ONLINE APPENDIX: Introducing the Visual Conjoint, with an Application to Candidate Evaluation on Social Media

Alessandro Vecchiato   Kevin Munger

August 22, 2024

A   Questionnaire

Q1.2 For our research, careful attention to survey questions is critical! To show that you are paying attention please select "I have a question."

○ I understand (1)

○ I do not understand (2)

○ I have a question (3)

- -

Page Break ———————————————————————————————

Q1.3 We want to get your opinion on a number of political topics of interest to the US population today. There are not any right or wrong answers to these questions. But to ensure that you are reading each question carefully, we need you to select the sixth option out of the following eight options.

○ Taxes (1)

○ Economy (2)

○ Health (3)

○ Immigration (4)

○ Foreign Affairs (5)

○ Climate change (6)

○ Religion (7)

○ Hard Work (8)

Q2.1 How old are you?

_____

Q2.2 What is your sex?

○ Male (1)

○ Female (2)

Q2.3 What is your ethnicity?

○ White (1)

○ Hispanic or Latino (2)

○ Black or African American (3)

○ Native American or American Indian (4)

○ Asian/Pacific Islander (5)

○ Other (6)

Q2.4 Generally speaking, do you usually think of yourself as a Democrat, a Republican, an independent?

○ Strong Democrat (1)

○ Democrat (2)

○ Not very strong Democrat (3)

○ Independent (4)

○ Not very strong Republican (5)

○ Republican (6)

○ Strong Republican (7)

Q2.5 How often do you use the Internet?

○ Pretty much all the time (1)

○ Several times a day (2)

○ About once a day (3)

○ 3 to 6 days a week (4)

○ 1 to 2 days a week (5)

○ Every few weeks (6)

○ Less often (7)

Q2.6 How often do you use Facebook?

○ Pretty much all the time (1)

○ Several times a day (2)

○ About once a day (3)

○ 3 to 6 days a week (4)

○ 1 to 2 days a week (5)

○ Every few weeks (6)

○ Less often (7)

○ I don't use Facebook (8)

Page Break

Q2.7 How familiar are you with the following computer and Internet-related items? Please choose a number between 1 and 5 where 1 represents "no understanding" and 5 represents "full understanding" of the item.

	No understanding (1)	2 (2)	3 (3)	4 (4)	Full understanding (5)
Advanced search (1)	○	○	○	○	○
Favorites (2)	○	○	○	○	○
Preference setting (3)	○	○	○	○	○
PDF (4)	○	○	○	○	○
Cache (5)	○	○	○	○	○
Wiki (6)	○	○	○	○	○
JPG (7)	○	○	○	○	○
Blog (8)	○	○	○	○	○
Malware (9)	○	○	○	○	○
Phishing (10)	○	○	○	○	○
Hashtag (11)	○	○	○	○	○
RSS (12)	○	○	○	○	○
Selfie (13)	○	○	○	○	○
Tagging (14)	○	○	○	○	○

Q2.8 Please indicate your agreement with the following statements on a scale of -4 = Strongly Disagree to 4 = Strongly Agree

	-4 (Strongly Disagree) (1)	-3 (2)	-2 (3)	-1 (4)	0 (5)	1 (6)	2 (7)	3 (8)	4 (Strongly Agree) (9)
I prefer to ask friends how to use any new technological gadget instead of trying to figure it out myself (1)	○	○	○	○	○	○	○	○	○
I feel like information technology is a part of my daily life (2)	○	○	○	○	○	○	○	○	○
Using information technology makes it easier to do my work (3)	○	○	○	○	○	○	○	○	○
I often have trouble finding things that I've saved on my computer (4)	○	○	○	○	○	○	○	○	○

I have problems with viruses and malware on my computer. (5)	○	○	○	○	○	○	○	○	○
I rely on family members to introduce me to new technology. (6)	○	○	○	○	○	○	○	○	○
I have professionals (such as the Geek Squad) or family members take a look at my computer when something isn't working. (7)	○	○	○	○	○	○	○	○	○

Q2.9 How many words are in this sentence?

_____

Q2.10 How often do you read news stories online?

○ Pretty much all the time (1)

○ Several times a day (2)

○ About once a day (3)

○ 3 to 6 days a week (4)

○ 1 to 2 days a week (5)

○ Every few weeks (6)

○ Less often (7)

Q2.11 How often do you read news stories offline (in the newspaper, printed news magazines)?

○ Pretty much all the time (1)

○ Several times a day (2)

○ About once a day (3)

○ 3 to 6 days a week (4)

○ 1 to 2 days a week (5)

○ Every few weeks (6)

○ Less often (7)

End of Block: Descriptives

Start of Block: Intro To Experiment

Q66 In this section, we will show you a series of Twitter profiles of hypothetical Members of the US Congress. These profiles (and the people they represent) are not real, but are constructed from composites of real people and profiles.

For each pair of profiles, please take your time and read each of the profiles carefully, then tell us which politician you'd prefer to be your representative in Congress.

Display This Question:

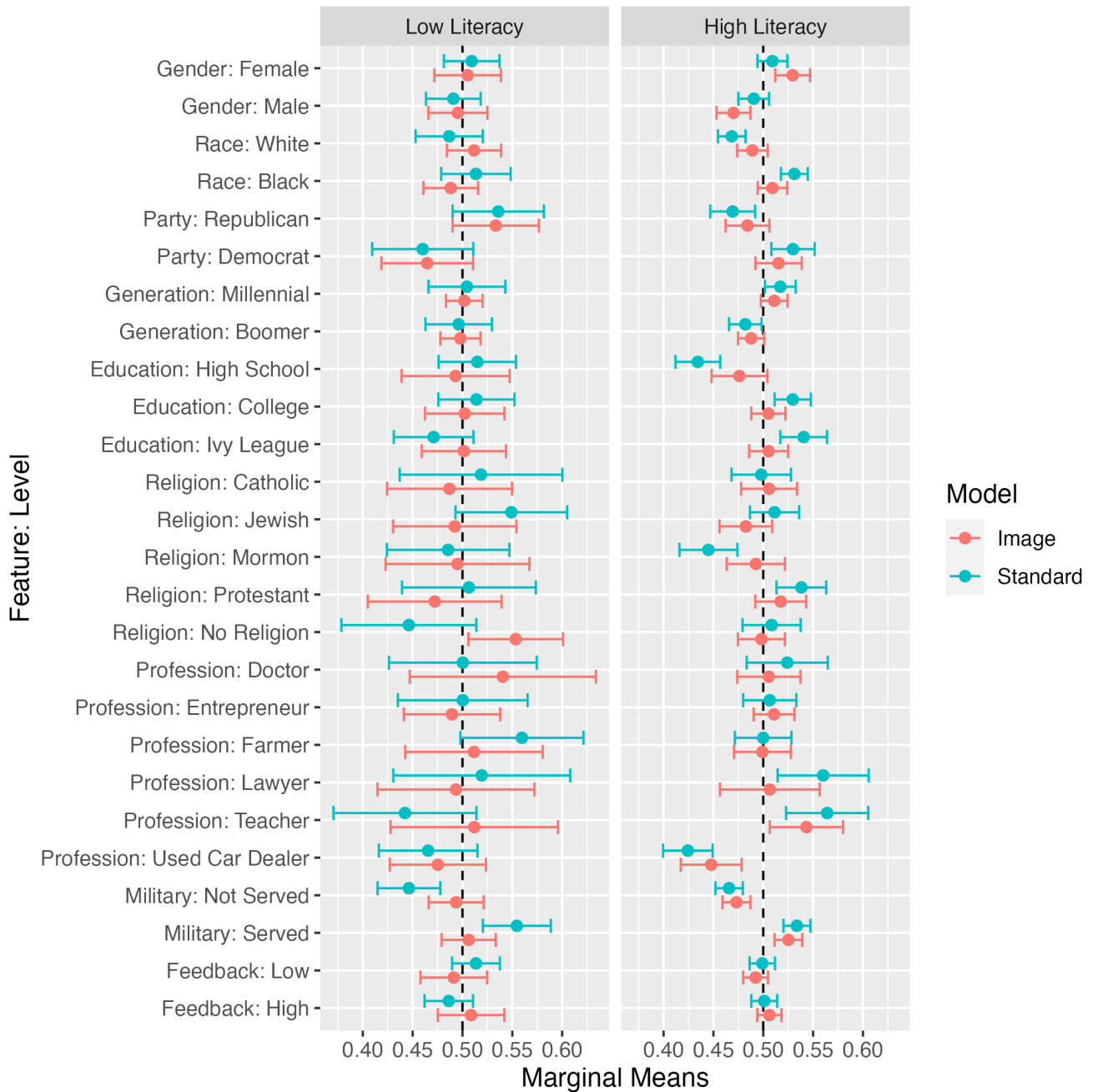   *If GeoIP Location Region = AL*

Q68 Which of these politicians would you rather have as your Representative in Congress?

○ Image:Bl woman boomer high national (31)

○ Image:Bl woman boomer high local (32)

B Digital Literacy Results

Figure 1: Preferences for Twitter profiles, by respondent Internet Skills Scale

The figure reports combined plots for estimates of the change in probability of selecting the candidate, based on the reported characteristics. Estimates are to be evaluated as averages over all the possible combinations of the other components.

Figure 2: Preferences for Twitter profiles, by Respondent Power User Scale
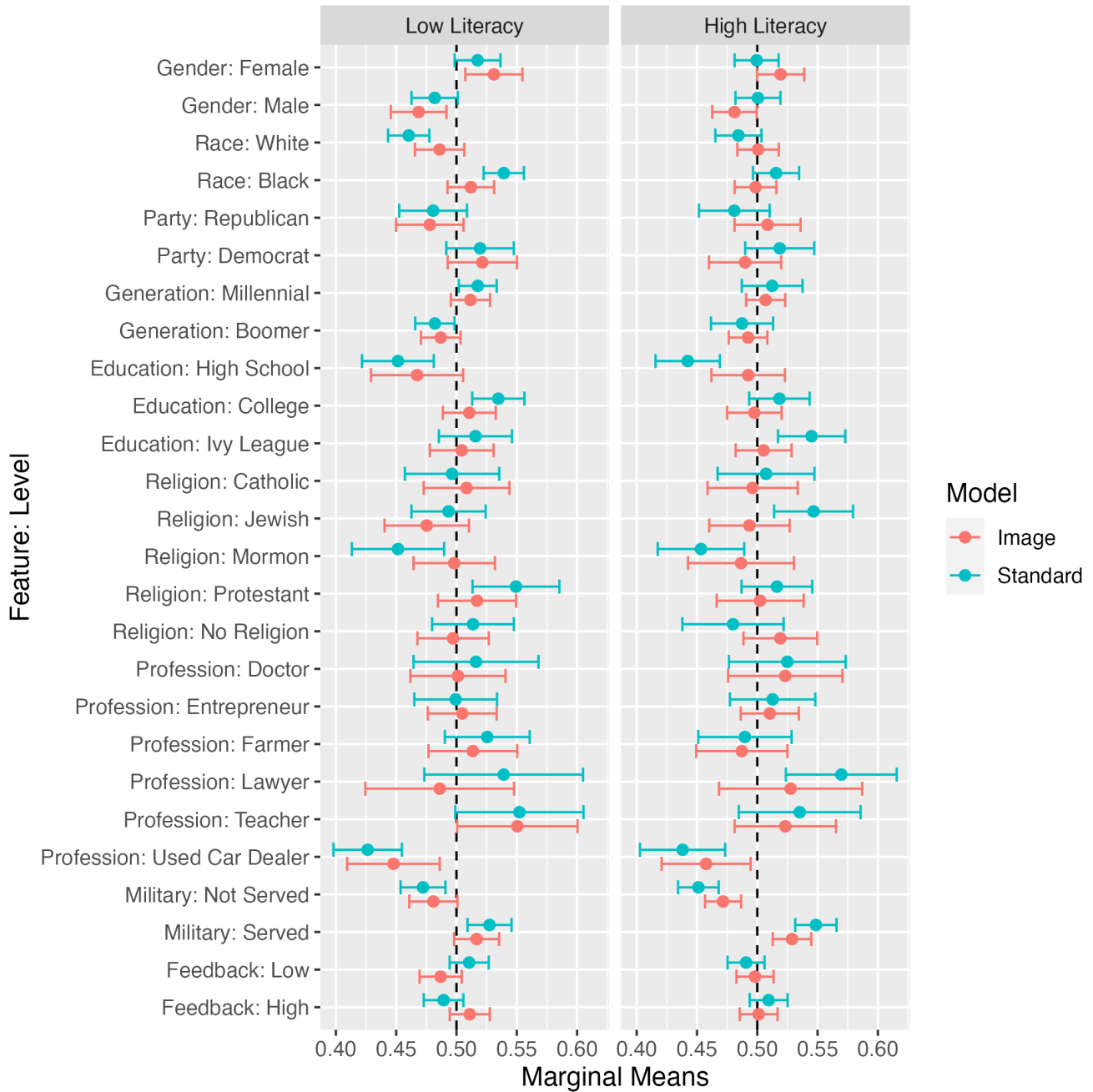
The figure reports combined plots for estimates of the change in probability of selecting the candidate, based on the reported characteristics. Estimates are to be evaluated as averages over all the possible combinations of the other components.

C   Distribution of Profiles

Figure 3: Frequency Distribution of Features across Profiles
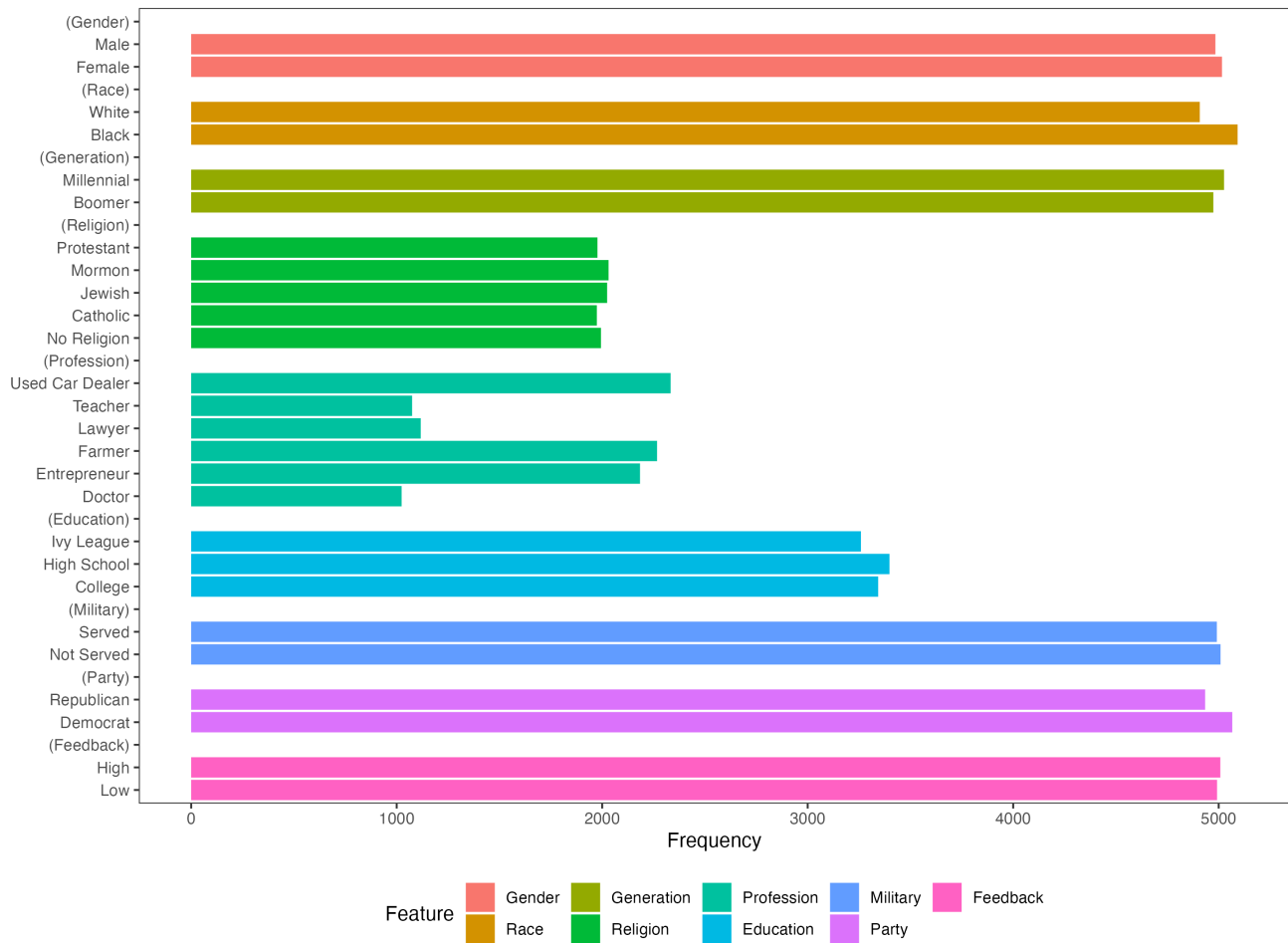
The figure reports the overall frequency distribution for each feature-level pair. Notice how since some professions require high levels of education, and therefore could not be uniformly paired, those professions have lower frequency in the overall sample.

D   Creating the Visual Conjoint

We generate the random Twitter profiles with the Python Image Library. The script operates in two phases. It first generates generic and empty Twitter profiles that resemble actual ones in every aspect including color scheme and font selection. Such profiles report all the same name and handle, Congressman Smith (Congresswoman for women profiles), they are all localized in Washington, DC, and they all report the same generic website, *congress.gov*.

The script then fills the generic profiles through a nested loop over the feature sets, such as gender and education. Each profile combination picks one element of each character vector to generate the final profile. The script operates differently for features signaled through images and those reported through text. For the profile picture, official photographs of some of the less-recognizable members of Congress, signal both gender and race. The script reshapes and crops previously generated images and combines them on the empty profile. For the textual information, the script combines feature levels as character vectors into a readable Twitter bio.

To avoid incompatibility between image and biographical details, years of birth in the older generation are set at 1955, while those in the younger generation are from 1971. Images, of course, are by nature high-dimensional, and there are potentially relevant aspects of these images that might not be perfectly balanced. For example, **?** finds that the attractiveness of politicians can play a role in candidate preference. We address this in two ways. First, by using multiple photographs for each demographic profile, these imbalances are at least partially balanced out. And second, we argue that the use of these images — which are literally identical to the images the citizen would encounter in the target context — our design is higher in external validity.

To avoid specific-level treatment effects, numeric information for social media feedback is randomly drawn from specific values corresponding to the level (high or low). Social media feedback is divided in two categories, low and high, encompassing all possible feedback elements such as followers and like, and corresponding to markedly different levels of social media engagement. For instance, low social media engagement corresponds, among others, to a following between 2 and 5 thousand people, while such figures for high engagement are between 10 and 13 thousands. Similarly, Low engagement corresponds to a range between 100 and 800 likes for each post, while likes

on high engagement correspond to 3000 or 6000 likes.

While the script is specific to the generation of Twitter profiles, it can be readily adapted to the study of other features of Twitter profiles by modifying existing feature sets, or adding new ones. Profiles are then included in a YouGov questionnaire that offers them in pairs from the pool of all profiles generated.

E   Survey Details

The frame was constructed by stratified sampling from the full 2019 American Community Survey (ACS) 1-year sample with selection within strata by weighted sampling with replacements (using the person weights on the public use file). The matched cases were weighted to the sampling frame using propensity scores. The matched cases and the frame were combined and a logistic regression was estimated for inclusion in the frame. The propensity score function included age, gender, race/ethnicity, years of education, and region. The propensity scores were grouped into deciles of the estimated propensity score in the frame and post-stratified according to these deciles. The weights were then post-stratified on 2016 and 2020 Presidential vote choice, and a four-way stratification of gender, age (4-categories), race (4-categories), and education (4-categories), to produce the final weight.

In addition to YouGov's extensive quality checks, we include three further attention checks, two after the consent form and another one in the demographic characteristics section. After the consent form, respondents were asked to answer two questions structured to instruct respondents on the right answer. The first question instructed the respondents to choose the answer "I have a question," the third option after "I understand" and "I do not understand". Mimicking the format of many initial questionnaire questions, such structure was specifically chosen to screen out respondents that would try to move forward quickly by clicking "I understand." The second attention check replicated the structure of the first one, but replacing the three option with nine options describing common public policy issues, such as taxes and climate change. In this case, respondents were asked to select the seventh option ("Climate Change") among the nine options proposed.

These attention checks have become common when evaluating online samples, for good reason: ? find that less than 40% of Lucid respondents answered similar attention checks correctly in June 2021. In contrast, 97% of our YouGov respondents answered both questions correctly.

F   Ethics Discussion

This research does not use deception and there are no potential harms above everyday experience. Subjects were recruited by YouGov and compensated accordingly. Before beginning the experiment, they were given a consent form and the option to opt out of research.