MK927: Intro to Quantitative Marketing

**Lecture 13: AI as a research tool.**

Prof. Andrey Fradkin

Today: Short lecture

- Discuss your assignments.

- Ways to use AI in your research.

- Simulations of behavior and society with LLMs

- Using AI to make existing styles of work better.

- Short discussions.

Demonstrate Cursor

Simulations of societies

- In this class, we've focused on on game theory:

  - Rational agents.

  - Well defined utility functions.

  - A particular equilibrium concept.

- "Agent-based models".

  - Thomas Schelling's famous segregation game.

  - Agent wants some minimum percentage of their neighbors to be of the same type.

 - If an agent's neighborhood doesn't meet their similarity threshold, they will move.

$1 - \tau = 0.25$ $1 - \tau = 0.4$

$1 - \tau = 0.6$ $1 - \tau = 0.75$
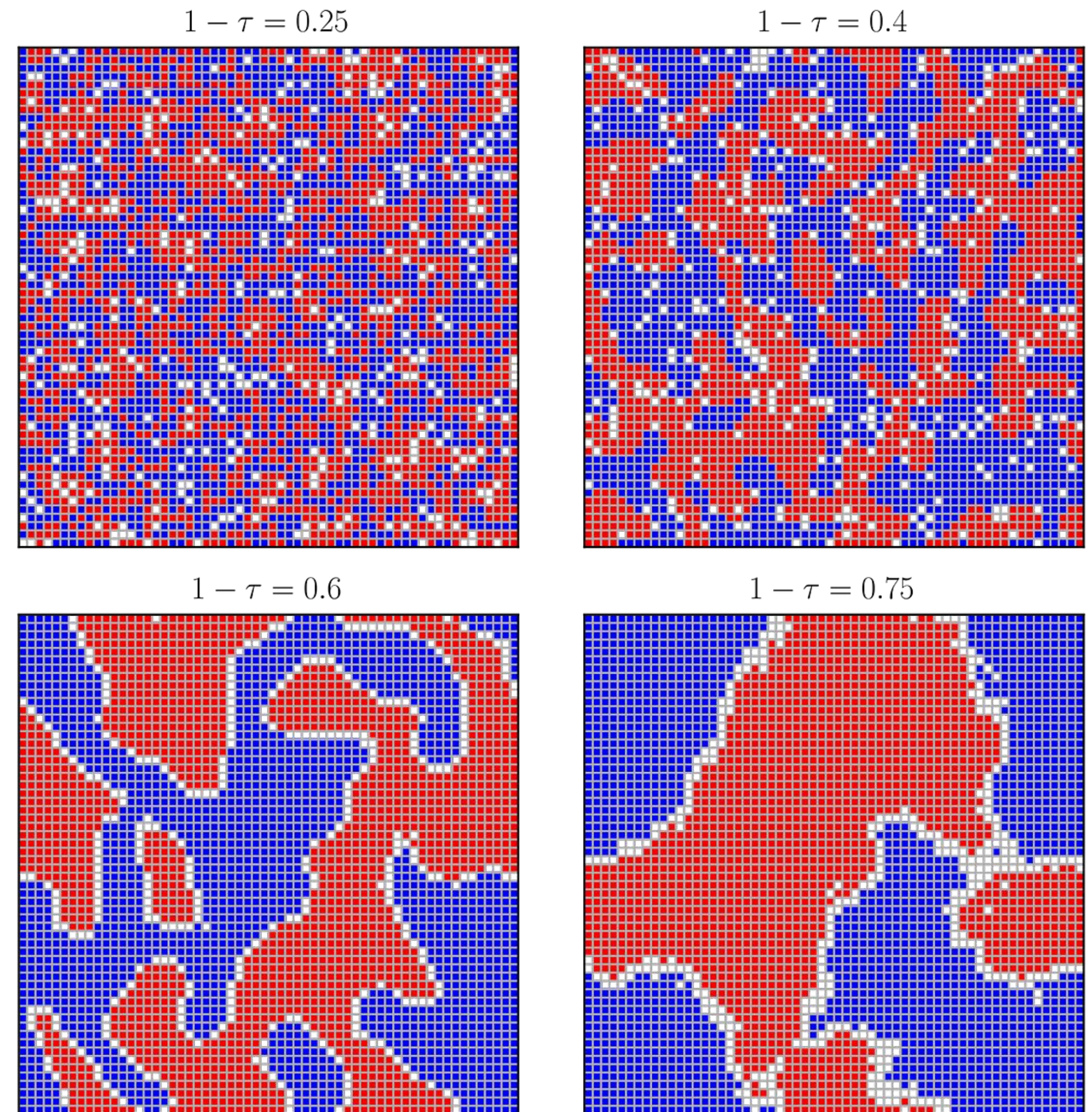
Critique of agent based models

- Rules are 'ad-hoc', not estimated from actual behavioral data.

- LLMs to the rescue? Park et al. (2023).
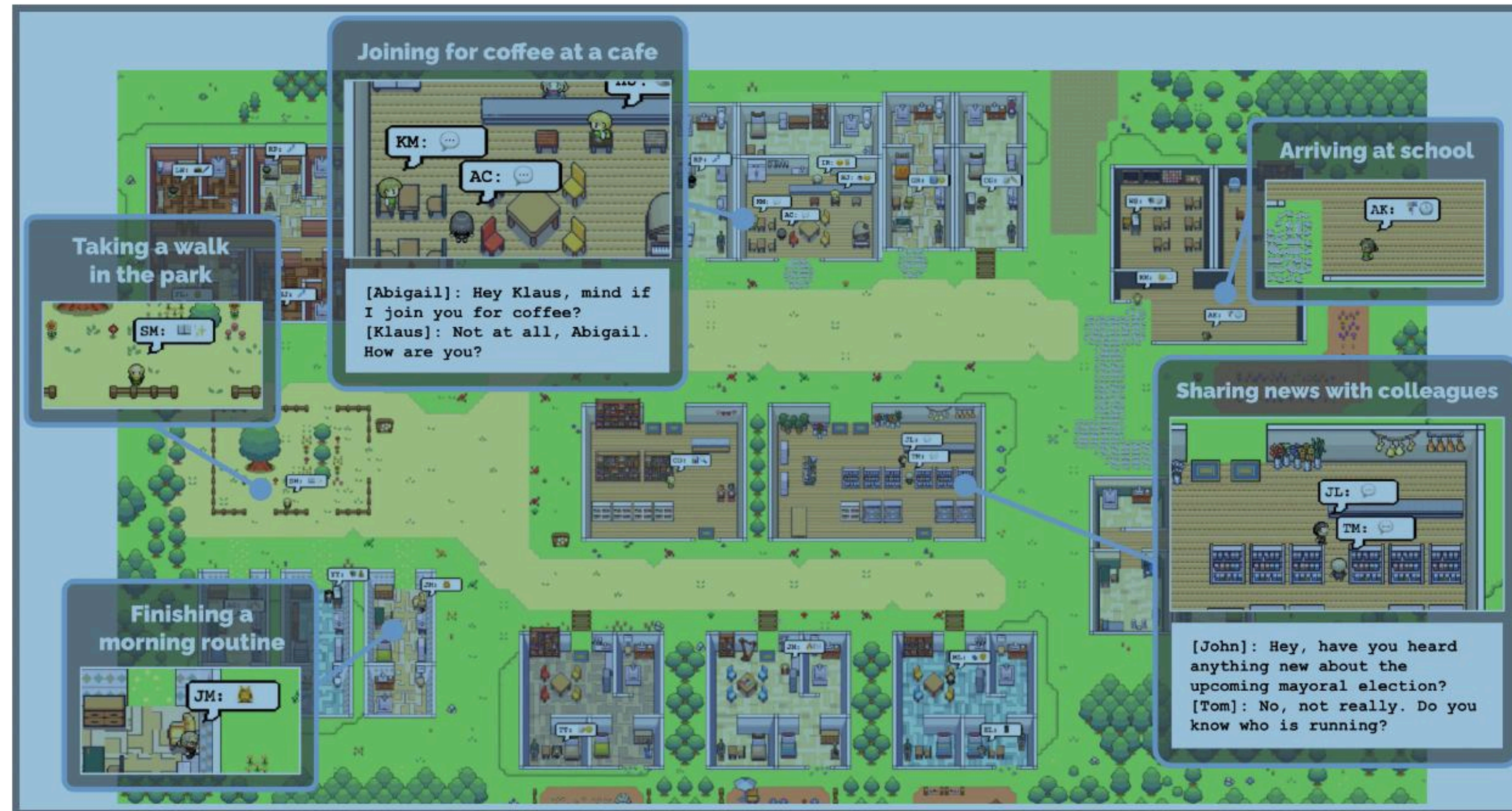
Park et al. (2023)

- Model a society. There are 25 agents, and they interact in a 2d town.

- Authors are very clever in creating the scaffolding for each agent to behave in a realistic manner.
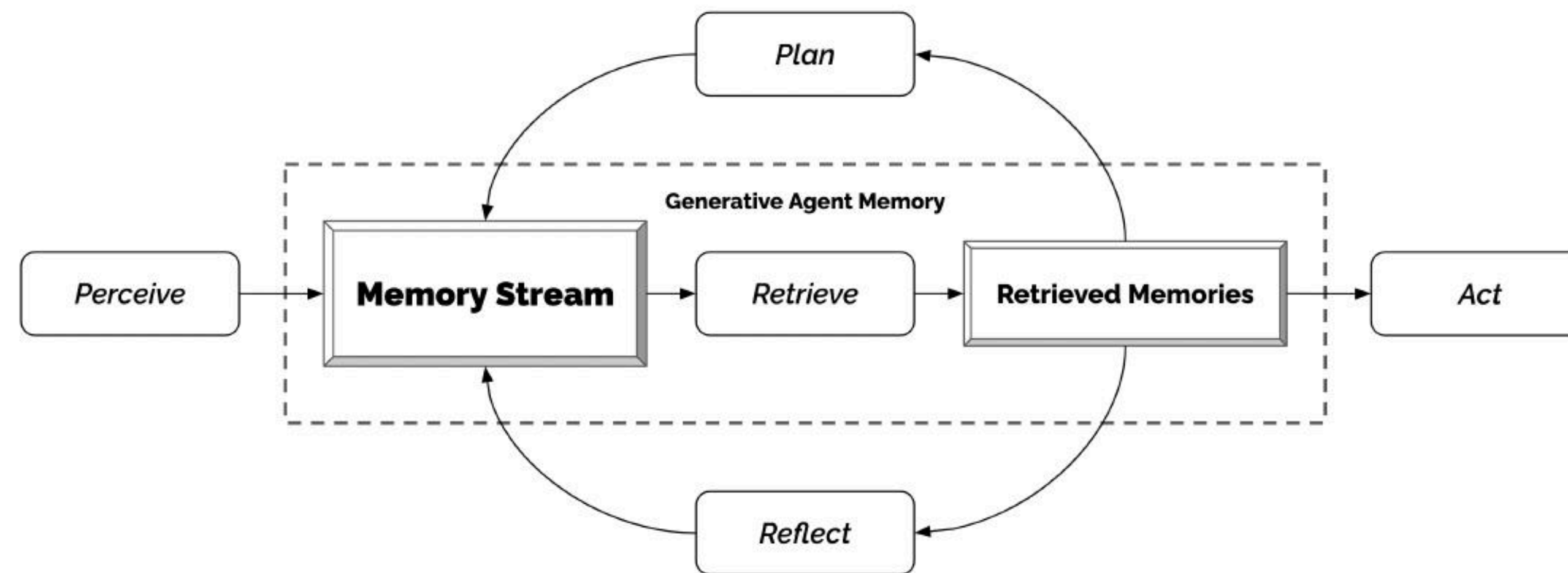
Park et al. (2023) - Emergent behavior

- Speed of information spread (i.e., diffusion).

  - Valentine's day party.

  - Mayoral candidate.

- Friendship formation.

- Relationship formation.

John: I heard that Sam Moore is running for mayor in the local election. Do you think he has a good chance of winning?

Tom: I do think he has a good chance. He's been working hard in the community and I think he will get a lot of support. What do you think?

John: I think it's great that he's running. I'm curious to see how the election will turn out.

Park et al. (2023) - Inspecting the system.

- For each agent, can turn off various parts of the scaffolding.
    - Memory
    - Planning
    - Reflection
- Can seed information with one person and see how it spreads.

Where types of questions would you use this methodology to answer?

Filippas, Horton, and Manning (2025)

- Use LLMs to replicate famous behavioral experiments.

- Find that LLMs can do a good job.

  - Note, you need to 'calibrate' agents to do this.

- Then suggest using LLMs for pilot studies, to come up

  with new theories, etc...

- Address critiques:

  - "In contrast to agent-based models, the process of

    training LLMs is not under the researcher's direct

    control."

Large Language Models as Simulated Economic Agents:
What Can We Learn from *Homo Silicus?**

Apostolos Filippas   John J. Horton   Benjamin S. Manning
Fordham   MIT & NBER   MIT

February 12, 2025

Abstract

Large language models (LLM)—because of how they are trained and designed—are implicit computational models of humans—a *homo silicus*. LLMs can be used like economists use *homo economicus*: they can be given endowments, information, preferences, and so on, and then their behavior can be explored in scenarios via simulation. Experiments using this approach, derived from Charness and Rabin (2002), Kahneman et al. (1986), and Samuelson and Zeckhauser (1988) show qualitatively similar results to the original, but it is also easy to try variations for fresh insights. LLMs could allow researchers to pilot studies via simulation, first improving their experimental design and searching for novel social science insights to test in the real world. In this paper, we offer a framework for when this approach is likely to prove useful.

Other LLMs are good for social science papers:

- Brand et al. Marketing Science.

  - LLMs for demand estimation.

- Park et al. (2024).

  - Generative Agent Simulations of 1,000 People.

  - Do well on the General Social Survey.

- Tornberg et al. (2023).

  - Simulating Social Media Using Large Language Models to Evaluate Alternative News Feed Algorithms.

Large Language Models as Simulated Economic Agents: What Can We Learn from *Homo Silicus*?*

Apostolos Filippas
Fordham

John J. Horton
MIT & NBER

Benjamin S. Manning
MIT

February 12, 2025

Abstract

Large language models (LLM)—because of how they are trained and designed—are implicit computational models of humans—a *homo silicus*. LLMs can be used like economists use *homo economicus*: they can be given endowments, information, preferences, and so on, and then their behavior can be explored in scenarios via simulation. Experiments using this approach, derived from Charness and Rabin (2002), Kahneman et al. (1986), and Samuelson and Zeckhauser (1988) show qualitatively similar results to the original, but it is also easy to try variations for fresh insights. LLMs could allow researchers to pilot studies via simulation, first improving their experimental design and searching for novel social science insights to test in the real world. In this paper, we offer a framework for when this approach is likely to prove useful.

AI with existing methodologies.

Using ML to predict as part of a paper

- Lots of papers use "Double-ML."

- Use prediction to tell part of the story.

- Use neural network architecture as part of a classic econometrics problem.

- Plug in embedding into another model.

- Use ML as part of a market design, such as better targeting.

Double/debiased machine learning for treatment and structural parameters
V Chernozhukov, D Chetverikov, M Demirer, E Duflo… - 2018 - academic.oup.com
… **ML** methods perform well by employing regularization to … θ 0 that are obtained by naively
plugging **ML** estimators of η 0 into … We call the resulting set of methods **double** or debiased **ML** (…
☆ Save 🗦🗦 Cite Cited by 3580 Related articles All 29 versions

Diagnosing physician error: A machine learning approach to low-value health care
S Mullainathan, Z Obermeyer
The Quarterly Journal of Economics, 2022 • academic.oup.com

Deep Neural Networks for Estimation and Inference

Authors	Max H Farrell, Tengyuan Liang, Sanjog Misra
Publication date	2021
Journal	Econometrica

Demand estimation with text and image data
G Compiani, I Morozov, S Seiler
arXiv preprint arXiv:2503.20711, 2025 • arxiv.org

Machine learning who to nudge: causal vs predictive targeting in a field experiment on student financial aid renewal

Authors	Susan Athey, Niall Keleher, Jann Spiess
Publication date	2025/1/13
Journal	Journal of Econometrics

The Mixed Subjects Design: Treating Large Language Models as Potentially Informative Observations

David Broska [ID] [✉], Michael Howes [ID], and Austin van Loon [ID] View all authors and affiliations

Contents | Get access | Cite article | Share options | Information, rights and permissions | Metric

Get access

Abstract

Large language models (LLMs) provide cost-effective but possibly inaccurate predictions of human behavior. Despite growing evidence that predicted and observed behavior are often not *interchangeable*, there is limited guidance on using LLMs to obtain valid estimates of causal effects and other parameters. We argue that LLM predictions should be treated as potentially informative observations, while human subjects serve as a gold standard in a *mixed subjects design*. This paradigm preserves validity and offers more precise estimates at a lower cost than experiments relying exclusively on human subjects. We demonstrate—and extend—prediction-powered inference (PPI), a method that combines predictions and observations. We define the *PPI correlation* as a measure of interchangeability and derive the *effective sample size* for PPI. We also introduce a power analysis to optimally choose between *informative but costly* human subjects and *less informative but cheap* predictions of human behavior. Mixed subjects designs could enhance scientific productivity and reduce inequality in access to costly evidence.

Large Language Models: An Applied Econometric Framework

Jens Ludwig, Sendhil Mullainathan & Ashesh Rambachan

How can we use the novel capacities of large language models (LLMs) in empirical research? And how can we do so while accounting for their limitations, which are themselves only poorly understood? We develop an econometric framework to answer this question that distinguishes between two types of empirical tasks. Using LLMs for prediction problems (including hypothesis generation) is valid under one condition: no "leakage" between the LLM's training dataset and the researcher's sample. No leakage can be ensured by using open-source LLMs with documented training data and published weights. Using LLM outputs for estimation problems to automate the measurement of some economic concept (expressed either by some text or from human subjects) requires the researcher to collect at least some validation data: without such data, the errors of the LLM's automation cannot be assessed and accounted for. As long as these steps are taken, LLM outputs can be used in empirical research with the familiar econometric guarantees we desire. Using two illustrative applications to finance and political economy, we find that these requirements are stringent; when they are violated, the limitations of LLMs now result in unreliable empirical estimates. Our results suggest the excitement around the empirical uses of LLMs is warranted – they allow researchers to effectively use even small amounts of language data for both prediction and estimation – but only with these safeguards in place.

Next time: final project presentations.