active updating models using our approach. These challenges are not fatal, but put some limitations on the conclusions we will be able to reach.

The first challenge is that our model is designed to allocate variance to each of the prior predictors rather than to assign probabilities to each data-generating process per se. Because of this, a few individuals making large persisting changes can inflate the ϕ estimate even if most individuals make small non-persisting changes. If we were willing to assume that persisting and non-persisting changes were drawn from the same distribution, we could make inferences about the relative proportion of people coming from each data generating process, but we have no reason to believe this assumption.

The second challenge is measurement error, which is a form non-persisting change. Because measurement error “looks the same” as non-persisting real change, estimates of ϕ will be biased toward .5. 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, these previous studies of measurement 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 this issue 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 exclusively on predicting wave 3. If for some reason individuals have a high likelihood of changing between waves 2 and 3, then our ability to predict responses at wave 3 with either wave 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. Then again, it may not be reasonable to consider this sort of change “persistent” in any meaningful way.

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 change in the overall 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.

4.2 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 (6) with a free estimate of ϕ to a model that constrains $\phi = .5$. We calculate the posterior probability that the model with the freed parameter fits better. If the model with the constraint is preferred, then 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.

Second, for variables that do show at least some evidence of active updating ($\phi > 0.5$), we ask whether the persistent change is disproportionately 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 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, with these findings in hand, we consider whether there are any meaningful patterns in the relative distribution of evidence for AUM across variables. Although, as we discussed above, previous work gives some indications about what we might expect, the approach here will necessarily be largely inductive.

4.3 Item Selection

We want to test our model on as broad a range of items tapping attitudes, beliefs, and social behaviors as possible. To do so, we estimate the model in equation 6 on data from the three different three-wave panel data sets of the General Social Survey, conducted from 2006 through 2014.

We sought measures of attitudes, beliefs, self-assessments, self-perceptions, and social behaviors that were asked in three waves. These questions tended to come from the “core” of the GSS, a set of questions asked in each wave. Rotating topical modules asked during the panels were only asked in select waves or were not asked consistently to the same people over time, leaving too few cases with complete data to analyze.

Since our theoretical framework focuses on attitudes, beliefs, and social behaviors, 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 also exclude questions that ask about an individuals’ childhood. While testing these questions with our method is possible, they are beyond the scope of our theoretical framework. To organize our presentation of such a large number of variables, we follow Hout and Hastings (2016) and group questions into 15 categories based on their 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 institutional leaders, questions about public spending, and questions about social life.

We also follow Hout and Hastings (2016) in recreating some commonly used scales designed to capture attitudes about gender roles, access to abortion, and social trust. This includes Rossi’s 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, Hermesen, 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 clergymen. 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 averaging five binary questions about the conditions under which individuals support police use of violence.

In total, we test the model on 184 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.

5 Results

Our model generates two parameter estimates of interest for each GSS item in our analysis: 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. If responses are generated through a 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 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. This is a highly unlikely scenario.

5.1 Evidence for Active Updating

To evaluate the evidence in favor of the active updating model, we compare for all 184 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 184 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.

The right side of Figure 2 shows that while the majority of items prefer the free parameter, 73 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 will tend to revert to their previously held position over time. 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.

Figure 3 plots the ϕ and β estimates for items colored by whether they prefer the constraint or not. Although we discuss the pattern of items below, we 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.

Items showing evidence for 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.

We will discuss how different items perform below. To answer our first question, however, we need only focus on the overall distribution. We find that about 40 percent of items show no evidence at all for active updating among GSS respondents. And among items that do show some evidence of persistent change, very few come close to approaching 1, meaning that for almost all items, measurement error or non-persistent change tends to be much more common than persistent changes. We can only be really confident in detecting persistent

²To 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 the Appendix.

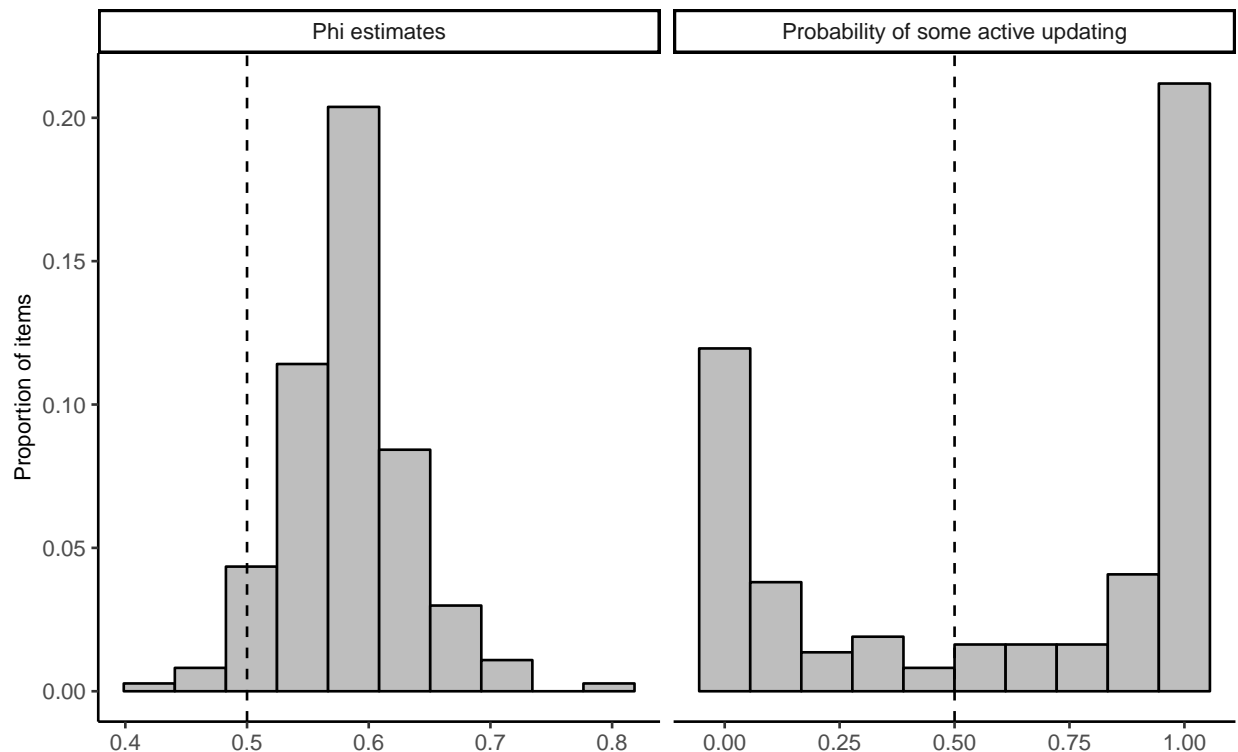


Figure 2: Distribution of phi estimates and probabilities that items show evidence of active updating.

change 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.

5.2 Age Heterogeneity

Our second research question is about whether there is evidence of cohort formation on some items. That is, we want to see if there is evidence that younger respondents are updating their views more than older respondents. Although it’s impossible for any item 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 111 items that showed evidence for active updating in the last section, 23 showed differential effects of age for more than 50 percent of cutoff ages we tested. To get a sense of the magnitude of difference between older and younger individuals, Figure 4 plots the estimates of these 23 items for individuals above or equal to or below 30.

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 categories 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 for the rest of life.

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 pass discrimination, all evidence of active updating disappears for individuals over 30. This suggests

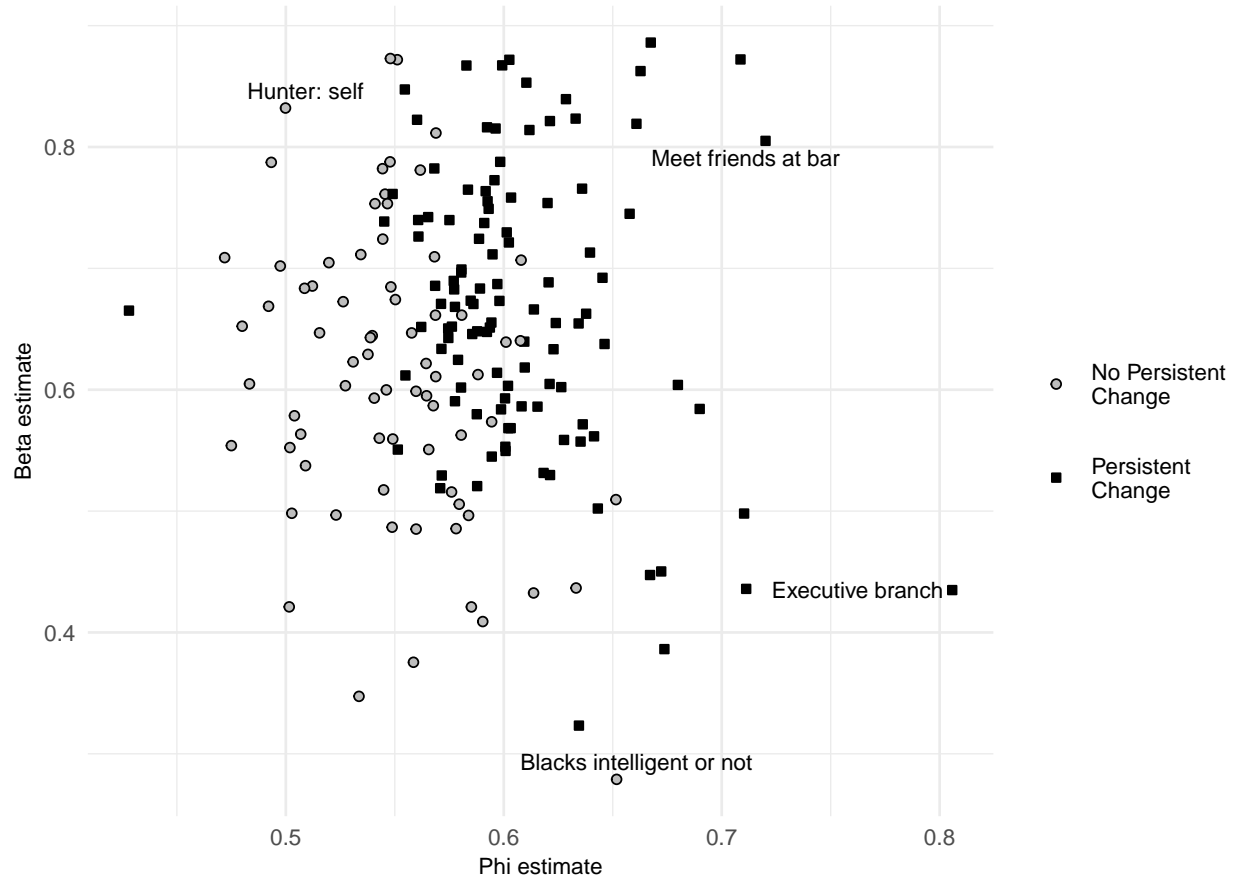


Figure 3: Distribution of beta and phi estimates for GSS items, by whether model prefers $\phi = .5$ constraint.

that these items follow an “impressionable years” pattern, where early adulthood is a time when these opinions are still malleable, but that over time they harden into “durable dispositions” (Vaisey and Lizardo 2016; J. A. 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 patterns is potentially consistent with the “increasing persistence” hypotheses.

Eight of these 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, views on whether aging parents should live with their children, and views on whether police should be allowed to use force against suspects who are verbally abusive toward them. Some of these items, such as views on whether aging parents should live with their children might be things young people are not forced to really consider until later in life and as a result do not form clear opinions until that exposure. This pattern where older individuals change their attitudes and behaviors at higher rates than younger individuals is relatively 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 previous theorizing has accounted for.

The remaining 89 items explored in this analysis (just under half of all items) show evidence for some active updating but do not show consistent evidence for age heterogeneity, again 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

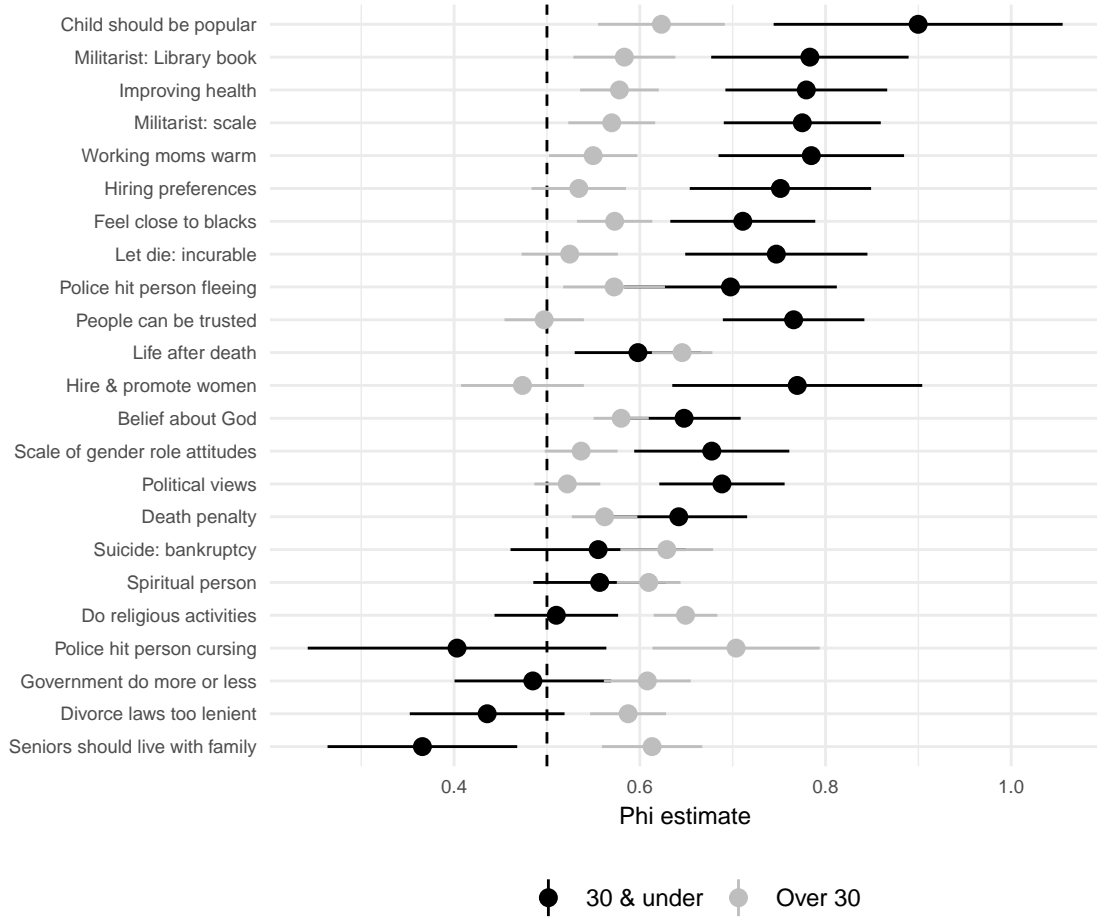


Figure 4: Comparison of ϕ estimates for individuals over and equal to to less than 30 years old.

small proportion of the population.

5.3 Item heterogeneity

Which items show the strongest evidence for active updating? There is no way we can possibly discuss all 184 items in detail without the discussion becoming tedious. Although we include all individual results in Appendix B, 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.

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 in absolute terms. 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 real or persistent but some combination of measurement error and short-term fluctuations.³

³While 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,

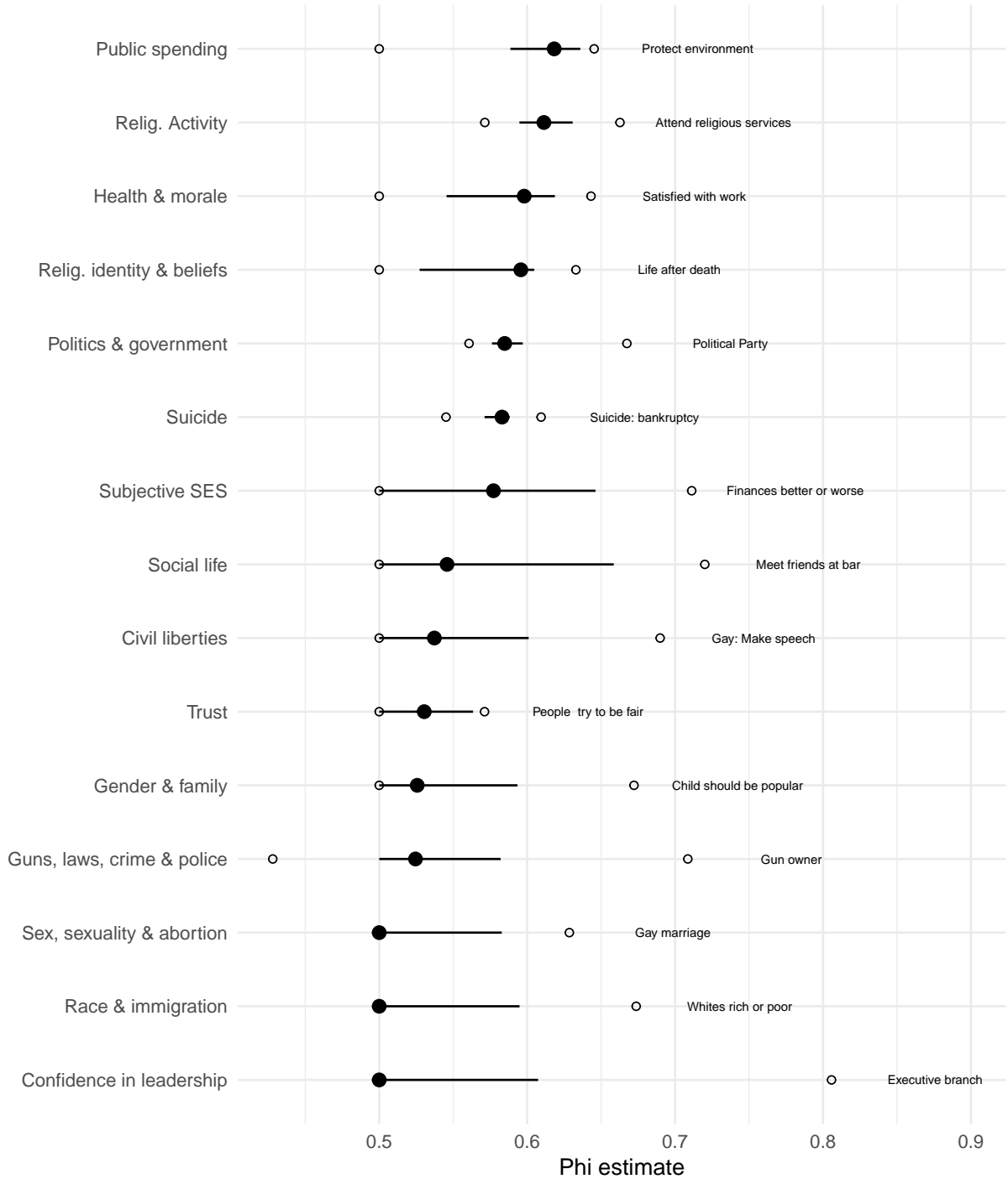


Figure 5: ϕ estimates for all items by topical group.

Consistent with the findings and theory of Vaisey and Lizardo (2016), we see that over half of items in the gender, family, race, sex, civil liberties, and confidence in institutions groups show no evidence for updating (see Appendix B). Views on these “core” issues are likely shaped by early socialization experiences and mostly

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.]

settled by the time a respondent reaches adulthood (and thus becomes eligible to participate in the GSS). This means that, for most of these items, social change must occur through cohort succession rather than through individual change.

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

Some of the 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 separate estimates 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, we see that 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 of 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 all spending 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 real external change is happening, we should see exactly this sort of pattern.

Some of the high- ϕ items rely on *external mechanisms* that likely help maintain them. If a person starts going to church or starts socializing with neighbors, she builds social networks that make this behavior more likely to continue. Switching political parties (which involves changing public registration) is a more persistent change than changing political ideology (which can happen privately in the mind). Buying a gun has a high ϕ value because a new physical object either enters or leaves the person’s possession.

Perhaps the most striking pattern in our findings is about 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. By comparison, abortion attitudes show almost no evidence for change.

The categories that are consistently higher than .5 are public spending and politics and government, areas where political parties make efforts to keep issues in the public eye and encourage people to update their views. Even so, most of these values are still low in absolute terms.

Finally, questions about religious activity, identity and (to a lesser extent) religious beliefs generally show some evidence for active updating. In some cases, this fits with the general pattern of external mechanisms (e.g., attendance and identification are largely public acts).

These patterns should be taken with a grain of salt since we began the study with only general notions of which type of items should show evidence for active updating. A story can always be told about any set of findings. We are encouraged, however, with the degree of fit between the findings here and our general expectations.

6 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 often implicit in analyses of belief change at the society level, but they are rarely examined 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, *are patterns of cultural change generally better described by the active updating model or the 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, and even those items that do show a relatively low degree of updating. Adult Americans are highly consistent in their attitudes from year to year, and those who change tend to bounce back to where they were. 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 appears 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 relatively insensitive to such shocks (see e.g., Bartels and Jackman 2014). Because the GSS begins with 18-year-olds, we may lack power to detect updating that occurs even earlier, so this may be an underestimate. 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?* Although we began with only very general expectations about patterns in updating, our results were consistent with those expectations. 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. Most items about gender, family, race, and institutions showed no evidence for updating, suggesting most people’s views on those topics are settled by the time they turn 18.

What are the broader implications of these findings, both for theories of cultural change and for empirical work in this area? We see two major implications, one theoretical and one methodological.

In the domain of theory, our findings support the view that most cultural change happens slowly, through the mechanism of cohort succession. The pattern of findings for gay rights show, however, that a high degree of public salience and social movements can accelerate change by encouraging some people to update their views who might not otherwise do so. But the baseline process 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.

The results also suggest much greater heterogeneity in the relationship between age and attitude change than what is emphasized in existing theories. The most prominent and well supported theories suggest a peak of 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, which are the those items most frequently studied, the pattern is not nearly as consistent for other kinds of attitudes. Some items, such views on the Bible, suggest perpetual openness to attitude change as individuals age. Others, such as views on most abortion questions, suggest that early adult socialization is so strong that opinions are hardened 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 actually becomes more common with age, a pattern that is not accounted for in any major hypothesis about the relationship between age and attitude change. Rather than supporting a single theory linking age to attitude change, our results call for more work linking attitude content to 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 social and psychological factors that that give rise to stable attitudes (Howe and Krosnick 2017).

Recent work that has attempted to provide social explanations for (mostly political) attitude stability in middle age is a strong start (Eaton et al. 2009; Visser and Mirabile 2004). However, this line of work should be cognizant of which attitudes stabilize in middle age and the relationship between social factors and these specific attitudes and explore how these factors link to attitudes that seem to change more frequently at these ages.

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 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 this kind of change is difficult. Even for items related to abortion, where there is almost no evidence of persistent change and the overall picture is one of stability of belief, individuals do bounce around from year-to-year.

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 younger individuals 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, and those that are repeated should be specifically targeted to topics that are believed to be changing broadly (e.g., politics, gay rights). Understanding the social origins of individuals’ attitudes requires greater focus on the “past conditions of production” that give rise to persistent beliefs in adulthood.

7 Appendix A: ϕ values for all variables

Figures 6 through 9 plot ϕ estimates for all items included in the analysis, grouped by subject material.

8 Appendix B: Coarsened variable estimates

ϕ estimates for coarsened items are presented in Figures 10 through 12. 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.

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

- 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.
- Bartels, Larry M., and Simon Jackman. 2014. "A Generational Model of Political Learning." *Electoral Studies* 33: 7–18.
- Bourdieu, Pierre. 1990. *The Logic of Practice*. Stanford, California: Stanford University Press.
- 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. Hermesen, 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.
- 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.
- Gross, Neil. 2009. "A Pragmatist Theory of Social Mechanisms." *American Sociological Review* 74: 358–79.
- 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.
- Krosnick, John A., and Duane F. Alwin. 1989. "Aging and Susceptibility to Attitude Change." *Journal of Personality and Social Psychology* 57 (3): 416–25.
- Lizardo, Omar. 2017. "Improving Cultural Analysis: Considering Personal Culture in Its Declarative and Nondeclarative Modes." *American Sociological Review* 82 (1): 88–115.
- Sears, David O., and Nicholas A. Valentino. 1997. "Politics Matters: Political Events as Catalysts for Preadult Socialization." *American Political Science Review* 97 (1): 45–65.
- Smith, Tom W. 1997. "Factors Relating to Misanthropy in Contemporary American Society." *Social Science Research* 26: 179–96.
- Vaisey, Stephen. 2009. "Motivation and Justification: A Dual-Process Model of Culture in Action." *American Journal of Sociology* 114: 1675–1715.
- 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.

Yang, Yang, and Kenneth C. Land. 2006. “A Mixed Models Approach to the Age-Period-Cohort Analysis of Repeated Cross-Section Surveys, with an Application to Data on Trends in Verbal Test Scores.” *Sociological Methodology* 36 (1): 75–97.

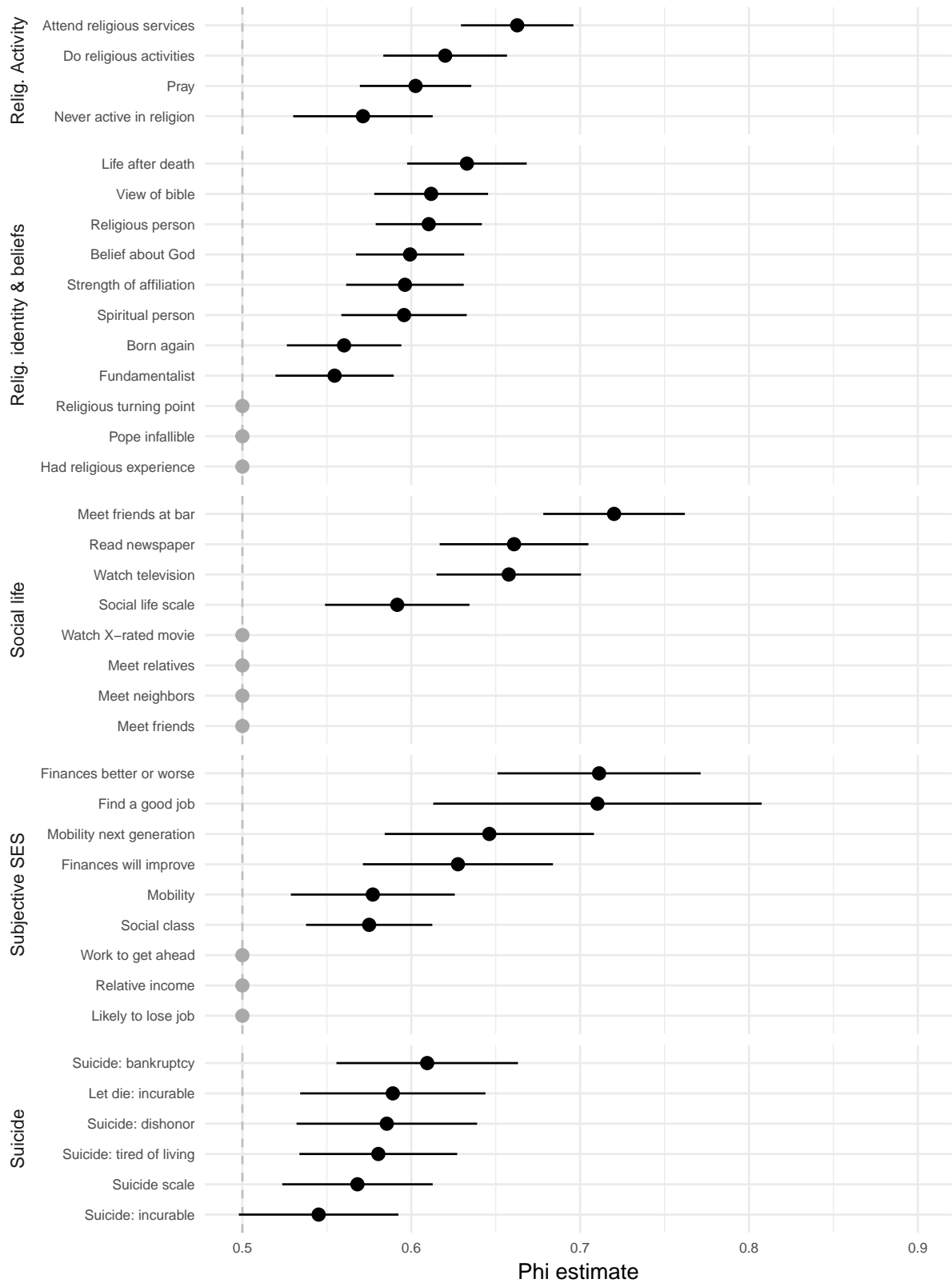


Figure 6: ϕ estimates for items about religious activity and beliefs, social life, subjective SES, and suicide.

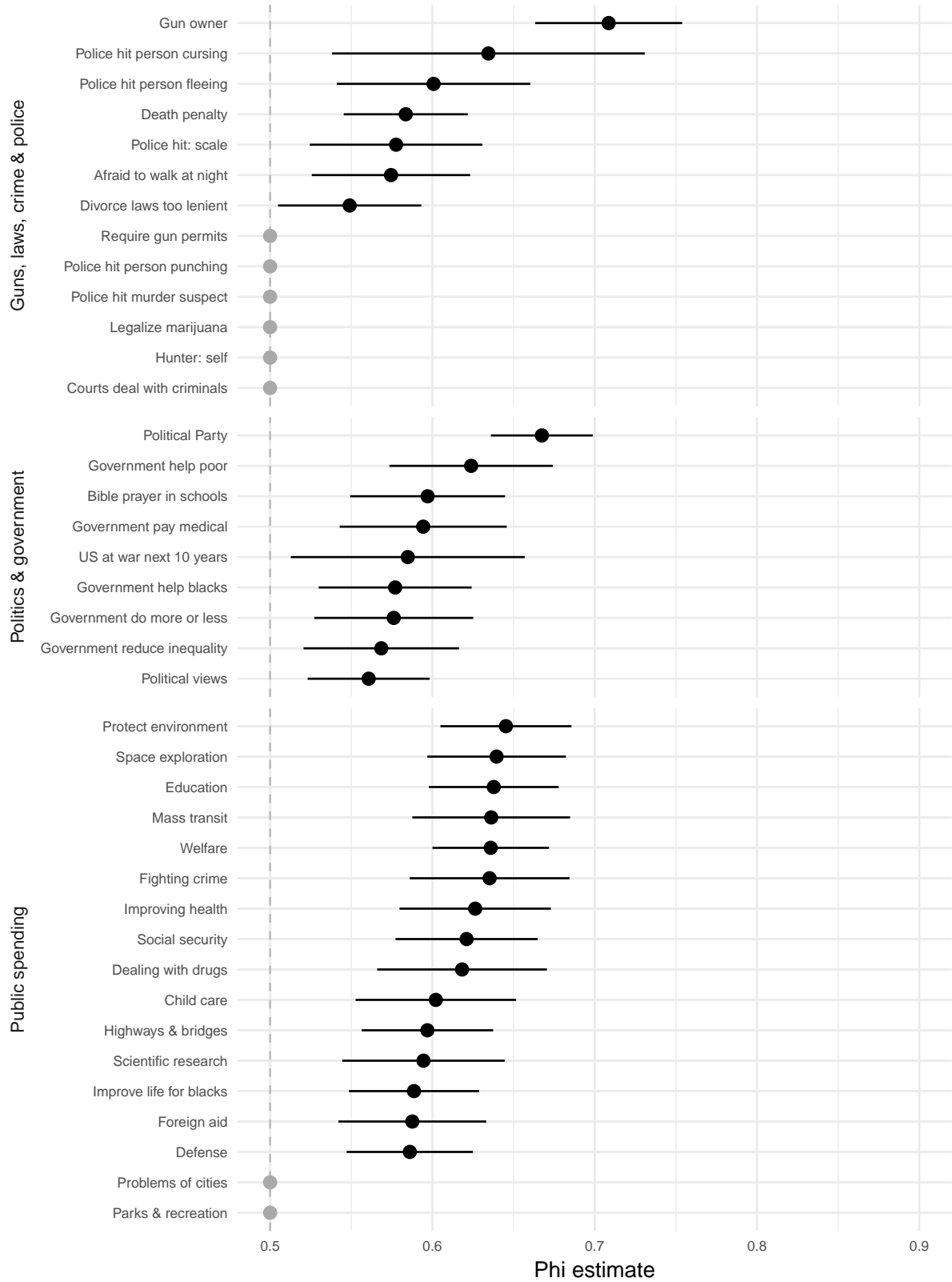


Figure 7: ϕ estimates for items about guns, law, crime and policing; politics and government; and public spending.

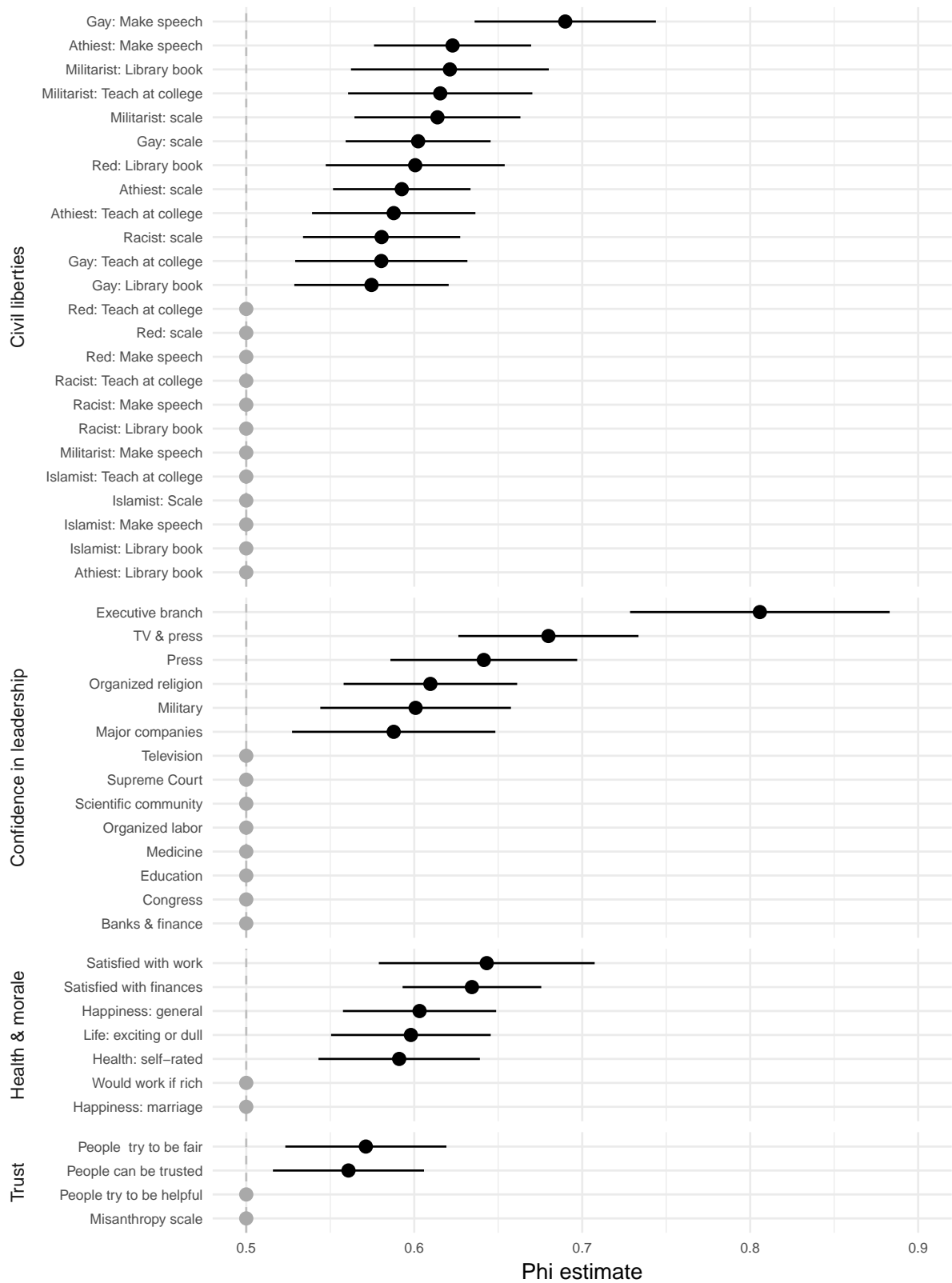


Figure 8: ϕ estimates for items about civil liberties, confidence in leadership, health, morale, and social trust.

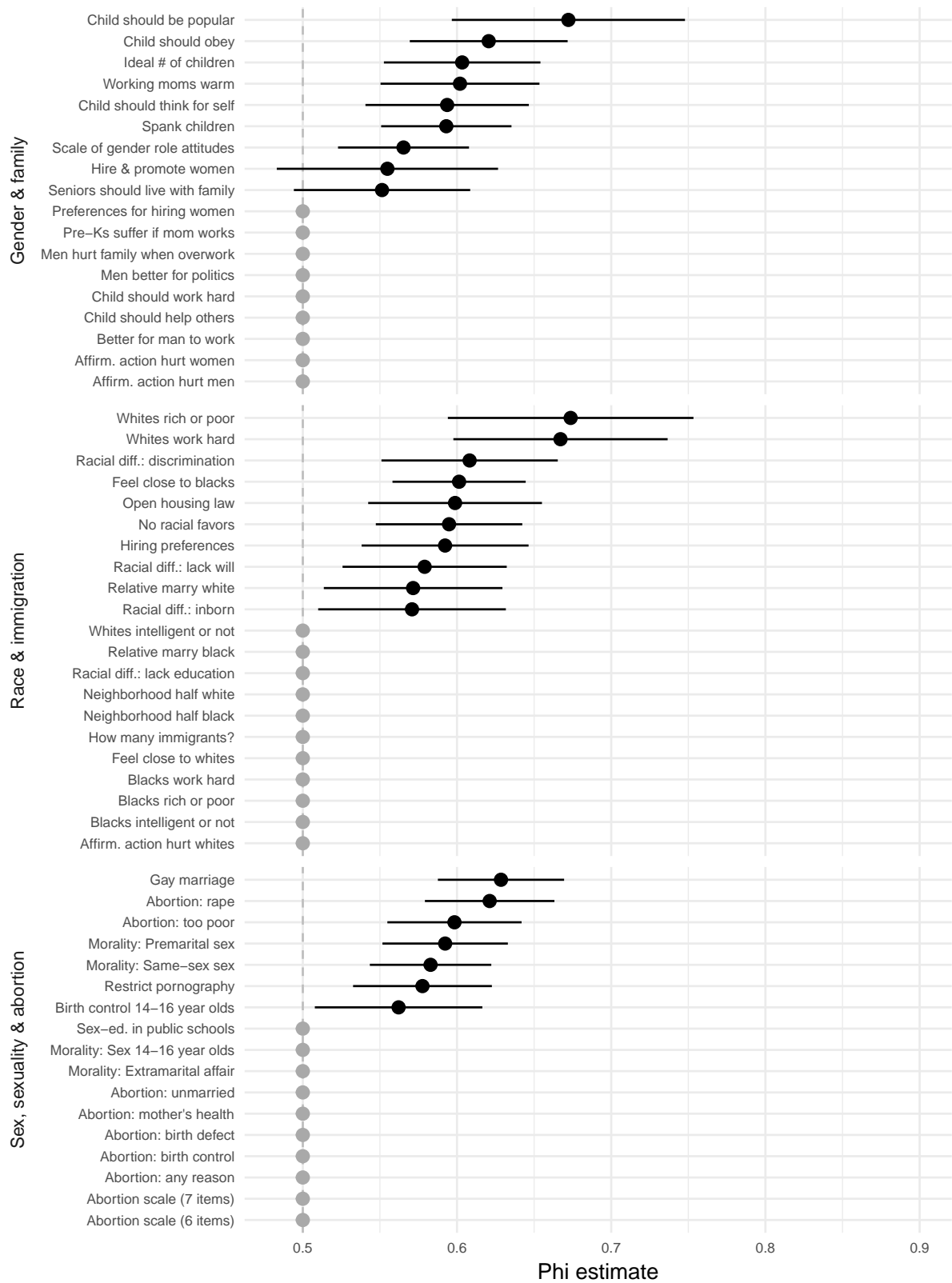


Figure 9: ϕ estimates for items about race, gender, sex, sexuality, and abortion.

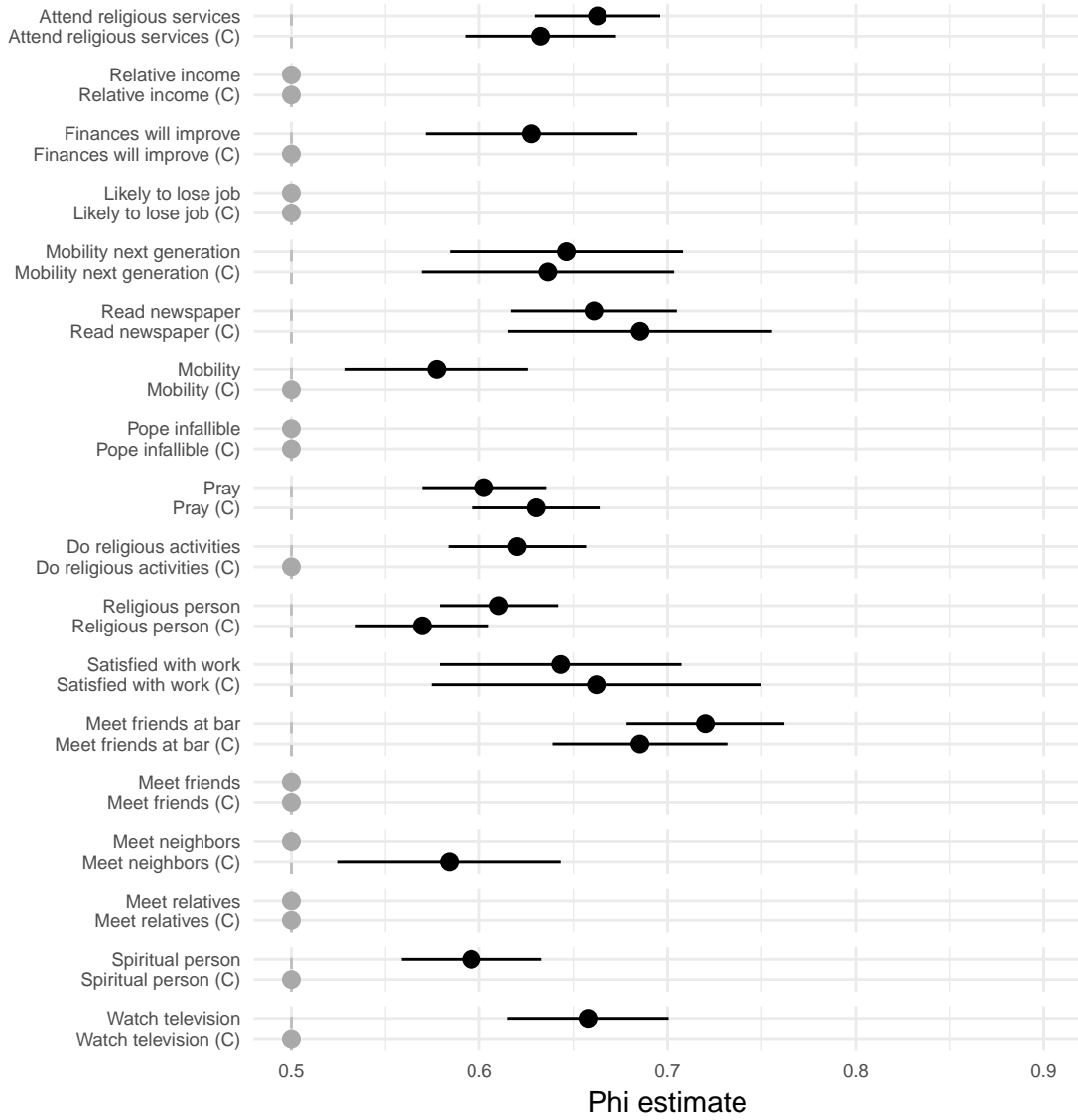


Figure 10: ϕ estimates for coarsened questions about religion, subjective socioeconomic status, and morale.

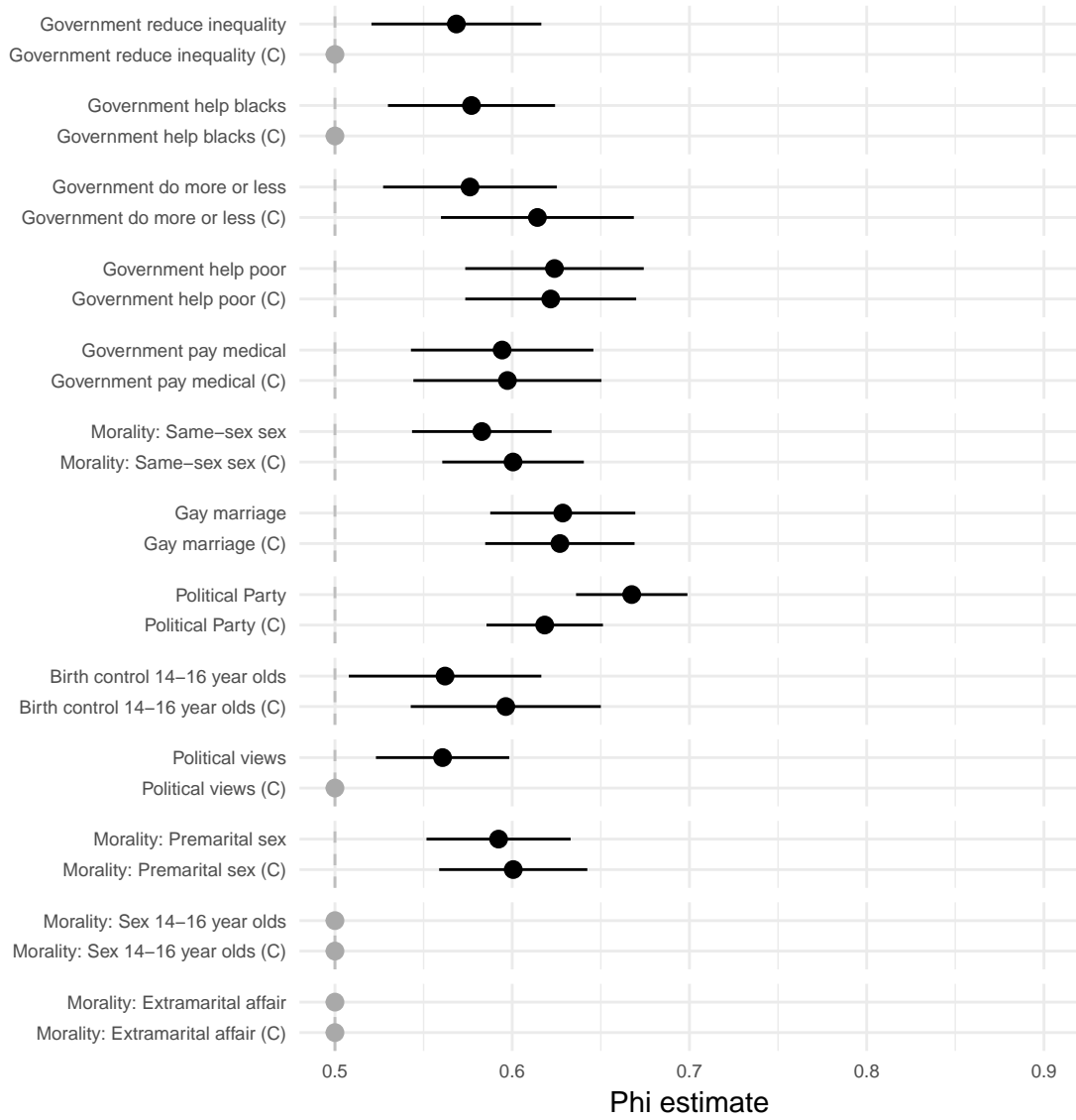


Figure 11: ϕ estimates for coarsened questions about politics, government, sex and sexuality.

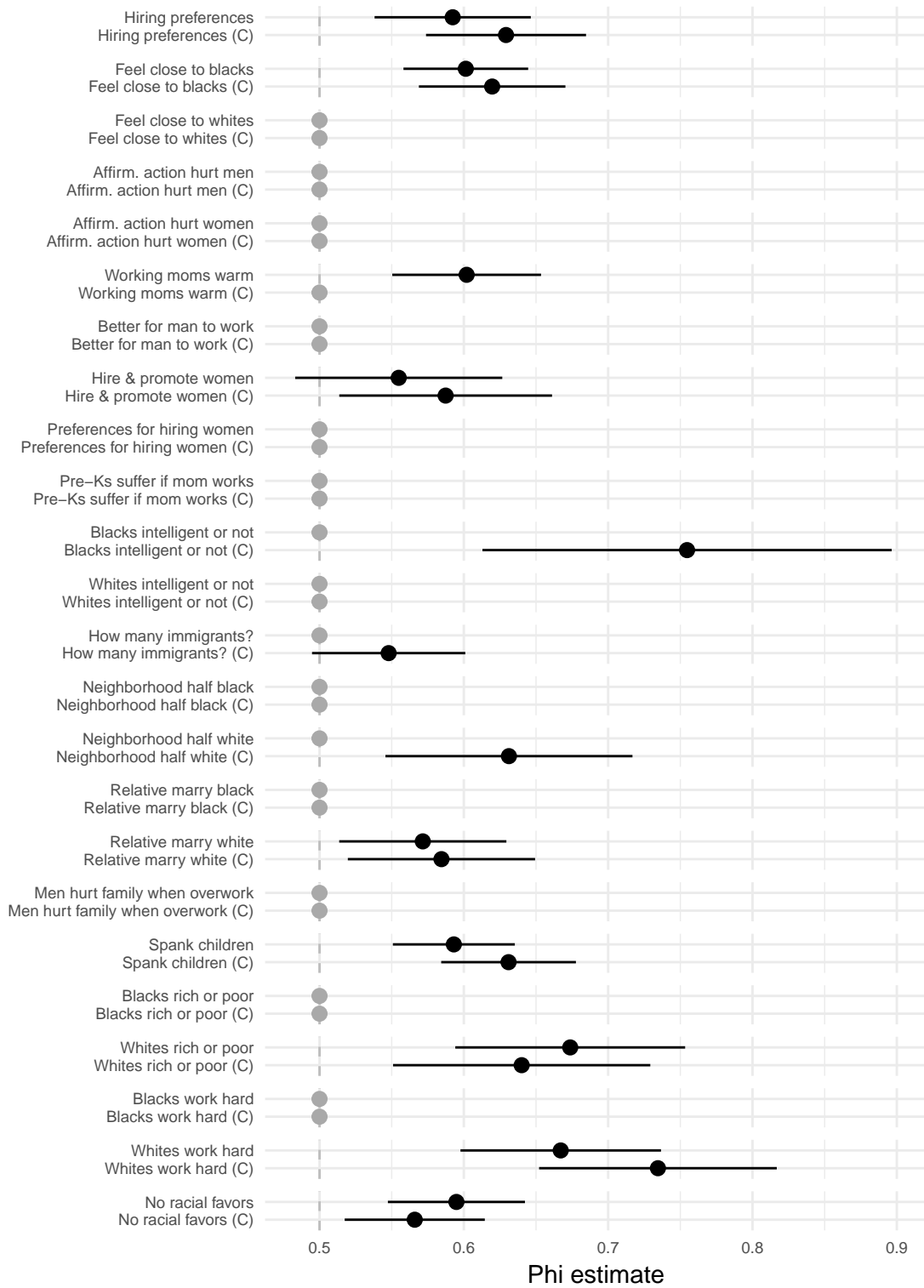


Figure 12: ϕ estimates for coarsened questions about race and gender.