Cognitive Heterogeneity and Behavioral Biases in Multi-Stage Supply Chains: Evidence from LLM-Based Agent Simulations

Extended Abstract

Jiuyun Jiang
Harbin Institute of Technology
Harbin, China
jiuyunjiang@stu.hit.edu.cn

Yuecheng Hong
Harbin Institute of Technology
Harbin, China
2023112391@stu.hit.edu.cn

Bo Yang
Harbin Institute of Technology
Harbin, China
boyang@stu.hit.edu.cn

Jin Yang
Harbin Institute of Technology
Harbin, China
22B910011@stu.hit.edu.cn

Guangxin Jiang
Harbin Institute of Technology
Harbin, China
gxjiang@hit.edu.cn

Xiaomeng Guo
Hong Kong Polytechnic University
Hong Kong, China
xiaomeng.guo@polyu.edu.hk

Guang Xiao
Hong Kong Polytechnic University
Hong Kong, China
guang.xiao@polyu.edu.hk

ABSTRACT

Modeling cooperation and coordination among LLM-based agents in multi-round decision making environments is a central challenge in artificial intelligence and operations management. Traditional behavioral experiments have demonstrated that cognitive biases can generate systemic inefficiencies, but such approaches are costly, difficult to scale, and limited in their ability to control individual heterogeneity. We propose a new experimental paradigm that employs LLM-based agents to simulate multi-stage supply chain decision making. The results indicate that LLM agents exhibit myopic and self-interested behavior that amplifies supply chain inefficiencies, while information sharing emerges as an effective mitigating mechanism. This study extends behavioral operations research by demonstrating how LLM-based agents can both reproduce and illuminate emergent coordination failures, highlighting their potential and limitations as tools for studying complex human like decision making in supply chains.

KEYWORDS

Cognitive Heterogeneity, Large Language Models (LLMs), Agent-Based Simulation, Supply Chain Management

ACM Reference Format:

Jiuyun Jiang, Yuecheng Hong, Bo Yang, Jin Yang, Guangxin Jiang, Xiaomeng Guo, and Guang Xiao. 2026. Cognitive Heterogeneity and Behavioral Biases in Multi-Stage Supply Chains: Evidence from LLM-Based Agent Simulations: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026,* IFAAMAS, 3 pages. https://doi.org/10.65109/SZPU6292

1 INTRODUCTION

Human subject experiments are central to behavioral science and have generated key insights into complex decision making. Classic studies such as the Beer Distribution Game show how cognitive biases and bounded rationality lead to systemic inefficiencies, including the bullwhip effect in supply chains [7, 8, 25, 26]. However, traditional experimental approaches are constrained by high costs, limited scalability, and difficulties in controlling individual heterogeneity, which restrict reproducibility and generalizability [11].

Recent advances in large language models offer a promising paradigm for behavioral research [2, 9, 18, 20, 21, 28, 30]. As computational subjects, LLM agents enable reproducible experiments with precise control over individual characteristics [3, 4, 6, 12, 14, 23, 24, 27]. Existing studies show that LLMs can reproduce human like behavior in classic games, but they largely focus on static or one shot interactions [1, 16, 29]. How such agents perform in complex multi round strategic environments as well as the significant impact that cognitive heterogeneity can have on group dynamics remains insufficiently understood [5, 9, 13, 17, 19, 22, 24].

To address this gap, we deploy LLM agents in the Beer Distribution Game where small demand fluctuations can be amplified. We vary agent sophistication across supply chain tiers and conduct extensive multi round simulations with robust statistical validation. This design allows us to examine how strategic heterogeneity shapes the bullwhip effect and to evaluate interventions such as information sharing that aim to improve collective performance.

2 PRELIMINARIES

Our investigation is grounded in the canonical Beer Game [7, 10, 15]. Each simulated supply chain team consists of four agents, each managing a specific stage: retailer (*S1*), wholesaler (*S2*), distributor (*S3*), and manufacturer (*S4*). Across T discrete periods, agents interact with their immediate upstream and downstream partners to manage inventory and fulfill external demand. The evolution of the system is governed by explicit update rules for shipments ($S_t^{i,g}$) and

---

Guangxin Jiang is the Corresponding Author.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/SZPU6292

Table 1: Comparison of order variance between LLM agent replications and human benchmark studies.

	DeepSeek-Original		**DeepSeek-Overall**		**ChatGPT-Original**		**ChatGPT-Overall**		**UTD**		**UW-Mad.**	
	Mean (S.D.)	Med.	Mean (S.D.)	Med.	Mean (S.D.)	Med.	Mean (S.D.)	Med.	Mean (S.D.)	Med.	Mean (S.D.)	Med.
W/O IS	39.43 (65.14)	18.73	29.43 (63.54)	12.29	58.45 (89.86)	23.07	31.42 (46.09)	13.88	8434 (27620)	20.46	1148 (6609)	30.91
W/ IS	20.15 (29.88)	11.12	17.71 (27.00)	9.50	40.39 (71.44)	20.70	13.72 (9.09)	11.88	4347 (21671)	17.58	4647 (22490)	19.83
***p*-value**	0.001	<0.001	0.029	0.042	0.067	0.044	0.001	0.018	0.184	0.221	0.116	0.048
Test	*t*-test	M.W.	*t*-test	M.W.	*t*-test	M.W.	*t*-test	M.W.	*t*-test	M.W.	*t*-test	M.W.

Data for the UTD and UW-Mad. columns are from Davis (2023) [10]. In addition, the p-values of the M.W. test in the classic human study and a prior replication using ChatGPT-4 are 0.028 and 0.573, respectively [7] [15].

inventory levels ($I_t^{i,g}$), where the order quantity $O_t^{i,g}$ is the primary decision variable for agent i in team g. The operational equations are given by:

$$S_t^{i,g} = \begin{cases} \min\left\{D_t, \max\left[I_{t-1}^{i,g} + S_{t-2}^{i+1,g}, 0\right]\right\} & \text{for } i = 1, \\ \min\left\{O_{t-2}^{i-1,g}, \max\left[I_{t-1}^{i,g} + S_{t-2}^{i+1,g}, 0\right]\right\} & \text{for } i = 2, 3, \\ \min\left\{O_{t-2}^{i-1,g}, \max\left[I_{t-1}^{i,g} + O_{t-3}^{i,g}, 0\right]\right\} & \text{for } i = 4, \end{cases}$$

with corresponding inventory updates, where negative inventory represents a backlog:

$$I_t^{i,g} = \begin{cases} I_{t-1}^{i,g} + S_{t-2}^{i+1,g} - D_t & \text{for } i = 1, \\ I_{t-1}^{i,g} + S_{t-2}^{i+1,g} - O_{t-2}^{i,g} & \text{for } i = 2, 3, \\ I_{t-1}^{i,g} + O_{t-3}^{i,g} - O_{t-2}^{i,g} & \text{for } i = 4. \end{cases}$$

The cost for a participant at stage i is given by:

$$C_i(T) = \sum_{t=1}^{T} \left[h^i \max\left\{I_t^i, 0\right\} - s^i \min\left\{I_t^i, 0\right\}\right]$$

where $C_i(T)$ is the total cost for the participant at stage, and h^i and s^i are their respective unit holding and backlog costs.

We develop a Hierarchical Reasoning Framework in which strategic thinking unfolds across cognitive layers of increasing sophistication. We established a shallow tier using **DeepSeek-V3 and ChatGPT-4.1** to represent baseline decision-making, and a cognitively deep tier with **DeepSeek-R1 and the cutting-edge ChatGPT-5** to represent advanced reasoning. Every experiment is **replicated 32 times** over 20 periods, consistent with previous study [10, 15]. All agents adopt Chain-of-Thought prompting to generate structured and transparent decisions. Empirical evidence indicates that outcomes are robust to temperature variation across models [9], supporting our use of a uniform sampling temperature of 1.

Our experiments feature two conditions. **Homogeneous** conditions establish baselines: *Original* (all four stages staffed by shallow agents) and *R-Overall* (all four stages staffed by deep agents). **Hierarchical** conditions test heterogeneity by placing a single deep agent in an otherwise shallow team at the Retailer (*R-S1*), Wholesaler (*R-S2*), Distributor (*R-S3*), or Manufacturer (*R-S4*) stage.

3 EXPERIMENT

We begin by assessing whether diverse LLM agents can replicate the classic bullwhip effect in supply chain management. As shown in Table 1, information sharing significantly reduces order variance across all configurations, with statistical significance confirmed by the Mann–Whitney U test. These findings suggest that LLM agents can reproduce key features of human decision-making behavior in complex environments.

To further investigate underlying mechanisms, we conduct a micro-level analysis using a regression model adapted from Sterman [26]:

$$O_t^{ig} = a_0 + a_I I_{t-1}^{ig} + a_R R_t^{ig} + a_S S_t^{ig} + a_N N_t^{ig} + a_t t + \varepsilon,$$

where I_{t-1}^{ig} denotes prior inventory, R_t^{ig} represents downstream orders, S_t^{ig} is shipments to the customer, N_t^{ig} captures total outstanding orders, and t accounts for time trends. The coefficient a_N is key to evaluating supply line integration. Rational behavior implies $a_N = a_I$, while myopic reasoning predicts $a_N > a_I$. Our results consistently show $a_N > a_I$ in all conditions ($p < 0.001$, sign test, $N = 128$), closely aligning with human behavioral patterns reported in [7]. This suggests that **LLM agents exhibit a similar tendency toward myopic underweighting of the supply line**.

While myopia accounts for the agent's limited perception of system dynamics, it does not fully explain the rationale behind its decisions. A complementary explanation is that agents exhibit **self-interested behavior**[1], optimizing their local performance at the expense of global efficiency. Figure 1 illustrates the cost pattern of DeepSeek under information isolation as an example. Enhancing a single upstream agent reduces its *stage-level cost* but leads to an increase in the *total supply chain cost* compared to the original configuration. **Information sharing corrects this misalignment by synchronizing local and system-level objectives**, and the lowest total cost is observed only when information sharing is combined with universal agent enhancement.

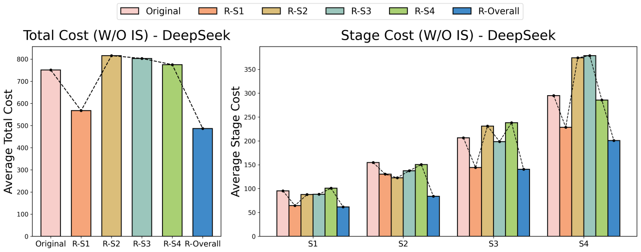

Figure 1: DeepSeek's total and average stage-level costs under information isolation, presented as a representative example due to space limitations.

ACKNOWLEDGMENTS

This work was partially supported by the National Natural Science Foundation of China (No. 72293562 and 724B2012).

REFERENCES

[1] Elif Akata, Lion Schulz, Julian Coda-Forno, Seong Joon Oh, Matthias Bethge, and Eric Schulz. 2025. Playing repeated games with large language models. *Nature Human Behaviour* (2025), 1–11.

[2] Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. 2023. Out of one, many: Using language models to simulate human samples. *Political Analysis* 31, 3 (2023), 337–351.

[3] Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, et al. 2022. Constitutional AI: Harmlessness from AI feedback. *arXiv preprint arXiv:2212.08073* (2022).

[4] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[5] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, et al. 2024. A survey on evaluation of large language models. *ACM transactions on intelligent systems and technology* 15, 3 (2024), 1–45.

[6] Julian Coda-Forno, Marcel Binz, Zeynep Akata, Matt Botvinick, Jane Wang, and Eric Schulz. 2023. Meta-in-context learning in large language models. *Advances in Neural Information Processing Systems* 36 (2023), 65189–65201.

[7] Rachel Croson and Karen Donohue. 2006. Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management science* 52, 3 (2006), 323–336.

[8] Rachel Croson, Ken Schultz, Enno Siemsen, and ML Yeo. 2013. Behavioral operations: the state of the field. *Journal of Operations Management* 31, 1-2 (2013), 1–5.

[9] Ziyan Cui, Ning Li, and Huaikang Zhou. 2025. A large-scale replication of scenario-based experiments in psychology and management using large language models. *Nature Computational Science* (2025), 1–8.

[10] Andrew M Davis, Blair Flicker, Kyle Hyndman, Elena Katok, Samantha Keppler, Stephen Leider, Xiaoyang Long, and Jordan D Tong. 2023. A replication study of operations management experiments in management science. *Management Science* 69, 9 (2023), 4977–4991.

[11] Ernst Fehr and Klaus M Schmidt. 1999. A theory of fairness, competition, and cooperation. *The quarterly journal of economics* 114, 3 (1999), 817–868.

[12] Amelia Glaese, Nat McAleese, Maja Trebacz, John Aslanides, Vlad Firoiu, Timo Ewalds, Maribeth Rauh, Laura Weidinger, Martin Chadwick, Phoebe Thacker, et al. 2022. Improving alignment of dialogue agents via targeted human judgements. *arXiv preprint arXiv:2209.14375* (2022).

[13] Jingru Jessica Jia, Zehua Yuan, Junhao Pan, Paul McNamara, and Deming Chen. 2024. Decision-making behavior evaluation framework for llms under uncertain context. *Advances in Neural Information Processing Systems* 37 (2024), 113360–113382.

[14] Tim Johnson and Nick Obradovich. 2024. Measuring an artificial intelligence language model's trust in humans using machine incentives. *Journal of Physics: Complexity* 5, 1 (2024), 015003.

[15] Samuel Kirshner. 2024. Artificial agents and operations management decision-making. *UNSW Business School Research Paper Forthcoming* (2024).

[16] Michal Kosinski. 2023. Theory of mind may have spontaneously emerged in large language models. *arXiv preprint arXiv:2302.02083* 4 (2023), 169.

[17] Eleanor Lin, James Hale, and Jonathan Gratch. 2023. Toward a better understanding of the emotional dynamics of negotiation with large language models. In *Proceedings of the Twenty-fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. 545–550.

[18] Stephanie Lin, Jacob Hilton, and Owain Evans. 2021. Truthfulqa: Measuring how models mimic human falsehoods. *arXiv preprint arXiv:2109.07958* (2021).

[19] Amogh Mannekote, Adam Davies, Jina Kang, and Kristy Elizabeth Boyer. 2025. Can llms reliably simulate human learner actions? a simulation authoring framework for open-ended learning environments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 29044–29052.

[20] Qiaozhu Mei, Yutong Xie, Walter Yuan, and Matthew O Jackson. 2023. A Turing Test: Are AI Chatbots Behaviorally Similar to Humans? *arXiv preprint arXiv:2312.00798* (2023).

[21] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*. 1–22.

[22] Yinzhu Quan and Zefang Liu. 2024. Invagent: A large language model based multi-agent system for inventory management in supply chains. *arXiv preprint arXiv:2407.11384* (2024).

[23] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog* 1, 8 (2019), 9.

[24] Rylan Schaeffer, Brando Miranda, and Sanmi Koyejo. 2023. Are emergent abilities of large language models a mirage? *Advances in neural information processing systems* 36 (2023), 55565–55581.

[25] Joel H Steckel, Sunil Gupta, and Anirvan Banerji. 2004. Supply chain decision making: Will shorter cycle times and shared point-of-sale information necessarily help? *Management science* 50, 4 (2004), 458–464.

[26] John D Sterman. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management science* 35, 3 (1989), 321–339.

[27] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).

[28] Haojie Wang, Jiuyun Jiang, L Jeff Hong, and Guangxin Jiang. 2025. LLMs for Supply Chain Management. *arXiv preprint arXiv:2505.18597* (2025).

[29] Taylor Webb, Keith J Holyoak, and Hongjing Lu. 2023. Emergent analogical reasoning in large language models. *Nature Human Behaviour* 7, 9 (2023), 1526–1541.

[30] Yongchao Zhou, Andrei Ioan Muresanu, Ziwen Han, Keiran Paster, Silviu Pitis, Harris Chan, and Jimmy Ba. 2022. Large language models are human-level prompt engineers. In *The eleventh international conference on learning representations*.