Language and Voting Behavior among Asian Americans: The Case of the 2024 US Presidential Elections

Amy H. Liu (Professor @ University of Texas) – *corresponding author:* amy.liu@austin.utexas.edu
Jangai Jap (Assistant Professor @ University of Georgia)
Keith Padraic Chew (Presidential Postdoctoral Fellow @ Arizona State University)
Meiying Xu (PhD Student @ University of Texas)
Seung Wook Ethan Yoo (PhD Candidate @ University of Texas)
Tamana Attai (Senior Compliance Associate @ CAF America)
Medea Badashvili (Associate Professor @ Tbilisi State University)
Olgahan Cat (PhD Candidate @ University of Texas)
Hiba Faruqi (Undergraduate Student @ University of Texas)
Nadya Hajj (Associate Professor @ Wellesley College)
Nivedita Jhunjhunwala (PhD Candidate @ University of Texas)
Maryanne Mendoza-Davé (Assistant Professor @ California State University at Pomona)
Ojashwi Pathak (PhD Candidate @ University of Maryland)
Md. Muhib Rahman (Postdoctoral Associate @ Cornell University)
Nhu Truong (Assistant Professor @ University of Wisconsin)
Delgerjargal Uvsh (Assistant Professor @ University of Texas)

Abstract: Survey data from before the 2024 election about Asian American voting behavior did not match exit poll numbers. What explains this discrepancy? In this note we argue there are methodological limitations – specifically with the data collection process. Existing surveys systematically under-sample *English-limited* Asian Americans – a population that is almost 44% of the community. Consider how less than 2% of the Asian respondents in the CMPS took the survey in an Asian language. Since linguistic proficiency is not randomly distributed, what this means is that research on Asian Americans is largely biased towards liberals and Democrats. We fielded an original survey that draws on extensive, targeted recruiting of the Asian American community (*N=4956*) – of which 28% of the respondents took the survey in a non-English language. Our results not only match exit polls but also highlight – on average – a 14% gap in Democrat support between English and non-English language speakers in the Asian American community. This note offers a workflow for future researchers to think about how to sample non-English speaking communities in the US with minimal financial barriers.

December 19, 2024
Word Count: 2715

In September 2024, survey data suggested 66% of Asian American voters planned to vote for Harris (AAPI Data 2024). Yet, exit polls showed a much lower figure: 54% (Zarsadiaz 2024). What explains this two-digit discrepancy? We argue the answer is not due to some Bradley effect; in fact, the downwards discrepancy suggests otherwise. Instead, it is due to methodological limitations; specifically, there is an over-reliance on the English language. On the one hand, survey efforts – from the September 2024 AAPI Data to the 2020 Collaborative Multiracial Post-Election Survey (CMPS) to the Pew's 2022-23 Asian American Survey – certainly acknowledge the linguistic diversity of the Asian American population. This is evident by the translations available. On the other hand, the translations are limited to only Chinese, Korean, Vietnamese – with a couple additional languages (Arabic and Urdu for CMPS; Hindi and Tagalog for Pew) – at the exclusion of some other linguistic communities that have low(er) English proficiency (e.g., Burmese, Mongolian, and Nepalese).

More importantly, during the actual data collection stage, the surveys are administered by third-party vendors (i.e., National Opinion Research Center, Pacific Market Research, and Westat) whose panel – by design – will have more English speakers. Consider that in the 2020 CMPS, only 1.55% of the Asian respondents took the survey in a non-English language. While the figure is higher for the 2022-23 Pew report (12.1%), this is still far from the actual population numbers.[1] Per the US Census (American Community Survey 2023), 43.8% of Asian Americans speak English "not well" – if none whatsoever. Moreover, just because someone can speak English "well" does mean they are comfortable with or competent in taking surveys in the English language. And since English proficiency is not randomly distributed, the systematic omission of this subset of the population means our inferences about Asian American voting behavior are systematically biased towards liberals and Democrats.

[1] Figures for September 2024 AAPI Data are not publicly available.

Importance of Language

Survey language matters (Liu et al. 2018; Pérez and Tavits 2022). For example, Pérez (2016) finds Latinx respondents have higher opinion levels of concepts (e.g., American identity) when tied to the interview language. Similarly, Lee and Pérez (2014) find the language of the survey – i.e., English or Spanish – affects how respondents report political facts. We build on this methodological approach but shift the focus to Asian Americans – recognizing the challenges that come with the linguistic diversity of the community.

Where we depart from the Pérez machinery is that we are not leveraging bilingual speakers to causally identify the difference between two languages. Instead, we are interested in the selection of the observed – i.e., which survey language. We contend language choice is a measure of some other covariate – and this covariate in turn affects voting behavior. And as such, without properly considering the survey language, our results are systematically biased towards Democrats.

Survey language choice can be a measurement of three general explanations. The first is about **individual attributes** that reflect cosmopolitanism – inclusive of the belief of liberal values. If English is a language of cosmopolitanism (Hu and Liu 2020), then individuals who are older, less educated, men, and/or poorer are both less likely to choose English as their survey language *and* vote Democrat.

The second explanation is about the **ancestral country.** There are political cultural narratives suggesting people who come from authoritarian countries (Inglehart 1988) or left-leaning regimes (Irizarry 2024) are more likely to tolerate – if not outright want – strongmen leaders. A derivative of this would suggest people who are from democracies are more likely to vote democrat. The argument is not simply that they do not crave the strong authority; instead, these countries – e.g., Japan, South Korea, Philippines, and Taiwan – were also the ones that had very strong

American state-building influence post-WW2 and were thus the first beneficiaries of the 1965 Immigration and Naturalization Act (Hsu 2015).

The third explanation is about **acculturation experiences** (Roman, Walker, and Barreto 2022; Roman 2023). The continued use of the Asian language can reflect either the inability to speak English or an intentional choice to continue using the alternative vernacular. In either case, what we are likely to see is very homogeneous social networks – from churches to social media – where the non-English language continues to be used. The lack of diversity means less inclusive values (Liu 2021; Tokeshi 2023).

We are agnostic as to which mechanism survey language choice is measuring. Moreover, we do not presume the mechanisms are mutually exclusive. However, we do contend that the use of a non-English language suggests individuals are less likely to vote Democrat. Specifically:

Hypothesis: *Respondents who take surveys in a non-English language are less likely to vote for Harris.*

Data Collection

We fielded an online survey between April 15 to October 15, 2024, to the Asian American community. The only exclusion criteria were age (18+) and that the respondent could trace their family ancestry to a country on the Aisa continent (inclusive of Russia and Turkey). The survey was available in English and **50 other languages:** Arabic, Armenian, Assamese, Azerbaijani, Bengali, Burmese, Cebuano, Chinese (simplified), Chinese (traditional), Dari, Dzongkha, Farsi, Filipino, Georgian, Gujarati, Hebrew, Hindi, Hmong, Ilocano, Indonesian, Japanese, Kannada, Kazakh, Khmer, Korean, Kurdish, Kyrgyz, Lao, Malay, Malayalam, Marathi, Mongolian, Nepali, Odia, Pashto, Punjabi, Russian, Sindhi, Sinhala, Tajik, Tamil, Thai, Telugu, Tibetan, Turkish, Turkmen, Urdu, Uyghur, Uzbek, and Vietnamese.

All translations underwent multiple rounds of verification. For languages that no one on the team spoke natively, we first ran the survey in English through Google Translate **(version 1).** Next, we hired native speakers in the Asian country through Upwork. We tasked the native speakers to edit the literal translation – i.e., making sure there were no major grammar and vocabulary errors. In addition, we asked them to minimize the clunky prose that happens with AI-generated translations **(version 2).** The average cost for version 2 was $20-$30 USD per language. Finally, we hired native speakers of the language again, but this time we restricted it to those who resided in the US. We recruited most of these individuals through our universities – from foreign language instructors to staff networks to international student offices. We tasked these native speakers to make sure the translation was more semantic – i.e., capturing the cultural context such as "national assault rifles", "affirmative action", and "reparations for African Americans" **(version 3).** The average cost for version 3 ranged anywhere from $50 to $5000 per language.

The survey questions mirrored many of the questions in the CMPS. For the purposes of this note, the question we are interested in is: "*If the election were being held today, would you be inclined to vote for Kamala Harris, Donald Trump, or some other candidate?*" Specifically, we are interested in whether respondents are likely to vote for Harris (*Harris=1*). And our key explanatory variable is whether the respondent took the survey in English (*English=1*) or one of the other 50 languages.

Survey Recruitment

Given the hard-to-reach nature of some of these communities, we collected as many surveys as possible with the assumption that a sample sufficiently large will have the same effects as one drawn representatively (Coppock, Leeper, and Mullinix 2018). We offered $100 gift cards through a lottery format (odds of winning: 1 in 500).

We employed three strategies for recruitment. The first strategy was **posting the survey link on social media platform** – in all 51 languages. While Facebook may not be the most popular social media site in the US (68%), its presence in some Asian countries is (near) monopolistic. In fact, in multiple countries the percentage of Internet users who use Facebook exceeds 100% (Statista 2024).[2] We expect these communities to continue using Facebook in the US. At the same time, we were also cognizant that Facebook penetration is not consistent across all groups. For example, the Chinese from the People's Republic of China (PRC) are much more likely to engage in WeChat. Likewise, Koreans are much more likely to use Kakao. And Telegram is popular in the Russian-speaking world. To this end, we also bought ad space – either on specific channels or with certain influencers.[3]

Our second strategy was **convenience and snowball sampling methods with our own personal and professional networks.** This included people in each of our respective ethnic communities, our alma maters (including high school), our current employing institutions, and our social (religious) circles. Note that collectively among the coauthors, our networks easily reached across at least 40 of the 50 states. We leveraged these networks not only to find survey takers but to also hire research assistants who then distributed the survey in their own networks.

The third strategy was an **aggressive cold-call email recruitment campaign.** We identified *four* targets of the campaign. The first was the *area studies centers* at universities and colleges. We sent the survey link to all relevant centers – e.g., Center for Asian Studies, Center for South Asian Studies, and Center for Korean Studies – at the top 200 universities *and* top 200 colleges per the *US News and World Report* (N>700). Next, we targeted *student organizations* representing an Asian community at any level of aggregation – e.g., from Asian Student Association to East

[2] Mongolia (116%), Philippines (115%), Cambodia (110%), and Vietnam (103%)

[3] Our efforts to post on Telegram were denied by the university compliance office.

Asian Student Association to Hong Kong Student Association – along any substantive interests – e.g., from law school to nursing, from dance to religion – at the top 200 universities (N>2000). Here, we focused only on the universities and not the liberal arts colleges given that former tend to have larger student populations.

Our third target was *community organizations*. The International Revenue Services maintains a list of all organizations registered in the US. We scraped the list searching for all organizations with "Asian" in the name, an international region within Asia (e.g., "South Asian"), an Asian country (e.g., "India"), an intranational region within an Asian country (e.g., "Punjab"), or an ethnic group within Asia (e.g., "Sikhs"). Doing so resulted in almost 15,500 organizations. We sent the survey link to each of these organizations with the recruitment message in both English and the relevant Asian language. And finally, our fourth target were the *K-12 educators* in the public schools (K-12) in the top 25 largest metropolitan areas in the US. For each school district, we identified administrators, teachers, and staff who were either of Asian descent or taught Asian language or culture.

Empirics

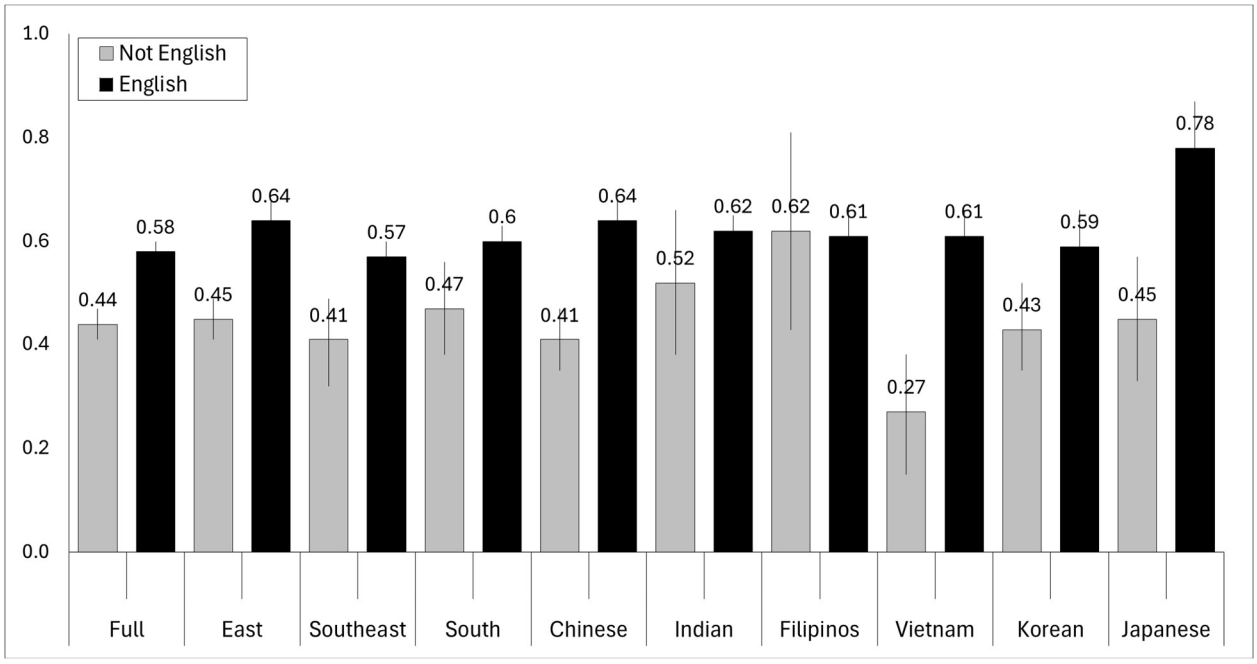
Given our recruitment strategy, we assign weights based on country of ancestral origin, state of residency, age, and gender. We do not weight by income given the higher non-answers for that question – and that missingness correlates very highly with language of survey: Respondents who took the survey in a non-English language were 25% more likely to *not* answer that question.

Let us first look at the demographic overview. When it comes to language, 27.7% of our respondents took the survey in a non-English language – a figure that is much higher than any other survey. And among the non-English languages, the most common were Chinese Simplified (6.5%), Korean (4.3%), Chinese Traditional (3.5%), Japanese (1.7%), and Vietnamese (1.4%). And when it comes to country of origin, the top five are India (17%), China – excluding Taiwan (15%), South

Korea (10%), Vietnam (8%), and the Philippines (6%). And overall, the likelihood of voting for Harris was 54% – a figure that matched the exit polls!

Figure 1 shows the likelihood of voting for Harris given the survey language based simply on means test. What is striking is that while those who took the survey in English were much more likely to vote for Harris (58%), the figure drops to 44% when we look at those who took the survey in another language. This 14-percentage point gap remains consistent even when we break it down by region (East Asia: 45% versus 64%; Southeast Asia: 41% versus 57%; and South Asia: 47% versus 60%) – suggesting the results are not being driven any one region. Furthermore, when we look at the top six countries, the results remain qualitatively the same, although the gap for Vietnam is terrifying: a 2.25-fold difference between the non-English survey takers (27%) and the English ones (61%). Interestingly, the observed differences across languages are less pronounced – if significant – for India and the Philippines. This is not wholly surprising given the official status of English as a language in both of those countries.

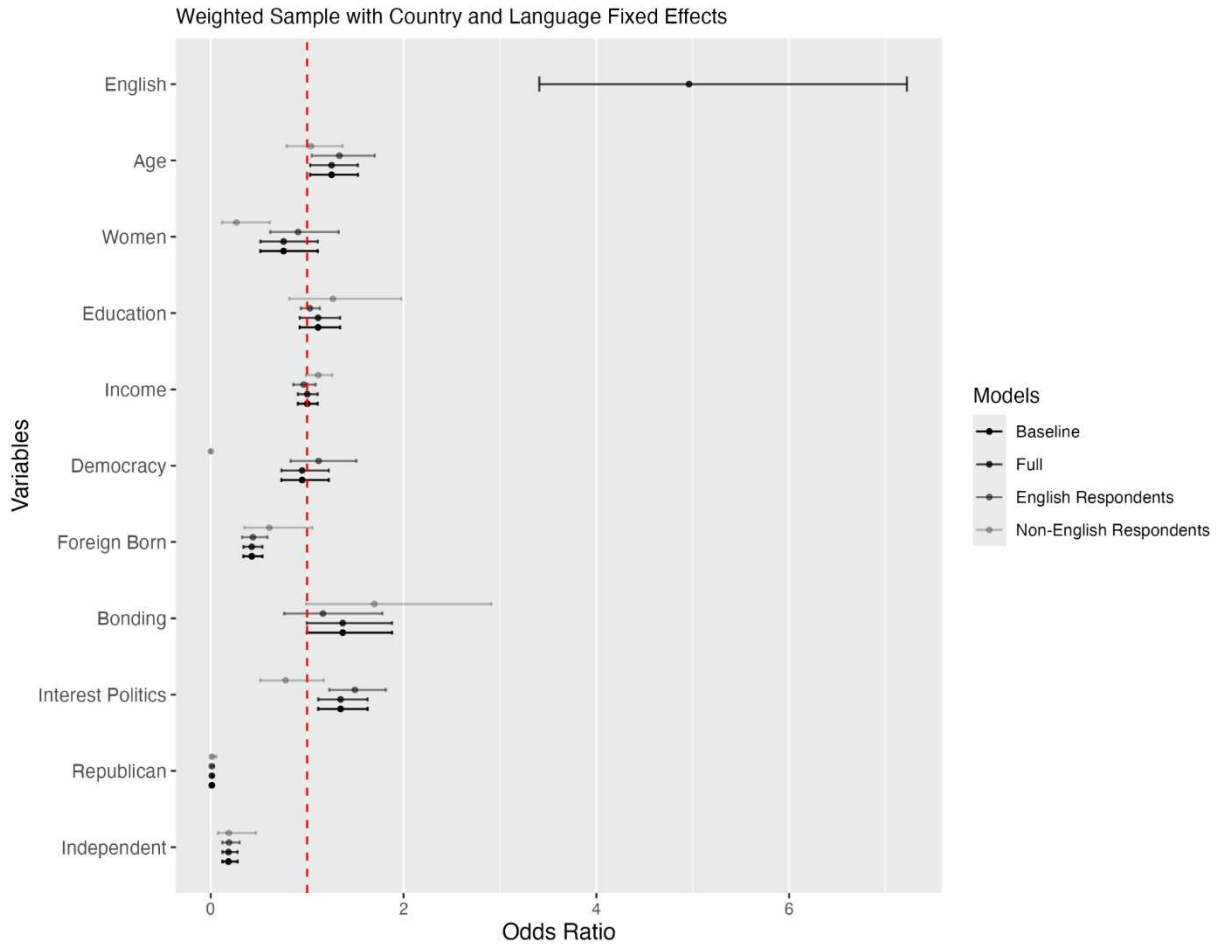
Figure 1 – Likelihood of Voting for Harris Based on Survey Language (Means Test)



Next, to identify what language is a measurement of, we look at **individual attributes** (age, education, gender, and income), **ancestral country** considerations (e.g., whether the country is a consolidated democracy today), and **acculturation experiences** in the US (e.g., whether the respondent was foreign born and the homogeneity of their social network). We estimate models with ancestral country and language fixed effects with standard errors clustered by ancestral country using a weighted sample.

Figure 2 shows the odds ratio from four models (see Appendix A1 for regression coefficients). The first model **(baseline)** includes all the variables without controlling for survey language. In the second model **(full),** we include the language of the survey. The effects for English are substantive

Figure 2 – Effects of Survey Language on Voting for Harris (Odds Ratio)



and significant – at an almost five-fold difference! Next, we consider the effects on two subsamples – those who took the survey in **English** versus **Non-English.** The results suggest the effects are statistically different across languages. What is particularly noteworthy is the effect of gender: Women are statistically less likely to vote for Harris – especially those who took the survey in a non-English language!

Discussion

While it is important to have nationally representative surveys, in this note we highlight how when it comes to immigrant communities, it is imperative that we do not exclude those who do not speak English. The 43.8% non-English proficient Asian Americans are far from a trivial population. And by no means is this an Asian American phenomenon. Even among the Latinx population, 40.6% do not speak English well (American Community Survey 2023). While it is certainly a *necessary* condition to have other languages on the survey, their mere presence is not *sufficient.* In this note, we suggest a workflow for reaching *and recruiting* English-limited respondents.

In addition, when it comes to immigrant communities, using third-party vendors usually means a price point that is simply not affordable for most researchers. In contrast, our workflow – while time-consuming – only cost us in the low five digits. What this suggests is that hard-to-reach communities need not be impossible to reach – and more importantly to study and understand – because of financial barriers.

References

AAPI Data. 2024. "September 2024 AAPI Voter Survey". URL: https://aapidata.com/surveys/sept-2024-aapi-voter-survey/ (Accessed December 18, 2024).

American Community Survey. 2023. "S1601 | Language Spoken at Home." URL: https://data.census.gov/table/ACSST1Y2023.S1601?q=S1601 (Accessed December 18, 2024).

Coppock, Alexander, Thomas J Leeper, and Kevin J Mullinix. 2018. "Generalizability of heterogeneous treatment effect estimates across samples." *Proceedings of the National Academy of Sciences* 115(49):12441-46.

Hsu, Madeline Y. 2015. *The Good Immigrants*. Princeton University Press.

Hu, Yue and Amy H. Liu. 2020. "The Effects of Foreign Language Proficiency on Public Attitudes." *Journal of East Asian Studies* 20(1):1-23.

Irizarry, Giovanni Castro. 2024. "The Influence of Country of Origin in the Process of Party Identification Acquisition." *Journal of Race, Ethnicity, and Politics* 9(1):80-98.

Lee, Taeku, and Efrén O Pérez. 2014. "The persistent connection between language-of-interview and Latino political opinion." *Political Behavior* 36:401-25.

Liu, Amy H. 2021. *The Language of Political Incorporation*. Temple University Press.

Liu, Amy H, Sarah Shair-Rosenfield, Lindsey R Vance, and Zsombor Csata. 2018. "Linguistic Origins of Gender Equality and Women's Rights." *Gender & Society* 32(1):82-108.

Pérez, Efrén O. 2016. "Rolling off the tongue into the top-of-the-head." *Political Behavior* 38:603-34.

Pérez, Efrén, and Margit Tavits. 2022. *Voicing Politics*. Princeton University Press.

Roman, Marcel F. 2023. "Living in the Shadow of Deportation." *Political Research Quarterly* 76(3):1460-74.

Roman, Marcel, Hannah Walker, and Matt Barreto. 2022. "How social ties with undocumented immigrants motivate Latinx political participation." *Political Research Quarterly* 75(3):661-75.

Statista. 2024. "Countries with the highest Facebook audience reach as of April 2024." URL: https://www.statista.com/statistics/278435/percentage-of-selected-countries-internet-users-on-facebook/ (Accessed December 18, 2024).

Tokeshi, Matthew. 2023. "Anti-black prejudice in Asian American public opinion." *Politics, Groups, and Identities* 11(2):366-89.

Zarsadiz, James. 2024. "Opinion: What did the Asian American vote this year tell us?" *Los Angeles Times* (November 10). URL: https://www.latimes.com/opinion/story/2024-11-10/election-2024-asian-american-voters (Accessed December 18, 2024).s