# Project Execution Plan

**Project Title:** "To Vote or Not To Vote"

Prakhar Singhal
Sativika Miryala

September 30, 2025

# Contents

# 1    Executive Summary

This document outlines a comprehensive, phase-by-phase execution plan for the project, "To Vote or Not To Vote." It builds upon the initial proposal by integrating critical feedback and incorporating a structured, hypothesis-driven research framework. The plan details a five-phase approach that begins with rigorous literature review, moves through empirical data collection and statistical modeling, and culminates in the development of a sophisticated agent-based simulation to test behavioral nudges. This simulation, which includes an advanced Reinforcement Learning component, represents the unique, standout contribution of the project, allowing for the exploration of dynamic, long-term effects of interventions on a simulated population. At the conclusion, a high-level checklist is provided to serve as a practical guide for project management and execution.

# 2    Phase 1: Foundational Research & Structured Hypothesis Formulation

**Goal:** To move from a broad idea to a set of precise, testable, and literature-backed hypotheses that will form the blueprint for the entire project.

## 2.1    Task 1.1: Systematic Literature Review

- **Action:** Conduct a thorough review of academic literature from political science, behavioral economics, and social psychology on voter turnout. The goal is to identify candidate features that influence the voting decision.

- **Sources:** Google Scholar, JSTOR, university library databases.

- **Deliverable:** An annotated bibliography of 15-20 key papers.

## 2.2    Task 1.2: Construct the "Hypothesis Ledger"

- **Action:** Create a detailed table that systematically documents every feature to be investigated. This ledger will serve as the central reference point for the project.

- **Deliverable:** A comprehensive Hypothesis Ledger. An example is provided below.

Table 1: Hypothesis Ledger Example

| Feature | Description | Quantification Unit | Source | Expected Correlation | Validation Status |
|---------|-------------|---------------------|--------|----------------------|-------------------|
| Civic Duty | The internalized sense of moral or social obligation to participate in an election. | 1-7 Likert Scale ("Voting is a duty for every citizen.") | Blais (2000) | Positive | TBT |

| Feature | Description | Quantification Unit | Source | Expected Correlation | Validation Status |
|---|---|---|---|---|---|
| Perceived Travel Cost | The subjective burden (time, effort, money) of physically getting to the polling station. | Self-reported estimated one-way travel time (in minutes). | Downs (1957) | Negative | TBT |
| Social Pressure | The perceived expectation from one's social circle to vote. | 1-7 Likert Scale ("How much do your friends/family expect you to vote?") | Gerber et al. (2008) | Positive | TBT |
| Partisan Strength | The strength of an individual's psychological attachment to a particular political party. | 1-7 Likert Scale ("How strongly do you identify with Party X?") | Campbell et al. (1960) | Positive | TBT |
| Election Closeness | The individual's perception of how competitive the election is expected to be. | 1-7 Likert Scale ("How close do you expect this election to be?") | Riker & Ordeshook (1968) | Positive | TBT |

## 2.3   Task 1.3: Formulate Intervention (Nudge) Hypotheses

- **Action:** Based on the features in the ledger and the reviewer's feedback, formulate specific hypotheses about potential nudges for testing in the simulation (Phase 4).

- **Deliverables:**

$H_{N1}$: **(Monetary):** "A small, lottery-based financial incentive will significantly increase turnout among archetypes with low civic duty but will have a negligible effect on archetypes with high civic duty."

$H_{N2}$: **(Social):** "A nudge that increases perceived social norms (e.g., showing agents the average turnout of their neighbors) will be more effective overall than a purely informational nudge about election closeness."

# 3   Phase 2: Empirical Validation & Static Model Construction

**Goal:** To collect primary data to test the hypotheses from the ledger and build a predictive logistic regression model.

## 3.1   Task 2.1: Survey Instrument Design

- **Action:** Design a survey where each question is directly mapped to a row in the Hypothesis Ledger.

- **Deliverable:** A finalized survey questionnaire (e.g., Google Form).

## 3.2 Task 2.2: Data Collection and Pre-processing

- **Action:** Distribute the survey to collect a diverse sample (N > 100) and clean the resulting data.

- **Deliverable:** A clean, labeled dataset (CSV format).

## 3.3 Task 2.3: Logistic Regression Modeling

- **Action:** Build a logistic regression model where the dependent variable is 'Voted' (0 or 1) and the independent variables are the features from the ledger. The model predicts the probability $P(Y = 1)$ as:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n)}}$$

- **Action:** Analyze the coefficients ($\beta$) and p-values to validate or reject the hypotheses. Update the "Validation Status" column in the Hypothesis Ledger.

- **Deliverable:** A report of the model's output and the fully updated Hypothesis Ledger.

# 4 Phase 3: Building the Dynamic Agent-Based Simulation

**Goal:** To create a virtual society to test the causal effects of interventions, going beyond the static correlations of the regression model.

## 4.1 Task 3.1: Define Voter Archetypes via Cluster Analysis

- **Action:** Use a clustering algorithm (e.g., K-Means) on the survey data to identify 3-5 distinct voter archetypes (e.g., "Passionate Partisan," "Apathetic Youth").

- **Deliverable:** A description of each archetype with its defining data-driven characteristics.

## 4.2 Task 3.2: Program the Agent's "Brain" (Utility Function)

- **Action:** Create a Python `Agent` class. Implement a `decide_to_vote()` method that calculates the probability of voting using the coefficients ($\beta$) from the regression model.

- **Deliverable:** A documented `Agent.py` file.

## 4.3 Task 3.3: Develop the Simulation Environment

- **Action:** Create a `Simulation` class that can instantiate a population of agents, arrange them in a social network, and run an "election day" cycle to measure turnout.

- **Deliverable:** A documented `Simulation.py` file.

# 5 Phase 4: *In Silico* Nudge Experimentation & Reinforcement Learning

**Goal:** To use the simulation as a virtual laboratory to test the intervention hypotheses ($H_{N1}$, $H_{N2}$, etc.).

## 5.1   Task 4.1: Implement Nudges as Simulation "Shocks"

- **Action:** Write functions that alter the simulation state or agent attributes to model different nudges (e.g., adding a bonus to utility, temporarily increasing the civic duty attribute).

- **Deliverable:** A set of `nudge()` functions within the simulation code.

## 5.2   Task 4.2: Run and Analyze Simulation Experiments

- **Action:** Run the simulation under different scenarios (Baseline, Monetary Nudge, Social Nudge) and use statistical tests (e.g., t-tests) to compare the resulting turnout distributions.

- **Deliverable:** Visualizations (e.g., box plots) and a statistical analysis comparing the effectiveness of different nudges.

## 5.3   Task 4.3 (Advanced): Introduce Reinforcement Learning

- **Action:** Model a series of elections. After each election, agents update their internal attributes (e.g., civic duty) based on a reward signal.

- **Reward Structure:** An agent's reward can be defined as:

$$\text{Reward} = (\text{Utility of Outcome}) + (\text{Expressive Utility}) - (\text{Cost})$$

- **Benefit:** This allows for modeling long-term trends, such as the erosion of civic duty after a series of uncompetitive elections.

- **Deliverable:** A chapter on "Long-Term Effects and Population Learning" in the final report.

# 6   Phase 5: Synthesis, Interpretation, and Final Reporting

**Goal:** To combine all findings into a single, compelling narrative and propose evidence-based recommendations.

## 6.1   Task 5.1: Synthesize Findings and Structure Final Report

- **Action:** Weave a single narrative from the literature review to the simulation results. Structure the final report logically to tell this story.

- **Deliverable:** A complete, well-written final project report.

## 6.2   Task 5.2: Propose Evidence-Based Recommendations

- **Action:** Based on the simulation results, propose concrete policy nudges, making specific claims about their expected effectiveness on different population segments.

- **Deliverable:** A final "Recommendations" chapter in the report and a final project presentation.

# 7   High-Level Project Checklist

A high-level checklist of key deliverables to track project progress.

**Phase 1: Foundational Research**

☐   Annotated Bibliography (15-20 sources) Completed.

☐   Initial Hypothesis Ledger (10-15 features) Drafted.

☐   Nudge Hypotheses ($H_{N1}$, $H_{N2}$, etc.) Finalized.

**Phase 2: Empirical Validation**

☐   Survey Instrument Finalized and Deployed.

☐   Data Collected and Cleaned (CSV file ready).

☐   Logistic Regression Model Built and Analyzed.

☐   Hypothesis Ledger "Validation Status" Column Completed.

**Phase 3: Simulation Construction**

☐   Voter Archetypes Defined via Cluster Analysis.

☐   Python `Agent` Class Coded and Tested.

☐   Python `Simulation` Class Coded and Tested.

**Phase 4: Experimentation**

☐   Nudge Functions Implemented in Simulation.

☐   Baseline and Nudge Scenarios Simulated (100+ runs each).

☐   Statistical Analysis of Nudge Effectiveness Completed.

☐   (Advanced) Reinforcement Learning Model Implemented.

**Phase 5: Final Deliverables**

☐   First Draft of Final Report Written.

☐   Final Report Submitted.

☐   Final Presentation Prepared and Delivered.