

A model-based method for measuring stability and change in personal culture with three-wave panel data

Stephen Vaisey, Duke University

Kevin Kiley, Duke University

12/16/2020

1 Introduction

Kiley and Vaisey (2020) recently published a method for assessing whether survey respondents appear to be changing their beliefs between waves or whether they instead appear to be repeating fixed responses with temporary local influences. This question is important because these processes reflect very different theoretical models of how “personal culture” operates (see Lizardo 2017). That is, if cultural beliefs are primarily public and responsive to external discourse, we should observe more updating as people respond to changes in their local environment. On the other hand, if cultural beliefs are primarily something learned early, then “settled dispositions” should be relatively resilient to local shocks (see also Vaisey and Lizardo 2016).

In this paper, we build on Kiley and Vaisey (2020) and introduce an alternative method for distinguishing between situations where respondents appear to be actively updating their beliefs and situations where respondents’ beliefs appear to be settled. This method, based on structural equation modeling, provides a close fit to the theoretical models outlined in Kiley and Vaisey (2020) and provides even stronger support for their claim that most cultural beliefs reflect settled dispositions developed relatively early in life.

2 Background

Kiley and Vaisey (2020, hereafter KV) distinguish between two main models of cultural change implied in the literature: the *active updating model* and the *settled dispositions model*. The assumptions of each model are represented in Figure 1. The active updating model assumes that shocks affecting a person’s response to a question at one point in time persist to the next time period, while the settled dispositions model assumes that shocks do not persist.

KV’s goal is to develop a method that can distinguish between the two processes. They proceed by estimating the following equation on three-wave panel data:

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

The intuition behind Equation (1) is that, if people are updating their beliefs, survey responses that are closer in time (e.g., wave 1 and wave 2) should be more similar than responses that are more distant in time (wave 1 and wave 3). The ϕ parameter indicates the relative weight given to the more proximate wave 2 response (ϕ) and the more distal wave 1 response ($1 - \phi$). As ϕ approaches 1, wave 2 becomes more informative about wave 3 and wave 1 becomes less informative about wave 3, suggesting persistent change between waves. As ϕ approaches .5, however, both prior waves are assigned equal weight, suggesting that the answers reflect a persistent disposition to answer the question in a particular way.¹

¹The β parameter reflects the extent to which wave 3 is predictable from *any* combination of wave 1 and wave 2. Questions with low β are simply “noisy.”

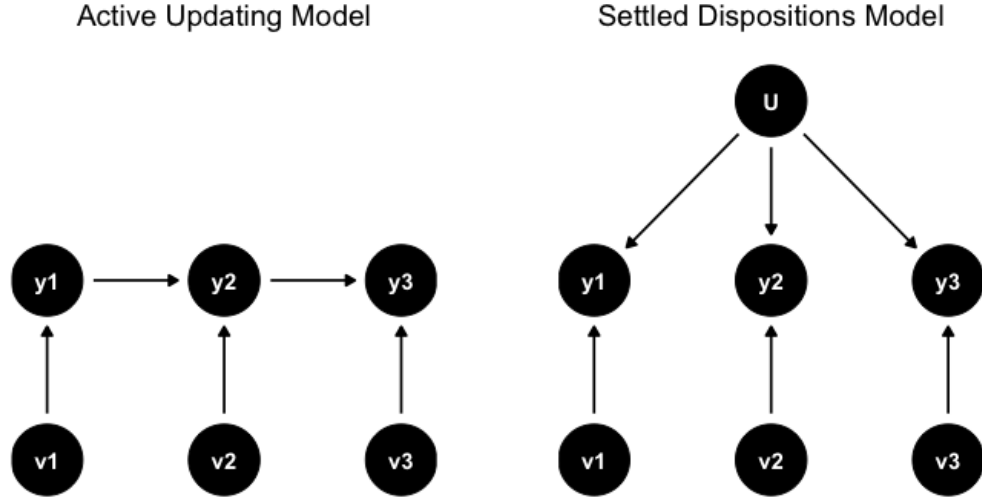


Figure 1: Theoretical Models

Estimating this model on 183 variables from the 2006-2014 GSS panels, KV find that about 40 percent of the variables show no evidence of active updating, and most of the remainder show only weak evidence of active updating. Although they point out several exceptions (e.g., some questions about gay rights, where attitudes do appear to be changing over the study period) they conclude that the settled dispositions model is a better default for thinking about public opinion among adults.

3 An alternative method

3.1 From theoretical to statistical models

The alternative method we propose uses structural equation modeling to implement the theoretical graphs more directly. By estimating the model this way, we can compare a wide variety of model specifications that have different theoretical implications (see Bollen and Brand 2010). For example, we could specify the core graph of the settled dispositions model as follows:

$$\begin{aligned} y_{i1} &= \alpha_1 + U_i + v_{i1} \\ y_{i2} &= \alpha_2 + U_i + v_{i2} \\ y_{i3} &= \alpha_3 + U_i + v_{i3} \end{aligned} \tag{2}$$

From this basic version, we can add further constraints. For example we can constrain $\alpha_1 = \alpha_2 = \alpha_3$ if we are willing to assume that the mean response is the same at all waves. And we could constrain $\sigma_{v1}^2 = \sigma_{v2}^2 = \sigma_{v3}^2$ if we are willing to assume that the variance of the responses does not change over time.

This is not the place for a general overview of structural equation modeling (Bollen 1989; see Kline 2015 for an accessible introduction). But the basic intuition in the three-wave case is that we are attempting to reconstruct 9 observed elements from the data (3 means, 3 variances, and 3 covariances) using a model with fewer than 9 parameters. For example, if we make the assumptions in Equation (2), including the equality constraints in the following paragraph, we can estimate a model that has only 3 parameters: one value of α (shared by all waves), one σ_v^2 (shared by all waves), and one σ_U^2 . Here U is a “latent variable,” which, in this specification, simply means a person-specific “error” that gets added equally to a person’s response at every

wave. U_i reflects an individual’s unchanging tendency to respond to a certain question the same way over the study period (a “settled disposition”).

Using this model will not work well to “predict” the data if the data were generated by a process that looks like the active updating model on the left of Figure 1. We could instead estimate the following model:

$$\begin{aligned} y_{i3} &= \alpha_3 + \rho y_{i2} + v_{i3} \\ y_{i2} &= \alpha_2 + \rho y_{i1} + v_{i2} \end{aligned} \tag{3}$$

Here we cannot make Y_1 a dependent variable because we don’t have its previous value (Y_0). Therefore we can only use it as a predictor. As above, we can make further simplifications if we want (e.g., $\alpha_2 = \alpha_3$ or $\sigma_{v2}^2 = \sigma_{v3}^2$). Using this fully simplified version, we are estimating a model with 5 parameters: one value of α (shared by waves 2 and 3), one σ_v^2 (shared by waves 2 and 3), the ρ autocorrelation parameter, and the mean and variance of Y_1 .

If Y_t were always measured perfectly, we could simply compare the penalized likelihoods (e.g., Bayesian Information Criteria) of the two models to see which is more likely to be the true model given the data (Bollen et al. 2014). Unfortunately, most GSS items likely contain some measurement error. Any measurement bias—the systematic tendency for an item to over- or underestimate a person’s “true” value—will induce error correlations between waves that have no basis in the causal process. Thus U_i in Equation (2) represents both settled dispositions and respondent-specific measurement bias.

For this reason, the most reasonable specification of the active updating model combines features of equations (2) and (3):

$$\begin{aligned} y_{i3} &= \alpha_3 + \rho y_{i2} + U_i + v_{i3} \\ y_{i2} &= \alpha_2 + \rho y_{i1} + U_i + v_{i2} \\ \text{Cov}(U, Y_1) &= \tau \end{aligned} \tag{4}$$

Adding U to equation (4) allows for systematic correlations between the wave 2 and wave 3 responses due to measurement bias or error.² Estimating τ , the covariance between U and Y_1 , reflects that U and Y_1 share common unobserved causes (e.g., systematic measurement error, values of the pre-survey Y_0).

The most straightforward³ way to compare models is to specify them as special cases of a more general model. Figure 2 shows the basic model.

We have three basic choices to make:

1. Do we allow wave-to-wave updating of responses (estimate ρ)?
2. Do we allow for settled dispositions and systematic bias (estimate σ_U^2 and τ)?
3. Do we allow wave-to-wave changes in the mean response (estimate α_2 and α_3 separately) or assume no aggregate change ($\alpha_2 = \alpha_3$)?

Crossing these three binary choices would normally lead to 8 candidate models. However, we exclude from consideration models that contain neither updating nor settled dispositions because no theoretical perspective argues for them. This leaves us with six models, shown below.

Model	σ_U^2, τ	ρ	α_t	# parameters
AUM1	0	free	=	5

²Unfortunately, in the three-wave case there’s no way to distinguish systematic measurement bias or error from true person-level stable differences as components of U without additional untestable assumptions.

³We could specify an even more parsimonious version of the settled dispositions model by estimating Equation (2) with the equality assumptions discussed in the text. This would effectively be estimating a confirmatory factor analysis on Y_t . This would only require three parameters. We will revisit this later.

Model	σ_U^2, τ	ρ	α_t	# parameters
AUM2	0	free	free	6
AUM3	free	free	=	7
AUM4	free	free	free	8
SDM1	free	0	=	6
SDM2	free	0	free	7

For any given three-wave panel, we can compare the fit of these models against each other to help determine which of the models is most likely to have generated the data.

3.2 Advantages of the approach

Kiley and Vaisey (2020) also compare estimates using Equation (1) with and without constraints to adjudicate between the active updating and settled dispositions accounts. The approach we outline here is essentially identical in spirit but a bit different in the details. The main difference is that KV compare models based solely on how well they “predict” the Wave 3 response. Using structural equation models allows comparing the fit of the model to all the observed data, not just the final response. For example, the KV model has no leverage in cases where respondents give the same response in waves 1 and 2, but a different response in wave 3, while the model presented here does.

To oversimplify slightly, the models presented here essentially adjudicate whether people are equally likely to deviate from their baseline response in all waves, which would be evidence of the settled dispositions model, or whether waves 1 and 3 are more likely than wave 2 to be deviations, which would be evidence of persisting changes and active updating.

3.3 Analytic strategy

We use the same 183 variables and three panels of the General Social Survey as Kiley and Vaisey (2020). Rather than pool the panels, we estimate models separately on each one. Most variables appear in all three panels but 6 are measured in only one panel and one is measured in two panels. This gives us 536 total three-wave data sets.

For each of the 536 datasets, we estimate all 6 of the candidate models. We then compare fits using the Bayesian Information Criterion (BIC, see Raftery 1995; Bollen et al. 2014). In order to make the best case possible for each theoretical model, we will compare the best-fitting of the four AUM models with the best-fitting of the two SDM models. In cases where the BIC difference between the two “finalist” models is less than 2, we consider the evidence inconclusive (Raftery 1995; Bollen et al. 2014).

The approach is quite conservative, favoring the active updating model over the settled dispositions model when there is meaningful evidence that some amount of people in the sample are making durable changes in opinion or behavior. This should not be taken as evidence that *many* people in the population are making durable changes, only that there is any evidence of durable change at all.

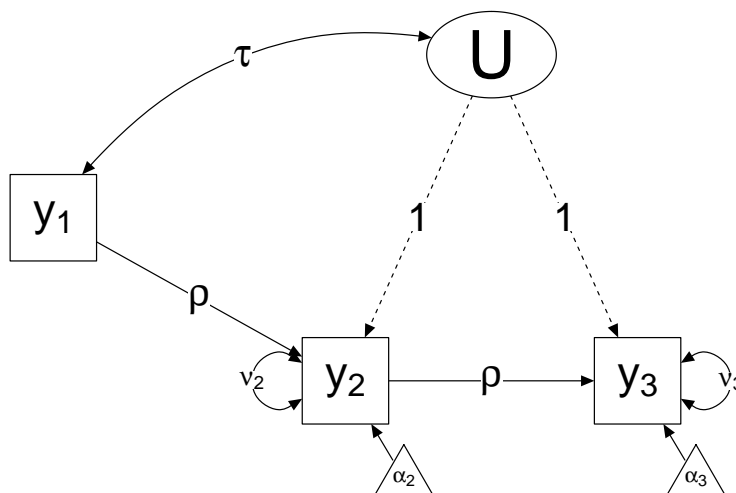


Figure 2: General Model

4 Results

4.1 BIC comparisons

Figure 3 shows the results of these comparisons for all 536 variable-panels. One of the two SDM models is preferred for 71% of cases, with one of the four AUM models preferred in 16% of cases. The remaining cases are inconclusive.

If we look at only variables measured in all three panels, we can also compare the number of variables where all three agree. There are 69 variables that unanimously point to the SDM and only 3 that unanimously favor the AUM. The three unanimous AUM variables are **news** (reading a newspaper), **owngun** (having a gun in your home), and **socbar** (going to a bar or tavern). The presence of **owngun** on this list is promising because we *know* that physical objects act according to an active updating model – when you buy a gun, it stays in your house until you get rid of it.

4.2 Goodness-of-fit

Almost all (99.5 percent) of the preferred SDM models have acceptable fits ($RMSEA < .08$). But we can put the SDM to an even stronger test by estimating a model that assumes no change and no updating across waves. This is the model implied by Equation (2) and is equivalent to a simple confirmatory factor analysis with just three parameters: one value of α (shared by all waves), one σ_v^2 (shared by all waves), and one σ_U^2 . In essence, this model assumes that each wave is a report of a fixed quantity determined before the time of the study.

Estimating this model shows that it fits 83.4 percent of the variable-panels acceptably ($RMSEA < .08$) and 53.4 percent of the variable-panels well ($RMSEA < .05$). Figure 4 shows the overall distribution of RMSEA values. We don't want to take these conventional cutoffs too literally (see Barrett 2007), but altogether these results indicate that the majority of GSS items are highly compatible with the settled dispositions model.

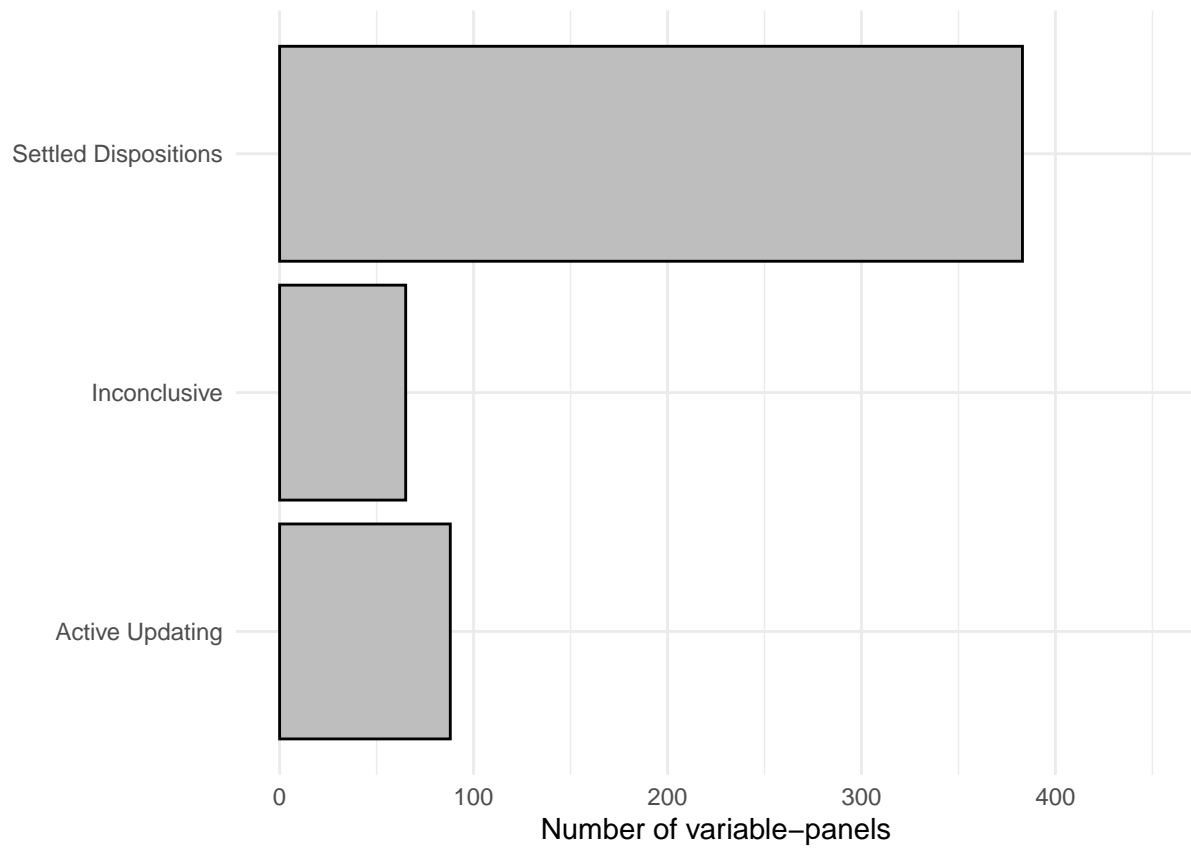


Figure 3: Preferred Models according to BIC Comparison

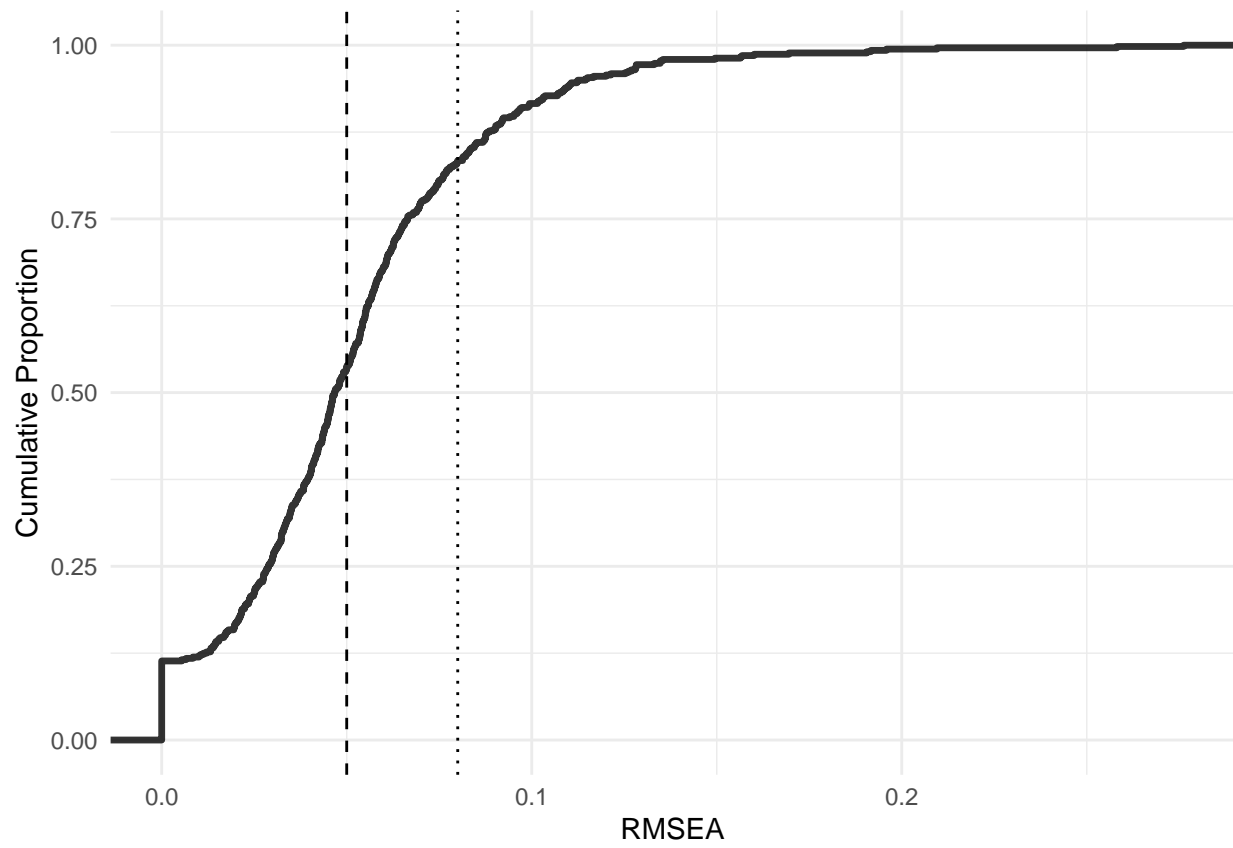


Figure 4: Distribution of Fit Values for Simple CFA

5 Discussion and Conclusions

Building on KV, we have presented a model-based approach to estimating whether the pattern of attitude and behavior change in the population should be best thought of as following a *settled dispositions* model, where people have stable baselines and change is temporary, or an *active updating* model, where changes tend to persist. Using a set of structural equation models that formalize these theoretical processes under a variety of assumptions, we compared the best-fitting settled dispositions model to the best-fitting active updating model for each variable in each panel, a total of 536 variable-panel pairs. The overwhelming majority of variable-panels prefer the settled dispositions model, with 69 variables preferring the settled dispositions model in all panels and only three variables preferring the active updating model in all panels.

While we cannot account for attitudes and behaviors that are not measured here, these results suggest, consistent with KV, that a wide range of attitudes and behaviors tend to be settled by the time people are old enough to participate in surveys like the GSS. While this is not surprising since we examine the same set of questions as KV, it is important that different statistical assumptions lead to similar substantive conclusions. People might change their responses from wave to wave, but this change does not tend to persist, meaning it reflects temporary changes in opinion or measurement error.⁴ There is little evidence of even a moderate amount of durable change in opinions in the medium-term (2-4 years).

On specific variables, the approach presented here occasionally suggests different conclusions than those presented in KV, as would be expected under models that make different assumptions, with the model presented here favoring the SDM model in more cases. One reason these differences emerge is because the model presented here uses slightly more data than the KV model. A person who gives the same response in waves 1 and 2 but a different response in wave 3 is mostly useless under the KV model but can factor into the calculation in the model presented here. Because variable-panels can switch from favoring the active updating model to favoring the settled dispositions model with the addition or subtraction of a few cases, this suggests that evidence of the AUM is very weak even when it is favored. In particular, when the AUM is favored, it is because a very small proportion of the population is making durable changes. It is difficult to draw broader conclusions about these discrepancies, as they cut across both question substance and question structure.

While it would be impossible to discuss all 183 variables or 536 variable-panels in depth, there are some broad patterns worth highlighting. Figure 5 plots the proportion of variable-panels preferring each model by subject material.⁵ KV's results suggest that issues related to civil liberties, gender, and social trust tend to be settled by the time people entered the the GSS. The method outlined above finds similar results, with 86 percent, 89 percent, and 91 percent of variable-panels in these categories, respectively, favoring the settled dispositions models. These are the three categories with the largest proportions of variable-panels favoring the settled dispositions model.

Similarly, KV suggested that public behaviors were more likely to demonstrate active updating, since these behaviors receive social reinforcement that facilitate durable change. In the results presented here, social behaviors such as socializing at a bar and religious behaviors such as church attendance are the two categories that show the highest rates of active updating, with 41 percent and 50 percent of variable-panels in these categories, respectively, showing evidence of active updating.

Another issue worth touching on is variation across panels in favoring the settled dispositions or active updating models within the same question. Of the 176 questions that appear in all three panels, 110 either favor the settled dispositions model or are inconclusive in all three waves (with 69 clearly favoring active updating in all panels). However, 58 variables favor each model in at least one panel.

Some of these discrepancies are likely due to real changes in the social environment that only affect members of one panel. For example, several of the questions with divergent findings pertain to whether the federal

⁴Within the settled dispositions model, our method cannot distinguish between change that occurs because people lack opinions and are more or less guessing at random (Converse 1964), change that occurs because people are influenced by short-term messaging and other considerations (Zaller 1992), or changes that occurs because of other forms of measurement error (Ansolabehere, Rodden, and Snyder 2008; Alwin 2007). We assume that these short-term changes are some combination of these forces, which might vary by question.

⁵It should be kept in mind when comparing across subject areas that the questions in subject areas often have different structures, and the questions asked in each category should not be thought of as a random sample of issues in that domain.

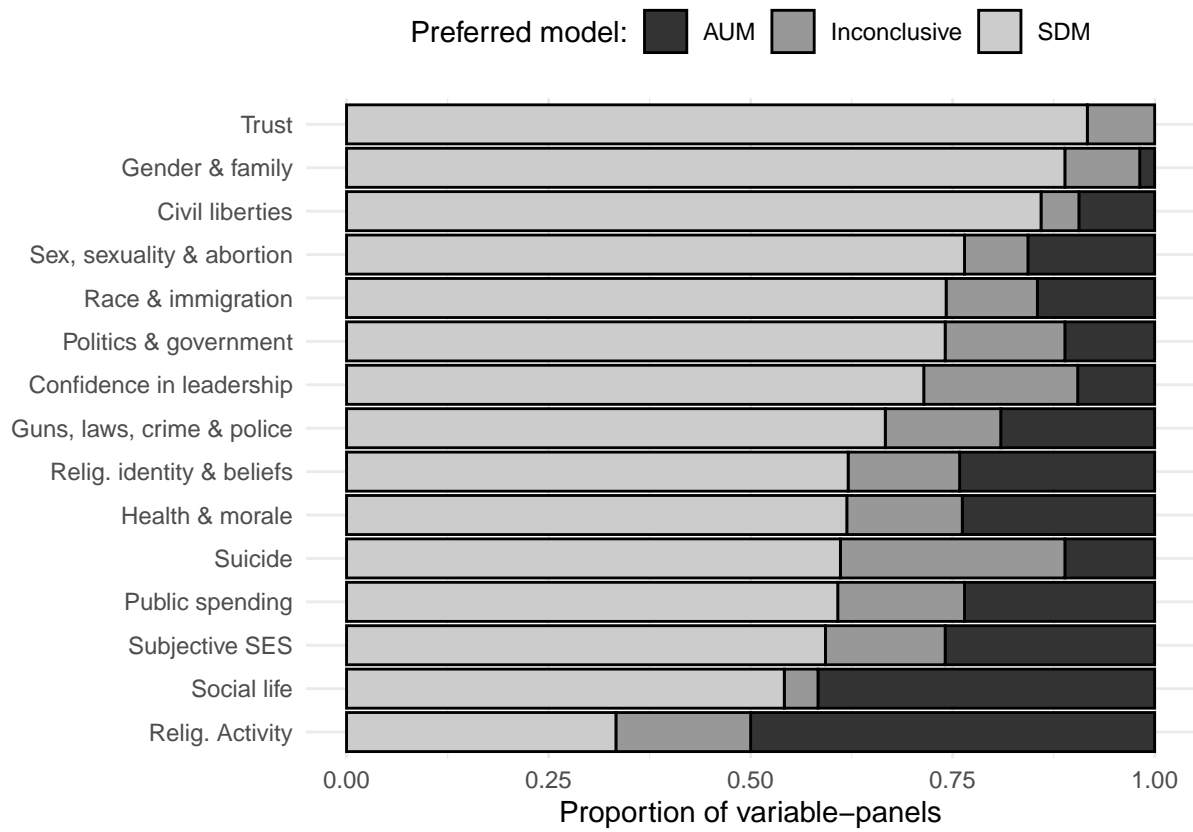


Figure 5: Proportion of preferred models, by subject category

government is spending too little, about right, or too much on various priorities. People’s responses might be changing in response to changes in federal spending, changes in who controls the federal government, or changes in public messaging around these issues, which might only happen in certain years.

For questions that demonstrate active updating in at least one panel but have no obvious changing external referent (such as church attendance, religious activity, partisan identification, frequency of prayer, and support for marijuana legalization), durable change might be happening but occur at such low rates that finite samples using imperfect measures do not capture enough respondents undergoing durable change in a particular panel to favor the active updating model. Again, even on generally stable issues we expect some people in the population to be actively updating their behavior – most people know an adult who has changed their opinion on some issue – but this should be thought of as the exception.⁶ Our results suggest that when we observe someone change their response to a survey question, the assumption should be that it will not be a permanent change.

This points to a limitation of both this approach and the one used by KV. Neither approach can quantify the proportion of the sample that is undergoing active updating. It is not clear whether, for the variable-panels that show evidence active updating, there is a large or small proportion of respondents making durable changes.

The two approaches share additional limitations. Both approaches assume linearity in the variable being measured. If an outcome is measured on an ordinal scale such as a four-point Likert scale (strongly agree, agree, disagree, strongly disagree), a change from agree to strongly agree is treated the same as a change from disagree to agree, which might or might not be a reasonable assumption. Both approaches also rely on weights to account for differential non-response over the course of the panel. If the mechanism of non-response and active updating are both driven by some unobserved variable, then our results might under-estimate the probability of durable change. That being said, we have no reason to assume that non-response is more likely to be related to active updating than settled dispositions.

The approach presented here and the approach outlined in KV attempt to adjudicate a similar theoretical question: are there more people reporting patterns that look like active updating than we would expect if people behaved according to a settled dispositions model? The difference between the two approaches are sets of assumptions about what counts as an active updating pattern and what counts as a settled dispositions pattern. The fact that different sets of assumptions reach the same substantive conclusion, even under tests very favorable to detecting active updating, strengthens the findings from KV and others (e.g., Vaisey and Lizardo 2016) that the settled dispositions model should be thought of as the dominant model for attitude behavior in adults.

⁶One reason people can recall examples of people in their social networks who have changed beliefs and behaviors is because it is rare and tends to stick out.

References

- Alwin, Alwin, Duane F. 2007. *Margins of Error: A Study of Reliability in Survey Measurement*. Hoboken, N.J.: John Wiley & Sons.
- Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102 (2): 215–32. <https://doi.org/10.1017/S0003055408080210>.
- Barrett, Paul. 2007. "Structural Equation Modelling: Adjudging Model Fit." *Personality and Individual Differences*, Special issue on Structural Equation Modeling, 42 (5): 815–24. <https://doi.org/10.1016/j.paid.2006.09.018>.
- Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.
- Bollen, Kenneth A., and Jennie E. Brand. 2010. "A General Panel Model with Random and Fixed Effects: A Structural Equations Approach." *Social Forces* 89 (1): 1–34.
- Bollen, Kenneth A., Jeffrey J. Harden, Surajit Ray, and Jane Zavisca. 2014. "BIC and Alternative Bayesian Information Criteria in the Selection of Structural Equation Models." *Structural Equation Modeling : A Multidisciplinary Journal* 21 (1): 1–19. <https://doi.org/10.1080/10705511.2014.856691>.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics (1964)." In *Ideology and Discontent*, edited by D. E. Apter, 18:206–61. New York: Free Press. <http://www.tandfonline.com/doi/abs/10.1080/08913810608443650>.
- Kiley, Kevin, and Stephen Vaisey. 2020. "Measuring Stability and Change in Personal Culture Using Panel Data." *American Sociological Review* 85 (3). SAGE Publications Inc: 477–506. <https://doi.org/10.1177/0003122420921538>.
- Kline, Rex B. 2015. *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Fourth edition. New York: The Guilford Press.
- Lizardo, Omar. 2017. "Improving Cultural Analysis: Considering Personal Culture in Its Declarative and Non-declarative Modes." *American Sociological Review* 82 (1): 88–115. <https://doi.org/10.1177/0003122416675175>.
- Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25 (3): 111–63.
- Vaisey, Stephen, and Omar Lizardo. 2016. "Cultural Fragmentation or Acquired Dispositions? A New Approach to Accounting for Patterns of Cultural Change." *Socius: Sociological Research for a Dynamic World* 2 (January): 2378023116669726. <https://doi.org/10.1177/2378023116669726>.
- Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.