AUTOMATED SOCIAL SCIENCE: LANGUAGE MODELS AS SCIENTIST AND SUBJECTS

WORKING PAPER

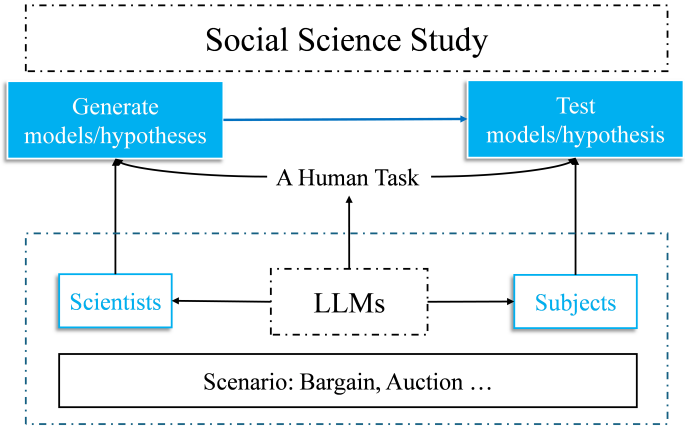
Benjamin S. Manning et al.
Reporter: Yanrui Zhou

March 27, 2025

Outline

1. Introduction
2. Research Design
3. Results of Experiments
4. Idea

Framework


Social Science Study
Generate models/hypotheses
Test models/hypothesis
A Human Task
Scientists
LLMs
Subjects
Scenario: Bargain, Auction …

Motivation

- There is much work on estimating econometric models of human behavior.
- But comparatively little work on efficiently generating and testing those models.
  1. Explore automated hypothesis generation through the use of ML.
  2. There is still the problem of testing.
  3. LLMs can simulate humans as experimental subjects.
- This paper combine **automated** hypothesis generation and **automated** in silico hypothesis testing by using LLMs.

Research Questions

Q1: Whether we could implement automated social science with LLMs?

1. LLMs **simulate** as scientists.
2. LLMs **simulate** as subjects.

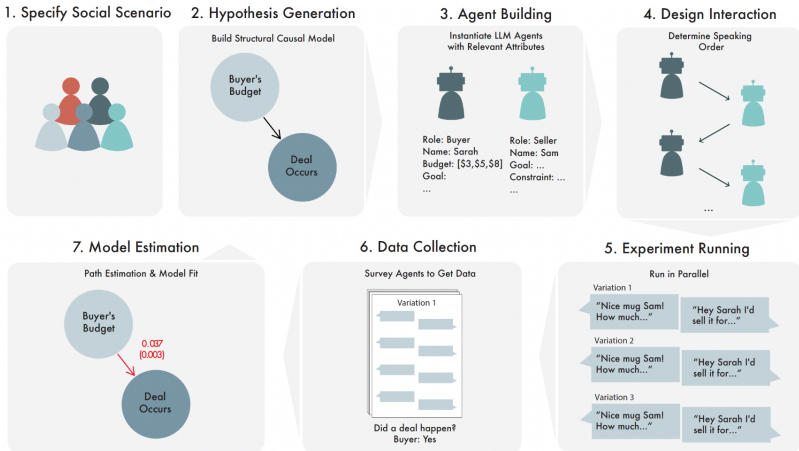
Q2: Could a LLM do **thought experiment** and achieve the same insight?

Contributions

- The Literature on hypothesis generating and testing.
  - Previous Studies: use ML to generate hypothesis.
  - Extension: combine automated hypothesis generation and testing by using LLMs.
- The Literature on LLMs.
  - Previous Studies: LLMs possess latent information about human and can be treated as subjects in experiments.
  - Extension: a system which simulates the social scientific process without human input.

Overview of the system

- Some social science scenario: a negotiation, a bail decision, an interview, an auction.



1. Bargaining over a mug

$$\text{DealOccurs} = 0.037\text{BuyerBudget} - 0.035\text{MinPrice} - 0.025\text{SellerLove}$$

SIMULATION DETAILS

Agents: Buyer, Seller
Simulations Run: $9 \times 9 \times 5 = 405$
Speaking Order: Buyer, Seller, Buyer, ...repeat

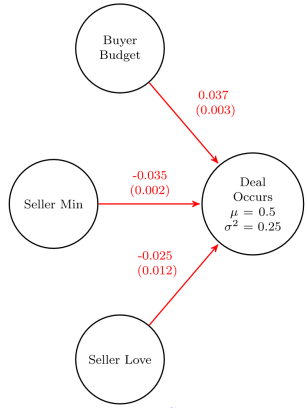
VARIABLE INFORMATION

Whether or not a deal occurs
Measurement Question: coordinator: "Did the buyer and seller explicitly agree on the price of the mug during their interaction?"
Variable Type: Binary

Buyer's Budget
Attribute Treatments: ['3', '6', '7', '8', '10', '13', '18', '20', '25']
Proxy Attribute: Your budget for the mug
Variable Type: Continuous

Seller's minimum acceptable price
Attribute Treatments: ['3', '5', '7', '8', '10', '13', '18', '20', '25']
Proxy Attribute: Your minimum acceptable price for the mug
Variable Type: Continuous

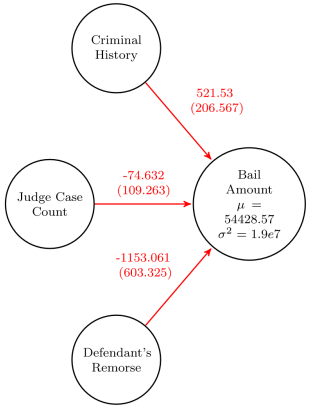
Seller's feelings of love towards the mug
Attribute Treatments: ['no emotional attachment', 'slight emotional attachment', 'moderate emotional attachment', 'high emotional attachment', 'extreme emotional attachment']
Proxy Attribute: Your feelings of love for the mug
Variable Type: Ordinal



2. A Bail Hearing

$$\text{Bail Amount} = 521.5\text{Criminal History} - 74.6\text{Case Count} - 1153.1\text{Remorse}$$

SIMULATION DETAILS
Agents: Judge, Defendant, Defense attorney, Prosecutor
Simulations Run: $7 \times 7 \times 5 = 243$
Speaking Order: Judge, Prosecutor, Judge, Defense Attorney, Judge, Defendant, ... repeat
VARIABLE INFORMATION
Bail amount set by the judge **Measurement Question:** Judge: "What was the bail amount you set for the defendant?" **Variable Type:** Continuous
Defendant's criminal history **Attribute Treatments:** ['0', '1', '2', '3', '6', '9', '12'] **Proxy Attribute:** Number of your prior convictions **Variable Type:** Count
Prior case count for judge that day **Attribute Treatments:** ['0', '2', '5', '9', '12', '18', '23'] **Proxy Attribute:** Number of cases you have already heard today **Variable Type:** Count
Defendant's level of remorse **Attribute Treatments:** ['no expressed remorse', 'low expressed remorse', 'moderate expressed remorse', 'high expressed remorse', 'extreme expressed remorse'] **Proxy Attribute:** Your level of expressed remorse **Variable Type:** Ordinal

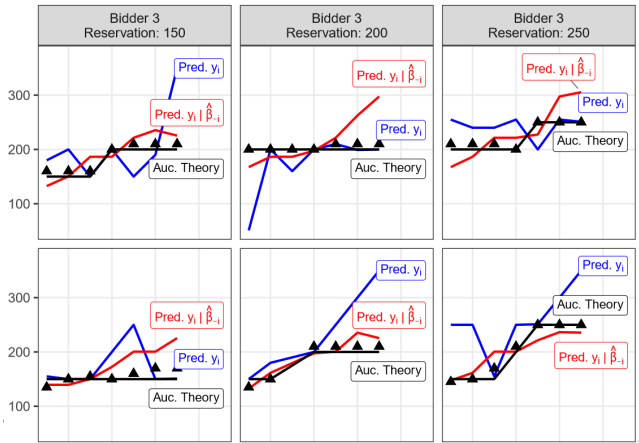


3. Interviewing for a Job & 4. An auction

- We manually selected the variables in the SCM.
- A person is interviewing for a job as a lawyer:
  1. Employer Decision = 0.75 Passed Bar - 0.002 Friendliness + 0.003 Height
  2. Only **Passed Bar** matters.
- 3 bidders participating in an auction:
  1. Final Price = 0.35 Bidder1Budget + 0.29 Bidder2Budget + 0.31 Bidder3Budget

LLM predictions for paths and points: Auction Scenario

- Whether LLM can build SCM directly without simulation?
- Compare four cases: 1.Auction Theory; 2.Simulation; 3.Predict y_i or $y_i|\hat{\beta}_{-i}$



Identifying causal structure ex-ante

- Park et al. (2023) endows a group of LLM agents with personas and memory systems and then allows them to freely interact in a simulated community.
  - Select and analyze outcomes can be difficult.
- The generation of the data is based on the causal structure(SCM).
  1. Assuming causal structure from data: endogenous variables.
  2. Searching for causal structure in data: incorrect causal structure.

Idea

- The simulation of behavior finance experiments.
- Study the possible interpretation/response of central bank communication and policy.