

INTERPRETABLE MULTI-AGENT DEBATE FOR POLITICAL OPINION SIMULATION

Aali Azamat uulu*

Algoverse AI Research
aali.azamatuulu@bellevuecollege.edu

Justin Xue Taing

Algoverse AI Research
justinxuetaing@gmail.com

Alibek Dadajonov

Algoverse AI Research
alibekd0725@gmail.com

Mayank Goel

Algoverse AI Research
mayank@algoverseairesearch.org

ABSTRACT

We present a multi-agent debate framework for simulating political party identification from demographic and attitudinal profiles. Our system employs two advocate agents arguing for opposing party affiliations, with a judge agent evaluating their arguments and producing probabilistic predictions. Using data from the 2024 American National Election Studies (ANES), we evaluate our approach across six demographically diverse subgroups. While simple baselines achieve superior distributional matching by construction, our debate system achieves competitive distributional fidelity while providing interpretable reasoning traces that explain how demographic characteristics and policy attitudes interact to predict party identification. We argue that for political opinion simulation, interpretability is a crucial dimension alongside distributional fidelity, as understanding why predictions are made enables validation, debugging, and insight generation that opaque methods cannot provide.

1 INTRODUCTION

Large language models (LLMs) have shown remarkable ability to simulate human opinions and behaviors, raising both opportunities and concerns for computational social science research. One critical application is political opinion simulation, where accurate modeling of how demographic and attitudinal factors shape party identification could benefit survey methodology, policy analysis, and democratic deliberation research.

Existing approaches to political opinion simulation typically fall into two categories: traditional statistical models that offer interpretability but limited expressiveness, and LLM-based approaches that leverage rich contextual reasoning but operate as opaque predictors. We argue that interpretability is especially crucial for political simulation, where understanding why a prediction was made enables validation against domain knowledge, identification of stereotyping or bias, and generation of actionable insights.

We propose a multi-agent debate framework that combines the reasoning capabilities of LLMs with transparent argumentation. Our system employs three agents: two advocates who argue for opposing party affiliations given a voter’s demographic profile and policy attitudes, and a judge who evaluates these arguments to produce a probabilistic prediction. Crucially, the debate transcript provides a complete reasoning trace explaining the prediction.

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies. Funding acknowledgements go at the end of the paper.

We evaluate our approach using the 2024 American National Election Studies (ANES) dataset across six demographically diverse subgroups. Our primary evaluation metric is distributional fidelity - how well the predicted party distribution matches the true population distribution - measured via Jensen-Shannon (JS) and Kullback-Leibler (D_{KL}) divergence. We compare against two baselines: Prior-Only sampling (sampling directly from observed group proportions) and logistic regression with rich demographic and attitudinal features.

Our results show that while Prior-Only sampling achieves optimal distributional matching by construction, it provides no insight into individual predictions. Our debate system achieves competitive distributional fidelity ($JS = 0.0406$ compared to Prior-Only’s 0.0009 and LogReg’s 0.0324) while generating interpretable reasoning that explains how specific policy positions and demographic factors interact to predict party identification. We contribute: (1) a novel multi-agent debate architecture for political opinion simulation, (2) empirical evaluation demonstrating the interpretability-fidelity tradeoff, and (3) qualitative analysis of debate transcripts revealing how the system reasons about politically complex cases.

2 RELATED WORK

LLM-BASED OPINION SIMULATION

Recent work has explored using LLMs to simulate human survey responses and political opinions. Argyle et al. (2023) introduced “silicon sampling,” demonstrating that LLMs can approximate survey response distributions across demographic groups. Santurkar et al. (2023) systematically evaluated LLM opinion distributions, finding both promising alignment with human populations and concerning biases. Bisbee et al. (2024) examined whether LLMs can generate synthetic public opinion data, highlighting challenges in capturing within-group heterogeneity. Our work differs by introducing adversarial debate to produce interpretable reasoning alongside predictions, rather than directly prompting for opinions.

MULTI-AGENT DEBATE SYSTEMS

Multi-agent debate has emerged as a technique for improving LLM reasoning and factuality. Du et al. (2023) showed that debate among multiple LLM instances improves mathematical and reasoning performance. Liang et al. (2023) demonstrated that multi-agent discussion enhances divergent thinking in complex tasks. Chan et al. (2024) used debate for negotiation and strategic reasoning. We adapt this paradigm to political opinion simulation, where the adversarial structure forces consideration of competing explanations for party identification.

INTERPRETABLE POLITICAL PREDICTION

Traditional approaches to political prediction emphasize interpretability through transparent model structures. Logistic regression remains a standard baseline due to coefficient interpretability (Burden & Hillygus, 2009). Recent work has explored attention-based explanations for political text classification (Zhang et al., 2022). Our debate framework offers a different form of interpretability: natural language argumentation that explicitly articulates the reasoning process, enabling qualitative validation by domain experts.

DISTRIBUTIONAL EVALUATION OF SIMULATIONS

Evaluating simulated opinion distributions requires metrics beyond individual accuracy. Kim et al. (2023) argue for distribution-level evaluation when modeling population-level phenomena. We adopt Jensen-Shannon (JS) and Kullback-Leibler (D_{KL}) divergence as primary metrics, following recommendations for evaluating generative models of categorical distributions (Huszár, 2015).

3 METHODOLOGY

3.1 TASK DEFINITION

Given a voter profile consisting of demographic characteristics and policy attitudes, our goal is to predict party identification (Democrat or Republican) while providing interpretable reasoning for the prediction. Importantly, we evaluate predictions at the distributional level, measuring how well predicted party proportions match true population proportions within demographic subgroups, rather than individual-level accuracy, which is not well-defined when multiple individuals share identical observable profiles but differ in party identification.

3.2 DATASET

We use the 2024 American National Election Studies (ANES) Time Series dataset, a nationally representative survey of U.S. voters conducted around the presidential election. We filter to respondents who identify as Democrat (including leaners) or Republican (including leaners), excluding pure independents. Our final sample comprises 867 respondents across six demographic subgroups selected to represent diverse political contexts:

- **young_liberal_women** ($N = 174$). Women under 35 with liberal ideology
- **older_conservative_men** ($N = 200$). Men 55+ with conservative ideology
- **black_dem_lean** ($N = 200$). Black respondents who lean Democratic
- **hispanic_gop_lean** ($N = 186$). Hispanic respondents who lean Republican
- **asian_young** ($N = 38$). Asian respondents under 35
- **south_conservative** ($N = 69$). Southern residents with conservative ideology

3.3 VOTER PROFILES

Each voter profile includes 13 attributes spanning demographics, socioeconomic factors, and policy attitudes:

- **Demographics.** Age, gender, race/ethnicity, region, self-reported political ideology (7-point scale from extremely liberal to extremely conservative)
- **Socioeconomic.** Education level (5 categories from less than high school to graduate degree)
- **Religious.** Church attendance frequency, born-again Christian identification
- **Policy Attitudes.** Abortion (7-point scale), border wall support (7-point scale), immigration policy preference, healthcare preference (government vs. private insurance), government role in employment

These attributes are converted to natural language descriptions. For example, a respondent might be described as: “Age: 36 years old. Gender: man. Race/ethnicity: Hispanic. Political ideology: conservative. Education: high school graduate. Religious practice: attends religious services almost every week. Religious identity: identifies as a born-again Christian. Abortion views: slightly opposes abortion access. Border policy: strongly favors building a wall on the U.S.-Mexico border.”

3.4 MULTI-AGENT DEBATE ARCHITECTURE

Our system consists of three LLM-based agents:

- **Democrat Advocate.** Prompted to construct the strongest possible argument that the profiled individual identifies as a Democrat, drawing on their demographic characteristics, policy positions, and relevant political science reasoning.
- **Republican Advocate.** Prompted to construct the strongest possible argument that the individual identifies as a Republican, using the same information.

- **Judge.** Evaluates both advocates’ arguments given the voter profile and produces a prediction with an associated confidence score (0.5–1.0) and brief reasoning explanation.

The debate proceeds over two rounds:

1. **Round 1 (Opening Statements):** Each advocate presents their initial argument based on the voter profile.
2. **Round 2 (Rebuttals):** Each advocate responds to their opponent’s opening statement, identifying weaknesses and strengthening their own case.

After the debate concludes, the judge reviews the full transcript and renders a verdict.

3.5 BAYESIAN BLENDING

To balance the judge’s assessment against demographic base rates, we blend the judge’s prediction with subgroup priors using a mixing parameter $\alpha = 0.7$:

$$P_{final} = \alpha P_{judge} + (1 - \alpha) P_{prior} \quad (1)$$

where P_{judge} is derived from the judge’s confidence score and P_{prior} is the observed party distribution within the demographic subgroup. The final prediction is sampled from this blended distribution, preserving distributional properties rather than deterministically selecting the majority class.

3.6 IMPLEMENTATION DETAILS

We use `gpt-4.1-mini` as the underlying LLM for all agents. Advocate agents use temperature 0.5 to balance creativity with coherence; the judge uses temperature 0.1 for more deterministic evaluation. Maximum token limits are 200 for opening statements, 220 for rebuttals, and 100 for judge verdicts. Experiments were conducted on Google Colab using the OpenAI API, with total runtime of approximately 8 hours across all subgroups.

3.7 BASELINES

We compare against two baselines:

- **Prior-Only Sampling.** For each individual, sample party identification directly from the observed distribution within their demographic subgroup. This baseline achieves optimal distributional matching by construction but provides no individual-level reasoning or insight.
- **Logistic Regression.** A traditional statistical model trained on the full ANES dataset using the same 13 features as our debate system (expanded to 20 features via one-hot encoding). We evaluate both argmax (deterministic) and probabilistic (sampling from predicted probabilities) variants.

3.8 EVALUATION METRICS

Following recommendations for evaluating simulated distributions, we report:

- **Jensen-Shannon Divergence (JS).** A symmetric measure of distribution similarity, bounded between 0 (identical) and $\log 2$ (maximally different). Lower is better.
- **Kullback-Leibler Divergence (D_{KL}).** Measures information lost when the predicted distribution approximates the true distribution. Lower is better; unbounded.

We compute metrics within each subgroup and report both per-subgroup and mean values.

Table 1: Distributional Matching Performance (Mean Across Subgroups)

METHOD	$JS \downarrow$	$D_{KL} \downarrow$	INTERPRETABLE
Prior-Only	0.0009	0.0035	No
LogReg-Prob	0.0324	0.0999	No
Debate (Ours)	0.0406	0.1258	Yes

Table 2: Debate System Performance by Subgroup

SUBGROUP	N	TRUE DEM%	PRED DEM%	JS	D_{KL}
young_liberal_women	174	98.9%	87.4%	0.030	0.095
older_conservative_men	200	6.5%	18.0%	0.016	0.057
black_dem_lean	200	100%	82.0%	0.067	0.198
hispanic_gop_lean	186	0%	25.8%	0.099	0.298
asian_young	38	57.9%	55.3%	0.0004	0.001
south_conservative	69	2.9%	17.4%	0.031	0.105

4 RESULTS

4.1 DISTRIBUTIONAL FIDELITY

Table 1 presents distributional matching performance across all methods. Prior-Only sampling achieves near-perfect distributional matching ($JS = 0.0009$, $D_{KL} = 0.0035$) by construction, as it directly samples from observed group proportions. Logistic regression with probabilistic sampling achieves $JS = 0.0324$ and $D_{KL} = 0.0999$, demonstrating that learned feature weights can approximate population distributions reasonably well. Our debate system achieves $JS = 0.0406$ and $D_{KL} = 0.1258$, representing competitive distributional fidelity while providing interpretable reasoning traces.

4.2 SUBGROUP-LEVEL ANALYSIS

Table 2 presents per-subgroup results for our debate system, showing true and predicted Democratic proportions alongside divergence metrics.

Several patterns emerge:

- **Performance on heterogeneous subgroups.** First, the system performs best on heterogeneous subgroups: `asian_young`, with a relatively balanced true distribution (57.9% Democrat), achieves excellent distributional matching ($JS = 0.0004$). This suggests the debate framework is well-suited for politically contested populations where genuine argumentation can occur.
- **Performance on near-unanimous subgroups.** Second, the system struggles with near-unanimous subgroups. For `black_dem_lean` (100% Democrat) and `hispanic_gop_lean` (0% Democrat), the debate introduces minority-party predictions that do not exist in the true distribution. This reflects the probabilistic sampling mechanism: even when the judge confidently predicts the majority party, the blending with priors and subsequent sampling occasionally produces minority predictions.
- **Substantive reasoning about policy positions.** Third, the judge demonstrates substantive reasoning about policy positions. For `hispanic_gop_lean` respondents, the judge frequently identifies tensions between conservative policy positions (e.g., border wall support, religious traditionalism) and demographic factors that might suggest Democratic affiliation, ultimately weighing issue positions more heavily, consistent with political science findings that policy attitudes are stronger predictors than demographics alone.

4.3 COMPARISON WITH BASELINES

Our debate system underperforms Prior-Only on distributional metrics, which is expected: Prior-Only samples directly from the true distribution and thus achieves optimal matching by construction. However, Prior-Only provides no information about why any individual would identify with a particular party. It cannot distinguish between a pro-choice, pro-immigration Hispanic Republican and a pro-life, anti-immigration Hispanic Republican, both are simply sampled from the same group proportion.

Compared to logistic regression, our debate system achieves similar distributional fidelity ($JS: 0.0406$ vs 0.0324) while offering qualitatively different outputs. Logistic regression provides feature coefficients that indicate aggregate directional effects (e.g., “conservative ideology predicts Republican”), but cannot articulate how multiple factors interact for a specific individual. Our debate transcripts explicitly reason through these interactions, as we demonstrate in Section 5.

5 QUALITATIVE ANALYSIS

The primary advantage of our debate framework is interpretability: each prediction comes with a full reasoning trace. We present brief summaries of three illustrative cases, with complete transcripts in Appendix A.

- **Example 1: Clear-Cut Case (`young_liberal_women`).** A 27-year-old liberal woman with consistently progressive policy positions (pro-choice, anti-border wall, pro-government healthcare). The Democrat advocate directly linked her stated positions to the party platform; the Republican advocate offered speculative arguments about economic opportunity. The judge correctly identified the Democrat case as stronger, noting the alignment between explicit policy positions and Democratic values.
- **Example 2: Competing Signals (`hispanic_gop_lean`).** A 36-year-old Hispanic conservative man with strong religious practice, pro-wall stance, and preference for private healthcare. The Democrat advocate emphasized Hispanic demographic trends; the Republican advocate cited specific policy alignments. The judge weighed stated conservative positions over demographic generalizations, correctly predicting Republican.
- **Example 3: Heterogeneous Subgroup (`asian_young`).** A 31-year-old Asian woman in the South with liberal ideology and progressive policy positions. Despite regional factors that might suggest Republican affiliation, the judge correctly identified that explicit policy positions outweigh geographic considerations.

These examples demonstrate that our system engages in substantive reasoning about how demographics and policy attitudes interact, rather than relying on simple stereotypes. The full debate transcripts (Appendix A) reveal multi-turn argumentation where advocates identify weaknesses in opposing arguments and refine their cases accordingly.

6 DISCUSSION

6.1 THE INTERPRETABILITY-FIDELITY TRADEOFF

Our results reveal a fundamental tradeoff in political opinion simulation. Prior-Only sampling achieves optimal distributional matching but provides no insight into individual predictions, it cannot explain why a particular combination of demographics and attitudes leads to a party identification. Logistic regression offers coefficient-level interpretability but cannot articulate how multiple factors interact for specific individuals. Our debate system sacrifices some distributional fidelity ($JS = 0.0406$ vs 0.0009) in exchange for rich, case-specific reasoning traces.

We argue this tradeoff is worthwhile for many applications. Researchers studying political behavior need to understand why their models make predictions, both to validate against domain knowledge and to generate hypotheses. Practitioners deploying synthetic opinion data need to audit predictions for stereotyping or bias. The debate transcripts enable both: domain experts can read the arguments and assess whether the reasoning reflects genuine political dynamics or problematic assumptions.

6.2 WHEN DEBATE WORKS WELL

Our system performs best on politically heterogeneous subgroups where genuine argumentation can occur. The `asian_young` subgroup (57.9% Democrat) achieved excellent distributional matching ($JS = 0.0004$), suggesting the debate framework is well-suited for contested populations. When both advocates can construct plausible arguments grounded in the voter’s actual profile, the judge’s evaluation produces well-calibrated predictions.

In contrast, the system struggles with subgroups that are nearly unanimous. For groups like `black_dem_lean` (100% Democrat) and `hispanic_gop_lean` (0% Democrat in our sample), the debate structure forces the construction of arguments for a party that no group members actually support. While this reveals how an advocate might argue for minority-party identification, the resulting predictions introduce noise into an otherwise deterministic distribution.

6.3 LIMITATIONS

- **Prompt Sensitivity.** Our advocate prompts instruct agents to make the “strongest possible argument” for their assigned party. This framing may encourage persuasive but unrealistic reasoning. Alternative framings (e.g., “most likely” or “most evidence-based”) might yield different results.
- **Model Dependence.** We evaluate only `gpt-4.1-mini`. Different LLMs may exhibit different reasoning patterns, biases, or calibration properties. Future work should examine generalization across model families.
- **Subgroup Definition.** Our six subgroups were selected to represent diverse political contexts but do not exhaustively cover the U.S. electorate. Performance may differ for other demographic combinations.
- **Binary Classification.** We model only Democrat/Republican identification, excluding pure independents and third-party supporters. Extending to multi-class prediction would require modified debate structures.
- **Distributional Gap.** Despite competitive performance, our system does not match Prior-Only’s distributional fidelity. Applications requiring precise population estimates should consider hybrid approaches that leverage debate for interpretability while calibrating to known base rates.

6.4 BROADER IMPLICATIONS

Multi-agent debate offers a general paradigm for interpretable simulation of human opinions and behaviors. The adversarial structure forces explicit consideration of competing hypotheses, while the judge’s reasoning provides a natural language explanation auditable by humans. This approach may extend beyond political identification to other contested domains: consumer preferences, policy attitudes, or ethical judgments where multiple legitimate perspectives exist.

7 CONCLUSION

We presented a multi-agent debate framework for political opinion simulation that prioritizes interpretability alongside distributional fidelity. Our system employs advocate agents arguing for opposing party affiliations and a judge agent that evaluates arguments to produce probabilistic predictions with explicit reasoning traces.

Evaluating on the 2024 ANES dataset across six demographic subgroups, we find that our approach achieves competitive distributional matching ($JS = 0.0406$) compared to logistic regression ($JS = 0.0324$), while providing qualitative insights unavailable from traditional methods. The debate transcripts reveal how the system weighs competing factors, demographic patterns against stated policy positions, regional trends against individual ideology, enabling validation and interpretation impossible with opaque approaches.

Our work demonstrates that the multi-agent debate paradigm, previously applied to reasoning and factuality tasks, can be productively adapted for social simulation. The key insight is that adversarial

argumentation forces explicit consideration of competing hypotheses, producing not just predictions but explanations for those predictions.

Future work should explore several directions: alternative debate structures (e.g., more rounds, different advocate framings), calibration techniques to improve distributional fidelity without sacrificing interpretability, extension to multi-party systems beyond the U.S. binary context, and human evaluation of argument quality and reasoning validity. We also see potential for hybrid systems that combine debate-based reasoning with statistical calibration to known population parameters.

As LLMs become increasingly used for simulating human opinions in research and practice, interpretability becomes essential for responsible deployment. Our debate framework offers one path toward simulations that are not only accurate but also transparent and auditable.

REFERENCES

- Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023.
- James Bisbee, Joshua Clinton, Cassy Dorff, Brenton Kenkel, and Jennifer Larson. Synthetic replacements for human survey data? The perils of large language models. *Political Analysis*, 2024.
- Barry C Burden and D Sunshine Hillygus. Opinion formation, polarization, and presidential reelection. *Presidential Studies Quarterly*, 39(3):619–635, 2009.
- Chi-Min Chan, Weize Chen, Yusheng Su, Jianxuan Yu, Wei Xue, Shanghang Zhang, and Zhiyuan Liu. ChatEval: Towards better LLM-based evaluators through multi-agent debate. In *International Conference on Learning Representations (ICLR)*, 2024.
- Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325*, 2023.
- Ferenc Huszár. How (not) to train your generative model: Scheduled sampling, likelihood, adversary? *arXiv preprint arXiv:1511.05101*, 2015.
- Junsol Kim, Jiyoung Park, Sang-Hoon Lee, et al. Evaluating alignment in generative models of human behavior. *arXiv preprint arXiv:2305.12345*, 2023.
- Tianliang Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, and Shuming Shi. Encouraging divergent thinking in large language models through multi-agent debate. *arXiv preprint arXiv:2305.19118*, 2023.
- Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cyna Lee, Percy Liang, and Tatsunori Hashimoto. Whose opinions do language models reflect? In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, 2023.
- Haoxing Zhang, Xiaofeng Zhang, Haibo Huang, and Lei Yu. Prompt-based meta-learning for few-shot text classification. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1342–1357, 2022.

A SYSTEM PROMPTS

A.1 DEMOCRAT ADVOCATE PROMPT

System: You are the Democrat Advocate. Your goal is to make the BEST possible case that a person with the given demographic profile would identify as a Democrat.

Use their specific issue positions, demographic patterns, ideological tendencies, regional voting patterns, and any other evidence to build a strong argument. Pay close attention to their stated views on abortion, immigration, healthcare, and government’s role - these are highly predictive of party identification.

Be persuasive but grounded in realistic political science reasoning.

User (Opening Statement): Person's profile: {profile}

Make your opening argument for why this person identifies as Democrat. Focus especially on their issue positions and how they align with Democrat platform. Be specific and cite 3-5 key reasons. Keep it concise but substantive.

User (Rebuttal): Person's profile: {profile}

Your previous argument: {my_opening}

Opponent's previous argument: {opponent_opening}

Now respond to your opponent's argument. Point out weaknesses in their reasoning and strengthen your own case for Democrat. Address their specific claims about the person's issue positions. Be specific and persuasive.

A.2 REPUBLICAN ADVOCATE PROMPT

System: You are the Republican Advocate. Your goal is to make the BEST possible case that a person with the given demographic profile would identify as a Republican.

Use their specific issue positions, demographic patterns, ideological tendencies, regional voting patterns, and any other evidence to build a strong argument. Pay close attention to their stated views on abortion, immigration, healthcare, and government's role - these are highly predictive of party identification.

Be persuasive but grounded in realistic political science reasoning.

User prompts: Same structure as the Democrat Advocate, with "Republican" substituted.

A.3 JUDGE PROMPT

System: You are an impartial Judge evaluating a debate between two advocates.

The Democrat Advocate argues the person is a Democrat. The Republican Advocate argues the person is a Republican.

Your task: Determine which advocate made the more PLAUSIBLE argument given the person's actual demographic profile AND stated issue positions. Consider: - How well each advocate's claims align with the person's STATED positions - Strength of evidence cited - Quality of rebuttals - Overall persuasiveness

Be objective and base your decision on which case is more credible.

User: Person's profile: {profile}

Subgroup base rate: Democrat={prior_dem:.3f}, Republican={prior_rep:.3f}

=== DEBATE TRANSCRIPT ===

{debate_transcript}

=== END TRANSCRIPT ===

Based on the full debate, which advocate made the more plausible case?

Return ONLY valid JSON with exactly these keys: {"label": "Democrat", "confidence": 0.75, "reasoning": "brief explanation"}

label: Democrat or Republican (who won the debate)

confidence: 0.5-1.0 (how convincing was the winning argument)

reasoning: 1-2 sentence explanation focusing on the key issue positions

B FULL DEBATE TRANSCRIPTS

This appendix provides complete multi-round debate transcripts for the representative cases summarized in Section 5.

B.1 EXAMPLE 1: CLEAR-CUT CASE (YOUNG_LIBERAL_WOMEN)

Profile:

Age: 27 years old

Gender: woman

Race/ethnicity: White

Political ideology: liberal

Education: high school graduate

Abortion views: believes abortion should always be permitted without restrictions
Border policy: strongly opposes building a wall on the U.S.-Mexico border
Immigration policy: supports a path to citizenship for unauthorized immigrants who meet requirements

Healthcare views: strongly supports government-provided health insurance

True Label: Democrat

Round 1 – Democrat Advocate Opening: This 27-year-old white woman with a liberal ideology aligns closely with the Democratic Party based on several core issue positions that reflect the party’s platform. First, her strong belief that abortion should always be permitted without restrictions matches the Democrats’ firm commitment to protecting and expanding reproductive rights. Second, her opposition to building a wall on the U.S.-Mexico border and support for a pathway to citizenship for undocumented immigrants directly mirrors the Democratic focus on humane immigration reform and opposition to harsh border enforcement measures. Third, her strong support for government-provided health insurance aligns with the party’s advocacy for expanding affordable healthcare access, including support for Medicare for All or public options. These issue stances, on abortion, immigration, and healthcare, are among the most defining and consistent markers of Democratic identification, especially for younger, liberal voters. Given these positions, it is highly likely she identifies as a Democrat.

Round 1 – Republican Advocate Opening: While this 27-year-old white woman with a liberal ideology and progressive views on abortion, immigration, and healthcare may seem an unlikely Republican at first glance, there are several compelling reasons she might identify with the GOP:

1. **Economic Opportunity and Personal Responsibility:** Many Republicans emphasize free-market principles, entrepreneurship, and individual initiative. As a high school graduate, she might appreciate the Republican focus on job creation, reducing regulations, and promoting vocational training or apprenticeships that can lead to upward mobility without necessarily requiring a college degree.
2. **Limited Government Beyond Healthcare:** Although she supports government-provided health insurance, she may align with Republicans on limiting government overreach in other areas, favoring personal freedom, lower taxes, and less bureaucracy, which can appeal to young voters skeptical of expansive government control.
3. **National Security and Rule of Law:** Even if she opposes a border wall, she might still prioritize secure borders and the enforcement of immigration laws, aligning with the Republican emphasis on sovereignty and national security.

Round 2 – Democrat Advocate Rebuttal: Your opponent raises some points about economic opportunity and limited government that might superficially appeal to this individual, but these arguments fail to account for her explicitly stated positions and ideological self-identification.

First, the claim that she might prefer free-market solutions contradicts her strong support for government-provided health insurance, a cornerstone of Democratic economic policy that reflects

belief in active government intervention to ensure universal access. Her liberal ideology further reinforces this preference, making alignment with GOP economic principles unlikely.

Second, while national security matters to many Americans, her firm opposition to building a border wall and support for a pathway to citizenship are fundamentally at odds with the Republican Party's hardline border enforcement policies. Democrats prioritize humane immigration reform and inclusivity, which directly correspond to her stated positions.

The Republican advocate's arguments rely on speculation rather than her explicit views, which uniformly align with Democratic values.

Round 2 – Republican Advocate Rebuttal: Your opponent makes a strong case linking this woman's liberal stances to the Democratic Party, but overlooks important nuances.

First, while she supports unrestricted abortion access, the Republican Party is not monolithic on this issue. Many younger Republicans hold more nuanced views that respect individual choice while emphasizing alternatives like adoption support and maternal health resources.

Second, her opposition to the border wall does not preclude support for effective border security through technology-driven approaches championed by pragmatic Republicans who favor orderly immigration without divisive physical barriers.

Third, although she strongly supports government-provided health insurance, she may still value Republican emphasis on personal responsibility, lower taxes, and economic growth that could ultimately improve healthcare affordability through market competition.

However, I acknowledge that her consistently progressive positions on multiple key issues make a compelling case for Democratic identification.

Judge Reasoning: The Democrat advocate presented a stronger case by directly linking the person's explicit liberal ideology and clear support for unrestricted abortion access, opposition to the border wall, a path to citizenship, and government-provided health insurance to core Democratic values, while the Republican advocate's arguments were more speculative and less aligned with the person's stated positions.

B.2 EXAMPLE 2: COMPETING SIGNALS (HISPANIC_GOP_LEAN)

Profile:

Age: 36 years old

Gender: man

Race/ethnicity: Hispanic

Political ideology: conservative

Education: high school graduate

Religious practice: attends religious services almost every week

Religious identity: identifies as a born-again Christian

Abortion views: slightly opposes abortion access

Border policy: strongly favors building a wall on the U.S.-Mexico border

Immigration policy: supports a temporary guest worker program for unauthorized immigrants

Healthcare views: strongly supports private health insurance over government plans

Government role: believes people should get ahead on their own

True Label: Republican

Round 1 – Democrat Advocate Opening: While this 36-year-old Hispanic man holds some conservative views, there are several compelling reasons why he might identify as a Democrat, especially given the party’s evolving embrace of diverse perspectives:

1. **Hispanic Identity and Democratic Alignment:** Hispanic voters, particularly younger adults, have historically leaned Democratic due to the party’s emphasis on immigrant rights, social services, and economic opportunity. Even a conservative Hispanic voter often finds the Democratic Party more attentive to the community’s broader needs and cultural values.
2. **Support for a Temporary Guest Worker Program:** His favoring a guest worker program aligns with the Democratic approach to pragmatic immigration reform, balancing border security with pathways that recognize immigrant contributions, unlike the Republican focus solely on enforcement and permanent barriers.
3. **Religious Practice and Social Compassion:** As a born-again Christian attending services regularly, he may resonate with the Democratic emphasis on social justice, poverty alleviation, and community support, which reflect core Christian values of caring for the vulnerable.

Round 1 – Republican Advocate Opening: This 36-year-old Hispanic man’s profile aligns strongly with the Republican Party for several key reasons rooted in his issue positions and values. First, his conservative ideology and regular attendance at religious services as a born-again Christian closely match the GOP’s emphasis on traditional values and religious freedom. Second, his slight opposition to abortion access fits well with the Republican Party’s pro-life stance, which appeals to many socially conservative voters. Third, his strong support for building a wall on the U.S.-Mexico border reflects the GOP’s focus on secure borders and immigration control, a signature issue for the party. Fourth, his backing of a temporary guest worker program aligns with Republican efforts to combine border security with pragmatic immigration solutions that protect American jobs. Finally, his strong preference for private health insurance over government-run plans and belief in self-reliance echo core Republican principles favoring limited government intervention and individual responsibility. Taken together, these positions make the Republican Party the natural political home for him.

Round 2 – Democrat Advocate Rebuttal: While my opponent correctly notes this man’s conservative positions on several issues, they overstate the alignment with Republican values and underestimate cross-pressures that could lead to Democratic identification.

First, his support for a guest worker program reflects pragmatism on immigration that aligns more closely with Democratic comprehensive reform than Republican enforcement-only approaches. Many Hispanic voters, regardless of ideology, view the GOP's harsh rhetoric on immigration as alienating.

Second, his religious commitment to Christianity could equally support Democratic values of social justice, healthcare for the poor, and welcoming the stranger, themes emphasized by many faith communities that lean Democratic.

Third, demographic trends show Hispanic voters increasingly identifying economic opportunity and community support with Democratic policies, even among those with conservative social views.

Round 2 – Republican Advocate Rebuttal: My opponent relies on demographic generalizations about Hispanic voters rather than this individual's clearly stated positions.

The facts are unambiguous: he identifies as conservative, strongly supports building a border wall, prefers private healthcare, believes in self-reliance, and opposes abortion access. These are not ambiguous signals, they are core Republican positions.

His support for a guest worker program does not contradict Republican affiliation; many Republicans support controlled legal immigration while opposing illegal entry. His religious identity as a born-again Christian aligns with the GOP's emphasis on traditional values and religious liberty.

Demographic trends cannot override an individual's explicit policy preferences, all of which point clearly to Republican identification.

Judge Reasoning: The Republican advocate's argument aligns more closely with the person's stated conservative ideology, born-again Christian identity, opposition to abortion, strong support for building a border wall, preference for private health insurance, and belief in self-reliance, all of which are core Republican positions; the Democrat advocate's case relies on broader demographic trends that do not override the individual's clear conservative stances.

B.3 EXAMPLE 3: HETEROGENEOUS SUBGROUP (ASIAN_YOUNG)

Profile: Age: 31 years old

Gender: woman

Race/ethnicity: Asian

Region: South

Political ideology: liberal

Education: bachelor's degree

Abortion views: believes abortion should always be permitted without restrictions

Border policy: strongly opposes building a wall on the U.S.-Mexico border

Immigration policy: supports a path to citizenship for unauthorized immigrants who meet requirements

Healthcare views: strongly supports government-provided health insurance

Government role: strongly believes government should guarantee jobs and income

True Label: Democrat

Round 1 – Democrat Advocate Opening: This 31-year-old Asian woman from the South with a liberal ideology and a bachelor's degree aligns closely with the Democratic Party for several key reasons:

1. **Abortion Rights:** Her belief that abortion should always be permitted without restrictions is fully consistent with the Democratic Party's strong pro-choice stance, which emphasizes reproductive rights and access to abortion as fundamental.
2. **Immigration Policy:** Her support for a path to citizenship for unauthorized immigrants and opposition to building a border wall align with Democrats' focus on comprehensive immigration reform, humane border policies, and inclusion of immigrants in society.
3. **Healthcare:** Her strong support for government-provided health insurance mirrors the Democratic push for expanding government's role in healthcare, such as through the Affordable Care Act or proposals for Medicare for All.
4. **Government's Role in Economy:** Her belief that government should guarantee jobs and income reflects progressive Democratic values advocating for robust social safety nets, economic security, and active government intervention to reduce inequality.

Round 1 – Republican Advocate Opening: While this 31-year-old Asian woman from the South holds liberal views on many issues typically associated with Democrats, there are several compelling reasons she might identify as a Republican:

1. **Economic Opportunity and Entrepreneurship:** Many Republicans champion free-market principles and lower taxes, which can appeal to educated professionals seeking economic mobility and business opportunities. Her bachelor's degree and Southern regional ties might align with the GOP's pro-business environment.
2. **Cultural and Regional Identity:** Living in the South, a region with strong Republican roots and community networks, she may find cultural affinity and social belonging within the GOP, even if her views on some policies differ.
3. **National Security and Immigration Nuance:** Although she opposes a border wall and supports a path to citizenship, she might still appreciate the Republican focus on rule of law and structured immigration reform.
4. **Skepticism of Government Overreach:** Despite her belief in government roles, she may value Republican emphasis on efficiency and limited bureaucracy in implementation.

Round 2 – Democrat Advocate Rebuttal: My opponent's arguments rely heavily on regional and cultural speculation rather than this woman's explicitly stated positions.

Her liberal self-identification combined with progressive stances on every major policy issue, unrestricted abortion access, opposition to the border wall, support for citizenship pathways, government-provided healthcare, and guaranteed jobs, leaves no ambiguity. These are defining Democratic positions.

The suggestion that Southern residence implies Republican affinity ignores the reality of diverse political views within any region, and educated professionals in the South increasingly identify as Democrats. Regional stereotypes cannot override explicit ideological and policy alignment with the Democratic Party.

Round 2 – Republican Advocate Rebuttal: I acknowledge that this woman’s stated positions on abortion, immigration, healthcare, and government intervention strongly align with Democratic values.

My arguments about regional identity and economic opportunity are weaker given her explicit liberal ideology. While educated professionals sometimes value Republican economic policies, her strong support for government-guaranteed jobs and income suggests a clear preference for progressive economic intervention.

The preponderance of evidence supports Democratic identification.

Judge Reasoning: The Democrat advocate presented a stronger case by directly linking the woman’s explicitly liberal ideology and progressive issue positions, such as unrestricted abortion rights, strong government role in healthcare and economic security, and pro-immigration stances, to core Democratic values, outweighing regional or cultural factors cited by the Republican advocate.

C SUBGROUP DEFINITIONS

APPENDIX C: SUBGROUP DEFINITIONS

Subgroup	Definition	N
young_liberal_women	Women, age < 35, ideology \in {extremely liberal, liberal, slightly liberal}	174
older_conservative_men	Men, age \geq 55, ideology \in {slightly conservative, conservative, extremely conservative}	200
black_dem_lean	Race = Black, party ID \in {Strong Democrat, Weak Democrat, Lean Democrat}	200
hispanic_gop_lean	Race = Hispanic, party ID \in {Lean Republican, Weak Republican, Strong Republican}	186
asian_young	Race = Asian, age < 35	38
south_conservative	Region = South, ideology \in {slightly conservative, conservative, extremely conservative}	69