Correction

Kevin Kiley Stephen Vaisey

12/14/2021

We recently discovered that an error in the R source code for a function we used in the original analysis for our 2020 ASR paper, "Measuring Stability and Change in Personal Culture Using Panel Data," led to erroneous results, and we would like to correct the record. While this error produces changes in how we classify some variables throughout our analyses in that paper, the error does not affect the overall pattern of results or our substantive interpretation of those results.

In our original paper, we estimated two non-linear least squares models – one that modeled changes within people over time as random departures from a settled baseline (a settled dispositions model) and one that modeled changes within people over time as durable updates (an active updating model) – for 183 variables from the General Social Survey's rotating three-wave panels. We then compared the fit of these two models using the Bayesian Information Criterion to evaluate which model better fit the observed data. We found the settled dispositions model was preferred for 75 questions, meaning there was no evidence of durable change for these items. We also found that even for questions where there was evidence of durable change, most change was not durable.

We conducted this analysis using R version 3.5.3. In versions of R prior to 4.0.3, there was an error in the way R calculated the log likelihood of non-linear least squares models, which resulted in an incorrect calculation of the Bayesian Information Criterion. This led to errors in how we judged the two models against each other.

In versions of R prior to 4.0.3, the log likelihood of non-linear least squares models was calculated as:

$$-\frac{N}{2}(log(2\pi) + 1 - log(N) - \sum_{i=1}^{N} log(w) + log(\sum_{i=1}^{N} [w^2 * (y - \hat{y})]^2))$$

where N is the number of observations, w are survey weights, y is the observed value of the outcome, and $\hat{y}$ is the predicted value of the outcome. The key issue for our analysis was that the formula applied survey weights to residuals twice (w^2 instead of w). The formula also did not divide the sum of the log of the weights by N. This latter issue produced incorrect log likelihood values, but because this term is constant across both models, it did not affect comparisons.

The corrected calculation of the log likelihood for non-linear least squares models, implemented in version 4.0.3,[1] is:

$$-\frac{N}{2}(log(2\pi) + 1 - log(N) - \frac{\sum_{i=1}^{N} log(w)}{N} + log(\sum_{i=1}^{N} [w * (y - \hat{y})]^2))$$

Because the error lies in how survey weights were applied to calculation of the BIC, there is no clear pattern in how results change when the correct formula is applied. Re-running our analyses with the corrected calculation produces slightly different results than what we presented in the published paper. Seventeen variables where we found a better fit for the stable disposition model prefer the active updating model under the corrected calculation. Fourteen variables where we found that the active updating model fit the data

[1] Documentation of the error can be found: https://bugs.r-project.org/show_bug.cgi?id=16100#c3

better prefer the settled dispositions model under the corrected calculation. This resulted in a net change of three more variables preferring the active updating model to the settled dispositions model (75, compared to 72 in the original analysis). The revised calculation also found fewer variables with a statistically significant difference between the ϕ estimates for people 30 and under and people over 30 (9 instead of 22), but the overall pattern of younger people showing more active updating is still consistent.

The items we labeled as "settled dispositions" under the incorrect calculation that are "active updating" under the corrected calculation are: abnomore, confinan, conjudge, contv, courts, fund, hapmar, misanthropy, neverrelact, polattak, reborn, socommun, socrel, spkath, spkrac, teensex, and xmovie.

The items we labeled as "active updating" under the incorrect calculation that are "settled dispositions" under the corrected calculation are: aged, colrac, divlaw, getahead, marwht, meovrwrk, pillok, polhitok, racdif2, socfrend, suicide1, trust, uswary, and workblks.

While fixing this calculation changes how these variables are classified, the erroneous calculation was unrelated to our estimates of ϕ or β, and the overall substantive conclusions of our paper, which we reproduced with different methods elsewhere (**vaisey2021?**), remain the same.

Corrected versions of Figures 2, 3, 4, 5, and A1-A4 are included below. Figures 1 and 6 were not affected by the change.

We thank Hugo Mercier for bringing the discrepancy to our attention. We thank Sebastian Meyer for patching the function in R and for explaining the change in personal communications.

Figures

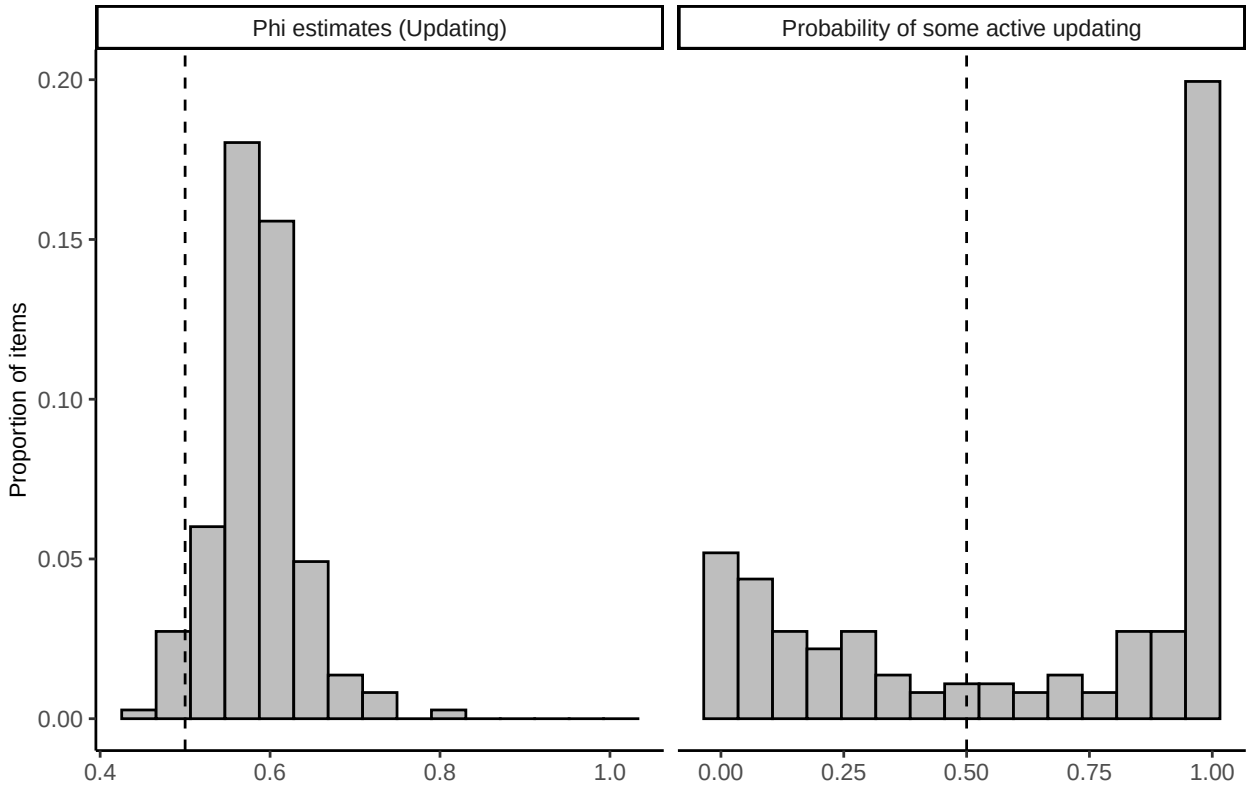

Figure 1: Distribution of phi estimates and probabilities that items show evidence of active updating.

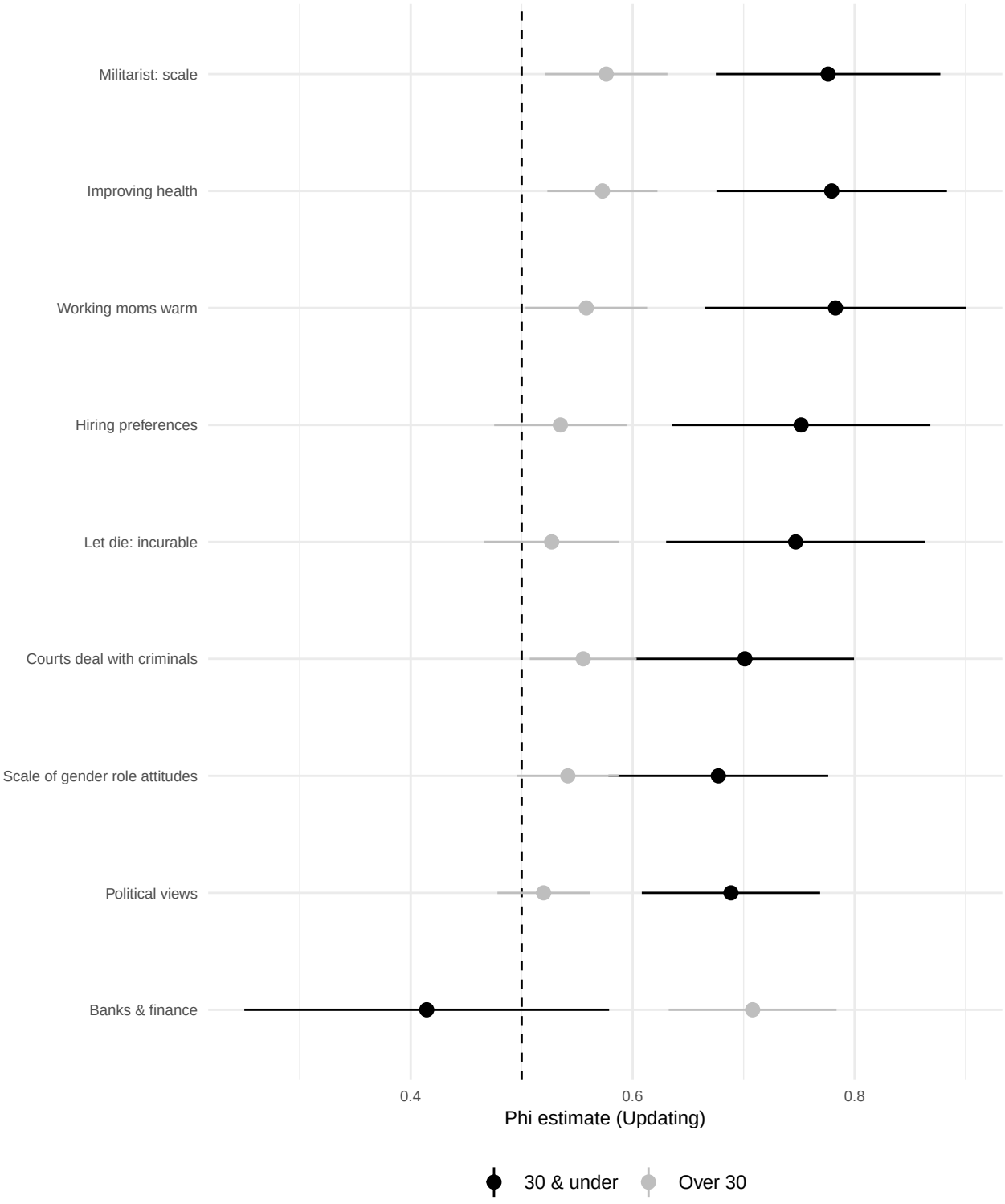

Figure 2: Comparison of phi estimates for individuals over and equal to or less than 30 years old, with 95 percent confidence intervals.

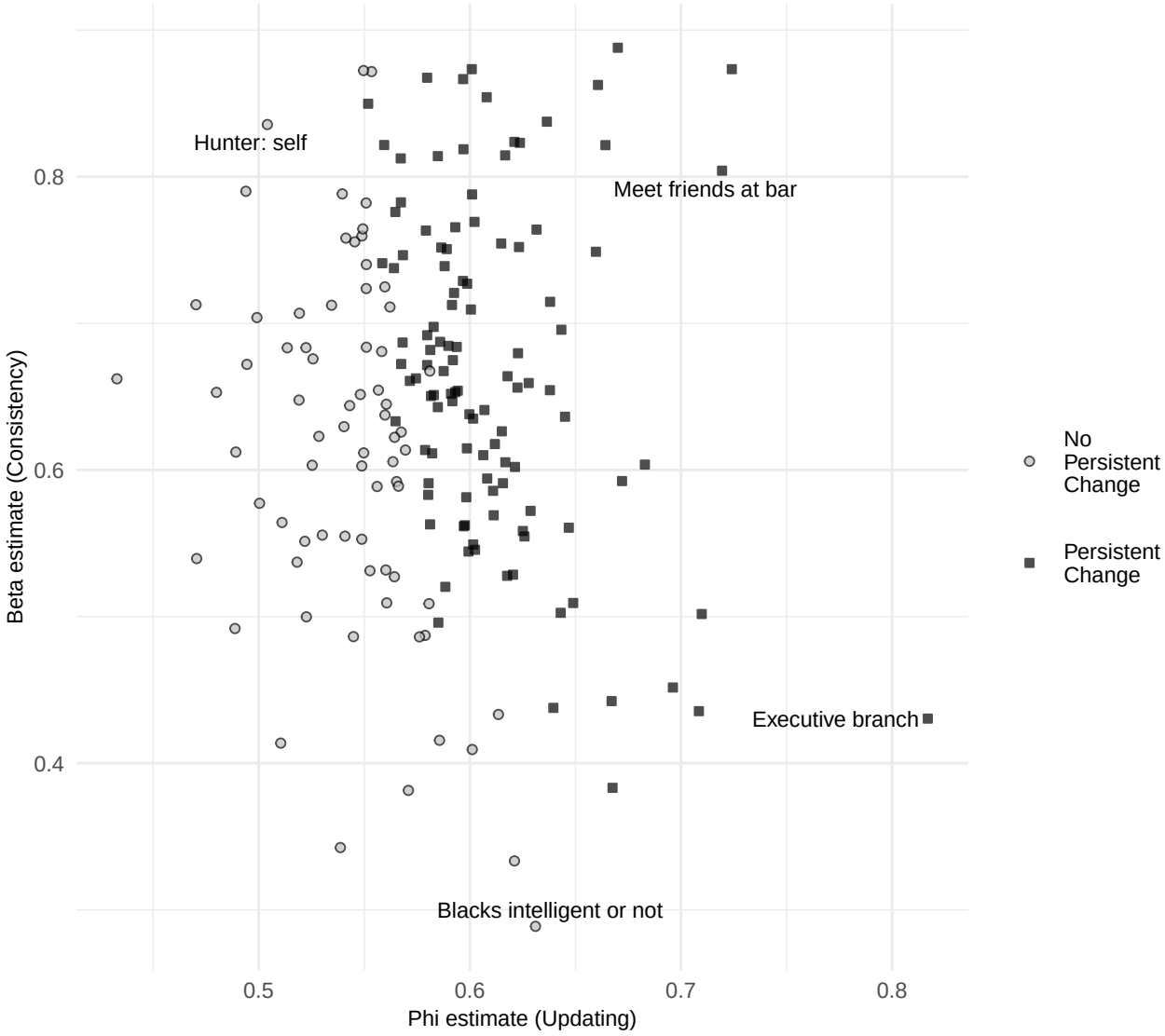

Figure 3: Distribution of beta and phi estimates for GSS items, by whether model prefers phi = .5 constraint.

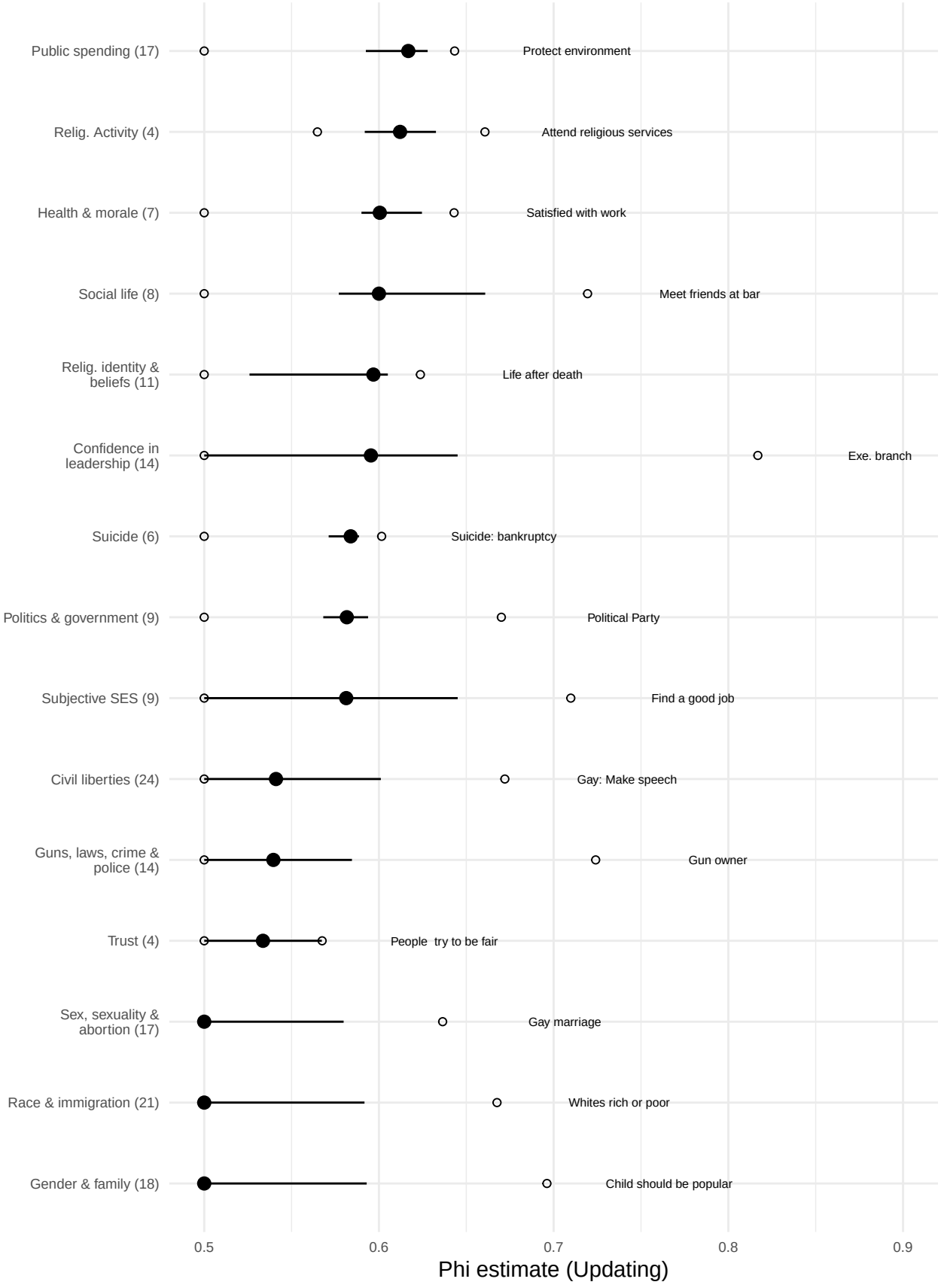

Figure 4: Summary of phi estimates for all items, by topical group. Numbers in parentheses indicate the number of items in each topical group.

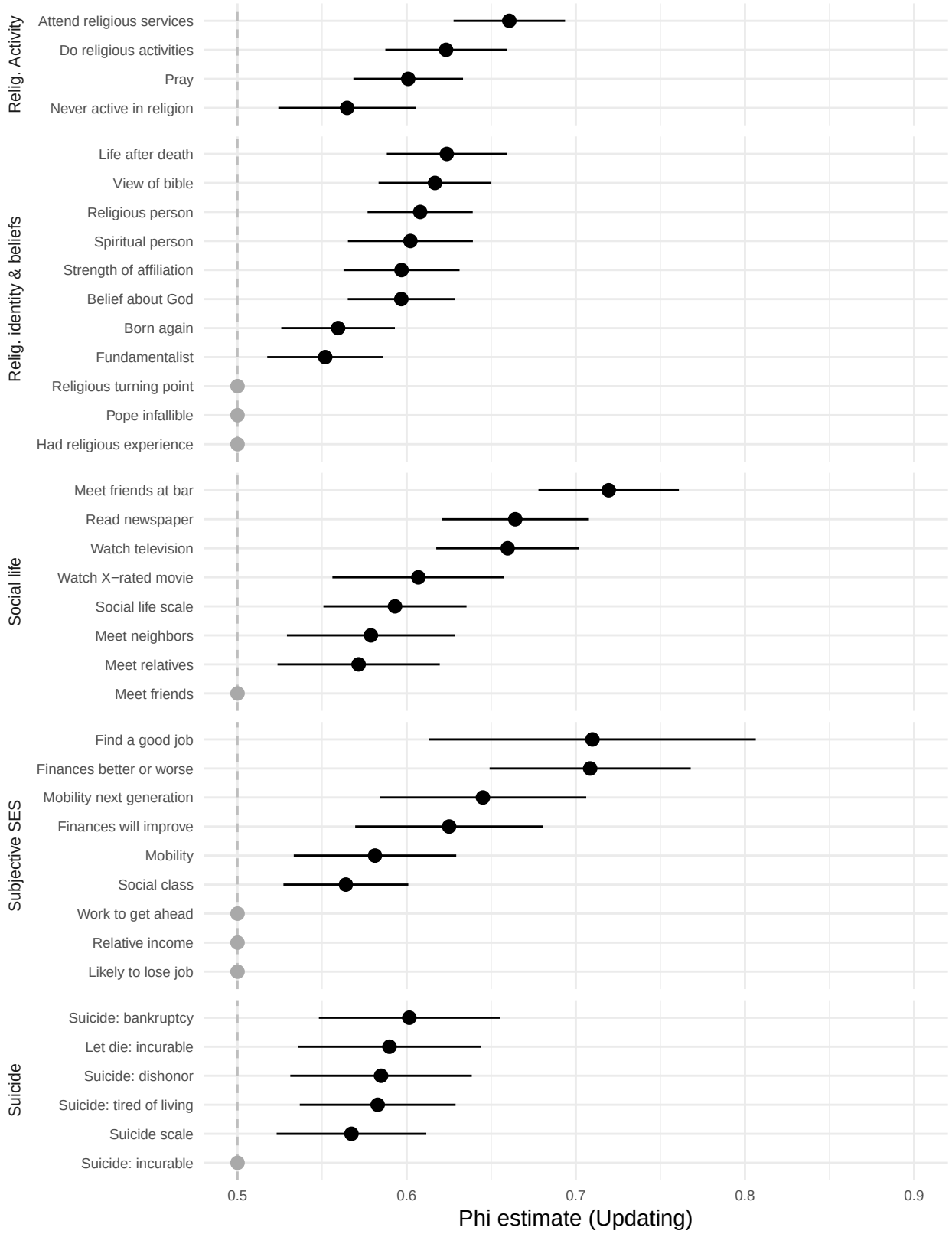

Figure 5: Phi estimates for items about religious activity and beliefs, social life, subjective SES, and suicide.

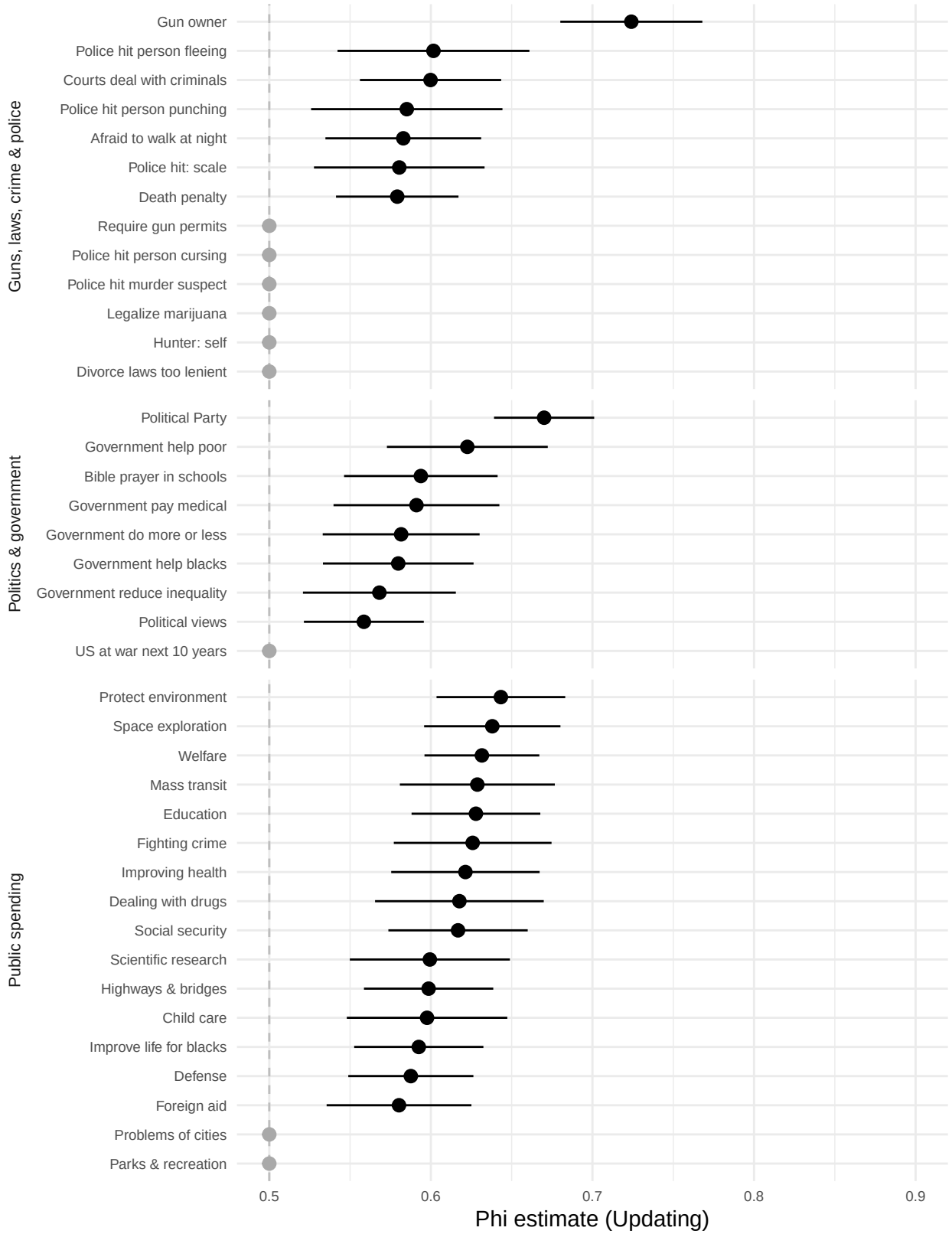

Figure 6: Phi estimates for items about guns, law, crime and policing; politics and government; and public spending. The item 'Police can hit citizens,' which has phi = .43, has been removed for ease of viewing

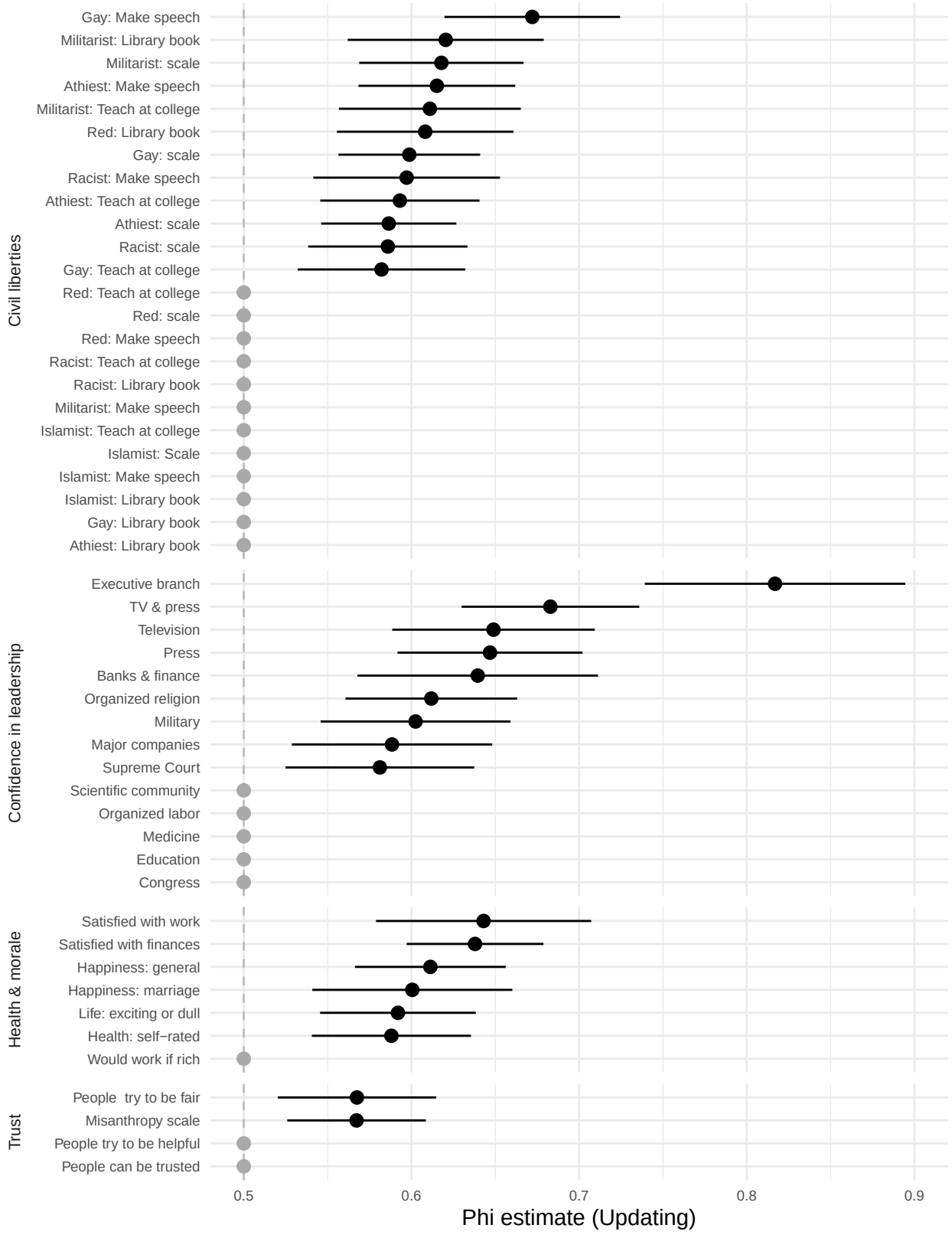

Figure 7: Phi estimates for items about civil liberties, confidence in leadership, health, morale, and social trust.

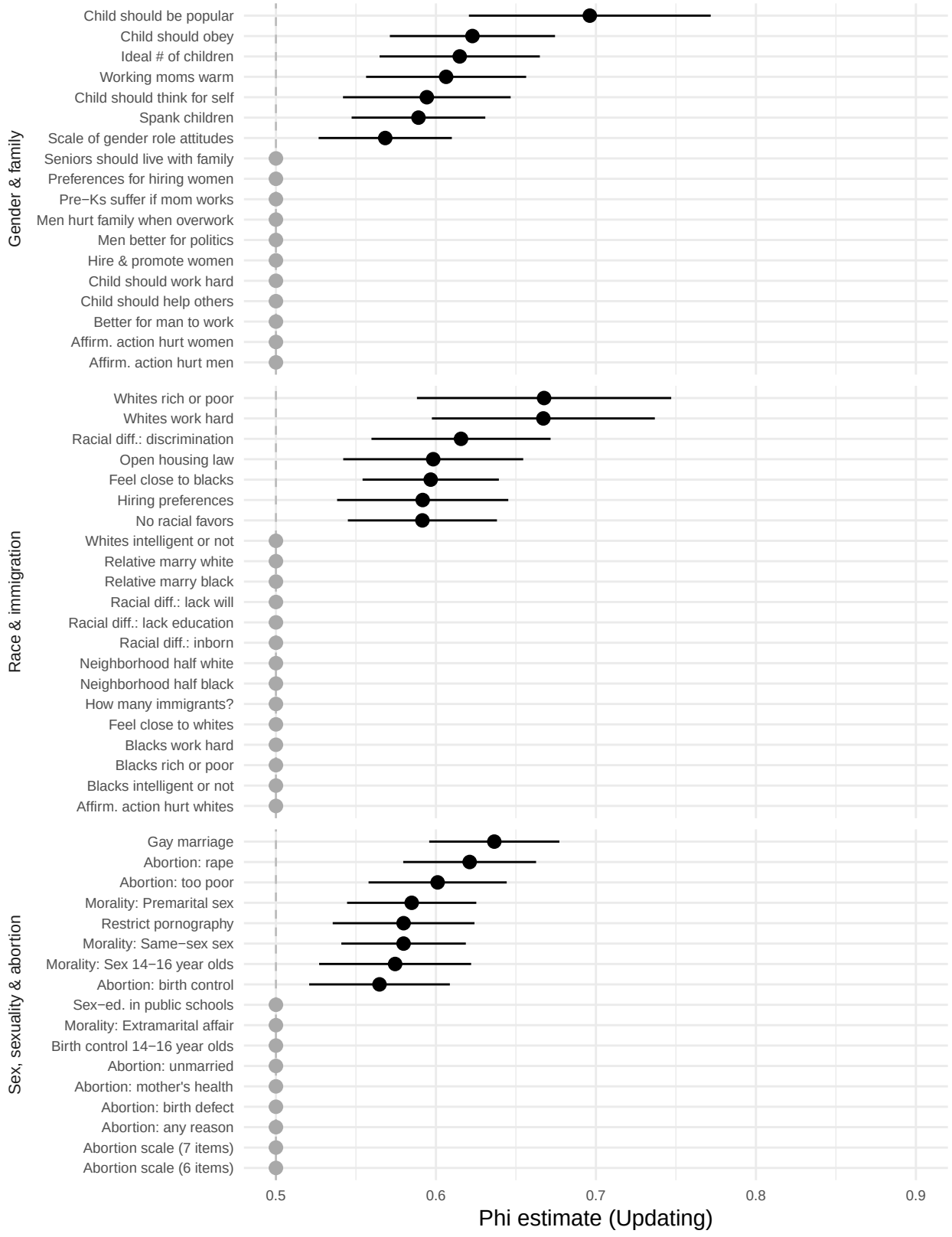

Figure 8: Phi estimates for items about race, gender, sex, sexuality, and abortion.