Appendix

Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments*

Libby Jenke[†‡] Kirk Bansak[†§] Jens Hainmueller[¶] Dominik Hangartner[‖]

March 31, 2020

*Authors' Note: The authors thank Avidit Acharya, Ken Scheve, seminar participants at Stanford University, and three anonymous reviewers for useful comments. The authors also thank Dianna Amasino, Khoi Vo, and Nicolette Sullivan for general advice about eye tracking. Replication materials are available in Jenke et al. (2020). The authors declare that they have no competing interests.

†These authors contributed equally to this work.

‡Assistant Professor, Department of Political Science, University of Houston, 3551 Cullen Boulevard, Room 447, Houston, TX 77204-3011, United States. E-mail: ljenke@uh.edu

§Assistant Professor, Department of Political Science, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States. E-mail: kbansak@ucsd.edu

¶Professor, Department of Political Science, 616 Serra Street Encina Hall West, Room 100, Stanford, CA 94305-6044. E-mail: jhain@stanford.edu

‖Associate Professor, Public Policy Group, ETH Zurich, Leonhardshalde 21, 8092 Zurich, Switzerland, and Department of Government, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, United Kingdom. Email: dominik.hangartner@gess.ethz.ch.

A Details of the Conjoint Survey Design

Table A.1 shows the list of attributes with their corresponding probability weights. While we assigned equal probability weights to most attributes, we adjusted the weights for more common characteristics to increase the ecological validity of the survey experiment.

Table A.1: Attributes and Corresponding Probability Weights

Attribute	Values	Weights
Age	36, 45, 53, 61, 77	(1/6, 1/4, 1/4, 1/6, 1/6)
Gender	Female, Male	(1/2, 1/2)
Race/Ethnicity	White, Hispanic/Latino, Black, Asian American, Native American	(1/2, 1/6, 1/6, 1/12, 1/12)
Previous Occupation	Business executive, College professor, Lawyer, Doctor, Activist	(1/4, 1/6, 1/4, 1/6, 1/6)
Military Service Experience	Did not serve, Served in the Army, Served in the Navy, Served in the Marine Corps, Served in the Air Force	(2/3, 1/12, 1/12, 1/12, 1/12)
Prior Political Experience	Mayor, Governor, U.S. Senator, U.S. Representative, No prior political experience	(1/6, 1/6, 1/4, 1/4, 1/6)
Party	Democrat, Republican, Independent	(1/3, 1/3, 1/3)
Religion	Catholic, Evangelical Protestant, Mainline Protestant Mormon, Jewish	(1/4, 1/4, 1/6, 1/6, 1/6))
Position on Same-Sex Marriage	Strongly support, Support, Strongly oppose, Oppose	(1/4, 1/4, 1/4, 1/4)
Position on Tax Raise for Wealthy	Strongly support, Support, Strongly oppose, Oppose	(1/4, 1/4, 1/4, 1/4)
Position on Gun Control	Strongly support, Support, Strongly oppose, Oppose	(1/4, 1/4, 1/4, 1/4)

Each subject completed a total of 120 decision tasks, consisting of 20 decision tasks for each of the six experimental conditions (five, eight or eleven attributes times two or three profiles). The experimental conditions were presented in random order.

B Subject Pool and Sample Characteristics

Our sample consists of 122 subjects recruited from the Duke Behavioral Research subject pool. The sample includes university students (39%), university employees (26%), as well as other members of the local community (35%). Table A.2 provides descriptive statistics on the socio-demographic composition and political leanings of the sample.

The experiment took place between July 5 and July 31, 2019 at the Fuqua Behavioral Lab of Duke University. The median completion time for the experiment, not including the socio-demographic survey, was 34.6 minutes. Subjects were compensated for completing the experiment with $15. The study was approved by the Institutional Review Boards of Stanford University (IRB Protocol 49988) and Duke University (IRB Protocol 2019-0328).

C Eye-Tracking Technology and Data Preprocessing

To track subjects' eye movements, we used a Tobii T60XL Eye-tracker, which records 60 gaze locations per second. The calibration system consisted of five predefined calibration points between which a dot moved, holding in a location for approximately three seconds. Subjects looked at each

Table A.2: Sample Demographics

Variable	Proportion
Age	
18 – 29	0.58
30 – 39	0.21
40 – 49	0.12
50 – 59	0.06
60 –	0.04
Income	
\$0 – \$49,000	0.46
\$50,000 – \$99,999	0.25
\$100,000 – \$150,000	0.13
\$150,000 –	0.16
Education	
High school degree	0.03
Some college	0.19
Bachelor's degree	0.43
Postgraduate degree	0.34
Gender	
Male	0.37
Female	0.63
Race	
White	0.57
African-American	0.15
Asian	0.18
Hispanic	0.06
Other	0.05
Party Identification	
Strong Democrat	0.39
Not-so-strong Democrat	0.22
Democratic leaner	0.25
Pure Independent	0.05
Republican leaner	0.03
Not-so-strong Republican	0.05
Strong Republican	0.01
Political Ideology	
Extremely liberal	0.23
Liberal	0.34
Slightly liberal	0.18
Slightly conservative	0.18
Conservative	0.04
Extremely conservative	0.03

location of the dot and the system detected their pupil and corneal reflection. If necessary, subjects were re-calibrated until the result was of high quality.

Fixations were detected using the Tobii I-VT Fixation Filter. There are several steps in pre-processing eye-tracking data when using this algorithm. First, the angular velocity of each gaze data point is calculated in visual degrees per second. This is calculated by taking the angular difference between two neighboring data points and dividing them by the time interval between them. To classify fixations, the velocity threshold value used was 30°/second, which is the value recommended by Komogortsev et al. (2010) and Over et al. (2007). The shortest saccades typically peak in velocity at about 30°/second (Holmqvist et al., 2011), meaning that this is an appropriate threshold to use. Those data points that were above 30°/second in velocity were classified as saccades, while those below the threshold value were classified as fixations.

Eye-tracking data is noisy because of participant features and factors inherent in the system. The I-VT Fixation Filter has several ways of dealing with such issues. First, a gap fill-in function linearly interpolates the placement of very short losses in data. Second, single fixations that have been incorrectly split into multiple fixations must be accounted for. Last, very short fixations must be discarded. These post-processing steps are necessary to deal with artefacts and noise events in eye-tracking data (Holmqvist et al., 2011).

The purpose of the gap fill-in function is to prevent a single fixation from being counted as two fixations because of a few missing samples of data over a short period of time, which can occur due to temporary reflections in subjects' glasses that occlude the eyes or issues with the eye-tracker's hardware such as temporary malfunctions or delays in data transfers within the hardware systems. These forms of data loss must be distinguished from data loss caused by legitimate separations of fixations such as the subject blinking or looking away from the screen. Since the minimum duration for a blink is 75 ms (Komogortsev et al., 2010), 75 ms was set as the maximum gap length to be filled in. Data was then filled in through linear interpolation, along straight lines between the valid data points on each side of the gap.

Even with a gap fill-in function, some fixations may still be incorrectly split into multiple fixations and must be merged. This can occur due to post-saccadic oscillations, square-wave jerks, and other involuntary saccadic intrusions. These will look like short, high velocity movements between fixations that are located close in time and space to one another. If the time between the end of one fixation and the beginning of the next was shorter than 75 ms (the normal blink duration), then the amplitude of the intervening saccade was determined. If its amplitude was less than 0.5°, then the fixations were merged. 0.5° is the typical amplitude of micro-saccades (Komogortsev et al., 2010), which are involuntary eye movements.

Last, fixations that are very short must be discarded because the brain requires time to process visual input, meaning that the visual input from very short fixations will not be cognitively registered. We used a minimum fixation duration of 60 ms because our screens involved reading rather than scene-viewing, and fixations of this duration can be processed when reading (Over et al., 2007).

D Additional Results

Figure A.1 shows the proportion of decision tasks in which an attribute was the primary focus, split by experimental condition.

Figure A.2 shows the mean proportion of fixations per attribute, pooled across all experimental conditions.

Figure A.3 shows the mean proportion of fixations per attribute, split by experimental condition.

Figure A.4 shows the distribution, by experimental condition, of Spearman correlations between subjects' attribute rankings as computed from marginal R^2 and rankings based on eye fixations.

Figure A.5 shows the AMCEs by experimental condition where we limit the sample to the tasks where respondents are shown all three policy positions for the candidates.

Figure A.6 shows the differences between the AMCEs estimated from each of the experimental conditions (blocks), as displayed in Figure 5. All pairwise comparisons are displayed.

Figure A.7 shows the differences between the AMCEs estimated from early decision tasks (1 - 10) vs. late decision tasks (11 - 20) within blocks, as displayed in Figure 8a.

Figure A.8 shows the AMCEs estimated from the early decision tasks (1 - 10) and late decision tasks (11 - 20) within the first block completed by each subject.

Figure A.9 shows the AMCEs estimated from four subsets of decision tasks across the subjects' full sequence of 120 decision tasks. The subsets are: decision task 1 on its own, tasks $2-5$, tasks $6-20$, and tasks $21-120$.

Figures A.10 and A.11 show replications of Figure 4 in Bansak et al. (2019) while subsetting their sample to Democratic and Republican respondents, respectively. Bansak et al. (2019) is based on a paired candidate conjoint that included four core attributes of interest and a randomly assigned number of additional filler attributes that were uncorrelated with the core attributes. The plots show the AMCEs of core attributes of interest as the number of filler attributes increases from zero to 35 for a sample from three survey waves administered on Amazon Mechanical Turk (see Bansak et al. (2019) for details). In our replication, we see that the stability of the AMCEs as the number of attributes increases is similar for Democrats and Republicans, which suggests that party identification does not moderate how respondents react to increased complexity in the conjoint table.

Figures A.12 and A.13 show replications of Figure 6 in Bansak et al. (2019) while subsetting their sample to Democratic and Republican respondents, respectively. The plots show the AMCEs of core attributes of interest as the number of filler attributes increases from zero to 35 for a sample of an online survey administered through Survey Sampling International (see Bansak et al. (2019)

for details). Again, party identification does not moderate how respondents react to increased complexity in the conjoint table.

Figure A.1: Mean Proportion of Fixations per Attribute, Split by Experimental Condition

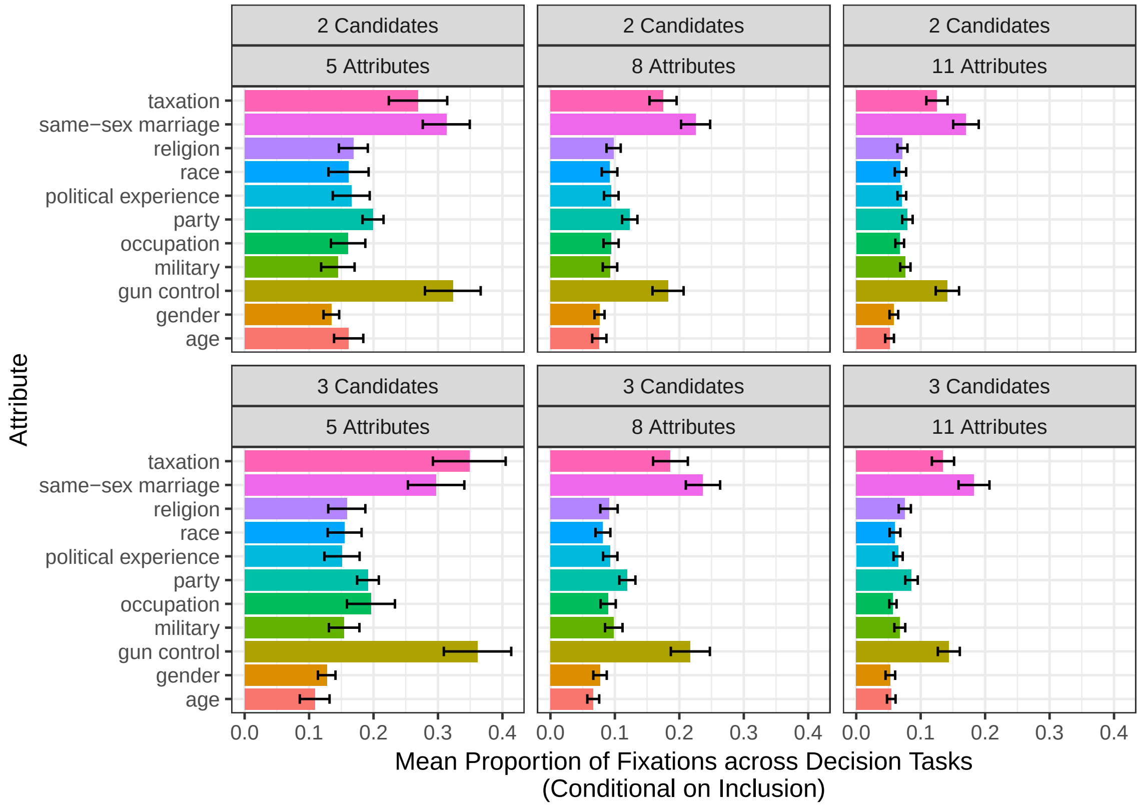


Figure A.2: Attribute Of Primary Focus (Pooled Data)

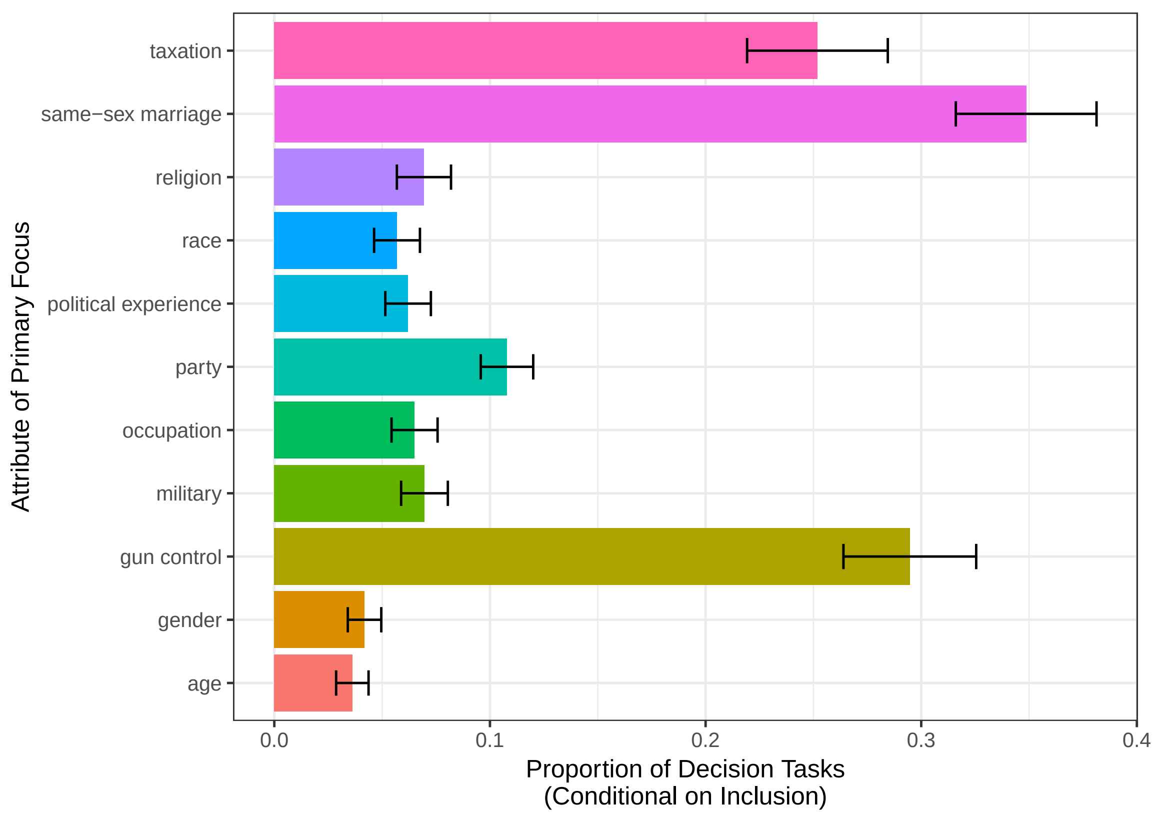


Figure A.3: Attribute of Primary Focus, Split by Experimental Condition

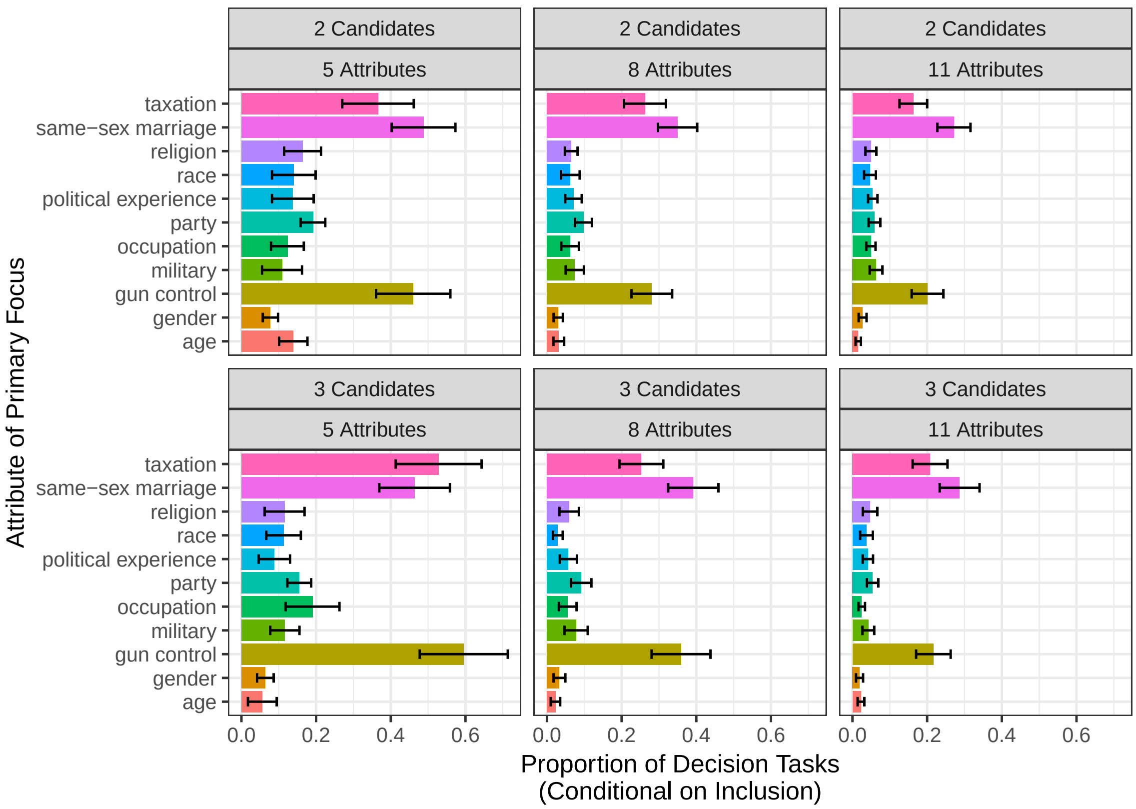


Figure A.4: Correlation between Attribute Importance in Choice and Eye-Tracking Data (by Experimental Condition)

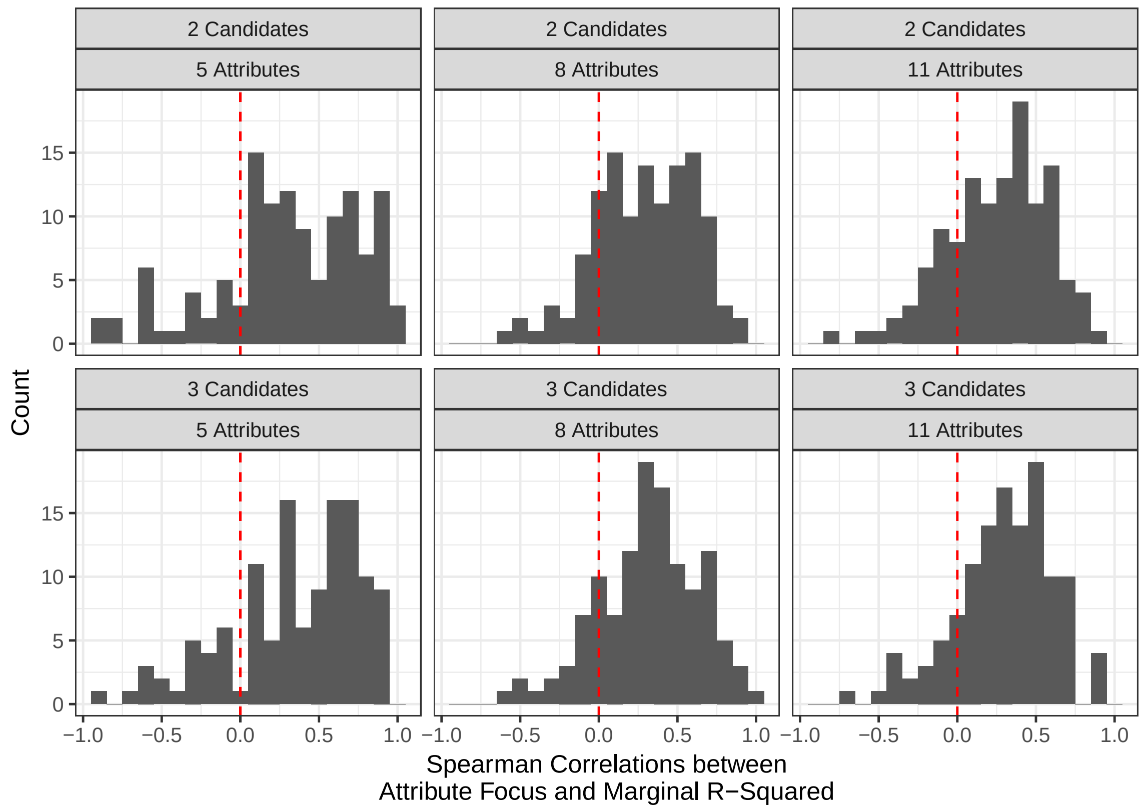


Figure A.5: AMCEs by Experimental Condition (Subset of All Policy Positions Shown)

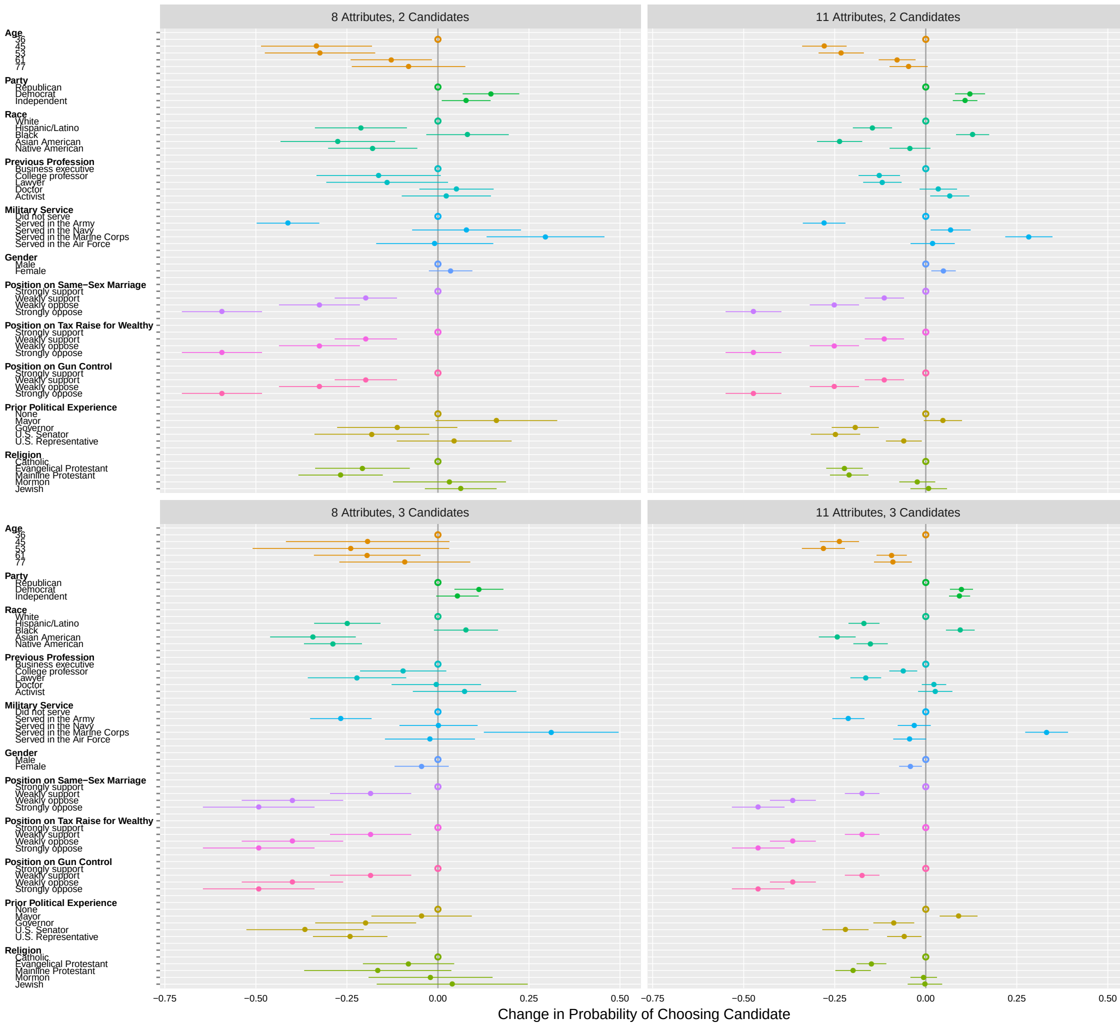


Figure A.6: Differences between AMCEs Estimated Across Experimental Conditions, All Pairwise Comparisons

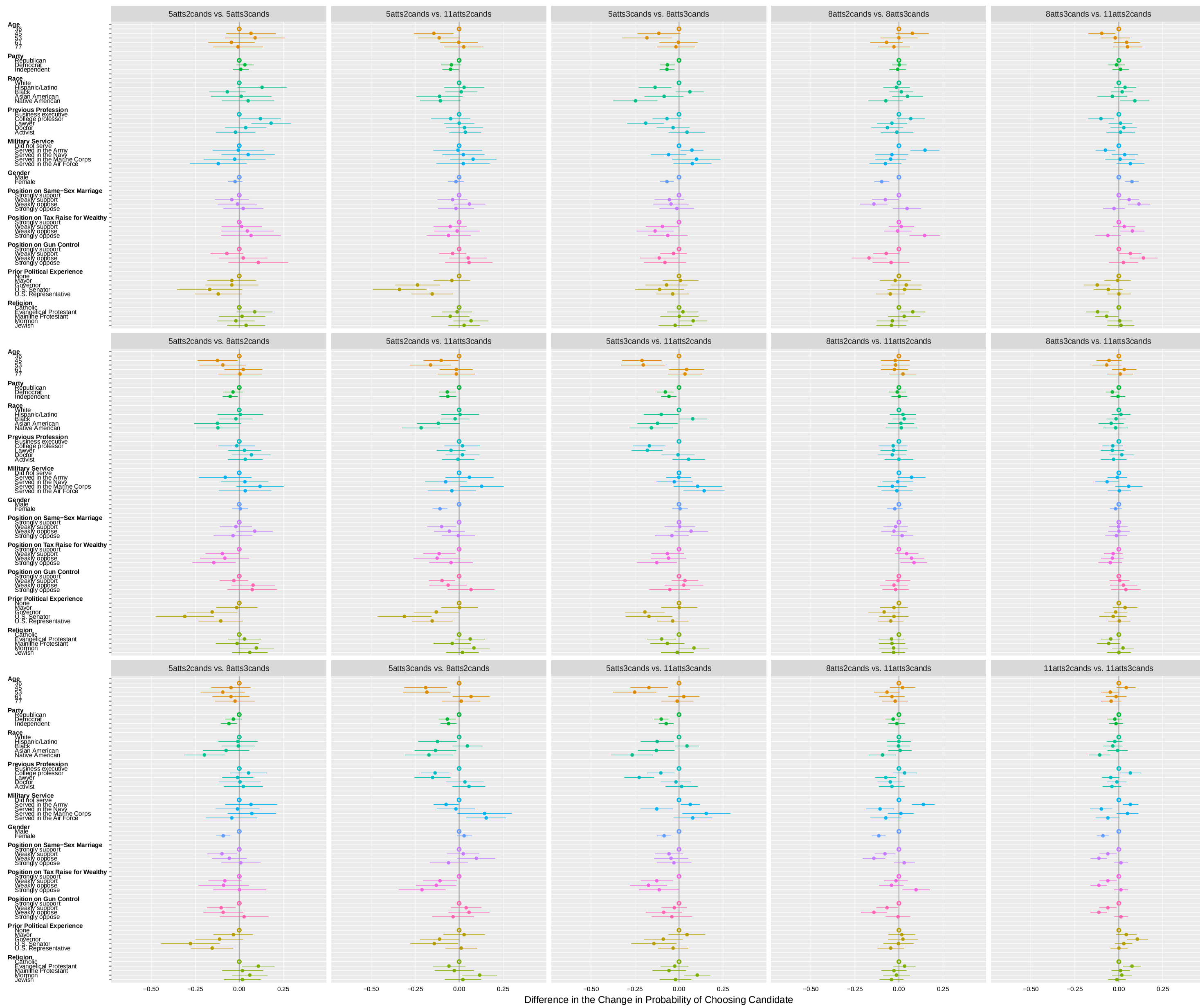


Figure A.7: Differences between AMCEs Estimated from Early Tasks (1 - 10) vs. Late Tasks (11 - 20) Within-Block, Pooled Across Blocks

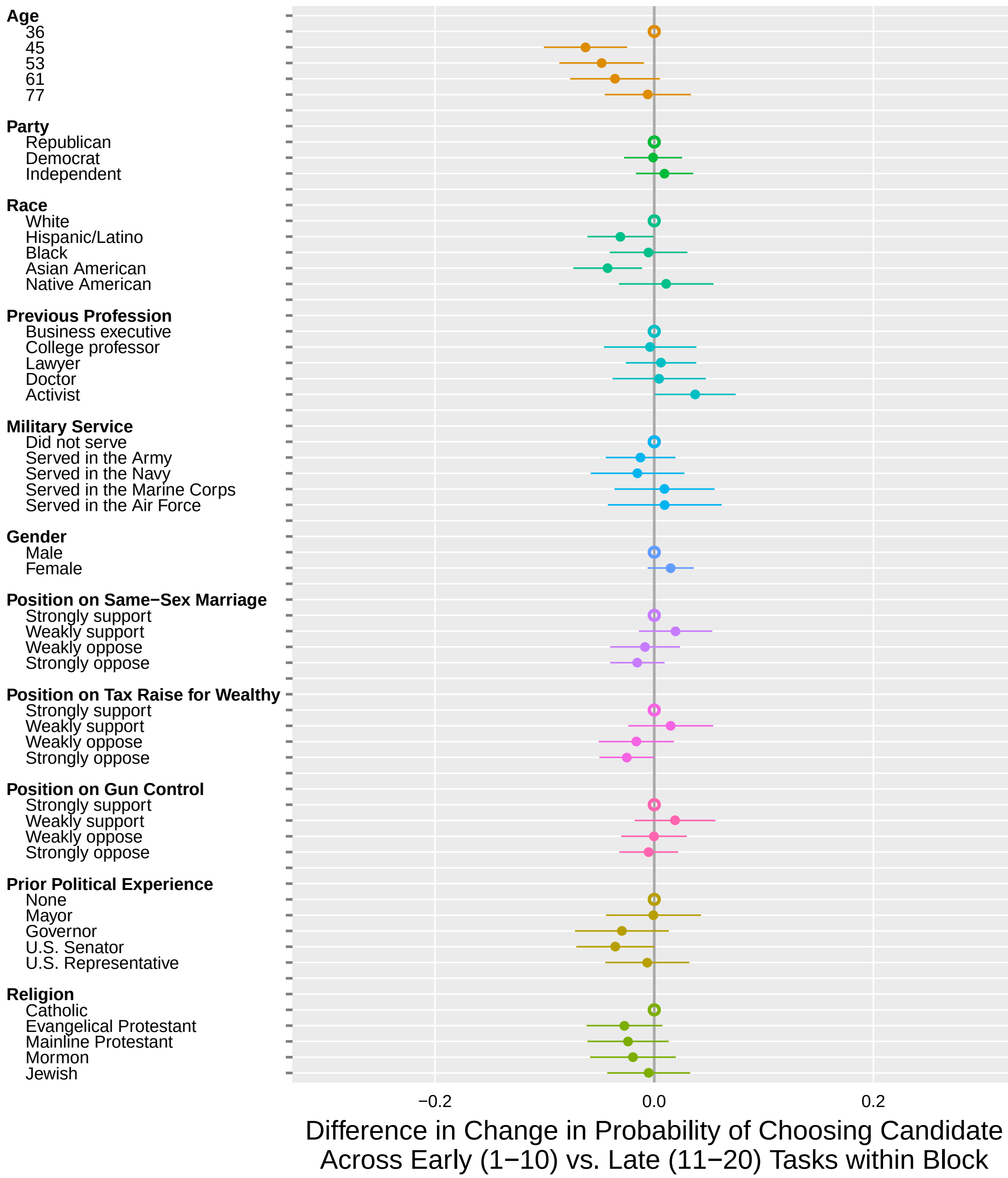


Figure A.8: AMCEs Across Tasks within Each Subject's First Block

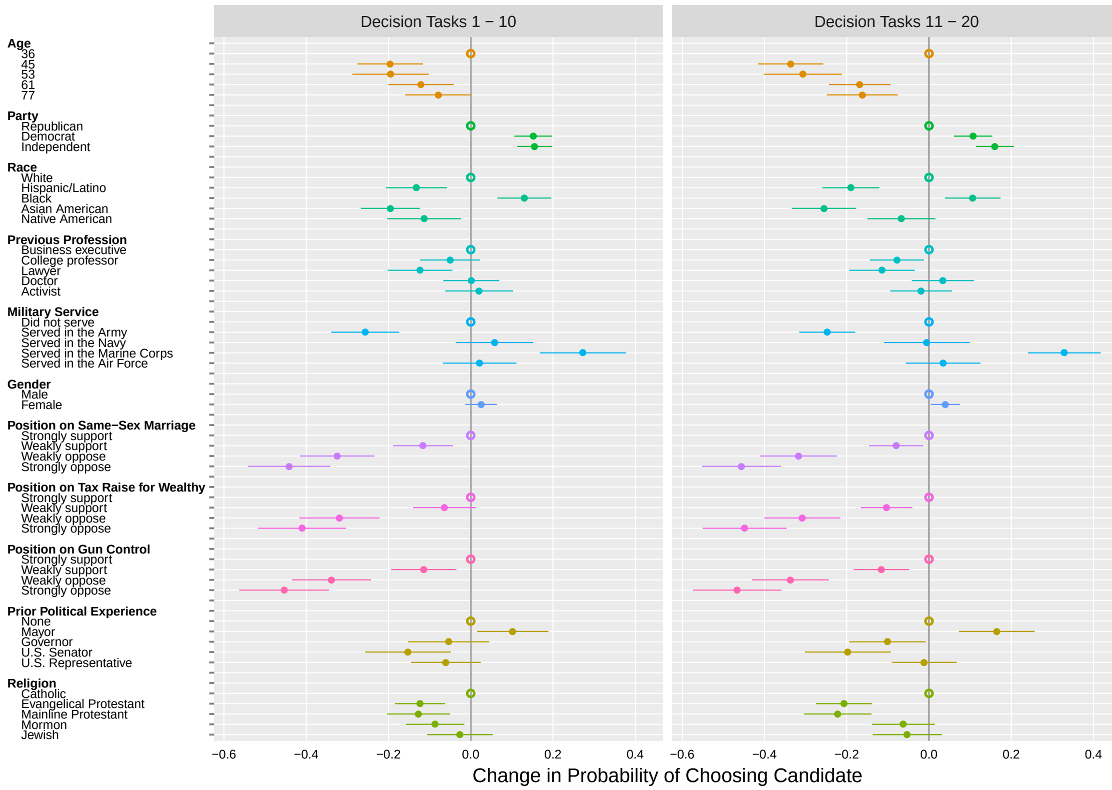


Figure A.9: AMCEs Across Sequence of Decision Tasks

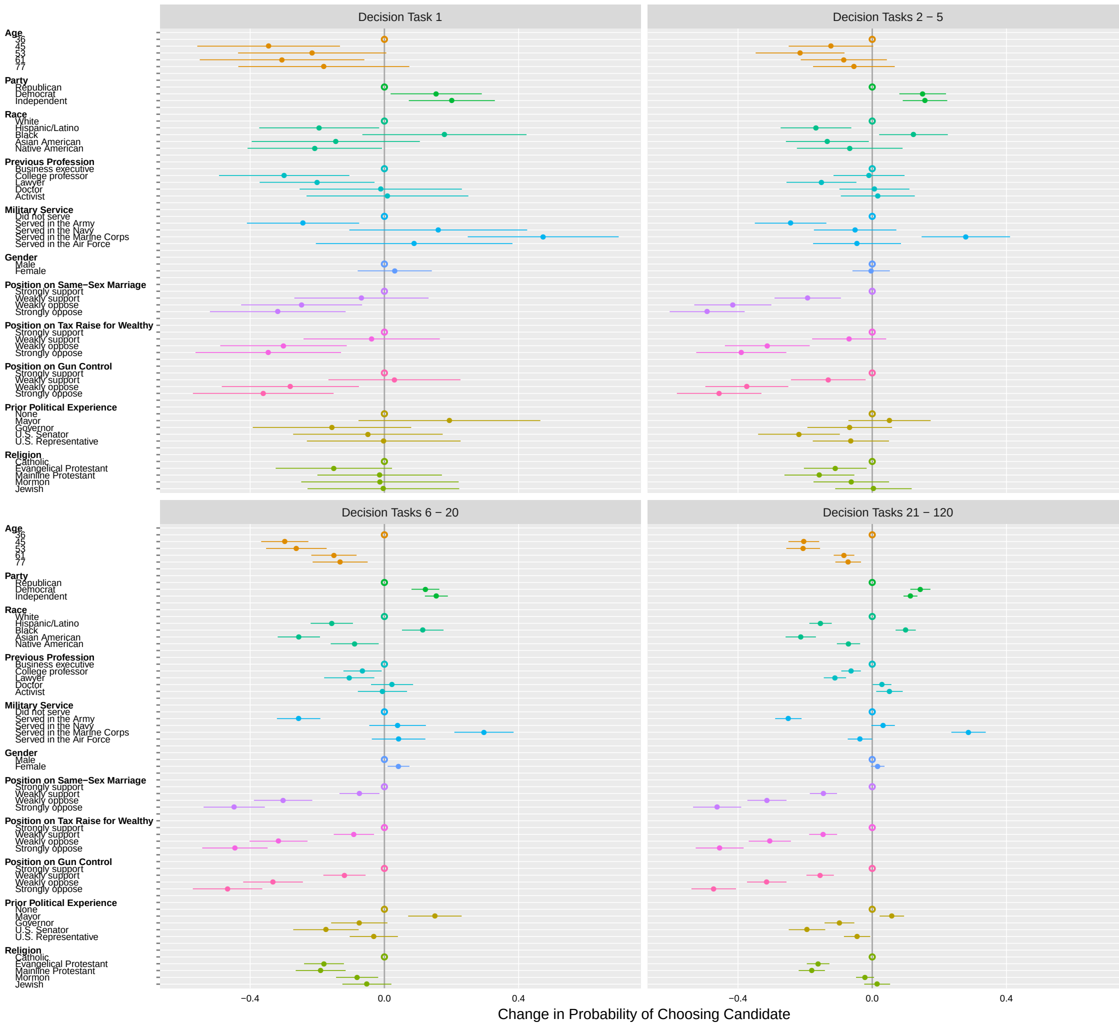


Figure A.10: Replication of Bansak et al. (2019) Figure 4 for Democratic Respondents: AMCEs of core attributes of interest from three Amazon Mechanical Turk survey waves as the number of filler attributes increases.

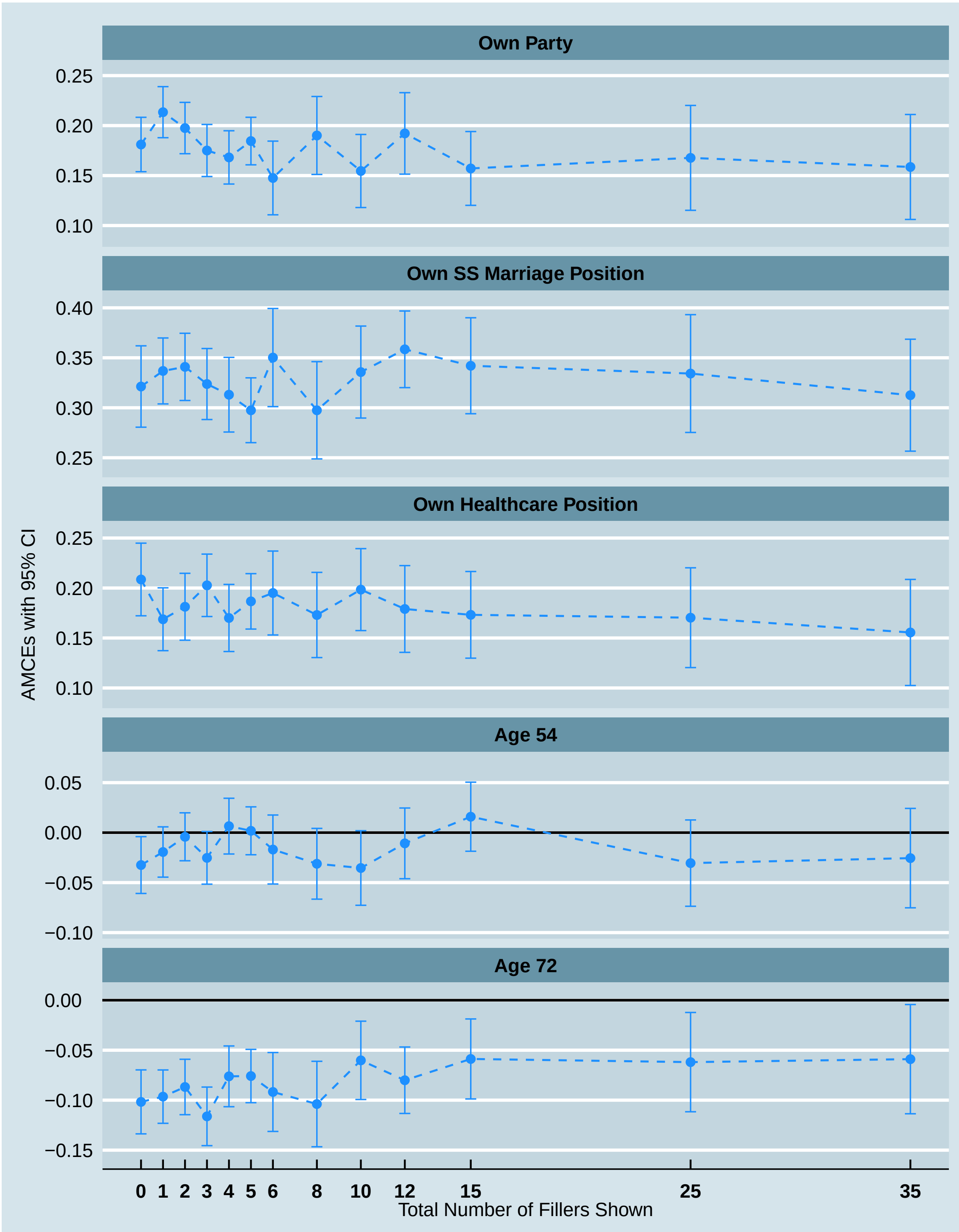


Figure A.11: Replication of Bansak et al. (2019) Figure 4 for Republican Respondents: AMCEs of core attributes of interest from three Amazon Mechanical Turk survey waves as the number of filler attributes increases.

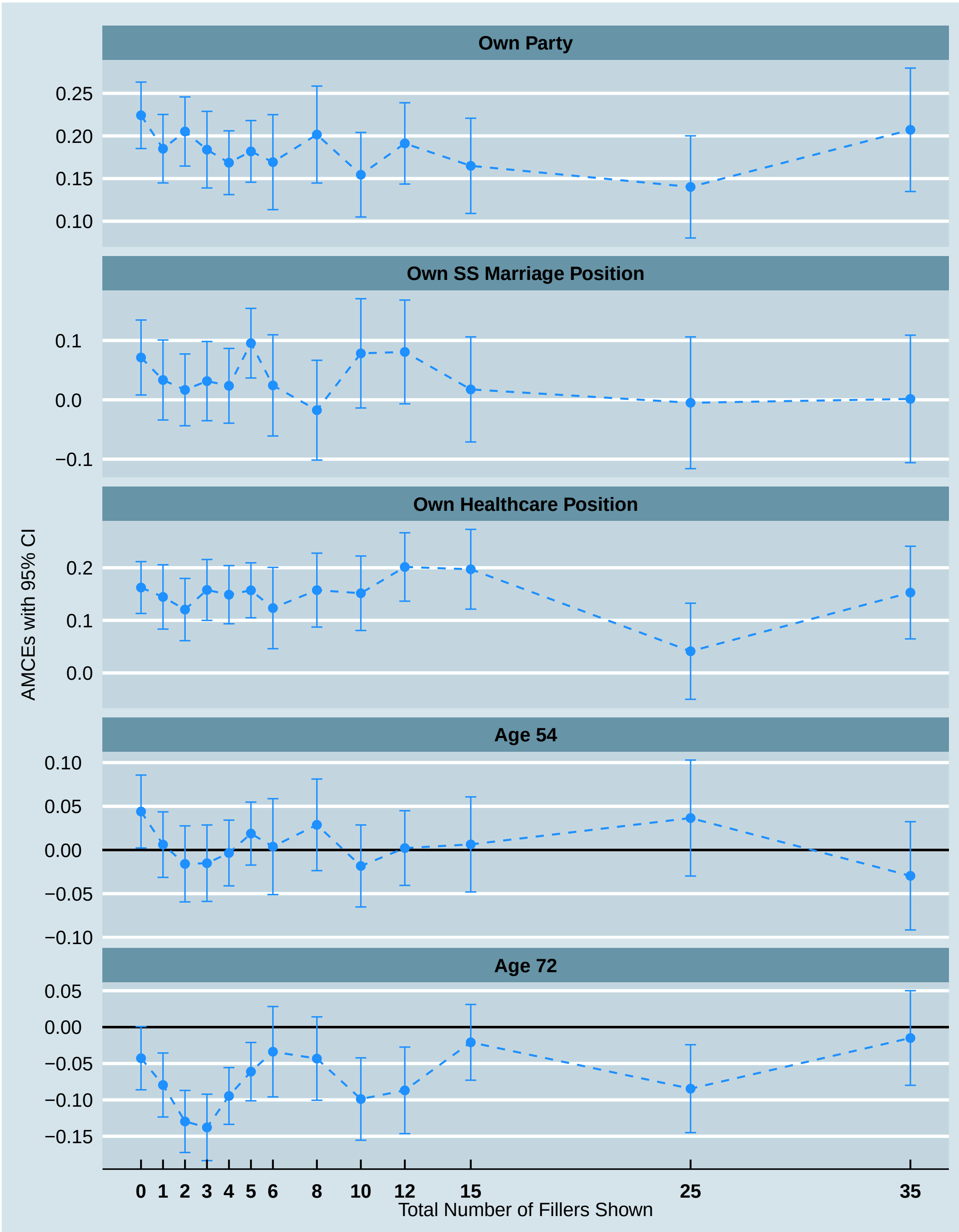


Figure A.12: Replication of Bansak et al. (2019) Figure 6 for Democratic Respondents: AMCEs of core attributes of interest from a Survey Sampling International survey as the number of filler attributes increases.

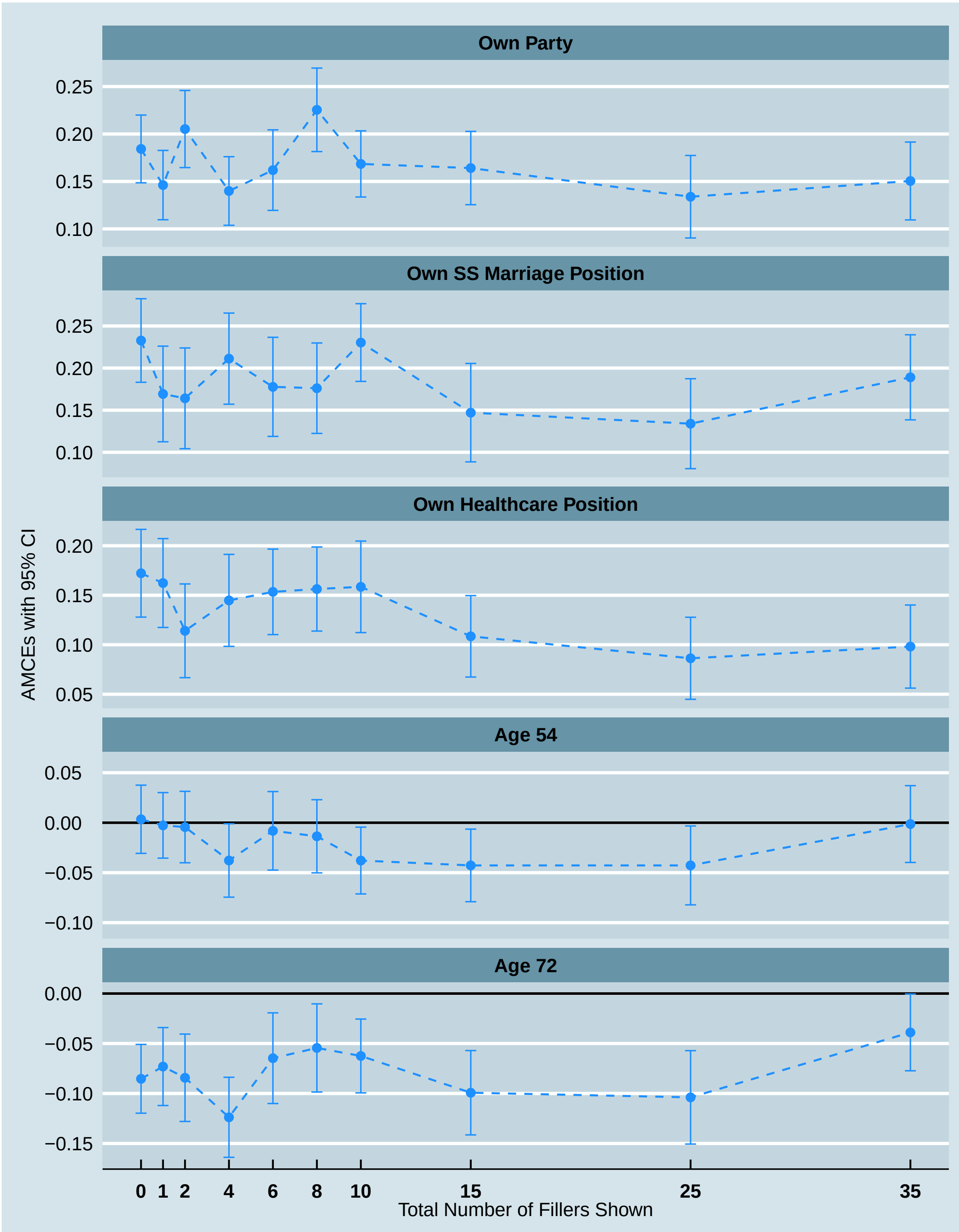
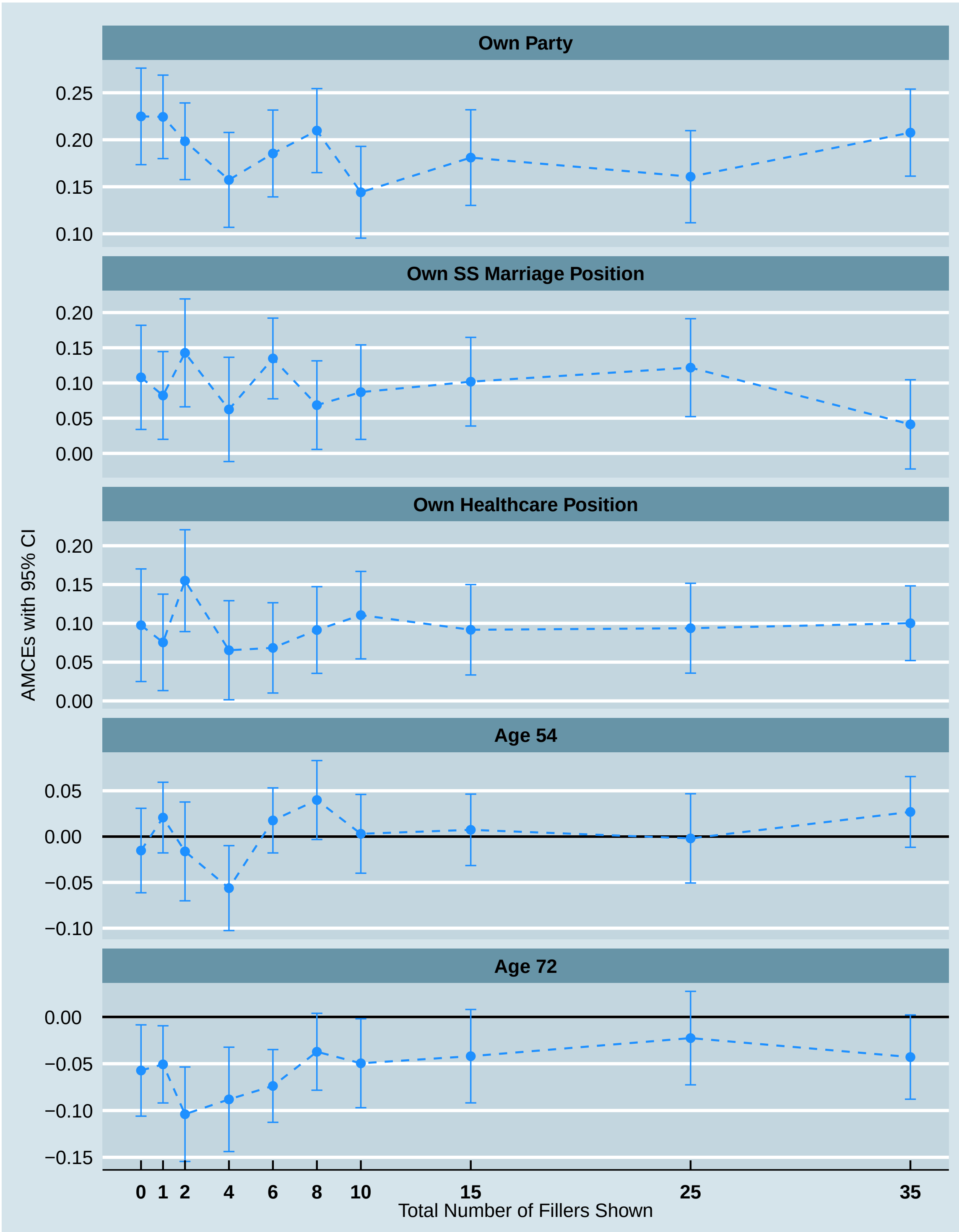


Figure A.13: Replication of Bansak et al. (2019) Figure 6 for Republican Respondents: AMCEs of core attributes of interest from a Survey Sampling International survey as the number of filler attributes increases.



References

Bansak, K., J. Hainmueller, D. J. Hopkins, and T. Yamamoto (2019). Beyond the breaking point? Survey satisficing in conjoint experiments. *Political Science Research and Methods*, 1–19.

Holmqvist, K., M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. OUP Oxford.

Jenke, L., K. Bansak, J. Hainmueller, and D. Hangartner (2020). Replication materials for: Using eye-tracking to understand decision-making in conjoint experiments. `https://doi.org/10.7910/DVN/TISLDL`, Harvard Dataverse.

Komogortsev, O. V., D. V. Gobert, S. Jayarathna, S. M. Gowda, et al. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering 57*(11), 2635–2645.

Over, E., I. Hooge, B. Vlaskamp, and C. Erkelens (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research 47*(17), 2272–2280.