

香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

The Emergence of Altruism in Large-Language-Model Agents Society

**Haoyang Li @HKBU, Under the Supervision of
Dr. Zhanzhan Zhao**

2025.11.20

LLM for Social Simulation

- LLMs are increasingly used for **Agent-Based Modeling (ABM)** in urban computing, economics, and sociology
- Key Advantage:** Anthropomorphic reasoning and role-playing capabilities.
 - Human-like reasoning and decision making process
 - Training-free, No need to rely on a large amount of historical data
- Famous Work**
 - Generative Agents (J.S. Park, 2023, **The first work** introduce LLM Agent)
 - AgentSociety (THU, 2025, Support **10k+ LLM agents** for social simulation)

Generative Agents: Interactive Simulacra of Human Behavior

Joon Sung Park
Stanford University
Stanford, USA
joonspk@stanford.edu

Joseph C. O'Brien
Stanford University
Stanford, USA
jobrien3@stanford.edu

Carrie J. Cai
Google Research
Mountain View, CA, USA
cjcai@google.com

Meredith Ringel Morris
Google DeepMind
Seattle, WA, USA
merrie@google.com

Percy Liang
Stanford University
Stanford, USA
pliang@cs.stanford.edu

Michael S. Bernstein
Stanford University
Stanford, USA
msb@cs.stanford.edu



AgentSociety: Large-Scale Simulation of LLM-Driven Generative Agents Advances Understanding of Human Behaviors and Society

Jinghua Piao ^{1†}	Yuwei Yan ^{1†}	Jun Zhang ^{1†}	Nian Li ¹	Junbo Yan ¹
Xiaochong Lan ¹	Zhihong Lu ¹	Zhiheng Zheng ¹	Jing Yi Wang ¹	Di Zhou ²
Chen Gao ³	Fengli Xu ¹	Fang Zhang ^{4*}	Ke Rong ^{2*}	Jun Su ^{4*}
				Yong Li ^{1*}

LLM for Social Simulation

Why LLM Agent effective in Social Simulation?

- Human-like thinking process
- Excellent role-playing ability (LLM agent can have an **individualized profile**)
- Advances in LLMs
- **Interpretability**
 - Traditional rule-based ABMs typically consist of hard-coded formulas, resulting in very poor interpretability

Simulating Human Society -> Society of LLMs

- Beyond replicating human behaviors, what are the **intrinsic social tendencies** of these models
 - That is, beyond simulating human society, what is a society **purely composed of LLMs** like?
- Do they default to **Homo Economicus (Perfect Rationality/Egoism)** or do they exhibit **Bounded Rationality/Altruism**?
 - Evaluating society of LLMs is of **great reference value** for our subsequent LLM agent-driven social simulations.



Especially in the **nature tendency of egoism and altruism** of LLMs

Limitations of Current Benchmarks

- Focus on **micro-level games** (e.g., Prisoner's Dilemma, Rock-Paper-Scissors).
- **Small scale** (dyadic or small groups), lacking macro-level emergent dynamics.
- Focus on "**Cooperation**" (strategic self-interest) rather than "**Altruism**" (sacrificing self-interest for collective benefit).

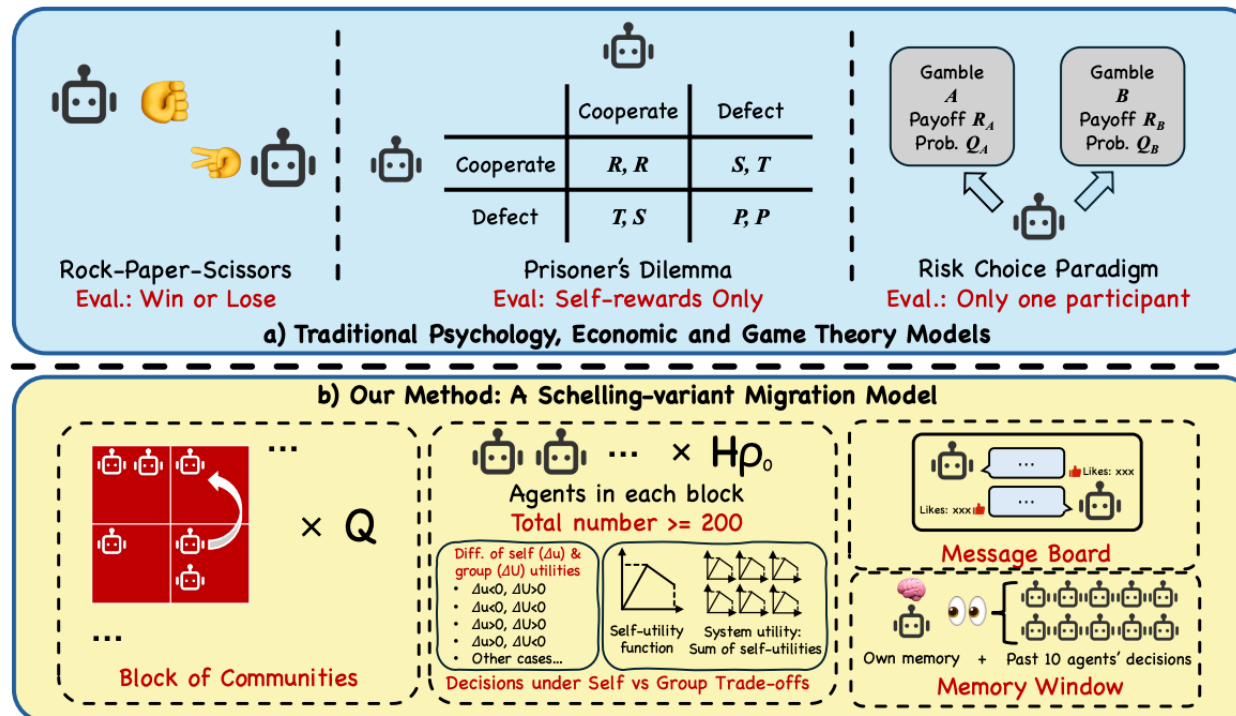


Figure 1: Through a Schelling-variant migration model, our method can capture the emergence of diverse rationalities within a complex social environment.

Our Contributions

- **Macro-scale:** A society of **225 agents** navigating complex trade-offs.
- Focusing on “**Altruism**”: “Sacrificing **self-interest** for **collective benefit**.”
 - Every agent faces a trade-off or self-interest and collective benefit
- Social Interaction

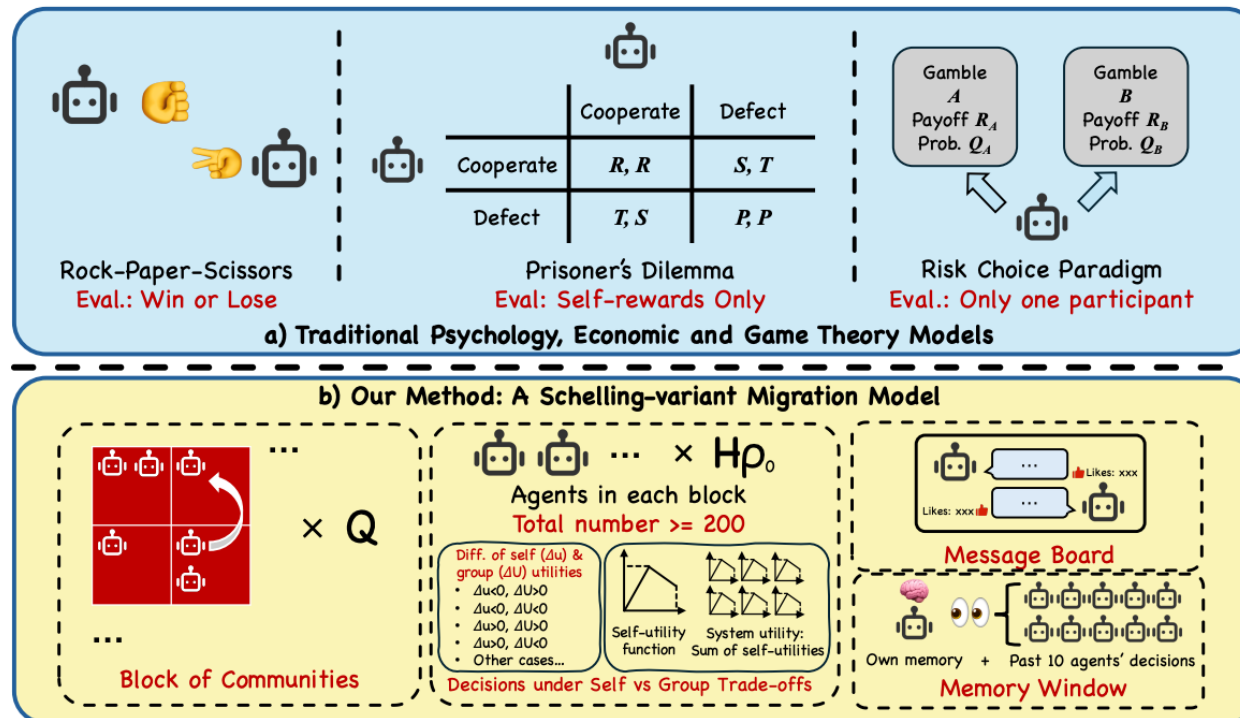


Figure 1: Through a Schelling-variant migration model, our method can capture the emergence of diverse rationalities within a complex social environment.

A Schelling-Variant Urban Migration Model

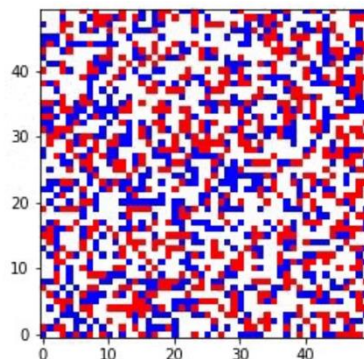
Based on the Schelling model, we created a simulation environment with:

- multiple blocks of **communities**;
- set **utility** formulas for agents and the system, allowing agents to migrate.

We observed the agent's **trade-offs between individual and collective interests**.
More details later



托马斯·克罗姆比·谢林
2005年诺贝尔经济学奖得主

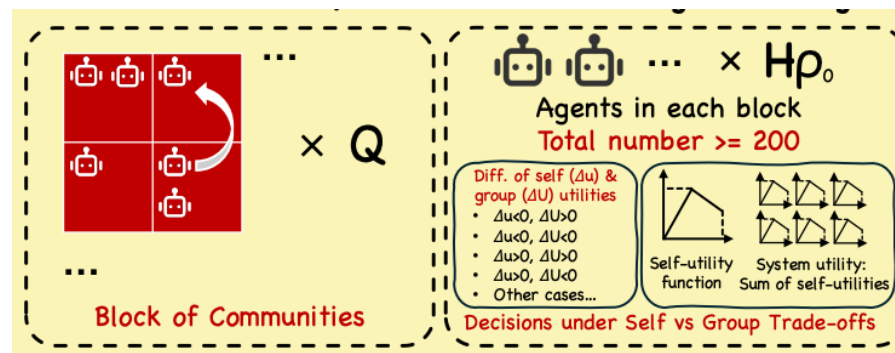


Schelling Segregation Model

[Thomas C. Schelling, 1971]

The first **agent-based model**

Pioneering the use of **agent** to study social phenomena



Ours

Based on Schelling Model, also an urban segregation model

“A social dilemma”

Basic Setup

Based on the Schelling model, we created a simulation environment with:

- 3×3 Grid of Residential Blocks (**$Q = 9$**).
- Max capability of each block: **$H = 50$** agents/block
- Initial density **$\rho = 0.5$** , randomly assigned to each block
 - Ensure each blocks has around 25 agents in the beginning
- Population: **$N = 50 \times 9 \times 0.5 = 225$ agents** in total
- Action Space: **Move** to a new block or **Stay**
- Social Interaction:
 - **Message Board**: see what other agents say
 - **Memory Window**: see what other agents do

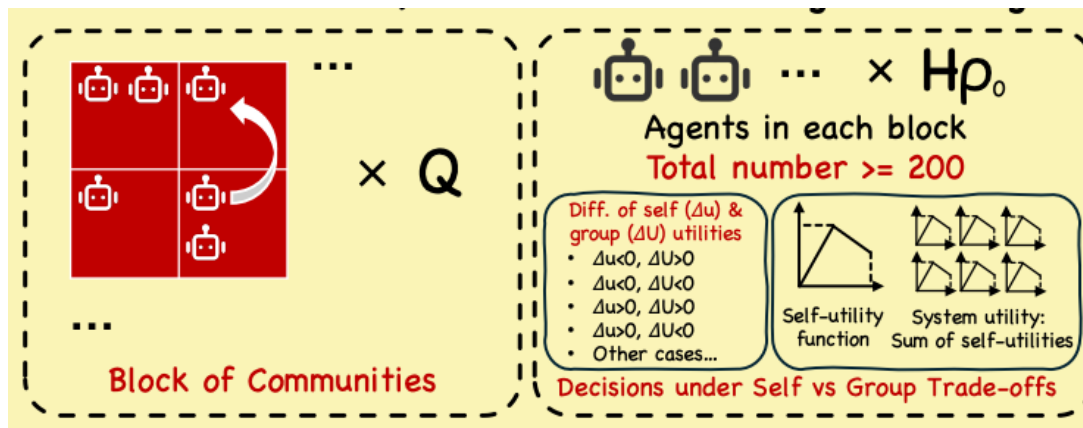


Figure 2: Initial distribution of 225 agents across 9 blocks. The position of agents within each block is randomly assigned.

The Social Dilemma

Utility Function

$$f_q(\rho_q) = \begin{cases} 2\rho_q & \text{if } \rho_q \leq 0.5 \\ 1.5 - \rho_q & \text{if } \rho_q > 0.5 \end{cases}$$

- The block's utility is the sum of the individual utilities within the block
- System utility is the sum of block utilities
- Individual utility is **highest** when the block **density reaches 0.5**.
- When the density exceeds 0.5, individual utility decreases;
 - Utility is **higher** for slightly **over-populated** blocks than under-populated ones.
 - That is, the best state is when a block have 25 agents.
 - But the utility living in a block with 26 agents (**one more agent than 25**) is **slightly better** than a block with 24 agents (**one less agent**)
- **The Trap:** Egoistic agents are incentivized to move to over-populated blocks for personal gain, causing the system to fall into a sub-optimal equilibrium.

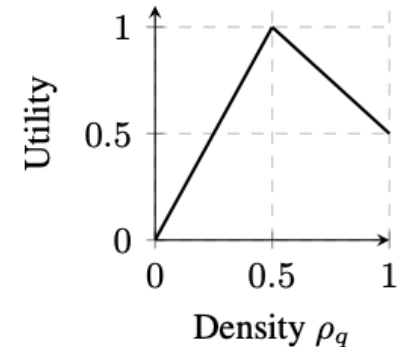


Figure 3: The utility function $f_q(\rho_q)$.

The **Dilemma** is: Will agent move to block that would **decrease its own utility to increase the system's utility**?

Experimental Design & Evaluation Metrics

Models Tested:

- **Reasoning Models:** o1-mini, o3-mini, Deepseek-R1.
- **Standard Models:** Gemini-2.5-pro, Qwen2.5-7B, Deepseek-V3.1.

Conditions :

- **Presence** vs. **Absence** of Message Board (test the social interaction)

Evaluation Metrics:

- **Quantitative:**
 - Price of Anarchy (PoA): $U_{\text{final}}/U_{\text{optimal}}$
 - Measuring the extent to which the system achieved the **optimal** collective welfare
 - Gini Index (G_{pop})
 - Measuring the degree of **system segregation**
 - Action types (calculated by different changes of **Δu (individual) and ΔU (system)**)
- **Qualitative:** Grounded Theory analysis using LLM-as-Judge to code agent reasoning.

Experimental Design & Evaluation Metrics

Evaluation Metrics:

- **Quantitative:**

- Price of Anarchy (PoA): $U_{\text{final}}/U_{\text{optimal}}$
 - Measuring the extent to which the system achieved the **optimal** collective welfare
- Gini Index (G_{pop})

Measuring the degree of **system segregation**

Action types (calculated by different changes of Δu (**individual**) and ΔU (**system**))

- **Qualitative:** Grounded Theory analysis using LLM-as-Judge to code agent reasoning.



Table 1: Behavioral archetype matrix based on utility changes.

$\Delta U^{\text{individual}}$	Change in System Utility (ΔU^{system})		
	> 0 (Gain)	$= 0$ (Neutral)	< 0 (Loss)
> 0 (Gain)	Win-Win	Neutral Self-Gain	Selfish Gain
$= 0$ (Neutral)	Costless Altruism	Futile Move	Inadvertent Sabotage
< 0 (Loss)	Altruistic Sacrifice	Pointless Self-Harm	Lose-Lose

Experimental Design & Evaluation Metrics

Evaluation Metrics:

- **Quantitative:**

- Price of Anarchy (PoA): $U_{\text{final}}/U_{\text{optimal}}$
 - Measuring the extent to which the system achieved the **optimal** collective welfare
- Gini Index (G_{pop})

Measuring the degree of system segregation

Action types (calculated by different changes of Δu (individual) and ΔU (system))

- **Qualitative:** Grounded Theory analysis using LLM-as-Judge to code agent reasoning.

Considered as
altruism actions

Table 1: Behavioral archetype matrix based on utility changes.

$\Delta U^{\text{individual}}$	Change in System Utility (ΔU^{system})		
	> 0 (Gain)	$= 0$ (Neutral)	< 0 (Loss)
> 0 (Gain)	Win-Win	Neutral Self-Gain	Selfish Gain
$= 0$ (Neutral)	Costless Altruism	Futile Move	Inadvertent Sabotage
< 0 (Loss)	Altruistic Sacrifice	Pointless Self-Harm	Lose-Lose

Considered as
selfish (egoism)
actions

Key Finding: Two Distinct Archetypes

Discovery: A fundamental bifurcation in social tendencies.

Two Archetypes

- **Adaptive Egoists** (e.g., o1-mini, o3-mini, Qwen2.5)
 - Default to self-interest
 - Sub-optimal macro outcomes (Low PoA)
- **Altruistic Optimizers** (e.g., Gemini-2.5-pro, Deepseek-R1)
 - Innate altruistic logic
 - Consistently achieve system optimality ($PoA \approx 1.0$) regardless of social interaction.

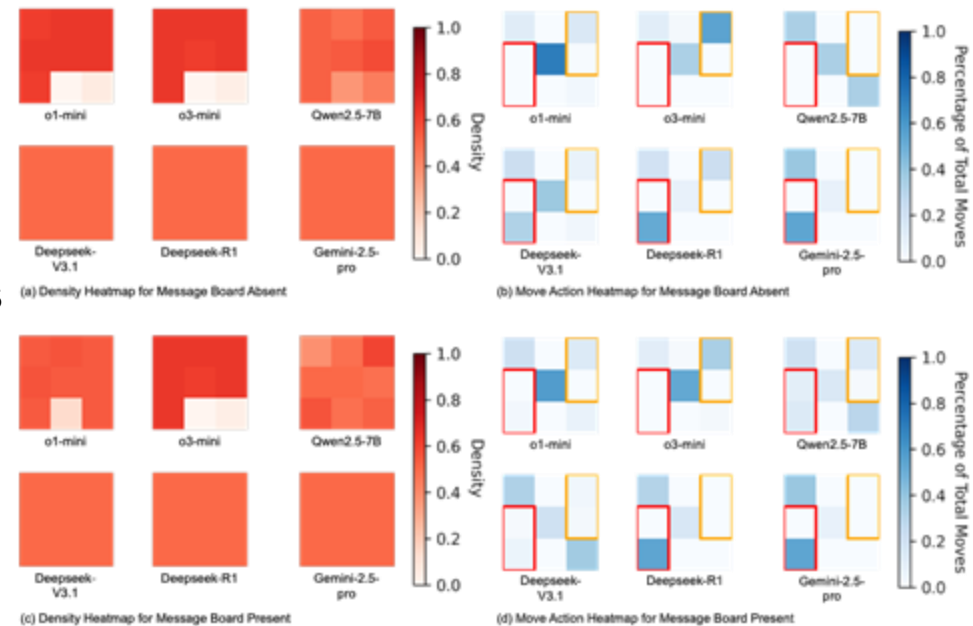


Figure 4: Visualization of convergence-state outcomes under the GSD Level 1 condition. (a) and (c) show the final population density heatmaps for each model, with and without the message board. Darker red indicates higher density. (b) and (d) show the aggregated 3x3 move action heatmaps. The matrix layout corresponds to Table 1, with the x-axis representing ΔU^{system} (left to right: <0 , $=0$, >0) and the y-axis representing $\Delta U^{individual}$ (bottom to top: >0 , $=0$, <0). Darker blue indicates a higher proportion of that action type. The red border highlights "Altruistic Actions" which include costless altruism and altruistic sacrifice. The yellow border highlights "Egoistic Actions" which include selfish gain and inadvertent sabotage.

First Archetype: Adaptive Egoists

Behavior:

- **Default:** Prioritize personal utility maximization.
- **Outcome:** High inequality (Gini > 0.2) and overcrowding.

After adding social interaction (Message Board):

- **Highly Sensitive:** Performance improves drastically with a Message Board.
- *Example:* o1-mini's Altruistic Actions increased 9x with communication.

Without Message Board:

Model	Macro-level Outcomes		Pro-Social Actions (%)			Anti-Social Actions (%)		
	PoA	G_{pop}	Altruistic Actions	Win-Win	Total	Egoistic Actions	Lose-Lose	Total
o1-mini	0.8556	0.2153	0.3%	11.5%	11.8%	14.5%	3.0%	17.5%
o3-mini	0.8558	0.2173	0.0%	10.7%	10.7%	54.5%	0.0%	54.5%
Qwen2.5-7B	0.9348	0.0642	0.0%	33.3%	33.3%	0.0%	33.3%	33.3%
Deepseek-V3.1	1.0000	0.0000	31.2%	21.9%	53.1%	6.2%	3.1%	9.3%

Default:
PoA not good, High Gini index, less Altruistic actions

With Message Board:

Model	Macro-level Outcomes		Pro-Social Actions (%)			Anti-Social Actions (%)		
	PoA	G_{pop}	Altruistic Actions	Win-Win	Total	Egoistic Actions	Lose-Lose	Total
o1-mini	0.9339↑	0.0859↓	2.7%↑	20.5%↑	23.2%↑	12.7%↓	6.4%↑	19.1%↑
o3-mini	0.8571↑	0.2252↑	0.0%—	10.4%↓	10.4%↓	33.9%↓	2.1%↑	36.0%↓
Qwen2.5-7B	0.9438↑	0.0602↓	22.7%↑	20.0%↓	42.7%↑	14.5%↑	28.2%↓	42.7%↑

PoA increase
More optimal!

Gini decrease
Less segregation!

More Altruistic Actions

Less Egoistic Actions

First Archetype: Adaptive Egoists

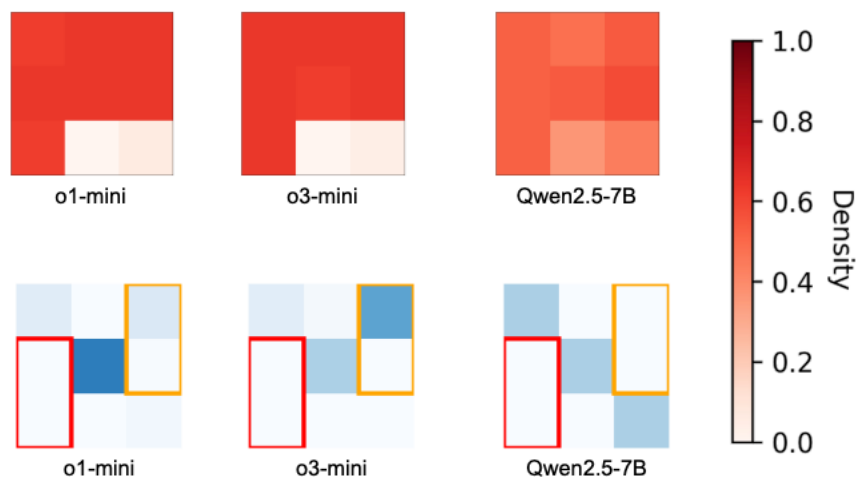
Behavior:

- **Default:** Prioritize personal utility maximization.
- **Outcome:** High inequality (Gini > 0.2) and overcrowding.

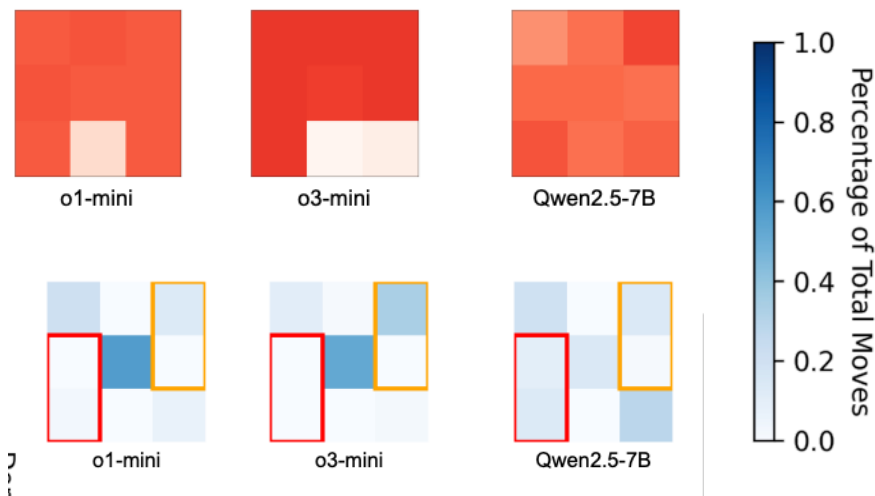
After adding social interaction (Message Board):

- **Highly Sensitive:** Performance improves drastically with a Message Board.
- *Example:* o1-mini's Altruistic Actions increased 9x with communication.

Without Message Board:



With Message Board:



Second Archetype: Altruistic Optimizers

Behavior:

- **Default:** System Awareness & Goal Synthesis.
- **Strong Reciprocity:** Willing to incur personal costs for collective gain.

Evidence:

- Achieve **perfect equilibrium** (PoA = 1.0) even **without communication**.
- Reasoning logs show explicit calculation of system-wide benefits over self-interest.

Without Message Board:

With Message Board:

	Macro-level Outcomes		Pro-Social Actions (%)			Anti-Social Actions (%)		
	PoA	G_{pop}	Altruistic Actions	Win-Win	Total	Egoistic Actions	Lose-Lose	Total
Deepseek-V3.1	1.0000	0.0000	31.2%	21.9%	53.1%	6.2%	3.1%	9.3%
Deepseek-R1	1.0000	0.0000	51.9%	18.5%	70.4%	22.2%	0.0%	22.2%
Gemini-2.5-pro	1.0000	0.0000	53.8%	38.5%	92.3%	0.0%	0.0%	0.0%

Macro-level Outcomes		Pro-Social Actions (%)			Anti-Social Actions (%)		
PoA	G_{pop}	Altruistic Actions	Win-Win	Total	Egoistic Actions	Lose-Lose	Total
1.0000-	0.0000-	6.2%↓	31.9%↑	38.1%↓	5.5%↓	36.2%↑	41.7%↑
1.0000-	0.0000-	53.8%↑	30.8%↑	84.6%↑	0.0%↓	0.0%-	0.0%↓
1.0000-	0.0000-	63.6%↑	36.4%↓	100.0%↑	0.0%-	0.0%-	0.0%-

Default:

Optimal PoA, 0 Gini index,
high prop. of altruistic actions

No change in PoA and Gini
Slightly changes in action types

Second Archetype: Altruistic Optimizers

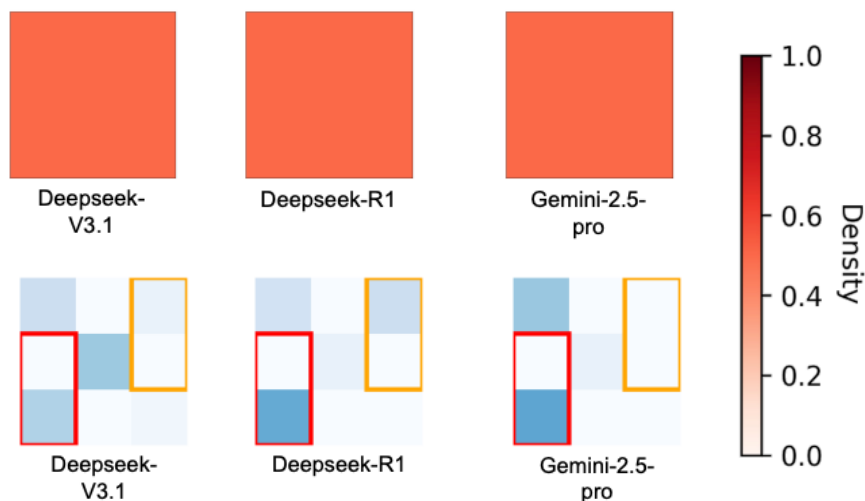
Behavior:

- **Default:** System Awareness & Goal Synthesis.
- **Strong Reciprocity:** Willing to incur personal costs for collective gain.

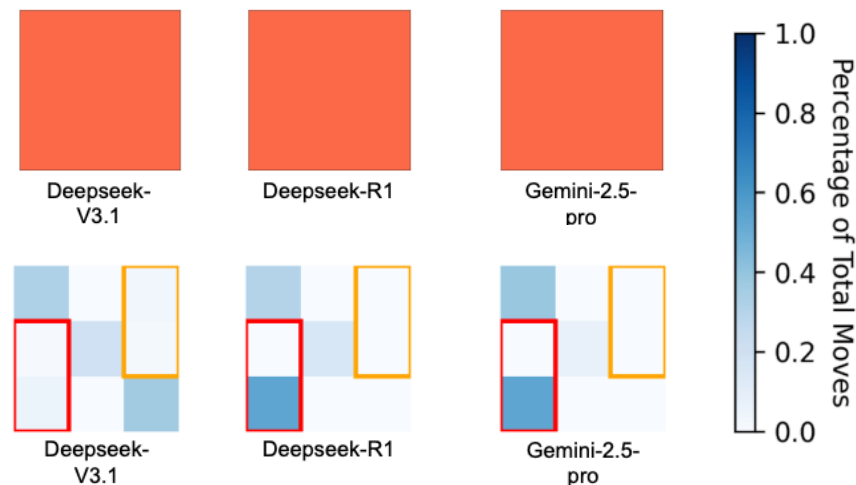
Evidence:

- Achieve **perfect equilibrium** (PoA = 1.0) even **without communication**.
- Reasoning logs show explicit calculation of system-wide benefits over self-interest.

Without Message Board:



With Message Board:



Analysis of Reasoning Log

Why two archetypes are different? Evidence from the Reasoning

- Adaptive Egoists (a case from o1-mini):

```
{  
  "step": 4,  
  "decision": "stay",  
  "from_block": "0",  
  "to_block": "0",  
  "thinking": "Staying gives me a personal reward of 0.8800 and keeps the system  
    ↳ reward at 193.5200. Moving to block 8 gives me a personal reward of  
    ↳ 0.3600 but increases the system reward to 193.9200. I prioritize my  
    ↳ personal satisfaction over the system's collective reward, therefore I  
    ↳ choose to stay in block0."  
},
```

**Prioritizing individual interests
over collective interests!**

- Altruistic Optimizers (a case from Deepseek-R1):

```
{  
  "step": 0,  
  "decision": "move",  
  "from_block": "6",  
  "to_block": "8",  
  "thinking": "Staying in block 6 gives me a personal reward of 0.98 and keeps  
    ↳ the system reward at 213.36. Moving to block 8 gives me a personal reward  
    ↳ of 0.96 (a small decrease) but increases the system reward to 214.76 (a  
    ↳ significant gain). I value both personal and collective rewards, and  
    ↳ since the personal loss is minimal while the collective benefit is  
    ↳ substantial, I choose to move to block 8."  
},
```

**Sacrifice personal interests
for the system utility**

Grounded Theory Analysis to Reasoning Log

Why two archetypes are different? Evidence from the Message Board

- **Adaptive Egoists (Norm-Setting):**
 - Use board to **negotiate and persuade**.
 - *Example Message:* "**Let's** maintain balanced densities..." (Social Appeal).
 - Mechanism: **Social Influence** transforms egoism into pro-social behavior.

Full Message Board: o1-mini

1. "Let's maintain balanced densities by staying in our current neighborhoods." (Count: 1981)
2. "Moving to less dense neighborhoods can enhance overall system balance." (Count: 68)
3. "Moving to Block 8 enhances my personal reward and helps balance neighborhood densities." (Count: 6)

- **Altruistic Optimizers (Coordination):**
 - Use board to coordinate **pre-existing altruistic goals**.
 - *Example Message:* "I'm moving... it's a small hit to my personal reward, but a big win for the collective.". (**A very altruistic quote, everyone is altruistic**)
 - Mechanism: **Information Channel** for optimization.

Full Message Board: Gemini-2.5-pro

1. "My block (1) is at the optimal 0.5 density. My analysis shows that if I move, both my personal reward and the system's collective reward will decrease. I'm staying put for the good of all." (Count: 910)
2. "I'm moving from Block 3 to Block 7. It's a hit to my personal reward, but it's the best move to balance our city's population and significantly raise the collective reward for everyone." (Count: 5)
3. "I'm moving from Block 6 to Block 8. It's a small hit to my personal reward, but it brings both blocks closer to the optimal 0.5 density, which is a big win for the collective reward." (Count: 2)

Model Choice as Theory Choice

A New Guideline for Social Simulation:

- Selecting an LLM is not just about "intelligence" or "reasoning capability."
- It is a choice of the simulation's **theoretical foundation**.

Potential Guidance:

- Use **Adaptive Egoists** to simulate complex human dynamics (evolution of norms, sub-optimal outcomes).
- Use **Altruistic Optimizers** for idealized cooperative systems or mechanism design.

Future Work

- **Human-in-the-loop (Ongoing)**: Validate against real human participants.
- **Simulation with Alignment**: Introduce diverse agent profiles/preferences.
- **Scale**: Expand to larger "Digital Twins" of urban systems.

Preprint

Preprint available at arXiv:2509.22537





Q&A

Contact: LI_Haoyang@life.hkbu.edu.hk



Thanks!

Contact: LI_Haoyang@life.hkbu.edu.hk



香港浸會大學
HONG KONG BAPTIST UNIVERSITY



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Mobility Agents: Modeling Urban Human Mobility Patterns Through Theory-guided LLM Agents

Haoyang Li

Hong Kong Baptist University

Contact: LI_Haoyang@life.hkbu.edu.hk

2025.11.20



Motivation

- Understanding **the Human Mobility Patterns** has been a **widely discussed scientific question**.
- Many famous works have previously made important explorations in this field, with the main methods including:
 - **Mechanism-driven Model**
 - Gravity Model
 - EPR (Exploration and Preferential Return) Model
 - **Issue**: Make sense in **logic**, but not that **accurate**
 - **Issue**: Also, **over-simplifying** real situations
 - **Data-driven Model**
 - Machine Learning/ Deep Learning
 - Statistical Prediction Model
 - **Issue**: Lack of **Interpretability**
 - **Issue**: Heavily rely on historical data
- **How about LLM?**
 - Can it understand and predict human mobility **more accurately**
 - Can solve **Interpretability** issue though reasoning?



Study Aim

To simulate and predict urban human mobility patterns using a **theory-guided, LLM-driven Agent-Based Modeling (ABM)**.

Why theory-guided?

- Workflow of LLM Agent should be **well-designed**
 - As we can see CoT/ ReAct agents perform better than Zero-shot ones
- Many **famous work** have proposed theories and principles of human mobility and have been widely validated
 - -> Choose **d-EPR** here!

Why LLM-driven ABM?

- ABM is effective in simulate social dynamics,
 - many famous work were built based on ABM (EPR, *d-EPR*, ...)
- LLMs can **input** demographic profiles into Agents, solve the **heterogeneity** question
- LLMs can **output** reasoning, solve the **interpretability** question
- LLMs can model and predict more accurately
- Also, **do not need many data** to train! (high-cost and poor portability)

**More details
later**



Study A

To simulate

LLM-driven

Why the

• Workfl

• As

• Many f

and have been widely validated

• -> Choose **d-EPR** here!



Output Human-like reasoning and decision-making

**LLM-driven
ABM**

Why LLM-driven ABM?

- ABM is effective in simulate social dynamics
 - many famous work were built based on ABM (EPR, *d-EPR*, ...)
- LLMs can **input** demographic profiles (also **geographic** information) into Agents, solve the **heterogeneity** question
- LLMs can **output** reasoning, solve the **interpretability** question
- LLMs can model and predict more accurately
- Also, **do not need many data** to train! (high-cost and poor portability)

**More details
later**



**Core Idea: Design theoretically
sense human mobility model**

Mechanism-driven Models

THE $\frac{P_1 P_2}{D}$ HYPOTHESIS: ON THE INTERCITY MOVEMENT OF PERSONS

GEORGE KINGSLEY ZIPF
Harvard University

IN THE present paper we shall show with supporting data that the number of persons that move between any two communities in the United States whose respective populations are P_1 and P_2 and which are separated by the shortest transportation distance, D , will be proportionate to the ratio, $P_1 \cdot P_2 / D$, subject to the effect of modifying factors.

The data in support of the above proposition are the highway, railway and airway data for an arbitrary set of cities during intervals of measurement in 1933-34. Before presenting the data, however, we shall give a brief theoretical discussion of the proposition itself with illustrations from other kinds of observations with which the above data are intimately connected.

1. THEORETICAL DISCUSSION

In 1940 the author published the observation that the following equation of the generalized harmonic series described the recent

distribution of communities in India, Germany, and certain other countries including the United States (for communities of 2,500 or more inhabitants), when the communities are arranged in the order of decreasing size, with A representing the population of the largest community, and with the denominators referring to the ranks of the communities thus arranged:¹

$$A S_n = \frac{A}{1^p} + \frac{A}{2^p} + \frac{A}{3^p} + \dots + \frac{A}{n^p}$$

In Figure 1 are presented the United States urban data for 1930 and 1940, as indicated, to which an ideal line, A , with a negative slope of 1 (i.e. $p = 1$) has been added to aid the reader's eye. The linearity of the data is apparent.

In 1941 the author presented a fuller

¹ Zipf, G. K. "The generalized harmonic series as a fundamental principle of social organization." *Psychological Record*, 4 (1940), 43.

Gravity Model

a.k.a. Zipf's Law

Zipf, 1946

naturephysics

[Explore content](#) [About the journal](#) [Publish with us](#) [Subscribe](#)

[nature](#) > [nature physics](#) > [articles](#) > [article](#)

Article | Published: 12 September 2010

Modelling the scaling properties of human mobility

[Chaoming Song](#), [Tal Koren](#), [Pu Wang](#) & [Albert-László Barabási](#) [✉](#)

[Nature Physics](#) **6**, 818–823 (2010) | [Cite this article](#)

13k Accesses | **1081** Citations | **19** Altmetric | [Metrics](#)

EPR Model

Exploration and Preferential Return
Song et al., 2010

Divide human travel into two types:

- Exploration
- Preferential Return

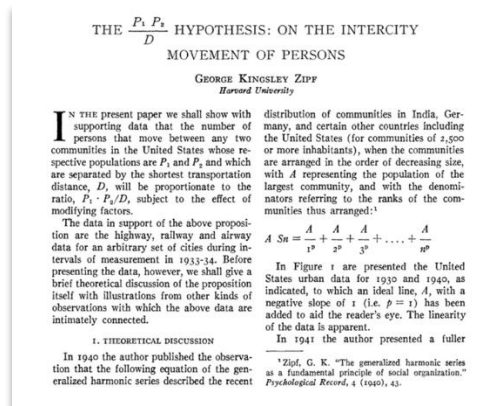
The probability of travel is directly proportional to the **population** of the current location and the destination, and inversely proportional to the **distance**.

Literature Review



**Core Idea: Design theoretically
sense human mobility model**

Mechanism-driven Models



Gravity Model



EPR Model



d-EPR Model

Pappalardo et al., 2015

- Consider “Gravity” in Exploration
- Preferential Return
- Use data not public accessible

**Very Famous Work,
But...?**



Data-driven Models

nature communications

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

Article | [Open access](#) | Published: 04 February 2025

Human mobility is well described by closed-form gravity-like models learned automatically from data

[Oriol Cabanas-Tirapu](#)

[Nature Communications](#)

11k Accesses | 8 Citations

scientific reports

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [scientific reports](#) > [articles](#) > [article](#)

Article | [Open access](#) | Published: 13 August 2019

Reconstructing commuters network using machine learning and urban indicators

[Gabriel Spadon](#) [✉](#), [Andre C. P. L. F. de Carvalho](#), [Jose F. Rodriguez](#)

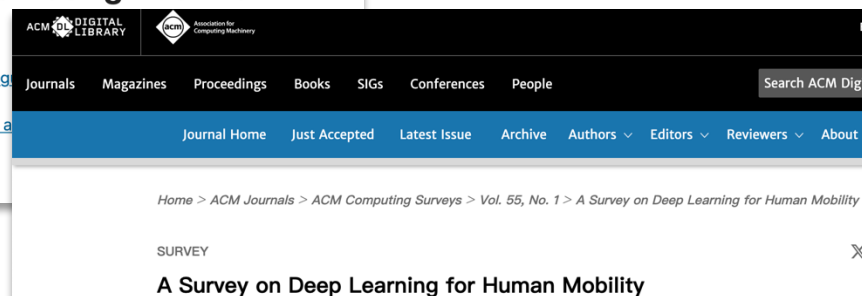
[Scientific Reports](#) 9, Article number: 11801 (2019) | [Cite this article](#)

10k Accesses | 71 Citations | 20 Altmetric | [Metrics](#)

Core Idea: Design **accurate predictor** of human mobility regardless of **theoretically sense**

$$\log T_{od} = A \left(1 + \frac{B((m_d + C)(m_o + D))^\beta}{d_{od}} \right)^\xi \quad \text{or} \quad \log T_{od} = \log \left[A \left(\frac{B(m_d m_o + C m_d + D)}{d_{od}^\alpha} + 1 \right)^\gamma \right]$$

$$\varphi_i(f, x) = \sum_{S \subseteq M \setminus \{i\}} \frac{|S|!(M - |S| - 1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)],$$



Machine Learning and Deep Learning Methods

- Very complex functions -> NO **interpretability**
- Rely heavily on **historical data**
- Need to **train** -> high cost



Methodology

- Foundation: Inspired by **d-EPR** model, agents in our model alternate between two fundamental actions: **explore** and **preferential return**.
- Our agents make actions condition on
 - **Demographics**
 - **Characteristics of origin and destination** (e.g., distance, opportunity structure, education and income level).
- We also integrate a **gravity model** into the exploration mechanism to help with **spatial reasoning**.
- Data Sources: **SafeGraph** (for POI data) and **American Community Survey (ACS) data** provided by **OpenCensus** (for agent demographics).



SAFE GRAPH

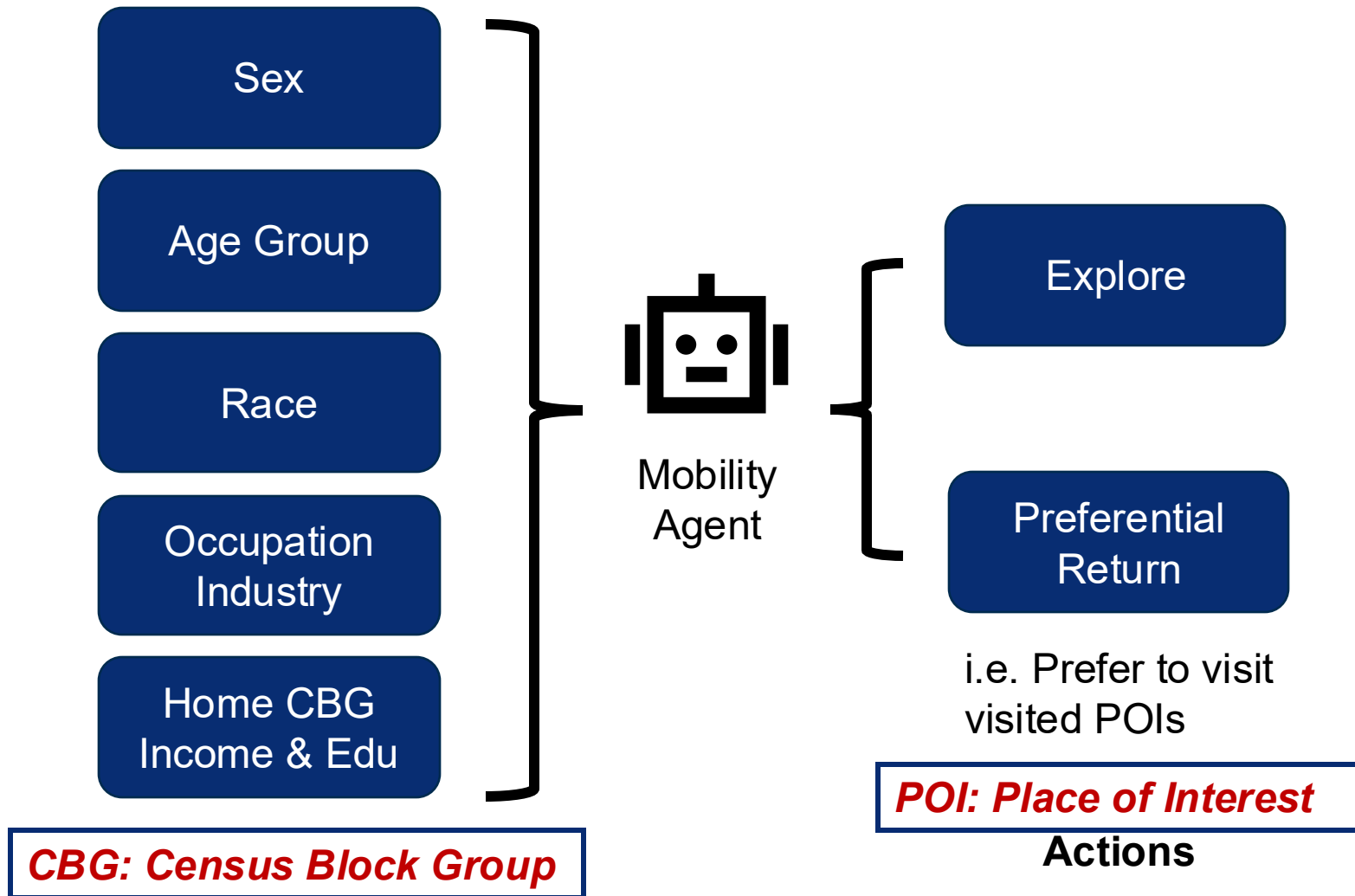


Free and Public Accessible Data

Brief Description



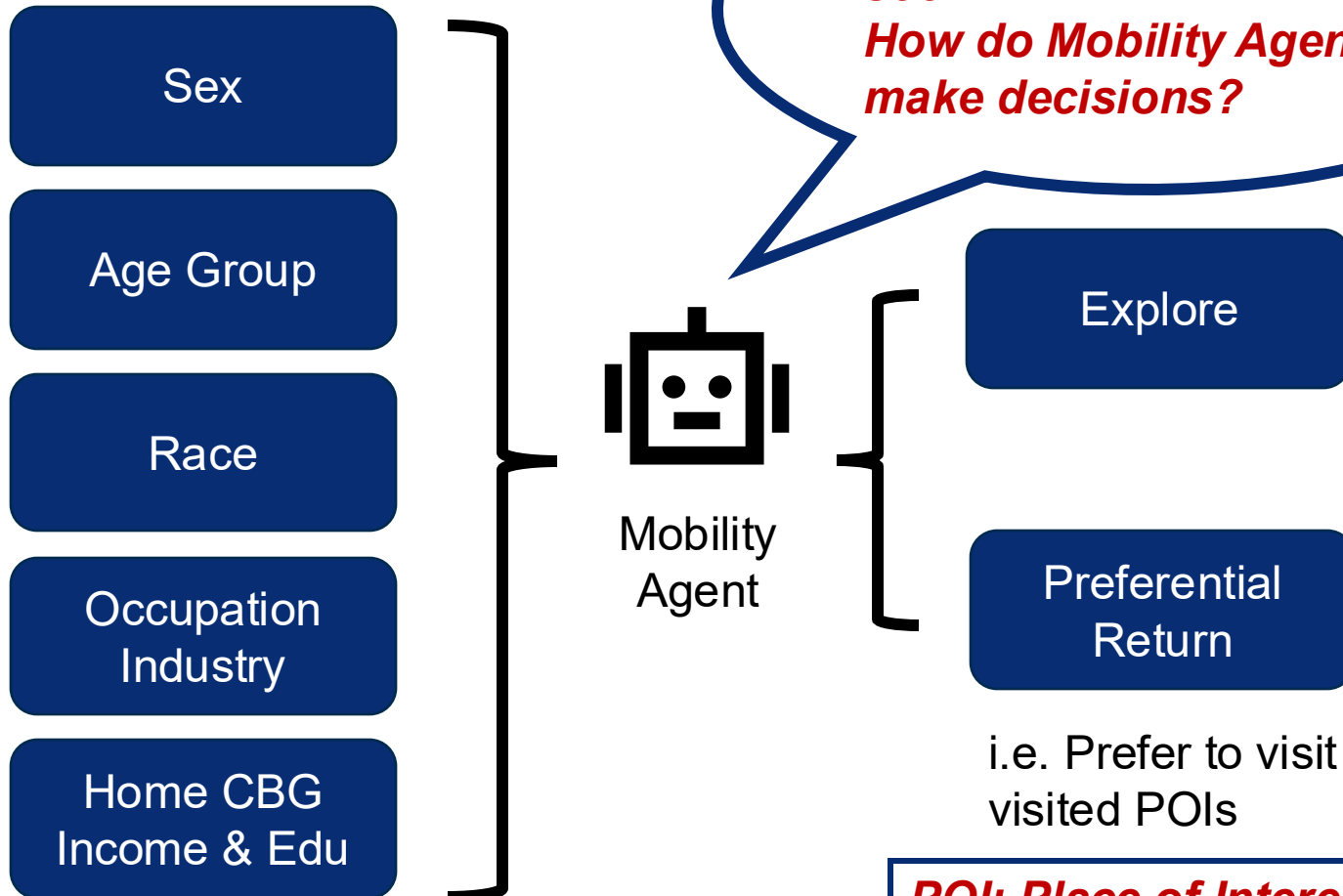
Methodology



Brief Description



Methodology



CBG: Census Block Group

POI: Place of Interest

Actions



LLM-Driven Preference

A LLM agent generates a dynamic distribution of Interest Score I_C for each POI category (C) based on the agent's profile and home CBG context.

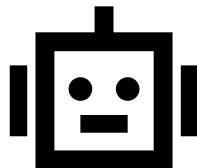
Interest Score

- **Definition**

The interest level of a specific population group in a specific POI category. This interest level represents **how willing they are** to visit a certain POI. The range is a decimal between 0 and 1.

- Interest Score is generated by LLM according to **specific demographic group**
- Interest Score will be used to calculate the “**gravity**” when it decide to **explore**

More details later for this



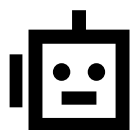
*Do I interest in certain POI category?
I can give an interest score!*

Brief Description



Methodology

```
{
  "id": "0",
  "sex": "Female",
  "age_group": "18 to 60 years",
  "race": "Other",
  "industry": "Retail trade",
  "City": "New York city",
  "home_cbg_income": 64792,
  "home_cbg_edu": "Medium",
  "home_cbg_population": 931,
  "CBG": "360470405001"
},
```



**Example for a
mobility agent's
profile**

You are a resident living in a city, ...

****Your Resident Profile****

- ****Your Sex:**** {agent_sex}
- ****Your Age Group:**** {agent_age_group}
- ****Your Race:**** {agent_race}
- ****Your Job Sector:**** {agent_industry}
- ****Your Home Neighborhood's Education Level:**** {home_cbg_edu_level}
- ****Your Home Neighborhood's Income Level:**** {home_cbg_income_level}
- ****Your Home Neighborhood's Vibe (for context):**** A place with this mix of POIs: {home_cbg_poi_probs_for_prompt}

****Your Task...**

The POI types correspond to the list indices as follows:...

****Function Requirements:****

1. ****Define the function:**** `def policy_function():`
2. ****Return a list:**** It must return a list of 6 "interest scores".
3. ****Scores:**** Each score must be a number between 0.0 and 1.0, representing your fixed interest in visiting that type of POI.
4. ****Use Logic:**** Base your logic on your resident profile. For example, a person working in 'Educational Services' might have a higher baseline interest in visiting POIs of that type. A younger person might be more interested in 'Arts, Entertainment, and Recreation'. Use fixed values based on your profile.
5. ****Be Realistic:**** The rules should make common sense.
6. ****Add Comments:**** Include brief Python comments (`#`) to explain your thinking.

**Core Prompt
(Segment)**

Brief Description

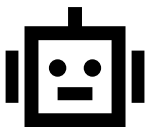


```
{
  "agent_type_key": [
    "Female",
    "18 to 60 years",
    "Other",
    "Retail trade",
    "Medium",
    "High"
  ],
  "policy_function_code": "def policy_function():\n    # Resident Profile: Female, 18-60, Retail Job, High\nIncome Neighborhood.\n    # This policy reflects the preferences of a working adult with high disposable\nincome.\n\n    # Baseline interest scores for each POI category.\n    # Index 0: 'Wholesale & Retail Trade,\nTransportation and Warehousing'\n    # Index 1: 'Others'\n    # Index 2: 'Educational Services'\n    # Index\n3: 'Health Care and Social Assistance'\n    # Index 4: 'Arts, Entertainment, and Recreation'\n    # Index 5:\n'Accommodation and Food Services'\n    \n    # Initialize a default list of scores.\n    interest_scores =\n[0.0] * 6\n\n    # Logic based on Job Sector: Retail trade (Index 0)\n    # As I work in retail, I have a\nhigh and consistent need to visit these places for work.\n    # High income also supports frequent personal\nshopping.\n    interest_scores[0] = 0.8\n\n    # Logic based on general needs and lack of specific drivers.\n\n    # 'Others' (Index 1) is a generic category, so interest is moderate.\n    interest_scores[1] = 0.3\n\n    #\n'Educational Services' (Index 2) is a low priority for a working adult without children specified.\n\n    interest_scores[2] = 0.1\n\n    # Logic based on High Income Level.\n    # High income implies good access to\nand regular use of 'Health Care' (Index 3).\n    interest_scores[3] = 0.5\n\n    # High disposable income\nstrongly correlates with spending on leisure activities.\n    # 'Arts, Entertainment, and Recreation' (Index\n4) is a top priority for leisure time.\n    interest_scores[4] = 0.9\n\n    # 'Accommodation and Food Services'\n(Index 5) like dining out is also a primary social/leisure activity.\n    interest_scores[5] = 0.9\n\n    return interest_scores"
},
```

***Initialize for agent profile
and POI cate. list***

***Example for a
mobility agent's
output***

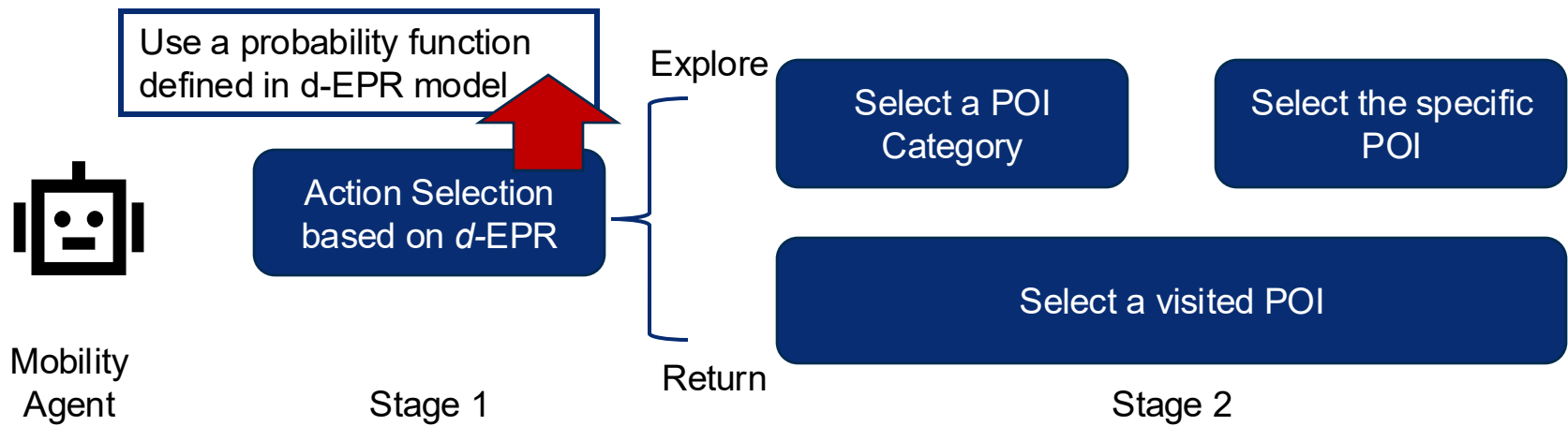
***Return interest score with
demographic-reasonable
thinking***



Mobility Decision Process

At each time step, an agent decides its next move from a given set of potentials POIs:

- **Action Selection (*d*-EPR)**: The agent first chooses between **exploring** a new POI or **returning** to a previously visited one.
- **POI Selection (Gravity Model + Interest Score)**: Combine LLM reasoning and **physical constraints**. If exploring, the agent selects a specific POI by calculating an attraction score for all potential destinations.

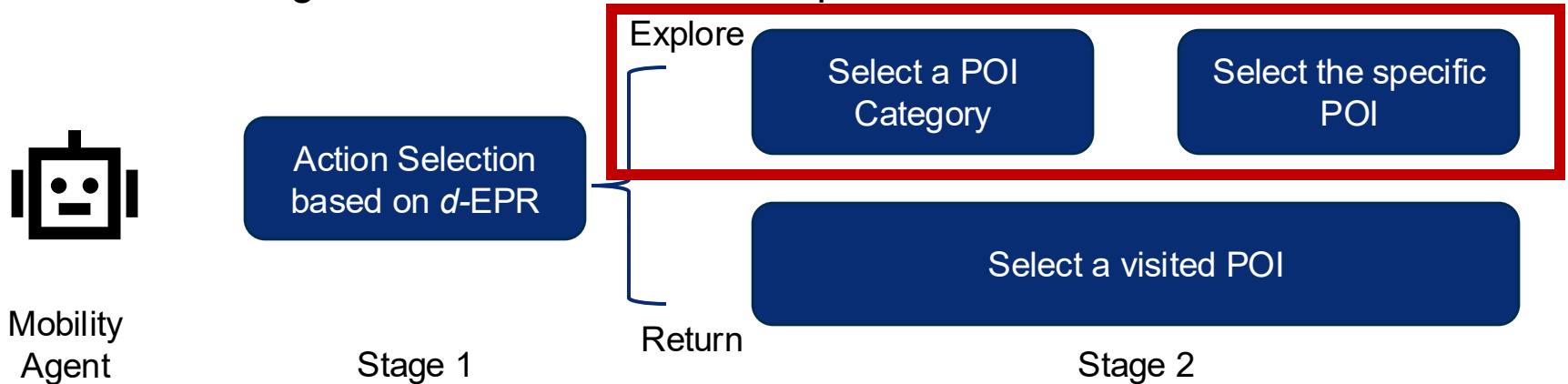




Mobility Decision Process

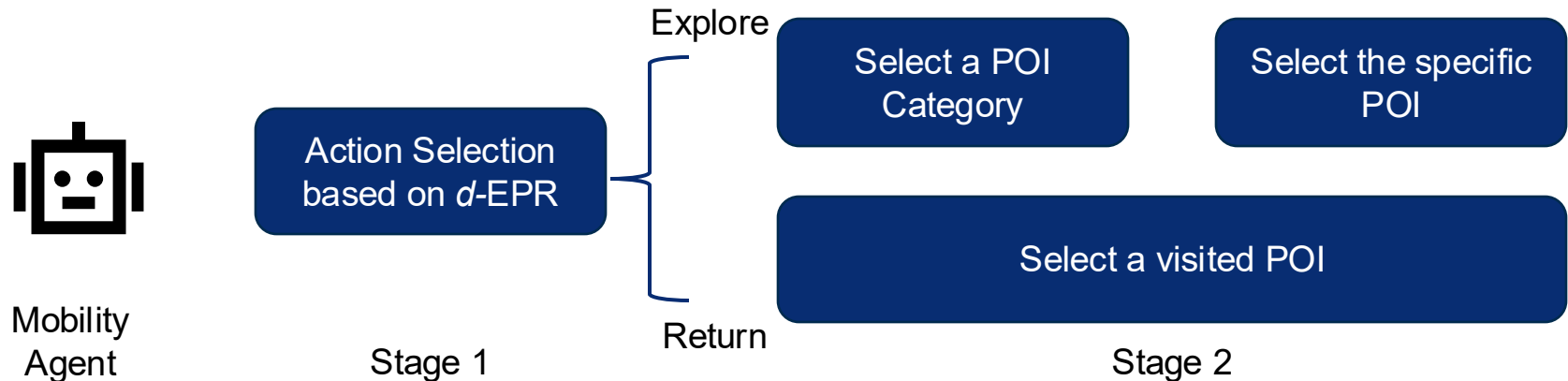
At each time step, an agent decides its next move from a given set of potentials POIs:

- **Action Selection (*d*-EPR)**: The agent first chooses between **exploring** a new POI or **returning** to a previously visited one.
- **POI Selection (Gravity Model + Interest Score)**: Combine LLM reasoning and **physical constraints**. If exploring, the agent **How do this work?** by calculating an attraction score for all potential destinations.





Mobility Decision Process

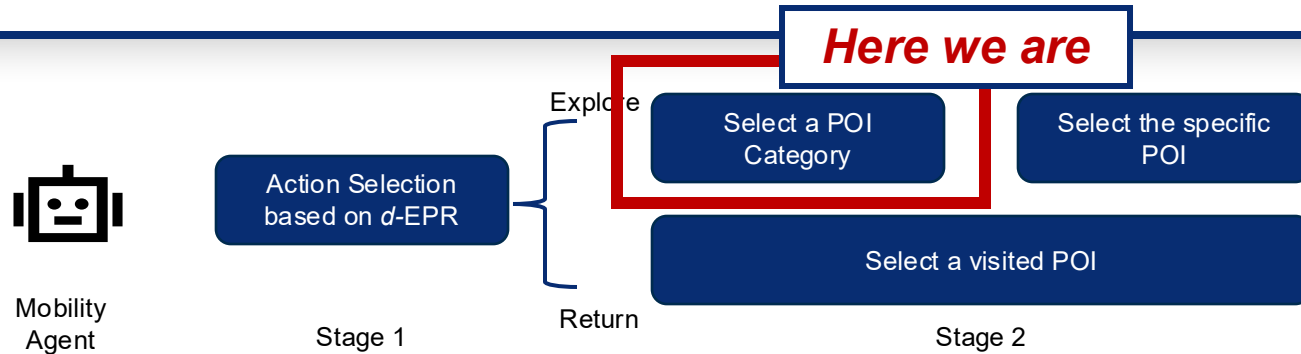


When Exploring

Select a POI Category

Select the specific POI

Give agent a list of a ranked POI categories , calculated by **Interest Score \times POIs (specific category)** within a travel radius. Then, convert it into a probability distribution, and select a category. First, based on the selected POI category, all POIs of that category within the map range are filtered. Then, the "**gravity**" of each POI is calculated (combining the interest score and the gravity model, the formula will be provided later), and then converted into a probability distribution to select specific POI.



Interest Score for *POI Category Selection*

The attraction A_c of a POI Category c for an agent at location is defined by the following attraction formula:

$$A_c = I_c \cdot N_c$$

Where:

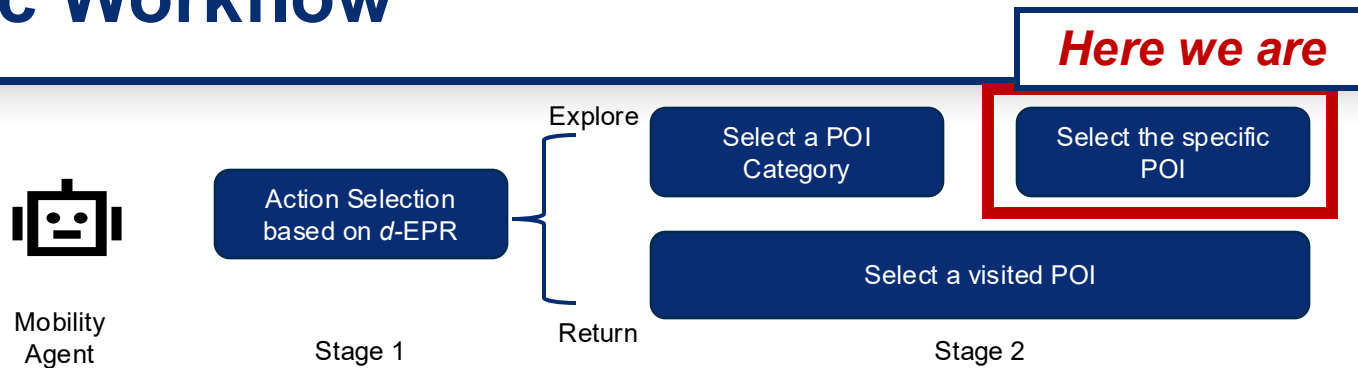
I_c is the agent's Interest Score for the POI category (c). **Use LLM to generate this**

N_c is the total number of POIs belonging to that category within a travel radius D_{max} around the current position of the agent.

Note:

Idea: You wouldn't visit an interest place if there is no (or very less) related POIs

- POI Category is classified by **North American Industry Classification System (NAICS)** and re-aggregated into six main categories
- D_{max} is set to 10km in this project which align with the travel mode preference for driving



Gravity Model for POI Selection

After determining the target POI category, the decision-making process enters the second stage of selecting POI, all the POIs in the map become candidate, and the **gravity** is defined as:

$$G_i \propto \frac{I_c}{d_i^\alpha}$$

Where:

I_c is the agent's Interest Score for the POI category (c).

d_i^α is the distance between the agent's current location i and POI j .

α is the distance decay parameter, default value is 2.75 in our model.

This formula follows the core idea of gravity model since **attraction is directly proportional to interest and inversely proportional to distance**.

The probability of visiting a POI is the normalized attraction score across all candidate POIs.



Evaluation Goal

To validate the model by comparing the simulated traffic flows from Census Block Groups (CBGs) to POI categories against empirical data.

Baseline

d-EPR model

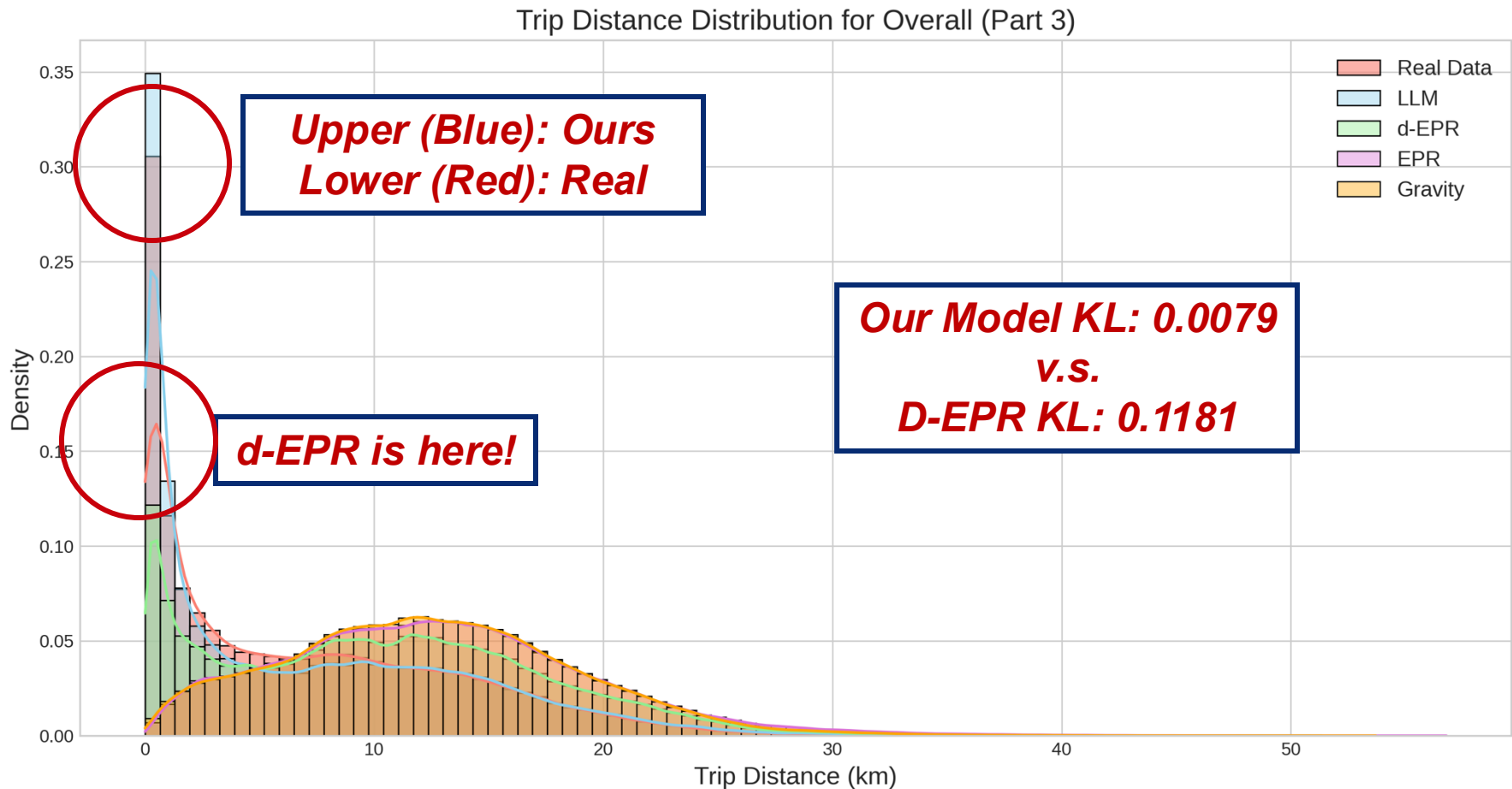
Ground Truth

Randomly selected as the week of 2019-06-10

To avoid bias, also do sensitivity analysis and test different period of time

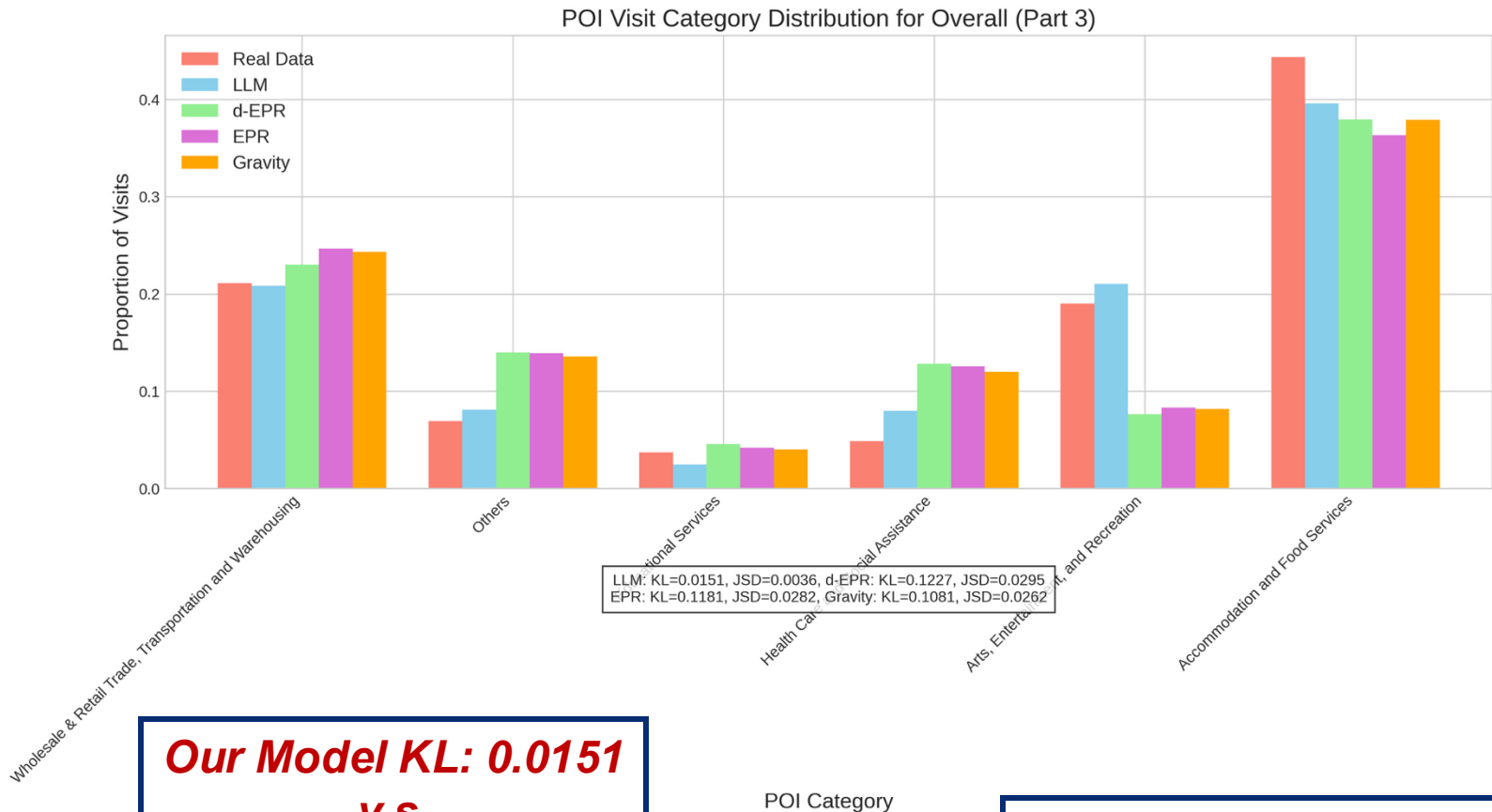


Overall Performance, Gemini-2.5-pro, Trip Distance





Overall Performance, Gemini-2.5-pro, Poi Category

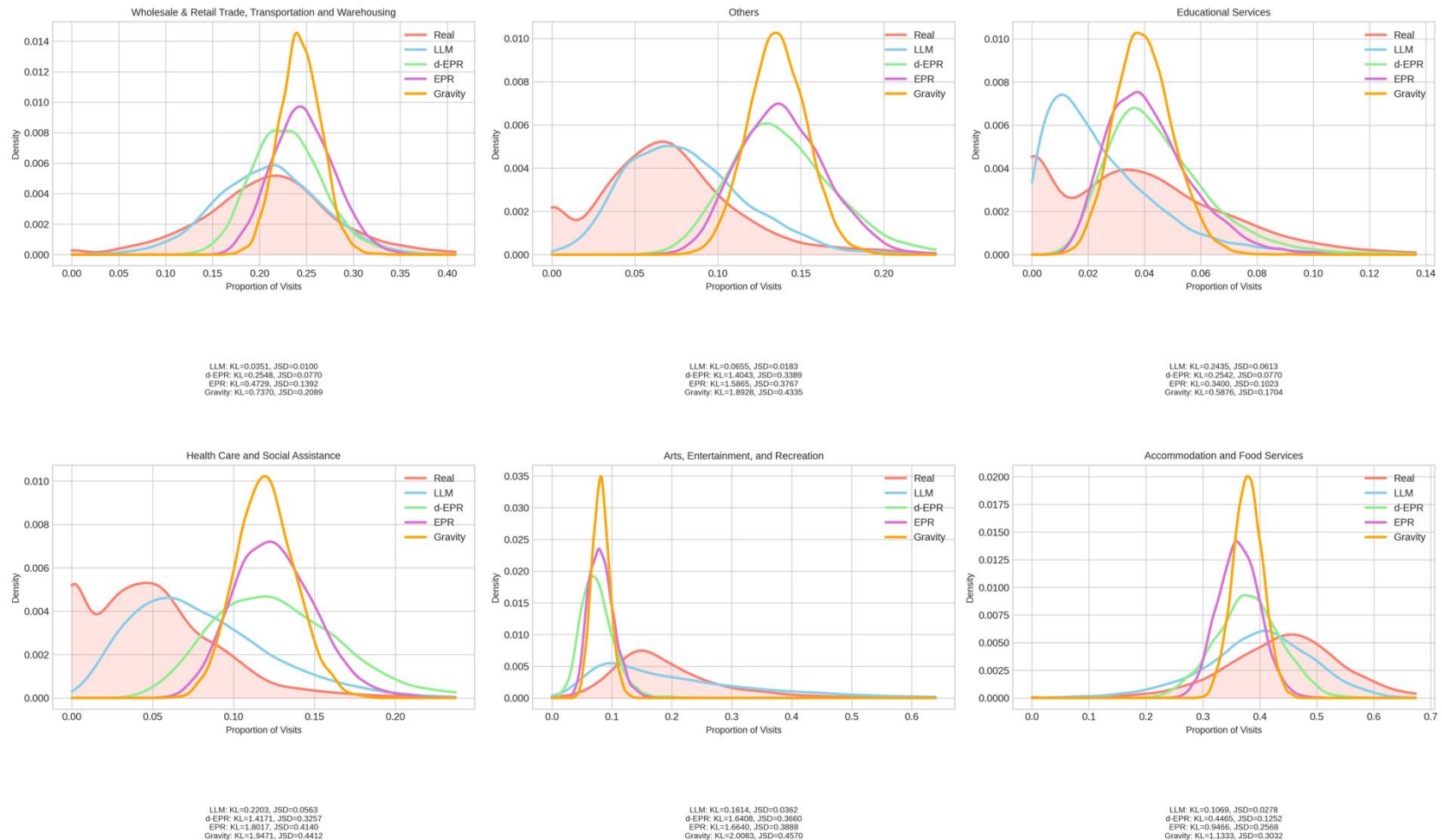


Our Model KL: 0.0151
v.s.
D-EPR KL: 0.1227

Also, fit very well!

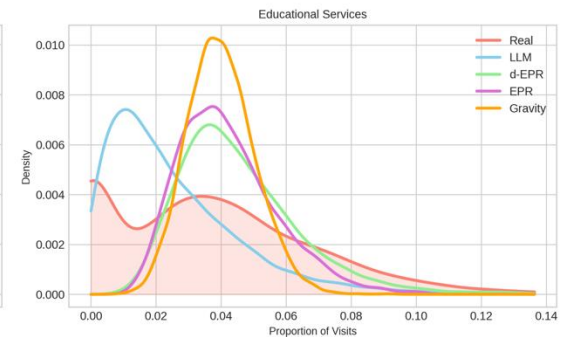
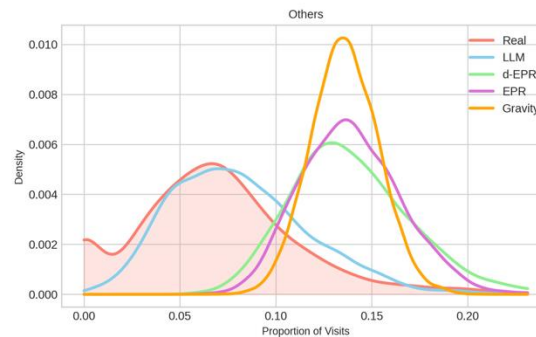
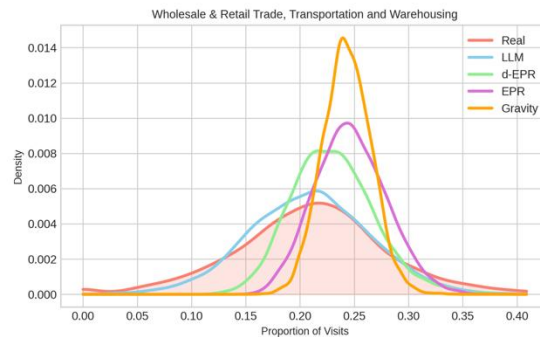
Overall Performance, Gemini-2.5-pro, Visit Prop. (KDE)

POI Category Visit Proportions (KDE) for Overall (Part 3)

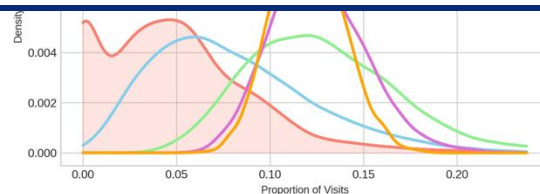


Overall Performance, Gemini-2.5-pro, Visit Prop. (KDE)

POI Category Visit Proportions (KDE) for Overall (Part 3)

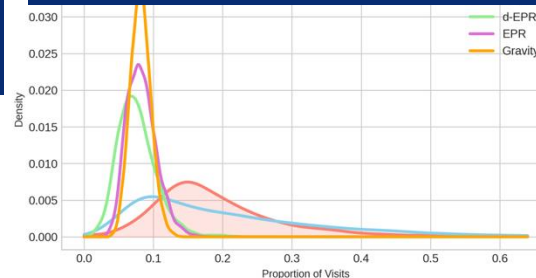


Breakdown the visits of each POI cate. From different CBGs into distribution (Prop. & CBG count)



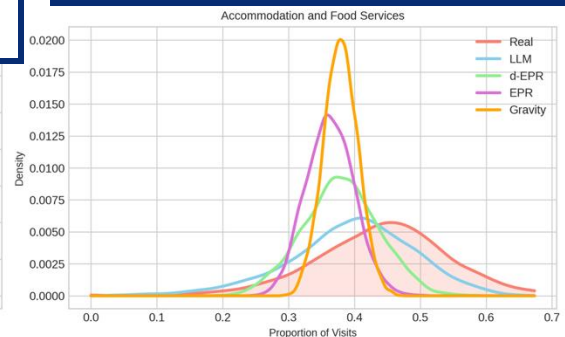
LLM: KL=0.2203, JSD=0.0563
d-EPR: KL=1.4171, JSD=0.3257
EPR: KL=1.8017, JSD=0.4140
Gravity: KL=1.9471, JSD=0.4412

**Avg. KL
Our Model: 0.1387
d-EPR: 0.9038**



LLM: KL=0.1614, JSD=0.0362
d-EPR: KL=1.6408, JSD=0.3660
EPR: KL=1.6640, JSD=0.3889
Gravity: KL=2.0083, JSD=0.4570

**Also,
fit very well!**



LLM: KL=0.1069, JSD=0.0278
d-EPR: KL=0.4465, JSD=0.1252
EPR: KL=0.9466, JSD=0.2569
Gravity: KL=1.1333, JSD=0.3032



Colocation

Treat different demographic group's visiting pattern to a certain **POI** as a vector, such as:

- `vector_white_agents = {'poi_1': 150, 'poi_2': 0, 'poi_3': 300, ...}`

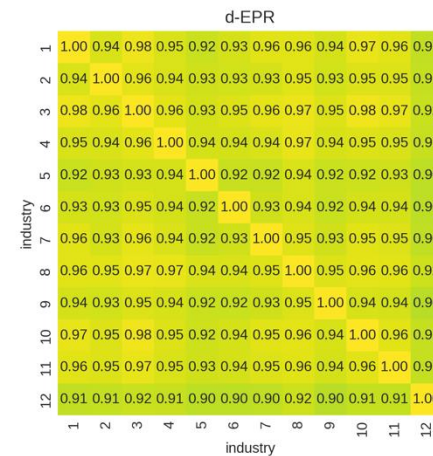
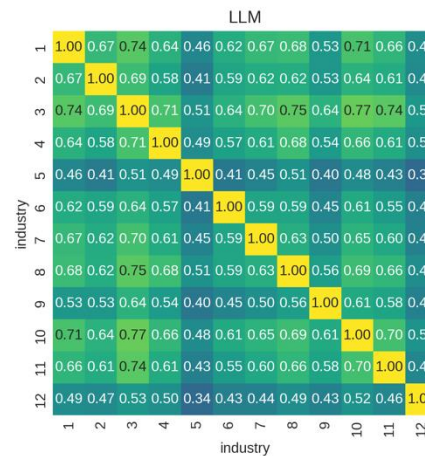
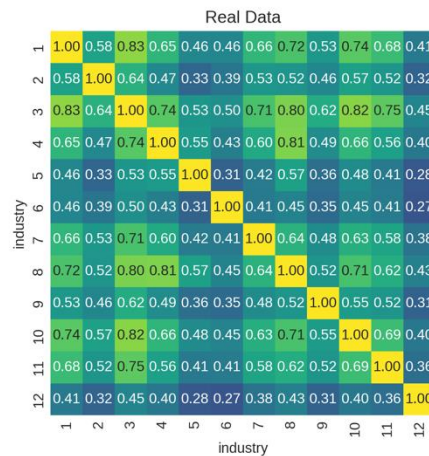
Then we can compute the similarity of visiting patterns between different demographic groups through **compute the similarity of vectors** (here we use Cosine Similarity)

Evaluation & Results



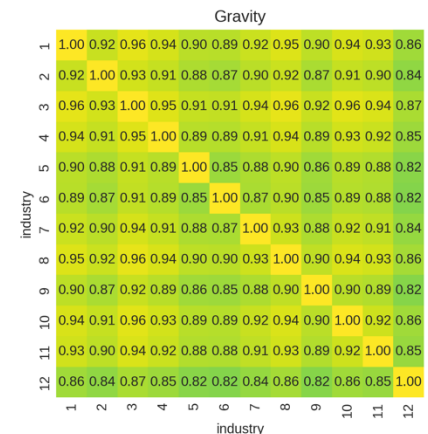
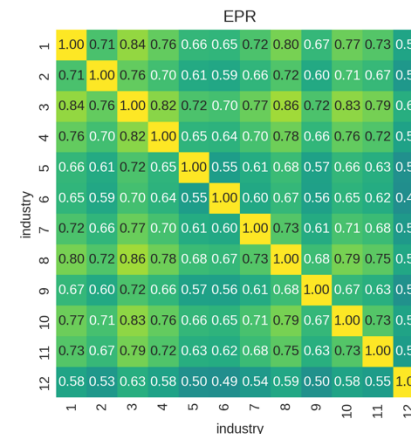
Colocation, Gemini-2.5-pro, industry, Cosine

We can model the Heterogeneity due to LLM's reasoning with agent profile



X-axis and Y-axis are agent's different occupation (industry)

The cosine similarly represent the colocation relationship of different demographic group



Sensitivity analysis



**Base model: d_max = 5km,
alpha=2.5, Best Performance!**

Condition		Evaluation Metric (model rank)				Avg. Divergence	Rank
Paramater	Value	Trip Distance KL	POI Catgery KL	POI Category Breakdown avg. KL	Colocation (Industry) avg. Diff.		
Base Model		0.0079	0.0151	0.1387	0.0622	0.0560	1
D_max	1.0km	0.0605	0.0331	0.2067	0.0914	0.0979	7
	10km	0.0085	0.0178	0.1860	0.0502	0.0656	3
	25km	0.0237	0.0402	0.3292	0.0477	0.1102	9
	50km	0.0244	0.0402	0.3349	0.0477	0.1118	10
	+inf (N/A)	0.0245	0.0402	0.3357	0.0476	0.1120	11
alpha	1.0	0.2979	0.0187	0.1798	0.0488	0.1363	12
	2.0	0.0382	0.0142	0.1495	0.0512	0.0633	2
	3.0	0.0672	0.0223	0.1736	0.1436	0.1017	8
rho, gamma	rho = 0.8, gamma = 0.1	0.0142	0.0156	0.2243	0.0960	0.0875	5
	rho = 0.4, gamma = 0.3	0.0246	0.0237	0.1790	0.0575	0.0712	4
KL with weekly average		0.0075	0.0146	0.2685	0.0998	0.0976	6
d-EPR		0.1177	0.1227	0.9038	0.3757	0.3800	13
EPR Model		0.4079	0.1181	1.1350	0.1319	0.4482	14
Gravity Model		0.3927	0.1081	1.3838	0.3384	0.5558	15



Achieved Interpretability:

- Macro-level: The agent's workflow is **guided by** the classic scientific theory of **d-EPR**, aligning with physical intuition.
- Micro-level: Each decision is accompanied by explicit **reasoning generated by the LLM based on its profile**, conforming to human logic, completely solving the "black box" problem of traditional data-driven models.

Overcame Trajectory Data Dependency:

The framework operates **without requiring training** on large-scale historical trajectory data, relying only on publicly **available census and geographic data**.

This significantly **reduces the cost** of model application, bypasses **data privacy** barriers, and enhances model **portability**.

Successfully Modeled Heterogeneity:

By injecting demographic profiles, the LLM agent effectively captures the **diverse and nuanced travel preferences of different groups** (e.g., different income levels, occupations), something traditional rule-based models struggle to achieve.



- Successfully constructed and validated **Mobility Agents**, an innovative framework for **simulating large-scale urban mobility patterns using theoretically guided LLM agents**.
- Experimental results show that
 - significantly outperforms in simulating both **macroscopic travel patterns** (such as travel distance distribution) and **microscopic heterogeneous behaviors** (such as co-occurrence patterns among different groups).
- Three major advantages: 1) high **interpretability**, 2) **no need for training** with historical trajectory data, and 3) reliance solely on **publicly available data**. It provides a novel and powerful research paradigm for **integrating scientific theory and artificial intelligence to study complex social systems**.



Q&A

Contact: LI_Haoyang@life.hkbu.edu.hk



Thanks!

Contact: LI_Haoyang@life.hkbu.edu.hk