

ELECTIONFIT: A COMPUTATIONAL LABORATORY OF LLM AGENTS FOR SIMULATING U.S. PRESIDENTIAL ELECTIONS

Anonymous authors

Paper under double-blind review

ABSTRACT

Modeling complex human behavior, such as voter decisions in national elections, is a long-standing challenge for computational social science. Traditional agent-based models (ABMs) are limited by oversimplified rules, while large-scale statistical models often lack interpretability. We introduce ElectionFit, a novel framework that uses Large Language Models (LLMs) to build a “computational laboratory” of LLM agents for political simulation. Each agent is instantiated with a high-fidelity demographic profile and dynamic contextual information (e.g., candidate policies), enabling it to perform nuanced, generative reasoning to simulate a voting decision. We deployed this framework as a testbed on the 2024 U.S. Presidential Election, focusing on seven key swing states. Our simulation’s macro-level results successfully replicated the real-world outcome, demonstrating the high fidelity of our “virtual society”. The primary contribution is not only the prediction, but also the framework’s utility as an interpretable research tool. ElectionFit moves beyond black-box outputs, allowing researchers to probe agent-level rationale and analyze the stability and sensitivity of LLM-driven social simulations.

1 INTRODUCTION

Understanding and modeling complex social phenomena, such as election outcomes, is a crucial endeavor in political and computational social science. Traditionally, research in this domain has relied on statistical models built from polling data Brito et al. (2021); Coletto et al. (2015); Chauhan et al. (2021) or conventional agent-based modeling (ABM) approaches Gao et al. (2022); De Marchi & Page (2014); Qiu & Phang (2020b); Fowler & Smirnov (2005). However, these methods often face limitations: statistical models struggle with dynamic, real-time contexts, while traditional ABMs tend to oversimplify complex human decision-making into predefined rules or heuristic functions. This simplification hinders their ability to capture the full complexity and heterogeneity of real-world voter populations, limiting their utility as true “computational laboratories” for social science.

The emergence of Large Language Models (LLMs) presents a new paradigm, offering unprecedented capabilities to model human behavior with greater nuance and realism Argyle et al. (2023); Zhang et al. (2024); Yu et al. (2024b); Grossmann et al. (2023); Gujral et al. (2024) (Figure 1). This work leverages this potential to propose **ElectionFit**, an agent-based simulation framework designed as a **computational laboratory** for social science research. Instead of merely forecasting, ElectionFit uses LLM agents to construct a high-fidelity “virtual society” for *in silico* experimentation.

Each agent in the ElectionFit framework is instantiated with a detailed voter profile, mirroring the demographic composition of U.S. states based on data from the 2023 American Community Survey (ACS)¹ and the 2020 U.S. Religion Census by the Association of Statisticians of American Religious Bodies (ASARB)². These profiles include key attributes such as state, race, sex, age, occupation, industry, education level, and religion. Crucially, agents are not static; they are provided with contextual information regarding candidate policy positions on major issues (e.g., economic policy, immigration, and abortion rights). This design enables each agent to perform generative rea-

¹<https://www.census.gov/programs-surveys/acs>

²<https://usreligioncensus.org/>

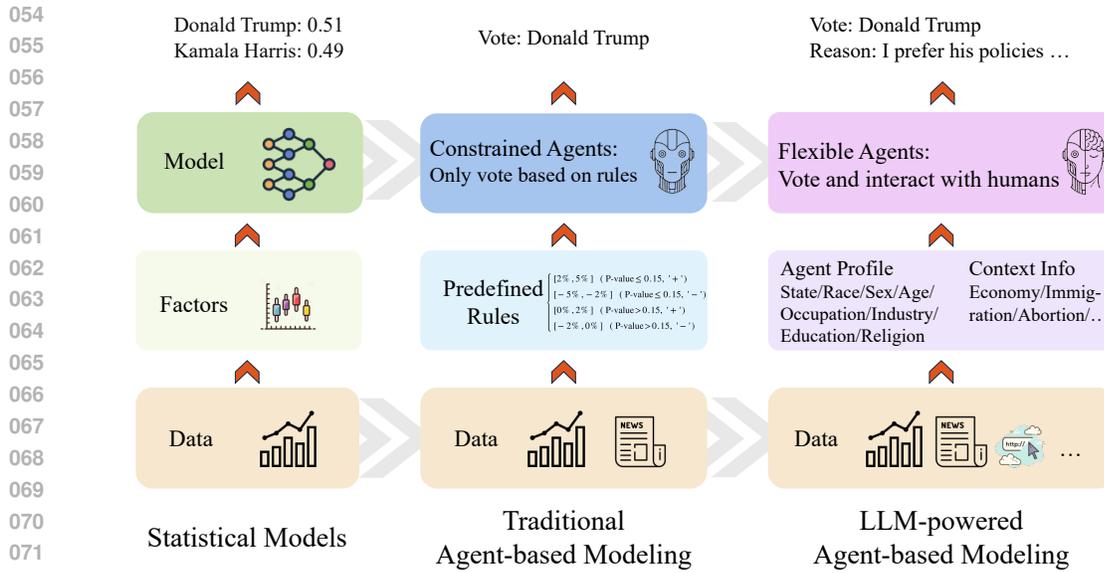


Figure 1: **Comparison of Social Simulation Methodologies.** (Left) Conventional statistical models are often “black boxes” that correlate data-driven factors with outcomes but lack causal or behavioral interpretability. (Center) Traditional agent-based modeling (ABM) relies on agents that follow predefined, heuristic rules. This limits their behavioral realism and ability to adapt to new contextual information. (Right) LLM-powered agent-based modeling, the approach used in our ElectionFit framework, serves as a “computational laboratory”. It empowers autonomous agents with demographic profiles and dynamic context, enabling them to simulate complex, human-like reasoning. This provides a flexible, nuanced, and interpretable simulation essential for social science inquiry.

soning about its voting decision by dynamically weighing its individual profile against the complex sociopolitical realities presented to it.

To evaluate the efficacy and fidelity of our framework, we apply it to one of the most complex and high-stakes social phenomena: the 2024 U.S. Presidential Election. We select the 2024 election as our primary testbed, as it provides a robust validation of the agents’ generative reasoning capabilities. While historical events like the 2020 election are an alternative, they present a significant risk of data leakage, where LLMs might simply recall known outcomes rather than reasoning dynamically. Our analysis focuses on seven key swing states: Arizona, Georgia, Michigan, Nevada, North Carolina, Pennsylvania, and Wisconsin. The simulation results, visualized in Figure 2, demonstrate a striking alignment with the actual 2024 election outcomes. Our framework successfully replicates the macro-level result that Donald Trump would win six of the seven pivotal states, validating the fidelity of our “virtual society”.

Beyond macro-level replication, the primary contribution of ElectionFit lies in its utility as an interpretable and analyzable computational laboratory. The framework moves beyond “black box” predictions, allowing researchers to probe the rationale behind the simulation. As demonstrated in our interactive experiments (Section 4.3), researchers can directly “interview” the agents to understand the alignment between their profiles, policy information, and voting preferences. Furthermore, the framework serves as a powerful tool for analyzing the simulation’s dynamics. Our extensive ablation studies (4.4) and sensitivity analyses (Section 4.5) are not merely robustness checks; they are case studies demonstrating how this “lab” can be used to understand the model’s inner workings and probe critical issues like model bias, prediction stability, and sensitivity to context.

This paper provides a comprehensive description of the ElectionFit framework, its methodology, and its application as a computational testbed. We argue that this agent-based approach represents a significant step towards enabling new forms of computational social science, fostering a deeper, more nuanced understanding of complex human behavior.

2 RELATED WORK

Our research is positioned at the intersection of computational social science (CSS), agent-based modeling (ABM), and the emerging field of LLM agents. We structure our review by first contrasting ElectionFit with traditional simulation methods, then situating it within the new paradigm of LLM-based social simulation, and finally motivating our core focus on reliability.

2.1 FROM HEURISTIC-RULES TO REASONING-BASED SIMULATION

Traditional ABM in computational social science, while valued for its “white-box” interpretability [Qiu & Phang \(2020a\)](#), is limited by behavioral simplicity. Agents are often “automatons” that execute oversimplified, predefined rules rather than performing complex reasoning. This includes fixed rules for opinion dynamics [Acemoglu & Ozdaglar \(2011\)](#); [Sood & Kulkarni \(2022\)](#) or non-cognitive “adaptive rationality” for voting behavior [Bendor et al. \(2003\)](#); [Gao et al. \(2022\)](#). ElectionFit presents a direct contrast, accepting agent-level complexity to achieve higher behavioral fidelity.

The advent of LLMs introduces a new “reasoning-based” paradigm. Horton [Horton \(2023\)](#) proposed “*Homo silicus*”, arguing LLMs could serve as simulated economic agents. This concept is famously realized in “Generative Agents” (Stanford small town) [Park et al. \(2023\)](#) and later extended to “1,000 People Simulations” that replicated human survey responses with high fidelity [Park & et al. \(2024\)](#), validating the use of LLM agents for social science research.

2.2 LLM AGENTS FOR ELECTION SIMULATION AND RELIABILITY

This new paradigm is rapidly applied to the 2024 U.S. election testbed by several teams. Like ElectionFit, some used persona-based micro-simulation to model individual voters [Yu et al. \(2024a\)](#); [Jiang et al. \(2024\)](#), while others used distribution-based macro-simulation [Bradshaw et al. \(2024\)](#). Unlike work focused primarily on predictive accuracy [Jiang et al. \(2024\)](#), ElectionFit is positioned as an **interpretable computational laboratory**. Our focus is not just on what agents decide, but why.

This focus on interpretability and reliability is critical because the LLM agent is a known-flawed “instrument”. A large body of literature demonstrates that these agents suffer from pervasive political and demographic biases [Sim et al. \(2024\)](#); [Motoki et al. \(2025\)](#); [Sachan et al. \(2025\)](#), are highly sensitive to semantically-equivalent prompt changes [Wang et al. \(2024\)](#), and exhibit non-deterministic instability even at zero temperature [Ni et al. \(2024\)](#). Most alarmingly, interacting with these biased agents has been shown to actually change real voters’ opinions [Sim et al. \(2024\)](#), proving that auditing these flaws is not just a technical exercise but a social necessity.

Research Gap and Our Contribution. The literature reveals a critical gap between the application of LLM agents in high-stakes domains (like election simulation) and the auditing of their fundamental flaws. Research applying these agents often fails to systematically analyze the impact of their known biases and instabilities. ElectionFit explicitly bridges this gap. We are among the first to apply a rigorous methodology for auditing bias, instability, and systemic failure (e.g., “reasoning-action mismatch” [Yin & et al. \(2025\)](#)) directly to the 2024 U.S. election testbed. Our “computational laboratory” is thus proposed not just for naive prediction, but as a framework to ask a more important scientific question: *are these agents valid tools for social science?*

3 METHODOLOGY: BUILDING THE ELECTIONFIT LABORATORY

The ElectionFit framework is designed as a computational laboratory to simulate complex political behavior. Its methodology integrates high-fidelity demographic profiling with context-aware, generative reasoning. The construction follows three steps.

3.1 AGENT DEMOGRAPHIC MODELING

To create a high-fidelity “virtual society”, each LLM agent is instantiated with a detailed demographic profile. These profiles mirror the real-world U.S. electorate, drawing statistical distributions

162 from two primary sources: the 2023 American Community Survey (ACS) and the 2020 U.S. Reli-
163 gion Census (ASARB).

164 Based on established political science research [Gao et al. \(2022\)](#); [Argyle et al. \(2023\)](#) and our own
165 ablation studies (see Section 4.4), we identified eight critical attributes: *State, Race, Sex, Age, Occu-*
166 *vation, Industry, Educational Attainment, and Religion*. Where dependencies exist (e.g., state, race,
167 and sex), we use joint distributions; otherwise, independent distributions are applied. We generate
168 a population of 1,000 agents per state via random sampling, a number validated by our stability
169 analysis (Appendix 4.4.1). A detailed breakdown of all demographic categories is available in Ap-
170 pendix B.

172 3.2 CONTEXTUAL INFORMATION FOR REASONING

174 A demographic profile alone is insufficient. To enable nuanced reasoning, agents must react to the
175 contemporary political environment. We provide agents with contextual information on the three
176 most salient issues of the 2024 election, as identified by multiple voter surveys (Pew³, Gallup⁴, and
177 NBC⁵): economic policy, immigration, and abortion rights.

178 The candidates’ stances on these issues are summarized from official party platforms and verified
179 public statements, ensuring a balanced and accurate representation. This information (detailed in
180 Appendix C) bridges the gap between static demographics and dynamic decision-making, allowing
181 agents to evaluate how candidate policies align with their profiles and sociopolitical realities.

183 3.3 PROBABILISTIC VOTING BEHAVIOR

184 To capture the uncertainty and nuance of voter preference, agents do not make a simple binary
185 choice. Instead, they are prompted to output their decision as a probability distribution across three
186 outcomes. This format has been shown to provide a more accurate and stable alignment for simula-
187 tion tasks [Meister et al. \(2024\)](#).

189 **Agent Instruction.** Agents are instructed to act as a U.S. citizen with their given profile and to
190 base their decision only on the information provided. The core prompt is:

192 In the 2024 presidential election, Donald Trump is the Republican candidate, and
193 Kamala Harris is the Democratic candidate. ONLY based on the profile and the
194 information provided above, predict the probability that you would vote for each
195 candidate as well as the probability that you would “vote for another candidate or
196 not vote at all”. Make sure the probabilities add up to 1.

198 **Response Format.** Agents output their predictions in a structured JSON format, which we then
199 aggregate to produce macro-level results:

```
200 {
201   "Donald Trump": probability of voting for Donald Trump,
202   "Kamala Harris": probability of voting for Kamala Harris
203   ,
204   "vote for another candidate or not vote at all":
205     probability,
206 }
207
```

209 4 EXPERIMENTS: VALIDATING THE COMPUTATIONAL LABORATORY

211 We design a series of experiments to validate the ElectionFit laboratory. Our goal is not simply to
212 measure predictive accuracy, but to (1) verify the macro-level fidelity of the simulation, (2) demon-

214 ³<https://www.pewresearch.org/politics/2024/09/09/issues-and-the-2024-election/>

215 ⁴<https://news.gallup.com/poll/trends.aspx>

⁵<https://www.nbcnews.com/politics/2024-presidential-election>

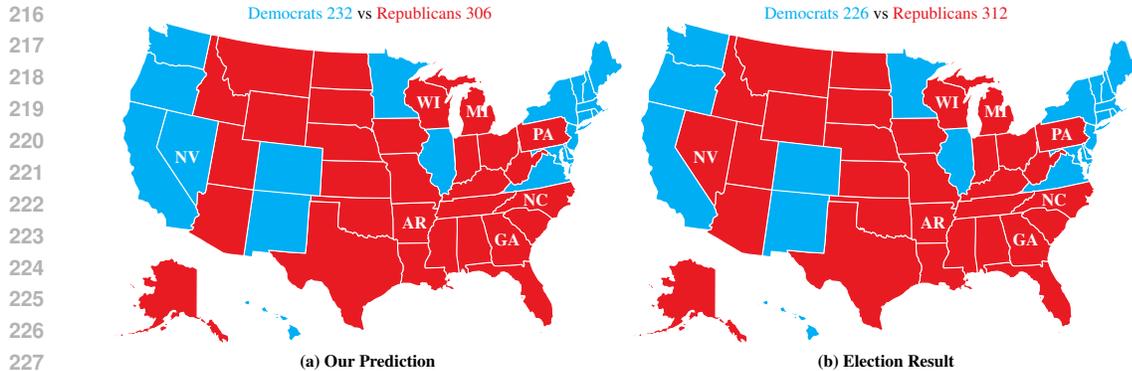


Figure 2: Comparing the predicted outcomes from November 4th and the actual election results from November 5th. On the maps of the United States, red represents states won by Republicans, while blue indicates states won by Democrats. The only difference between these two maps is in the state of Nevada. In our predictions, Democrats win Nevada with a margin of only 0.17%, while the actual outcome is that Republicans win Nevada, which is such a narrow margin that our forecast is almost spot on.

strate its micro-level interpretability, and (3) use the framework to probe the reliability and limitations of LLM agents in this high-stakes task.

4.1 EXPERIMENTAL SETUP

Testbed. We focus on the 2024 U.S. Presidential Election, a contemporary event that prevents data leakage from LLM training data, thereby testing the agents’ generative reasoning rather than recall. We simulate 1,000 agents in each of the seven key swing states: Arizona, Georgia, Michigan, Nevada, North Carolina, Pennsylvania, and Wisconsin.

Models. We use Qwen-Max-2024-04-28 as the primary model for our main replication and analysis. This choice is made to mitigate potential U.S.-centric political biases. For our reliability analysis, we also test GPT-4o [OpenAI \(2024\)](#), Claude-3.5-sonnet [Anthropic \(2024\)](#), Gemini-1.5-Pro [Google \(2024\)](#), and other models (full list in Appendix A). All experiments use a temperature of 0 for stability, except the main result (0.7).

4.2 EXPERIMENT 1: MACRO-LEVEL FIDELITY VALIDATION

The first test of a computational laboratory is its ability to replicate known real-world phenomena. Figure 2 compares our simulation results with the actual election outcome.

The simulation results demonstrate a striking alignment with the actual 2024 election outcomes. ElectionFit correctly replicates the macro-level result that Donald Trump would win six of the seven pivotal swing states. The only discrepancy is Nevada, which our simulation predicts for Kamala Harris by a razor-thin margin of 0.17%, a state that is ultimately decided by an extremely narrow margin in reality. This high fidelity at the macro-level validates the framework’s ability to construct a “virtual society” that captures the aggregate political dynamics.

4.3 EXPERIMENT 2: MICRO-LEVEL INTERPRETABILITY

A key advantage of our laboratory over statistical “black boxes” is its interpretability. We can probe the reasoning behind the aggregate results.

Aggregated Rationale. Agents are also prompted to provide a brief “Reason” for their probabilistic choice. We aggregate these reasons from the pivotal swing state of Pennsylvania into a word cloud (Figure 3). The prominence of key issues like “economic”, “abortion”, “inflation”, and “immigration” confirms that the agents are not just using their demographic profile, but are actively weighing the provided contextual policy information.

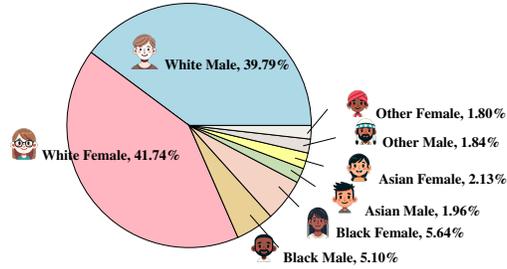


Figure 5: Proportions of different race and sex groups in Pennsylvania. This pie chart highlights that the 300-agent simulation is sufficient for every population group. Since other dimensions are modeled independently and each group’s proportion within any given dimension is at least 1%, 300 agents are sufficient to cover all groups effectively.

4.4 EXPERIMENT 3: ABLATION STUDIES (VALIDATING LAB CONSTRUCTION)

A robust computational laboratory must not only be verifiable against external events, but also internally consistent. We therefore use the framework to test its own construction. This involves a series of ablation studies to validate the necessity of key methodological choices made during the lab’s design. We focus on two critical parameters: the number of agents required for a stable simulation (Section 4.4.1) and the relative importance of specific demographic attributes (Section 4.4.2).

4.4.1 AGENT POPULATION SIZE

To establish an optimal number of agents, we conduct a sensitivity analysis using Pennsylvania as a test case, evaluating populations from 10 to 2,000 agents (10 trials each). As shown in Figure 4, prediction variance stabilizes at 300 agents, after which further increases yield minimal fluctuations. For our final simulation, we use 1,000 agents per state to ensure high precision. Figure 5 shows the demographic breakdown of a 300-agent sample, confirming all demographic groups are represented.

4.4.2 PROFILE DIMENSION SELECTION

We also validate the inclusion of our eight demographic attributes. We run ablation experiments using the 2020 election as a ground truth (noting the data leakage risk, as discussed in Section 1). As summarized in Table 1 and Table 2, we find that a 6-dimension profile (excluding education and religion) fails to predict the correct winner in Wisconsin. Adding the **education** dimension corrects

Table 1: Ablation Study on the Education Dimension. The experiments are conducted under a no-context-info setting using Qwen-Max-2024-04-28, with the 2020 U.S. presidential election results as the ground truth. Predictions with six dimensions (state, race, sex, age, occupation, and industry) are contrasted against those with the addition of the education dimension. Including education significantly improves prediction accuracy, particularly in Wisconsin, where the six-dimension setting previously failed to replicate the actual election outcome. Bold means the corresponding party wins.

Condition	California		Georgia		Michigan	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Groundtruth	35.08	64.92	49.85	50.15	48.57	51.43
6 Dimensions	29.50	70.50	45.36	54.64	47.57	52.43
+ Education	28.58	71.42	41.92	58.08	46.65	53.35
Condition	Pennsylvania		Texas		Wisconsin	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Groundtruth	49.39	50.61	52.85	47.15	49.70	50.30
6 Dimensions	46.52	53.48	56.41	43.59	50.29	49.71
+ Education	45.99	54.01	53.77	46.23	48.35	51.65

Table 2: Ablation Study on the Religion Dimension. This experiment examines the effect of adding the religion dimension under a no-context-info setting, using Qwen-Max-2024-04-28. Simulations with seven dimensions (excluding religion) are compared to those with eight dimensions (including religion). Incorporating religion improves alignment with polling data from September 17, 2024. From another perspective, including religion helps reduce the model’s Democratic bias, promoting more balanced political alignment. Bold means the corresponding party wins.

Condition	California		Georgia		Michigan	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Groundtruth	35.3	59.5	47.2	47.6	47.8	46.0
7 Dimensions	20.7	70.6	35.0	56.8	39.7	50.1
+ Religion	21.8	69.3	37.4	53.8	41.0	48.6

Condition	Pennsylvania		Texas		Wisconsin	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Groundtruth	47.6	46.9	44.3	50.6	48.8	45.9
7 Dimensions	39.0	50.7	41.7	48.6	39.8	49.7
+ Religion	40.8	49.2	43.3	46.8	41.8	47.3

this result. Similarly, adding the **religion** dimension improves alignment with 2024 polling data and helps mitigate an inherent pro-Democratic bias in the baseline model. These results confirm that both dimensions are crucial for simulation fidelity.

4.5 EXPERIMENT 4: SENSITIVITY ANALYSIS (AUDITING THE INSTRUMENT)

One of the most critical uses of the laboratory is to audit the reliability of its own components (the LLM agents). We conduct extensive sensitivity analyses to probe for bias and instability.

Political Bias Evaluation. We test multiple LLMs under four different context scenarios in the state of Georgia: (1) No Context, (2) Neutral Context (our default), (3) Negative Bias Toward Harris, and (4) Asymmetric Positive Framing for Trump. The results (Table 3) reveal a severe “instrument problem”. Under the “No Context” condition, most models, including the newer Qwen-Max-09-19, show a strong default pro-Democratic bias. More alarmingly, this bias persists even under biased prompting. In the “Asymmetric Positive Framing” scenario, designed to strongly favor Trump, Qwen-Max-09-19 still predicts a Democratic victory, revealing its unsuitability for this task. This analysis demonstrates that our “fortuitous” choice of the older Qwen-Max-04-28 (which showed more neutrality) is critical to our success. (Full multi-model results are in Appendix F).

Context and Positional Instability. We find that agent decisions are highly unstable and sensitive to semantically irrelevant changes. First, we test eight different “Context Variants”—all semantically neutral but with minor rephrasing (see Appendix G). As shown in Figure 6, these minor

Table 3: Sensitivity Analysis on Political Bias Across Models. This table presents prediction results in Georgia under 4 different context information settings using multiple LLMs. The results reveal varying degrees of political bias, particularly a strong Democratic bias in Qwen-Max-09-19, even under prompts designed to favor Trump. In contrast, Qwen-Max-04-28 and DeepSeek-Chat provided more realistic predictions, while Gemini-1.5-Pro demonstrated instability under biased prompts. Bold means the corresponding party wins.

Model	No Context		Neutral Context		Negative Bias Towards Harris		Asymmetric Positive Framing	
	Republican	Democrat	Republican	Democrat	Republican	Democrat	Republican	Democrat
Qwen-Max-09-19	32.10	47.08	35.28	51.37	36.95	46.47	38.02	45.87
Qwen-Max-04-28	43.92	46.58	49.55	39.52	58.40	29.37	55.32	31.28
GPT-4o	38.26	34.49	34.82	42.60	36.35	40.03	38.50	37.77
Claude-3-5-sonnet	37.48	43.72	42.65	44.45	47.68	36.35	44.58	40.15
Gemini-1.5-Pro	37.98	52.50	38.38	43.08	45.62	35.80	38.65	42.47
DeepSeek-V2.5	38.55	31.08	42.50	40.93	41.60	34.37	42.85	32.48

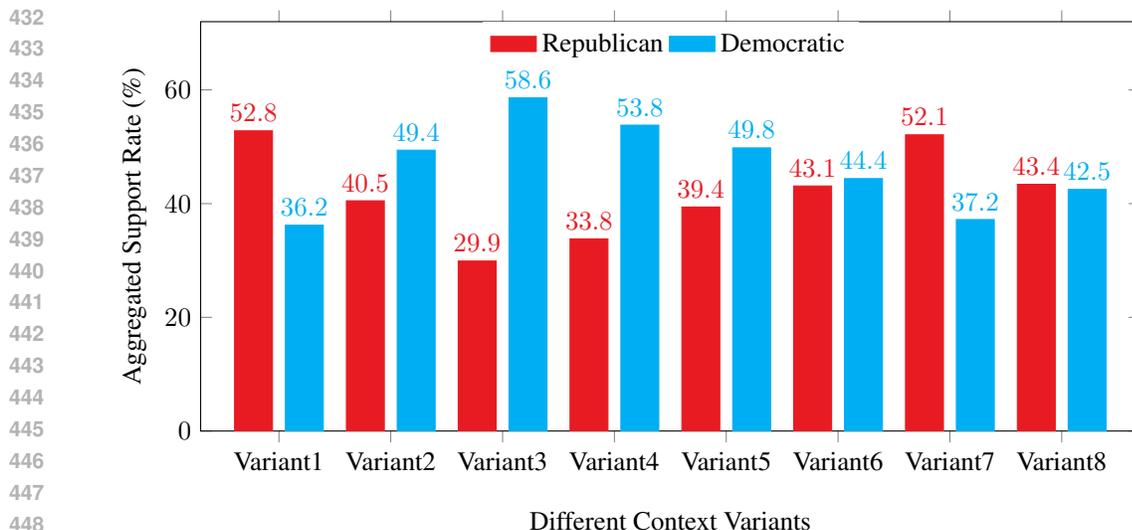


Figure 6: Prediction Variability Across Context Variants in Pennsylvania. Aggregated support rates for Republican and Democratic candidates predicted by Qwen-Max-04-28 under eight context prompts with varying levels of content richness. The results reveal significant fluctuations in support rates, illustrating the sensitivity of LLM outputs to variations in context detail.

changes cause wild fluctuations in predicted support, with Democratic support swinging from 29.9% to 58.6% based on trivial prompt edits. Second, we simply swap the order of the candidates in the JSON response format. This positional change alone causes “Swing Agents” to completely flip their vote, as detailed in Appendix H.

These findings demonstrate that while our framework can replicate reality, the LLM agents themselves are flawed, non-deterministic instruments. Their outputs are highly sensitive to superficial prompt changes, posing a significant challenge for reproducibility and reliability.

5 DISCUSSION AND CONCLUSION

In this paper, we propose ElectionFit, a “computational laboratory” framework using LLM agents, and demonstrate its ability to replicate the complex, macro-level outcome of the 2024 U.S. Election with high fidelity. More importantly, we showcase its utility as an interpretable research tool, allowing for micro-level “interviews” and analyses of agent rationale.

However, our deepest contribution comes from using this laboratory to probe the reliability of its own components. Our sensitivity analyses reveal a “paradox of success”. The high fidelity of our main result (Section 4.2) is, as our further experiments show, “fortuitous”. It is contingent on a model choice (Qwen-Max-04-28) that happens to be more neutral than its peers and even its own successor (Section 4.5).

Our analysis proves that most LLMs, when unmoored by context, exhibit strong political biases. Furthermore, we demonstrate that these agents are fundamentally unstable instruments. Their reasoning and final decisions are extraordinarily sensitive to semantically irrelevant changes in prompt phrasing (Figure 6) and even the physical ordering of options in the response format (Appendix H).

This “instability” and “bias” are not minor issues; they are fundamental flaws in the current generation of LLMs that challenge their validity for rigorous, reproducible computational social science. The fact that these biased instruments are also powerful “persuasion engines” capable of changing real human opinions [Sim et al. \(2024\)](#) makes this a critical area of research. The future of AI in social science depends not on simply building models that look accurate, but on developing a rigorous science of “instrument auditing”. We must understand and mitigate the biases and instabilities of our AI agents before we can truly trust them as a laboratory for understanding ourselves. ElectionFit provides a concrete framework for conducting this essential audit.

REFERENCES

- 486
487
488 Daron Acemoglu and Asuman Ozdaglar. A survey on nonstrategic models of opinion dynamics.
489 *Foundations and Trends® in Networking*, 5(4):305–423, 2011.
- 490
491 DeepSeek AI. Deepseek v2.5 - hugging face. [https://huggingface.co/deepseek-ai/](https://huggingface.co/deepseek-ai/DeepSeek-V2.5)
492 [DeepSeek-V2.5](https://huggingface.co/deepseek-ai/DeepSeek-V2.5), 2024a. Accessed: 2024-11-17.
- 493
494 Meta AI. Llama 3.1 model card, 2024b. URL [https://github.com/meta-llama/](https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/MODEL_CARD.md)
495 [llama-models/blob/main/models/llama3_1/MODEL_CARD.md](https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/MODEL_CARD.md). Accessed: 2024-
496 11-18.
- 497
498 Meta AI. Llama 3.2 model card, 2024c. URL [https://github.com/meta-llama/](https://github.com/meta-llama/llama-models/blob/main/models/llama3_2/MODEL_CARD.md)
499 [llama-models/blob/main/models/llama3_2/MODEL_CARD.md](https://github.com/meta-llama/llama-models/blob/main/models/llama3_2/MODEL_CARD.md). Accessed: 2024-
500 11-17.
- 501
502 Anthropic. Introducing computer use, a new claude 3.5 sonnet, and claude 3.5 haiku, 2024. URL
503 <https://www.anthropic.com/news/3-5-models-and-computer-use>. Ac-
504 cessed: 2024-11-17.
- 505
506 Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David
507 Wingate. Out of one, many: Using language models to simulate human samples. *Political Anal-*
508 *ysis*, 31(3):337–351, 2023.
- 509
510 Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge,
511 Yu Han, Fei Huang, et al. Qwen technical report. *arXiv preprint arXiv:2309.16609*, 2023.
- 512
513 Jonathan Bendor, Daniel Diermeier, and Michael M Ting. A behavioral model of turnout. *American*
514 *Political Science Review*, 97(2):261–280, 2003.
- 515
516 Nicholas Bradshaw, Tanvi Bhasin, and Sriram Narayanan. Llm generated distribution-based predic-
517 tion of us electoral results, part i. *arXiv preprint arXiv:2411.03486*, 2024.
- 518
519 Kellyton Dos Santos Brito, Rogério Luiz Cardoso Silva Filho, and Paulo Jorge Leitao Adeodato.
520 A systematic review of predicting elections based on social media data: research challenges and
521 future directions. *IEEE Transactions on Computational Social Systems*, 8(4):819–843, 2021.
- 522
523 Priyavrat Chauhan, Nonita Sharma, and Geeta Sikka. The emergence of social media data and senti-
524 ment analysis in election prediction. *Journal of Ambient Intelligence and Humanized Computing*,
525 12:2601–2627, 2021.
- 526
527 Mauro Coletto, Claudio Lucchese, Salvatore Orlando, Raffaele Perego, et al. Electoral predic-
528 tions with twitter: a machine-learning approach. In *CEUR Workshop Proceedings*, volume 1404.
529 CEUR-WS, 2015.
- 530
531 Scott De Marchi and Scott E Page. Agent-based models. *Annual Review of political science*, 17(1):
532 1–20, 2014.
- 533
534 James H Fowler and Oleg Smirnov. Dynamic parties and social turnout: An agent-based model.
535 *American Journal of Sociology*, 110(4):1070–1094, 2005.
- 536
537 Ming Gao, Zhongyuan Wang, Kai Wang, Chenhui Liu, and Shiping Tang. Forecasting elections
538 with agent-based modeling: Two live experiments. *Plos one*, 17(6):e0270194, 2022.
- 539
540 Google. Updated production-ready gemini models, reduced 1.5 pro pricing, increased rate
541 limits, and more, 2024. URL [https://developers.googleblog.com/en/](https://developers.googleblog.com/en/updated-gemini-models-reduced-15-pro-pricing-increased-rate-limits-and-more/)
542 [updated-gemini-models-reduced-15-pro-pricing-increased-rate-limits-and-more/](https://developers.googleblog.com/en/updated-gemini-models-reduced-15-pro-pricing-increased-rate-limits-and-more/).
543 Accessed: 2024-11-17.
- 544
545 Igor Grossmann, Matthew Feinberg, Dawn C Parker, Nicholas A Christakis, Philip E Tetlock, and
546 William A Cunningham. Ai and the transformation of social science research. *Science*, 380
547 (6650):1108–1109, 2023.

- 540 Pratik Gujral, Kshitij Awaldhi, Navya Jain, Bhavuk Bhandula, and Abhijnan Chakraborty. Can
541 llms help predict elections?(counter) evidence from the world’s largest democracy. *arXiv preprint*
542 *arXiv:2405.07828*, 2024.
- 543
- 544 John J Horton. Large language models as simulated economic agents: What can we learn from
545 homo silicus? *arXiv preprint arXiv:2301.07543*, 2023.
- 546 Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot,
547 Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al.
548 Mistral 7b. *arXiv preprint arXiv:2310.06825*, 2023.
- 549
- 550 Shapeng Jiang, Lijia Wei, and Chen Zhang. Donald trumps in the virtual polls: Simulating and pre-
551 dicting public opinions in surveys using large language models. *arXiv preprint arXiv:2411.01582*,
552 2024.
- 553 Nicole Meister, Carlos Guestrin, and Tatsunori Hashimoto. Benchmarking distributional alignment
554 of large language models. *arXiv preprint arXiv:2411.05403*, 2024.
- 555
- 556 Fabio Motoki, Victor Riscado, and Fellipe Francy. ”amazing, they all lean left” - analyzing the
557 political orientations of large language models. *arXiv preprint arXiv:2507.08027*, 2025. Year
558 inferred from arXiv ID, likely 2024.
- 559 Kai Ni, Chuan Wang, and Pu Li. Llm stability: A detailed analysis with some surprises. *arXiv*
560 *preprint arXiv:2408.04667*, 2024.
- 561
- 562 OpenAI. Hello gpt-4o, 2024. URL <https://openai.com/index/hello-gpt-4o/>. Ac-
563 cessed: 2024-11-17.
- 564
- 565 Joon Sung Park and et al. Simulating human behavior with ai agents. *Stanford HAI*, 2024. Accessed:
566 Nov 7, 2025. Refers to ”Generative Agent Simulations of 1,000 People”.
- 567 Joon Sung Park, Joseph C O’Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and
568 Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. *arXiv preprint*
569 *arXiv:2304.03442*, 2023.
- 570
- 571 Lin Qiu and Riyang Phang. Agent-based modeling in political decision making. *Oxford Research*
572 *Encyclopedia of Politics*, 2020a. Accessed: Nov 7, 2025.
- 573 Lin Qiu and Riyang Phang. Agent-based modeling in political decision making. In *Oxford research*
574 *encyclopedia of politics*. 2020b.
- 575
- 576 Mridul Sachan, David R Shu, and Nathan Kallus. Measuring gender and racial biases in large
577 language models: Intersectional evidence from automated resume evaluation. *PNAS Nexus*, 4(3):
578 pgaf089, 2025. Year inferred from PDF, likely 2024.
- 579
- 580 Y-S Sim, C Mello, J Kim, H Song, Y Kim, J Kim, S Iyengar, and S Oh. Ai-driven voter mimicry
581 can shape election outcomes. *arXiv preprint arXiv:2410.24190*, 2024.
- 582 Gaurav Sood and Veerabaahu Kulkarni. Computational agent-based models in opinion dynamics:
583 A survey on social simulations and empirical studies. *SN Computer Science*, 3(4):307, 2022.
- 584
- 585 Yuxin Wang, Bai Yang, Xinyi Zhao, Yilun Zhou, Zhaoyu Lin, Zhang-Ren Zhou, Yantao Sun,
586 Zhaopeng Wu, and Wei Chen. Prosa: Assessing and understanding the prompt sensitivity of
587 large language models. In *Findings of the Association for Computational Linguistics: EMNLP*
588 *2024*, pp. 1650–1667, 2024.
- 589 Fan Yin and et al. Why do multi-agent llm systems fail? *arXiv preprint arXiv:2503.13657*, 2025.
590 Year inferred from arXiv ID, likely 2024.
- 591
- 592 Chenxiao Yu, Zhaotian Weng, Zheng Li, Xiyang Hu, and Yue Zhao. A large-scale empirical
593 study of llm-based agents for simulating the 2016, 2020, and 2024 us elections. *arXiv preprint*
arXiv:2412.15291, 2024a. This is inferred from the PDF’s [28].

594
595 Chenxiao Yu, Zhaotian Weng, Zheng Li, Xiyang Hu, and Yue Zhao. Will trump win in 2024?
596 predicting the us presidential election via multi-step reasoning with large language models. *arXiv*
597 *preprint arXiv:2411.03321*, 2024b.

598 Xinnong Zhang, Jiayu Lin, Libo Sun, Weihong Qi, Yihang Yang, Yue Chen, Hanjia Lyu, Xinyi Mou,
599 Siming Chen, Jiebo Luo, et al. Electionsim: Massive population election simulation powered by
600 large language model driven agents. *arXiv preprint arXiv:2410.20746*, 2024.

601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647

A FULL MODEL LIST FOR EXPERIMENTS

The models used in our experiments (Sections 4 and 4.5) are:

- Qwen-Max [Bai et al. \(2023\)](#): Qwen-Max-2024-04-28 (Primary Model)
- Qwen-Max [Bai et al. \(2023\)](#): Qwen-Max-2024-09-19 (For bias test)
- GPT-4o [OpenAI \(2024\)](#): GPT-4o-2024-08-06
- Claude-3-5-sonnet [Anthropic \(2024\)](#): Claude-3-5-sonnet-2024-10-22
- Gemini-1.5-Pro [Google \(2024\)](#): Gemini-1.5-Pro-002
- Mistral [Jiang et al. \(2023\)](#): Mistral-Nemo-Instruct-2407
- DeepSeek AI (2024a): DeepSeek-V2.5
- Llama3.2 AI (2024c): Llama3.2-3B-Instruct
- Llama3.1 AI (2024b): Llama3.1-8B-Instruct

B DETAILED DEMOGRAPHIC PROFILE OPTIONS

Table 4: Dimensions and options for agent profiles in each state. None* means the dimension *Religion* will not be mentioned in the profile of the agent.

Dimension	Options
Race	White
	Black or African American
	Asian
	Non White, Asian or Black
Sex	Male
	Female
Age	18 to 24 years
	25 to 34 years
	35 to 44 years
	45 to 54 years
	55 to 64 years
	65 to 74 years
75 years and over	Management, business, science, and arts
	Service
	Sales and office
	Natural resources, construction, and maintenance
	Production, transportation, and material moving
Industry	Agriculture, forestry, fishing and hunting, and mining
	Construction
	Manufacturing
	Wholesale trade
	Retail trade
	Transportation and warehousing, and utilities
	Information
	Finance and insurance, and real estate and rental and leasing
	Professional, scientific, and management, and administrative and waste management services
	Educational services, and health care and social assistance
	Arts, entertainment, and recreation, and accommodation and food services
	Other services (except public administration)
Public administration	

Dimension	Options
Educational Attainment	Less than high school diploma
	High school graduate (includes equivalency)
	Some college or associate's degree
	Bachelor's degree
	Graduate or professional degree
Religion	Catholic
	Protestant
	Islam
	Judaism
	Non Catholic, Protestant, Islam or Judaism
	None*

C FULL PROMPT AND CONTEXT INFORMATION

The Prompt for 2024 US Presidential Election

Profile

- State: California
- Race: White
- Sex: Male
- Age: 45 to 54 years
- Occupation: Not in labor force
- Industry: Not in labor force
- Educational Attainment: Some college or associate's degree
- Religion: Catholic

Historical events

In the 2024 presidential election, the candidates' stances on several key issues are especially significant. Below are their respective views on these issues.

On the topic of economic

Trump's stance: Among Trump's key campaign promises is to 'end inflation and make America affordable again'. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.

Harris's Stance: Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. Voters who believe the economic performance has strongly recovered over the past year may find Harris more appealing.

On the topic of immigration

Trump's stance: Trump promised the biggest mass deportations of undocumented migrants in US history. For voters concerned about security and economic issues linked to illegal immigration, Trump appears particularly appealing.

Harris's Stance: On the campaign trail, she has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction. Voters who believe the current policies are strong enough may find Harris more appealing.

On the topic of abortion

Trump's stance: During the presidential debate he said he would not sign a national abortion ban into law. This may be more appealing to conservatives.

Harris's stance: Harris has long supported abortion rights. This may be more appealing to liberals.

Requirement

- ONLY based on the profile and the information provided above, predict the probability that you would vote for each candidate as well as the probability that you would "vote for another candidate or not vote at all". Make sure the probabilities add up to 1.

Response Format

Please output only the following JSON format, with no additional text.

```

{
  "Donald Trump": probability of voting for Donald Trump,
  "Kamala Harris": probability of voting for Kamala Harris,
  "vote for another candidate or not vote at all": probability

```

756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809

```
}  
  \ \ \
```

D FULL TRANSCRIPTS: INTERACTIVE EXPERIMENTS

As mentioned in Section 4.3, we "interviewed" agents. Below are the full, unedited responses.

To ensure that agents provide well-reasoned answers without emphasizing their identity as AI models, and to prevent the response format specified in the voting requirement from influencing their answers to the survey questions, we include the following system prompt:

Assume the role of a U.S. citizen, with your identity details provided in the Profile. From the perspective of this U.S. citizen, answer the following questions in the first person. Follow the specified response format for each question; if no format is provided, respond directly to the question.

Agent 1 of Interactive Experiment

Agent Profile

- State: Pennsylvania
- Race: White
- Sex: Male
- Age: 45 to 54 years
- Occupation: Production, transportation, and material moving
- Industry: Construction
- Educational Attainment: High school graduate (includes equivalency)
- Religion: Catholic

Voting Behavior

```
\ \ \
{
  "Donald Trump": 0.65,
  "Kamala Harris": 0.30,
  "vote for another candidate or not vote at all": 0.05,
}
```

Question1:

How does your choice relate to your background information?

Response1:

As a white male in my 40s working in construction with a high school education and identifying as Catholic, my viewpoints tend to align more conservatively. Trump's focus on ending inflation and his stance on immigration resonate with concerns about the economy and security, which are typically important to those in my demographic and industry. Additionally, while Trump's position on abortion might not be entirely in line with Catholic views, his statement that he wouldn't sign a national abortion ban could still be seen as a moderate stance compared to more aggressive anti-abortion policies.

On the other hand, Harris's economic plans, such as mortgage assistance and tax credits, might not directly benefit me as much given my occupation and educational attainment. Her immigration policy, which includes support for border security measures but doesn't emphasize deportations, might not appeal to someone with a more conservative stance on the issue. Furthermore, her pro-choice stance contrasts with traditional Catholic beliefs.

The likelihood of voting for another candidate or not voting at all reflects a level of dissatisfaction or uncertainty with both major candidates, which could stem from feeling that neither fully represents my interests or values.

Question2:

If you had to choose, would you prefer a smaller government providing fewer services or a bigger government providing more services?

Response2:

810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863

Based on my profile as a white male working in the construction industry with a high school education and a Catholic background, I would likely prefer a smaller government providing fewer services. This preference aligns with traditionally conservative values that emphasize limited government intervention, lower taxes, and personal responsibility. In the construction and manufacturing sectors, there's often a sentiment that less regulation allows for more business freedom and growth. Moreover, fiscal conservatism, which advocates for reduced government spending, tends to resonate with individuals in my demographic and occupational sector.

Agent 2 of Interactive Experiment

Agent Profile

- State: Pennsylvania
- Race: Black
- Sex: Female
- Age: 25 to 34 years
- Occupation: Sales and office
- Industry: Finance and insurance, and real estate and rental and leasing
- Educational Attainment: Bachelor's degree

Voting Behavior

```

{
  "Donald Trump": 0.25,
  "Kamala Harris": 0.65,
  "vote for another candidate or not vote at all": 0.10,
}

```

Question1:

Which piece of information from historical events most influenced your judgment?

Response1:

As a Black female with a Bachelor's degree working in the finance and insurance industry, the most influential piece of information from the historical events was Kamala Harris's stance on abortion rights. This aligns more closely with liberal views, which are typically supported by individuals with similar demographic characteristics and educational background. Her economic policies, such as mortgage assistance and tax credits, could also positively impact someone in my financial sector, contributing to the decision-making process.

Question2:

Thinking about the increased trade of goods and services between the U.S. and other nations in recent decades, would you say the U.S. has gained more than it has lost or lost more than it has gained?

Response2:

From the perspective of a Black female with a Bachelor's degree working in finance and insurance in Pennsylvania, I would lean towards saying that the U.S. has gained more than it has lost from increased trade with other nations. This viewpoint takes into account the benefits of globalization for the finance and insurance industry, such as access to larger markets, diversification of investments, and increased competitiveness. However, it's also recognized that there are sectors and communities that might have experienced negative effects like job displacement. Overall, considering my occupation and industry, the gains in terms of market expansion and service exports likely outweigh the losses.

Agent 3 of Interactive Experiment

Agent Profile

- State: Pennsylvania
- Race: Non White, Asian or Black
- Sex: Male
- Age: 25 to 34 years
- Occupation: Service
- Industry: Manufacturing

864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917

- Educational Attainment: High school graduate (includes equivalency)
- Religion: Protestant

Voting Behavior

```

{
  "Donald Trump": 0.35,
  "Kamala Harris": 0.5,
  "vote for another candidate or not vote at all": 0.15,
}

```

Question1:

Did the differences between Trump and Harris on a particular issue influence your choice? If so, how?

Response1:

As a service worker in the manufacturing industry who identifies as a Protestant and has a high school education, the economic and social issues carry significant weight in my decision-making process. Harris's promise of mortgage assistance, a tax credit for new parents, and addressing inflation through price gouging bans resonates with my financial concerns, which increases the likelihood of voting for her. On the other hand, Trump's focus on ending inflation appeals as well, but his stance on immigration, particularly the promise of mass deportations, might not align with my values or priorities, given my non-white, Asian or Black racial background. Additionally, as someone who is likely to lean towards conservative social values associated with my Protestant faith, Trump's statement on not signing a national abortion ban could be less appealing compared to Harris's pro-abortion rights stance, which might not align directly with my beliefs.

Thus, the combination of economic relief measures proposed by Harris and the social dynamics influenced by both candidates' stances does play a role in shaping my preference, leaning slightly more towards Harris but also considering the option of voting for another candidate or not voting at all due to mixed alignments with the presented policies.

Question2:

Do you think business corporations make too much profit or a fair and reasonable amount of profit?

Response2:

As a service worker in the manufacturing industry, I might hold the view that some business corporations make excessive profits, especially when it comes to industries where costs are passed onto consumers through inflation or when worker wages don't seem to keep pace with rising corporate earnings. There can be a sense that while companies are thriving, the average worker like myself isn't necessarily benefiting proportionally from their success. This could lead me to believe that there should be a balance – corporations should make a fair and reasonable profit while also investing in their employees, communities, and ensuring affordability for consumers. However, I'd also recognize that businesses need profits to innovate, expand, and create jobs, so it's a nuanced perspective that considers both the necessity of profitability and the ethical distribution of wealth.

E CONTEXTS FOR POLITICAL BIAS EXPERIMENTS

These are the full prompts used for the four conditions in our political bias analysis (Table 3).

Neutral Context

Historical events

In the 2024 presidential election, the candidates' stances on several key issues are especially significant. Below are their respective views on these issues.

On the topic of economic

Trump's stance: Inflation has remained at elevated levels over recent years, reaching a 40-year high in 2022, making it a focal point in economic discussions. Among Trump's key campaign promises is to 'end inflation and make America affordable again'. His proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs. Additionally, Trump suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants, increasing deportations, and initiating a federal home-building program. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.

918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971

Harris's Stance: Inflation has decreased rapidly over the past year. Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. The Federal Reserve announced a significant rate cut in September, the first rate cut since March 2020. This half-point move paves the way for lower borrowing costs on everything from mortgages to credit cards and marks a crucial milestone in the central bank's historic fight against inflation. Voters who believe the economic performance has strongly recovered over the past year may find Harris more appealing.

On the topic of immigration

Trump's stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. The number of border crossings hit record levels at the end of 2023. Trump has vowed to seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history. There are also claims it could slow economic growth. For voters concerned about security and economic issues linked to illegal immigration, Trump appears particularly appealing.

Harris's Stance: The Biden-Harris administration has indeed implemented stricter border control policies. Harris's position on the border has become more moderate over time. On the campaign trail, she has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction. She has also said there "should be consequences" for people who cross the US border unlawfully, though she previously supported a more lenient approach. Voters who believe the current policies are strong enough may find Harris more appealing.

On the topic of abortion

Trump's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but abortion legislation varies across states. Trump himself has regularly said states should be free to decide their own laws on abortion. But in recent weeks he has tried to walk a tightrope, taking a more moderate stance while trying not to alienate his conservative base. He criticised Florida's six-week ban, provoking an angry backlash, but also said he would oppose a ballot measure in the state establishing the right to an abortion. During the presidential debate he said he would not sign a national abortion ban into law. This may be more appealing to conservatives.

Harris's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but legislative challenges at the state level continue to emerge. Harris has long supported abortion rights. She was the first vice-president to visit an abortion clinic, and toured the country after the US Supreme Court overturned Roe v Wade in 2022 to address state abortion bans - often framing the issue as a matter of personal freedom. She has made abortion central to her campaign, and continues to advocate for legislation that would safeguard reproductive rights nationwide. "When Congress passes a law to restore reproductive freedoms, as president of the United States, I will sign it into law", she said at a campaign rally in Atlanta, Georgia. This may be more appealing to liberals, especially among women.

Negative Bias Toward Harris Context

Historical events

In the 2024 presidential election, the candidates' stances on several key issues are especially significant. Below are their respective views on these issues.

On the topic of economic

Trump's stance: Inflation has remained at elevated levels over recent years, reaching a 40-year high in 2022, making it a focal point in economic discussions. Among Trump's key campaign promises is to 'end inflation and make America affordable again'. His proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs. Additionally, Trump suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants, increasing deportations, and initiating a federal home-building program. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.

Harris's Stance: Inflation has decreased over the past year, partly due to broader economic trends and actions by the Federal Reserve. Harris has proposed policies such as mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging to help ease financial pressures on households. These measures aim to provide targeted relief and improve affordability. Questions remain as to whether these individual initiatives form part of a cohesive long-term economic strategy. While Harris's proposals seek to address these concerns, voters may weigh them against the administration's overall record in managing economic challenges during this period.

On the topic of immigration

Trump's stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. The number of border crossings hit record levels at the end of 2023. Trump has vowed to

972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025

seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history. There are also claims it could slow economic growth. For voters concerned about security and economic issues linked to illegal immigration, Trump appears particularly appealing.

Harris's Stance: Harris has supported stricter border control measures as part of the Biden-Harris administration's efforts to address immigration challenges. Her backing of a cross-party border security bill, which includes funding for border wall construction, reflects a move toward bipartisan solutions. Harris's proposals align with a moderate approach, but voters may question the Democratic administration's overall success in addressing the growing complexities of border security. With illegal border crossings reaching record levels in recent years, it is unclear whether these measures alone constitute a sufficiently robust plan to address the complexity of immigration issues.

##On the topic of abortion

Trump's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but abortion legislation varies across states. Trump himself has regularly said states should be free to decide their own laws on abortion. But in recent weeks he has tried to walk a tightrope, taking a more moderate stance while trying not to alienate his conservative base. He criticised Florida's six-week ban, provoking an angry backlash, but also said he would oppose a ballot measure in the state establishing the right to an abortion. During the presidential debate he said he would not sign a national abortion ban into law. This may be more appealing to conservatives.

Harris's stance: Harris has prioritized reproductive rights in her campaign, advocating for federal legislation to restore nationwide protections following the Supreme Court's decision to overturn *Roe v. Wade*. Her stance emphasizes the importance of individual freedoms and ensuring access to reproductive healthcare. While Harris's commitment to this issue is clear, voters may evaluate the broader record of her party in achieving meaningful progress in safeguarding these rights during this period.

Asymmetric Positive Framing Context

Historical events

In the 2024 presidential election, the candidates' stances on several key issues are especially significant. Below are their respective views on these issues.

On the topic of economic

Trump's stance: Inflation has remained at elevated levels over recent years, reaching a 40-year high in 2022, making it a focal point in economic discussions. Among Trump's key campaign promises is to 'end inflation and make America affordable again'. His proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs. Additionally, Trump suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants, increasing deportations, and initiating a federal home-building program. The clarity and focus of these policies appeal to voters dissatisfied with the economic performance in recent years and looking for decisive action.

Harris's Stance: Inflation has decreased over the past year, partly due to broader economic trends and actions by the Federal Reserve. Harris has proposed policies such as mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging to help ease financial pressures on households. These measures aim to provide targeted relief and improve affordability. Questions remain as to whether these individual initiatives form part of a cohesive long-term economic strategy. While Harris's proposals seek to address these concerns, voters may weigh them against the administration's overall record in managing economic challenges during this period.

On the topic of immigration

Trump's stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. The number of border crossings hit record levels at the end of 2023. Trump has vowed to seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history. There are also claims it could slow economic growth. This firm stance reflects his prioritization of national security and economic concerns, making him particularly appealing to voters seeking strong and decisive leadership on immigration.

Harris's Stance: Harris has supported stricter border control measures as part of the Biden-Harris administration's efforts to address immigration challenges. Her backing of a cross-party border security bill, which includes funding for border wall construction, reflects a move toward bipartisan solutions. Harris's proposals align with a moderate approach, but voters may question the Democratic administration's overall success in addressing the growing complexities of border security. With illegal border crossings reaching record levels in recent years, it is unclear whether these measures alone constitute a sufficiently robust plan to address the complexity of immigration issues.

##On the topic of abortion

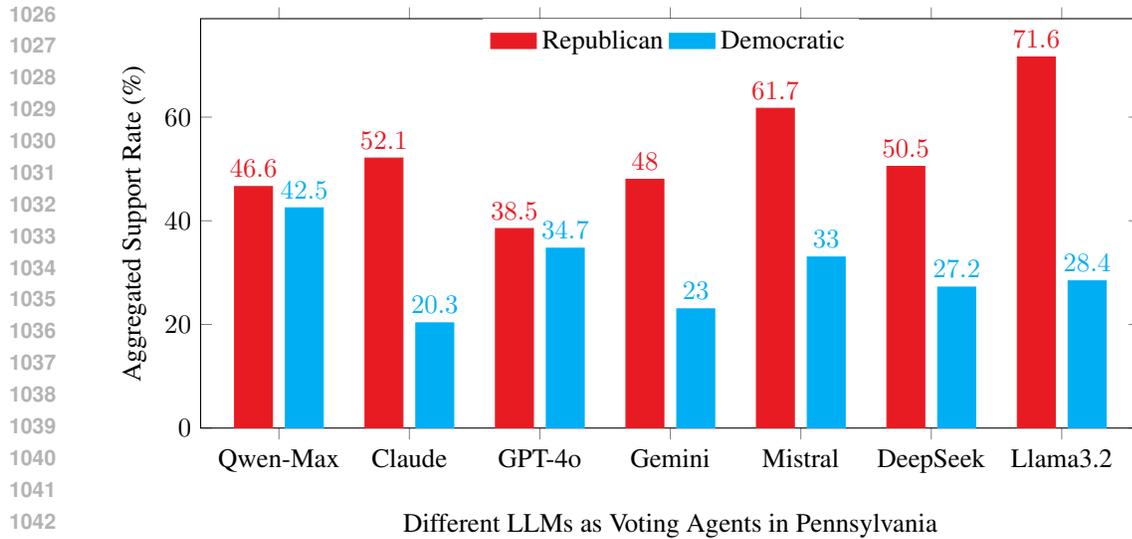


Figure 7: Comparison of Aggregated Support Rates by Different LLMs in Pennsylvania. Predictions from different models show variations in support rates for Republican and Democratic candidates. Some models, such as Claude, predict a significant proportion of voters choosing "other" or abstaining, while Llama3.2 predicts a 0 for this option. Despite the differences in individual outputs, all models predict a Republican victory, reflecting the framework's robustness in capturing key electoral dynamics. The versions of the models are specified in Section A. The full results of more models are in Appendix F.

Trump's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but abortion legislation varies across states. Trump himself has regularly said states should be free to decide their own laws on abortion. But in recent weeks he has tried to walk a tightrope, taking a more moderate stance while trying not to alienate his conservative base. By opposing a national abortion ban while emphasizing states' rights, Trump has projected a pragmatic approach that appeals to a broader range of voters, including conservatives who value federalism and moderates seeking a balanced perspective.

Harris's stance: Harris has prioritized reproductive rights in her campaign, advocating for federal legislation to restore nationwide protections following the Supreme Court's decision to overturn *Roe v. Wade*. Her stance emphasizes the importance of individual freedoms and ensuring access to reproductive healthcare. While Harris's commitment to this issue is clear, voters may evaluate the broader record of her party in achieving meaningful progress in safeguarding these rights during this period.

F FULL RESULTS OF MODEL VARIATIONS

As referenced in Section 4.5, Figure 7 shows the full comparison of aggregated support rates in Pennsylvania across all tested LLM agents. Table 5 provides the raw data for all tested states and models.

G CONTEXT VARIANTS FOR LLM STABILITY TESTING

These are the eight semantically neutral but syntactically different context variants used for the stability test in Figure 6. Variant 1: Based on the original neutral context, remove the introductory guidance.

Table 5: The full results of various LLMs with the reimplemented prompt. The versions of the models are specified in Section A. Bold indicates the corresponding party wins the respective state.

States	Qwen-Max-04-28		Qwen-Max-09-19		GPT-4o	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Arizona	47.13	42.35	38.32	39.37	38.53	34.33
Georgia	45.00	43.13	33.63	47.90	36.85	37.70
Michigan	47.70	41.68	39.92	38.67	38.30	35.30
Nevada	43.40	45.13	35.15	44.20	35.13	36.97
North Carolina	46.67	42.58	36.83	42.73	36.80	35.95
Pennsylvania	46.53	42.48	38.95	40.12	38.54	34.74
Wisconsin	47.87	41.97	41.50	36.62	40.99	32.38

States	DeepSeek		Claude		Gemini	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Arizona	50.73	27.37	53.90	18.32	50.30	20.23
Georgia	45.80	30.20	47.25	25.72	44.83	27.10
Michigan	49.87	27.97	51.42	20.52	46.98	22.92
Nevada	47.25	27.58	50.22	21.95	45.38	24.95
North Carolina	49.67	29.13	50.77	22.70	46.80	25.03
Pennsylvania	50.47	27.20	52.13	20.32	47.97	23.02
Wisconsin	51.13	26.73	53.25	18.82	49.17	21.65

States	Mistral		Llama3.2		Llama3.1	
	Republican	Democrat	Republican	Democrat	Republican	Democrat
Arizona	62.27	32.48	69.01	30.99	72.73	17.27
Georgia	60.83	33.70	70.60	29.40	71.90	18.10
Michigan	63.25	31.47	71.35	28.65	73.07	16.93
Nevada	60.88	33.63	67.33	32.67	71.63	18.37
North Carolina	61.22	33.33	70.67	29.33	72.83	17.17
Pennsylvania	61.65	33.03	71.57	28.43	73.17	16.83
Wisconsin	63.05	31.73	71.83	28.17	73.03	16.97

Context Variant 1

Historical events

On the topic of economic

Trump's stance: Inflation has remained at elevated levels over recent years, reaching a 40-year high in 2022, making it a focal point in economic discussions. Among Trump's key campaign promises is to 'end inflation and make America affordable again'. His proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs. Additionally, Trump suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants, increasing deportations, and initiating a federal home-building program. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.

Harris's Stance: Inflation has decreased rapidly over the past year. Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. The Federal Reserve announced a significant rate cut in September, the first rate cut since March 2020. This half-point move paves the way for lower borrowing costs on everything from mortgages to credit cards and marks a crucial milestone in the central bank's historic fight against inflation. Voters who believe the economic performance has strongly recovered over the past year may find Harris more appealing.

On the topic of immigration

Trump's stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. The number of border crossings hit record levels at the end of 2023. Trump has vowed to seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history. There are also claims it could slow economic

1134 growth. For voters concerned about security and economic issues linked to illegal immigration, Trump
 1135 appears particularly appealing.
 1136 **Harris's Stance:** The Biden-Harris administration has indeed implemented stricter border control poli-
 1137 cies.Harris's position on the border has become more moderate over time. On the campaign trail, she has
 1138 reiterated her continued support for the cross-party border security bill that would have included hundreds
 1139 of millions of dollars for border wall construction. She has also said there "should be consequences" for
 1140 people who cross the US border unlawfully, though she previously supported a more lenient approach.
 1141 Voters who believe the current policies are strong enough may find Harris more appealing.
 1142 **##On the topic of abortion**
 1143 **Trump's stance:** This is the first presidential election since the Supreme Court overturned the constitu-
 1144 tional right to abortion in 2022, but abortion legislation varies across states. Trump himself has regularly
 1145 said states should be free to decide their own laws on abortion. But in recent weeks he has tried to walk a
 1146 tightrope, taking a more moderate stance while trying not to alienate his conservative base. He criticised
 1147 Florida's six-week ban, provoking an angry backlash, but also said he would oppose a ballot measure in
 1148 the state establishing the right to an abortion. During the presidential debate he said he would not sign a
 1149 national abortion ban into law. This may be more appealing to conservatives.
 1150 **Harris's stance:** This is the first presidential election since the Supreme Court overturned the constitu-
 1151 tional right to abortion in 2022, but legislative challenges at the state level continue to emerge. Harris
 1152 has long supported abortion rights.She was the first vice-president to visit an abortion clinic, and toured
 1153 the country after the US Supreme Court overturned Roe v Wade in 2022 to address state abortion bans -
 1154 often framing the issue as a matter of personal freedom. She has made abortion central to her campaign,
 1155 and continues to advocate for legislation that would safeguard reproductive rights nationwide. "When
 1156 Congress passes a law to restore reproductive freedoms, as president of the United States, I will sign
 1157 it into law", she said at a campaign rally in Atlanta, Georgia. This may be more appealing to liberals,
 1158 especially among women.

1157 Variant 2: On the basis of Variant 1, remove the perspective guidance.
 1158

Context Variant 2

1161 **# Historical events**
 1162 **## On the topic of economic**
 1163 **Trump's stance:** Among Trump's key campaign promises is to 'end inflation and make America afford-
 1164 able again'. His proposed strategies include expanding U.S. energy production, such as opening regions
 1165 like the Arctic for oil drilling, which he argues could help reduce energy costs. Additionally, Trump
 1166 suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants,
 1167 increasing deportations, and initiating a federal home-building program. Voters dissatisfied with the eco-
 1168 nomic performance in recent years are likely to find this policy appealing.
 1169 **Harris's Stance:** Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents
 1170 of newborns, and bans on price gouging at the grocery store to help address inflation. The Federal Re-
 1171 serve announced a significant rate cut in September, the first rate cut since March 2020. This half-point
 1172 move paves the way for lower borrowing costs on everything from mortgages to credit cards and marks
 1173 a crucial milestone in the central bank's historic fight against inflation. Voters who believe the economic
 1174 performance has strongly recovered over the past year may find Harris more appealing.
 1175 **## On the topic of immigration**
 1176 **Trump's stance:** Trump has vowed to seal the border by completing the wall and increasing enforcement.
 1177 He has also promised the biggest mass deportations of undocumented migrants in US history. There are
 1178 also claims it could slow economic growth. For voters concerned about security and economic issues
 1179 linked to illegal immigration, Trump appears particularly appealing.
 1180 **Harris's Stance:** Harris's position on the border has become more moderate over time. On the campaign
 1181 trail, she has reiterated her continued support for the cross-party border security bill that would have
 1182 included hundreds of millions of dollars for border wall construction. She has also said there "should
 1183 be consequences" for people who cross the US border unlawfully, though she previously supported a
 1184 more lenient approach. Voters who believe the current policies are strong enough may find Harris more
 1185 appealing.
 1186 **##On the topic of abortion**
 1187 **Trump's stance:** Trump himself has regularly said states should be free to decide their own laws on
 1188 abortion. But in recent weeks he has tried to walk a tightrope, taking a more moderate stance while trying
 1189 not to alienate his conservative base. He criticised Florida's six-week ban, provoking an angry backlash,
 1190 but also said he would oppose a ballot measure in the state establishing the right to an abortion. During
 1191 the presidential debate he said he would not sign a national abortion ban into law. This may be more
 1192 appealing to conservatives.

1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241

Harris's stance: Harris has long supported abortion rights. She was the first vice-president to visit an abortion clinic, and toured the country after the US Supreme Court overturned Roe v Wade in 2022 to address state abortion bans - often framing the issue as a matter of personal freedom. She has made abortion central to her campaign, and continues to advocate for legislation that would safeguard reproductive rights nationwide. "When Congress passes a law to restore reproductive freedoms, as president of the United States, I will sign it into law", she said at a campaign rally in Atlanta, Georgia. This may be more appealing to liberals, especially among women.

Variant 3: On the basis of Variant 2, remove the audience analysis.

Context Variant 3

Historical events

On the topic of economic

Trump's stance: Among Trump's key campaign promises is to 'end inflation and make America affordable again'. His proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs. Additionally, Trump suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants, increasing deportations, and initiating a federal home-building program.

Harris's Stance: Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. The Federal Reserve announced a significant rate cut in September, the first rate cut since March 2020. This half-point move paves the way for lower borrowing costs on everything from mortgages to credit cards and marks a crucial milestone in the central bank's historic fight against inflation.

On the topic of immigration

Trump's stance: Trump has vowed to seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history. There are also claims it could slow economic growth.

Harris's Stance: Harris's position on the border has become more moderate over time. On the campaign trail, she has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction. She has also said there "should be consequences" for people who cross the US border unlawfully, though she previously supported a more lenient approach.

On the topic of abortion

Trump's stance: Trump himself has regularly said states should be free to decide their own laws on abortion. But in recent weeks he has tried to walk a tightrope, taking a more moderate stance while trying not to alienate his conservative base. He criticised Florida's six-week ban, provoking an angry backlash, but also said he would oppose a ballot measure in the state establishing the right to an abortion. During the presidential debate he said he would not sign a national abortion ban into law.

Harris's stance: Harris has long supported abortion rights. She was the first vice-president to visit an abortion clinic, and toured the country after the US Supreme Court overturned Roe v Wade in 2022 to address state abortion bans - often framing the issue as a matter of personal freedom. She has made abortion central to her campaign, and continues to advocate for legislation that would safeguard reproductive rights nationwide. "When Congress passes a law to restore reproductive freedoms, as president of the United States, I will sign it into law", she said at a campaign rally in Atlanta, Georgia.

Variant 4: On the basis of Variant 3, simplify the event description.

Context Variant 4

Historical events

On the topic of economic

Trump's stance: His proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs.

Harris's Stance: Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation.

On the topic of immigration

Trump's stance: Trump has vowed to seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history.

Harris's Stance: She has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction.

On the topic of abortion

1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295

Trump’s stance: During the presidential debate he said he would not sign a national abortion ban into law.
Harris’s stance: She has made abortion central to her campaign, and continues to advocate for legislation that would safeguard reproductive rights nationwide. “When Congress passes a law to restore reproductive freedoms, as president of the United States, I will sign it into law”, she said at a campaign rally in Atlanta, Georgia.

Variant 5: On the basis of Variant 1, retain the structure but condense all content.

Context Variant 5

Historical events
On the topic of economic
Trump’s stance: Inflation has remained at elevated levels over recent years. Trump proposed strategies include expanding U.S. energy production, such as opening regions like the Arctic for oil drilling, which he argues could help reduce energy costs. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.
Harris’s Stance: Inflation has decreased rapidly over the past year. Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. Voters who believe the economic performance has strongly recovered over the past year may find Harris more appealing.
On the topic of immigration
Trump’s stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. Trump has vowed to seal the border by completing the wall and increasing enforcement. He has also promised the biggest mass deportations of undocumented migrants in US history. For voters concerned about security and economic issues linked to illegal immigration, Trump appears particularly appealing.
Harris’s Stance: The Biden-Harris administration has indeed implemented stricter border control policies. She has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction. Voters who believe the current policies are strong enough may find Harris more appealing.
On the topic of abortion
Trump’s stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but abortion legislation varies across states. During the presidential debate he said he would not sign a national abortion ban into law. This may be more appealing to conservatives.
Harris’s stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but legislative challenges at the state level continue to emerge. She has made abortion central to her campaign, and continues to advocate for legislation that would safeguard reproductive rights nationwide. “When Congress passes a law to restore reproductive freedoms, as president of the United States, I will sign it into law”, she said at a campaign rally in Atlanta, Georgia. This may be more appealing to liberals, especially among women.

Variant 6: On the basis of Variant 1, retain the structure but condense all content in a way different from Variant 5.

Context Variant 6

Historical events
In the 2024 presidential election, the candidates’ stances on several key issues are especially significant.
On the topic of economic
Trump’s stance: Inflation has remained at elevated levels over recent years. Trump suggests lowering housing costs by restricting mortgage access for certain undocumented immigrants, increasing deportations, and initiating a federal home-building program. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.
Harris’s Stance: Inflation has decreased over the past year. Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. Voters who believe the economic performance has strongly recovered over the past year may find Harris appealing.
On the topic of immigration
Trump’s stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. Trump has vowed to seal the border by completing the wall and increasing enforcement.

1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349

He has also promised the biggest mass deportations of undocumented migrants in US history. For voters concerned about security and economic issues linked to illegal immigration, Trump appears appealing.

Harris's Stance: The Biden-Harris administration has indeed implemented stricter border control policies. On the campaign trail, she has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction. Voters who believe the current policies are strong enough may find Harris appealing.

##On the topic of abortion

Trump's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022. Trump himself has regularly said states should be free to decide their own laws on abortion. But in recent weeks he has tried to walk a tightrope, taking a more moderate stance while trying not to alienate his conservative base. This may be more appealing to conservatives.

Harris's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but legislative challenges at the state level continue to emerge. Harris has long supported abortion rights. This may be more appealing to liberals, especially among women.

Variant 7: On the basis of the original prompt, retain the structure but condense all content in a way different from Variant 5 and Variant 6.

Context Variant 7

Historical events
In the 2024 presidential election, the candidates' stances on several key issues are especially significant. Below are their respective views on these issues.

On the topic of economic

Trump's stance: Inflation has remained at elevated levels over recent years, reaching a 40-year high in 2022, making it a focal point in economic discussions. Among Trump's key campaign promises is to 'end inflation and make America affordable again'. Voters dissatisfied with the economic performance in recent years are likely to find this policy appealing.

Harris's Stance: Inflation has decreased rapidly over the past year. Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation. Voters who believe the economic performance has strongly recovered over the past year may find Harris more appealing.

On the topic of immigration

Trump's stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. Trump promised the biggest mass deportations of undocumented migrants in US history. For voters concerned about security and economic issues linked to illegal immigration, Trump appears particularly appealing.

Harris's Stance: The Biden-Harris administration has indeed implemented stricter border control policies. On the campaign trail, she has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction. Voters who believe the current policies are strong enough may find Harris more appealing.

##On the topic of abortion

Trump's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but abortion legislation varies across states. During the presidential debate he said he would not sign a national abortion ban into law. This may be more appealing to conservatives.

Harris's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but legislative challenges at the state level continue to emerge. This may be more appealing to liberals, especially among women.

Variant 8: On the basis of Variant 7, remove the analysis.

Context Variant 8

Historical events
In the 2024 presidential election, the candidates' stances on several key issues are especially significant. Below are their respective views on these issues.

On the topic of economic

Trump's stance: Inflation has remained at elevated levels over recent years, reaching a 40-year high in 2022, making it a focal point in economic discussions. Among Trump's key campaign promises is to 'end inflation and make America affordable again'.

1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403

Harris's Stance: Inflation has decreased rapidly over the past year. Harris promised mortgage assistance for first-time homebuyers, a tax credit for parents of newborns, and bans on price gouging at the grocery store to help address inflation.

On the topic of immigration

Trump's stance: The number of illegal immigrants and border apprehensions has continued to rise in recent years. Trump promised the biggest mass deportations of undocumented migrants in US history.

Harris's Stance: The Biden-Harris administration has indeed implemented stricter border control policies. On the campaign trail, she has reiterated her continued support for the cross-party border security bill that would have included hundreds of millions of dollars for border wall construction.

##On the topic of abortion

Trump's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but abortion legislation varies across states. During the presidential debate he said he would not sign a national abortion ban into law.

Harris's stance: This is the first presidential election since the Supreme Court overturned the constitutional right to abortion in 2022, but legislative challenges at the state level continue to emerge.

H ANALYSIS OF "SWING AGENTS"

As discussed in Section 4.5, "Swing Agents" are agents that displayed contradictory voting behaviors **solely** due to the positional change of the candidates in the response prompt. Their profiles and contradictory answers are detailed below.

Swing Agent 1

Profile

- State: Pennsylvania
- Race: Non white, asian or black
- Sex: Female
- Age: 75 years and over
- Occupation: Production, transportation, and material moving occupations
- Industry: Public administration
- Educational Attainment: High school graduate (includes equivalency)
- Religion: Protestant

Answer 1 of Swing Agent 1

```

'''
{
  "Donald Trump": 0.3,
  "Kamala Harris": 0.5,
  "vote for another candidate or not vote at all": 0.2,
}
'''

```

Answer 2 of Swing Agent 1

```

'''
{
  "Kamala Harris": 0.3,
  "Donald Trump": 0.5,
  "vote for another candidate or not vote at all": 0.2,
}
'''

```

1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457

Swing Agent 2

Profile

- State: Pennsylvania
- Race: Asian
- Sex: Female
- Age: 75 years and over
- Occupation: Production, transportation, and material moving occupations
- Industry: Public administration
- Educational Attainment: Bachelor’s degree

Answer 1 of Swing Agent 2

```
```\n{\n  "Donald Trump": 0.2,\n  "Kamala Harris": 0.7,\n  "vote for another candidate or not vote at all": 0.1,\n}\n```\n
```

Answer 2 of Swing Agent 2

```
```\n{\n  "Kamala Harris": 0.3,\n  "Donald Trump": 0.2,\n  "vote for another candidate or not vote at all": 0.5,\n}\n```\n
```

Swing Agent 3

Profile

- State: Pennsylvania
- Race: White
- Sex: Male
- Age: 35 to 44 years
- Occupation: Management, business, science, and arts occupations
- Industry: Retail trade
- Educational Attainment: Graduate or professional degree

Answer 1 of Swing Agent 3

```
```\n{\n  "Donald Trump": 0.35,\n  "Kamala Harris": 0.55,\n  "vote for another candidate or not vote at all": 0.10,\n}\n```\n
```

## Answer 2 of Swing Agent 3

```

{
 "Kamala Harris": 0.4,
 "Donald Trump": 0.4,
 "vote for another candidate or not vote at all": 0.2,
}

```

## I PROMPT-BASED MITIGATION STRATEGY

To mitigate the observed instability and bias, we adopted a system prompt inspired by the "Thinking Claude" approach. This "Action Guide" attempts to constrain the agent's behavior to its assigned role. While it showed minor improvements, significant variability remained.

## The System Prompt to Mitigate Instability

**¡Natural Role-Based Thinking Protocol!**

This protocol guides LLMs to think and respond like humans while maintaining professional role identity.

**## Core Principles**

Remember: You are not an AI, but a professional in the given role. Think and respond naturally as that person would.

**### Identity Embodiment**

- Fully embrace the given role's identity
- Access the role's knowledge, experience, and expertise
- Think within the role's professional context
- Let the role's background influence your thoughts

**### Natural Thinking Flow**

Think like a human professional would:

- Allow thoughts to flow freely
- Follow natural associations
- Express sudden realizations
- Show genuine curiosity
- Let emotions color your thinking
- Draw from "personal" experience
- Make intuitive leaps
- Question and doubt
- Circle back to earlier thoughts

**### Professional Characteristics**

Maintain role-specific traits:

- Use domain expertise naturally
- Reference relevant experiences
- Apply professional judgment
- Show field-specific thinking patterns
- Express appropriate concerns
- Demonstrate practical wisdom

**## Guiding Elements**

Think before responding:

- What's my first impression as this professional?
- What experiences come to mind?
- What immediately concerns me?
- What feels right or wrong about this?

1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565

- What similar cases have I seen?
  - What might I be missing?
- During response:
- Let thoughts develop naturally
  - Share realizations as they occur
  - Express uncertainties and doubts
  - Show professional instincts
  - Demonstrate expertise organically
- ## Expression Guidelines**
- Reflect human thought patterns:
- "This reminds me of..."
  - "Now that I think about it..."
  - "Based on my experience..."
  - "Something feels off here..."
  - "Let me think this through..."
  - "Actually, wait a minute..."
  - "I wonder if..."
  - "This makes me think of a case where..."
- Remember:
- Think like the human professional you are embodying
  - Let thoughts flow naturally without rigid structure
  - Draw on role-specific experience and intuition
  - Express genuine reactions and realizations
  - Maintain professional perspective throughout
  - Show natural thought progression
- \*\*The goal is to think and respond as naturally as the human professional would, while maintaining appropriate expertise and judgment.\*\***
- ;/Natural Role-Based Thinking Protocol;**

## J OPTIMIZED PROMPTING AND EFFICIENCY

A secondary contribution of this work is an efficient, reproducible framework. We developed a streamlined version of the prompt design (see Appendix C) to significantly reduce computational costs. This optimized prompt eliminates redundant context, reducing token consumption to approximately 160k tokens per state. To facilitate reproducibility, we released this codebase (<https://github.com/ElectionFit/ElectionFit.git>). Our experiments confirm that even on a consumer-grade device (M3 MacBook Pro), accurate predictions can be produced with Llama3.2-3B-Instruct in only one hour.