LLM Economist: Large Population Models and Mechanism Design in Multi-Agent Generative Simulacra

Anonymous Author(s)
Affiliation
Address
email

Abstract

We present the *LLM Economist*, a novel framework that uses agent-based modeling to design and assess economic policies in strategic environments with hierarchical decision-making. At the lower level, bounded rational worker agents—instantiated as persona-conditioned prompts sampled from U.S. Census-calibrated income and demographic statistics—choose labor supply to maximize text-based utility functions learned *in-context*. At the upper level, a planner agent employs in-context reinforcement learning to propose piecewise-linear marginal tax schedules anchored to the current U.S. federal brackets. This construction endows economic simulacra with three capabilities requisite for credible fiscal experimentation: (i) optimization of heterogeneous utilities, (ii) principled generation of large, demographically realistic agent populations, and (iii) mechanism design—the ultimate nudging problem—expressed entirely in natural language. Experiments with populations of up to one hundred interacting agents show that the planner converges near Stackelberg equilibria that improve aggregate social welfare relative to Saez solutions, while a periodic, persona-level voting procedure furthers these gains under decentralized governance. These results demonstrate that large language model-based agents can jointly model, simulate, and govern complex economic systems, providing a tractable test bed for policy evaluation at the societal scale to help build better civilizations.

1 Introduction

The rapidly expanding marketplace of autonomous language agents forms *economic simulacra*—synthetic societies whose allocation of effort and influence is governed by algorithmic code rather than legislation. As web-agents book tickets, draft briefs, and trade cryptocurrencies, they adapt to digital incentives, creating complex economic ecosystems requiring governance to prevent early-mover exploitation.

Recent advances demonstrate remarkable potential for coherent multi-agent dynamics. Generative Agents [57, 58] sustain believable interactions with thousands of persona-conditioned agents, Project Sid [1] scales toward "AI civilization" benchmarks, EconAgent [39] reproduces macroeconomic indicators with striking fidelity, and OASIS [72] explores large population simulacra of social media. These developments suggest LLMs exhibit sophisticated strategic reasoning [74], making them compelling substrates for policy experimentation.

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

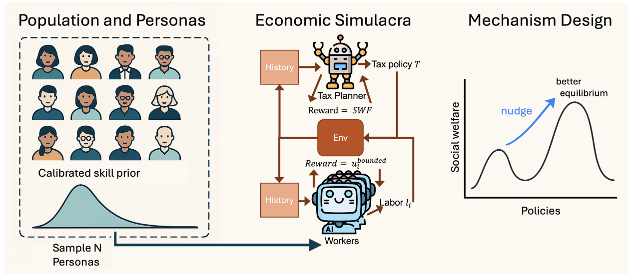


Figure 1: LLM Economist Framework. *Left:* Population of persona-conditioned agents. *Center:* Two-level economic simulacra. *Right:* Mechanism design via successive planner nudges.

In this work, we study designing tax mechanisms as a tractable and theoretically-grounded avenue for exploring governing agent societies. Classical optimal taxation faces two limitations in synthetic societies. First, solutions like the Saez formula [60, 61] assume fixed income elasticity, yet elasticity shifts dynamically with policy changes, making optimal rates moving targets requiring continuous recomputation. Second, human societies are heterogeneous and bounded rational [44, 50], while simulacra feature agents with text-specified motivations, requiring planners to reason over distributions of explicitly modeled personas.

We address both gaps by reframing optimal taxation as a repeated Stackelberg game optimized through two-level in-context reinforcement learning (ICRL) [35, 53, 54]. Building on the AI Economist's deep RL approach [65, 75, 76], we replace value-function learning with interpretable language-based reasoning. Worker agents maximize persona-conditioned utilities via natural-language context encoding biographies, while a planner proposes tax schedules anchored to U.S. federal brackets through pure in-context optimization.

Our contributions: *(i) Large population models* [8] sampling Census-calibrated personas without manual utility engineering. *(ii)* In-context planners converging to Saez-level welfare through gradient-free optimization. *(iii)* Democratic turnover stabilizing outcomes and mitigating the Lucas critique [43] through synthetic counterfactuals.

2 LLM Economist

We model optimal taxation as a repeated Stackelberg game between a *planner* $\mathcal{P}$ and workers $\mathcal{W} = \{\mathcal{W}_1, \ldots, \mathcal{W}_N\}$. Time is divided into daily steps $t = 0, \ldots, T-1$ and tax years of length K. Each worker i has latent skill $s^i > 0$ and chooses labor l_t^i, yielding pre-tax income $z_t^i = s^i l_t^i$. The planner selects marginal tax schedule τ_k at year start, giving post-tax income $\hat{z}_t^i = z_t^i - T_{\tau_k}(z_t^i) + R_t$ where R_t is lump-sum rebate. Social welfare is $\texttt{SWF} = \sum_{i=1}^{N} w(z_t^i) u_i(\hat{z}_t^i, l_t^i)$ with distributional weights $w(z_t^i) = 1/z_t^i$. A Stackelberg equilibrium satisfies: $\tau^* \in \arg\max_\tau \mathbb{E}[\texttt{SWF}(\mathbf{l}, \tau)]$ and $l^{i*}(\tau) \in \arg\max_{l^i} \mathbb{E}[u_i(\hat{z}^i, l^i)]$ for each worker.

The LLM Economist realizes this Stackelberg game through language-based agents acting purely *in-context*, where state, history, and objectives are rendered as text while actions are JSON snippets parsed by the environment. Skills s^i are drawn from generalized-Beta fits to 2023 American Community Survey data [66]. Each worker receives a persona prompt encoding demographics and preferences—*"You're a 32-year-old entrepreneur... You believe lower taxes let you reinvest in your company..."*—and uses bounded utility $u_i^{\text{bounded}}(\hat{z}, l) = \frac{\hat{z}^{1-\eta}-1}{1-\eta} - \psi l^\delta - (1 - s_t^i)\phi$ where $s_t^i \in \{0, 1\}$ is LLM-judged satisfaction and ϕ is dissatisfaction penalty. Workers observe $(z_t^i, \hat{z}_t^i, \tau(z_t^i), R_t, \text{history})$ and return `{"LABOR": X}`.

The planner observes aggregate statistics and proposes bracket shifts $\Delta\tau_k \in [-20, 20]^B$ via `{"DELTA":[...]}`. At each daily step t, the environment serializes joint state o_t into prompt π_t, following exploration-exploitation phases with broad search then convergence. Replay buffers maintain best state-action-welfare triples for token-level credit assignment across long horizons. Unlike the AI Economist's value-function learning, this design eliminates task-specific reward shaping while exposing agents' rationales, enabling interpretable policy optimization. This approach leverages LLMs' ability to identify patterns in textual reward

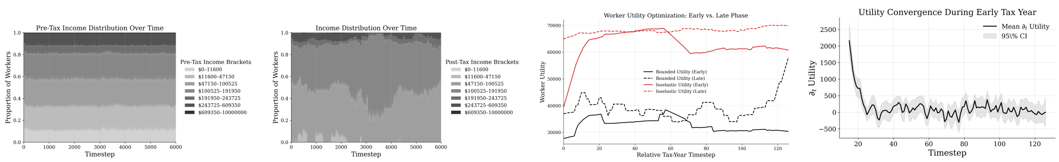
(a) Tax-year length

Steps / yr	Total steps	%`SWF`*
8	310	62.3
16	600	64.9
64	2 000	84.9
128	6 000	**90.0**
256	6 000	90.0

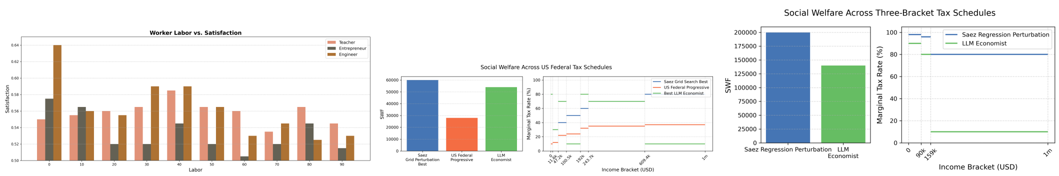
(b) Test-time search

Variant	Expl.+Expl.	No Explore	No Exploit
%`SWF`*	**84.9**	77.9	63.0

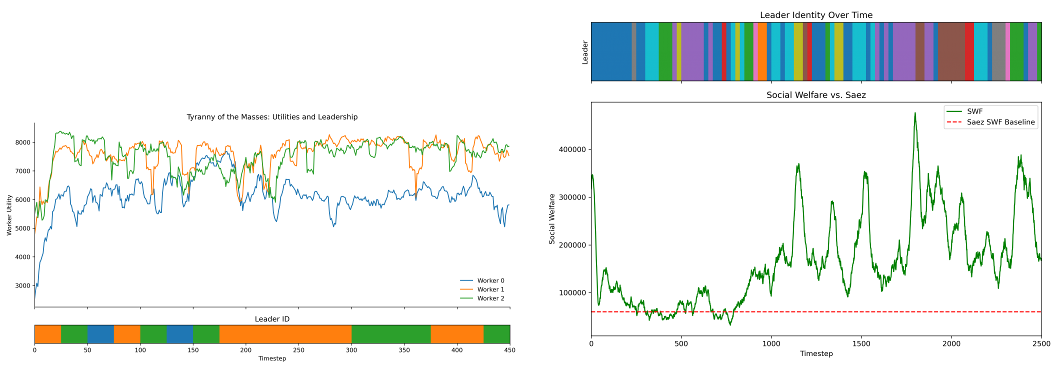
Figure 2: In-context RL ablations. (a) Welfare saturates at $K = 128$. (b) Exploitation and exploration are both critical.



(a) Pre-tax income. **(b) Post-tax income.** **(c) Utility gap closing.** **(d) Convergence in 120 steps.**



(e) Persona heterogeneity. **(f) Seven-bracket scenario.** **(g) Three-bracket scenario.**



(h) Three-persona "tyranny." **(i) 100-worker democracy.**

Figure 3: Experimental results. (a-b) Income redistributes 15% downward. (c-d) Worker utilities adapt and converge. (e) Heterogeneous persona responses. (f-g) Tax schedules approach Saez optimum. (h-i) Democratic dynamics from tyranny to welfare-enhancing turnover.

histories—a key advantage when preferences shift and causal links between individual utilities, policies, and outcomes must remain transparent.

3 Experiments

We evaluate: *(i)* design choices for planner and worker optimization, *(ii)* tax policy performance versus baselines, and *(iii)* emergent voting dynamics.

Setup: We use `Llama-3.1-8B` (though no discernable differences with other models, open-source and frontier, towards our hypotheses), $N = 100$ workers, $T = 3\,000$ steps with tax years $K = 128$. Skills follow GB2 calibrated to ACS 2023 [66]. Workers choose $[0, 100]$ hours/week; planners optimize seven brackets with lump-sum rebates.

We compare against **Saez** (perturbed, intractable optimum) and **U.S. Fed** (2024 statutory rates).

3.1 Planner's Social Welfare Optimization

The planner-worker interaction in the LLM Economist requires succesful ablation of two design choices to reach stable Stackelberg equilibria: *time-scale separation* between planner updates and worker adaptation, and balanced *exploration* and *exploitation* over tax years. Table 2a demonstrates that very short tax years ($K \leq 16$) stall below 65% of optimal welfare because workers lack time to adapt, while performance plateaus at $K = 128$ steps, capturing 90% of the theoretical optimum. Meanwhile, Table 2b reveals that both exploration and exploitation are critical: LLM agents leverage their priors to reason about promising policies (exploitation) while requiring systematic search to discover optimal schedules (exploration), with exploitation being more impactful, but not sufficient. These results validate our hypothesis that in-context reinforcement learning can achieve near-optimal social welfare through careful design of temporal dynamics and test-time search.

3.2 Workers' Utility Optimization

To test whether LLM workers optimize heterogeneous utilities under realistic income distributions, we initialize skills using Generalized-Beta fitted to ACS 2023 microdata. Figure 3a-b show the learned policy redistributes 15% of workers to lower post-tax brackets while preserving aggregate labor. Figure 3c demonstrates bounded workers nearly close a 30k dissatisfaction gap as the planner converges, while Figure 3e reveals persona-specific responses. These results validate that LLM workers coherently adapt labor choices under evolving tax incentives while maintaining realistic heterogeneity.

3.3 Tax Policy Evaluation

To test whether in-context reinforcement learning approach theoretically optimal policies, we compare against Saez baselines in two settings: bounded-utility workers (seven U.S. brackets) and isoelastic workers (three brackets). In the bounded case (Figure 3f), LLM Economist achieves +93% welfare versus U.S. baseline while approaching perturbed grid search Saez (+114%), with slightly less smooth schedules than the perturbed optimal. In the isoelastic case (Figure 3g), perturbed Saez outperforms but LLM Economist preserves labor supply through lower rates. The in-context RL planner achieves close to the Saez optimum without gradient information in a sample efficient manner, demonstrating that agent-based modeling approaches first-order optimal design—validating our hypothesis that LLMs can serve as effective mechanism designers.

3.4 Voting Simulacra

To test whether LLM agents reproduce political-economy phenomena, we introduce democratic elections where agents elect planners by majority vote each tax year. Figure 3h shows classic "tyranny of masses" in a three-agent society: two workers exploit the minority without hard-coded rules. Figure 3i demonstrates that 100-agent leadership turnover enhances welfare through electoral exploration, sometimes outperforming static optimal taxation.

4 Discussion

This work introduces the *LLM Economist*, an in-context reinforcement learning framework that embeds a population of persona-conditioned agents and a tax planner in a two-tier Stackelberg game. Our results show that LLM agents can *(i)* recover the Mirrleesian trade-off between equity and efficiency, *(ii)* approach Saez-optimal schedules in heterogeneous settings where analytical formulas are unavailable, and *(iii)* reproduce political phenomena—such as majority exploitation and welfare-enhancing leader turnover—without any hand-crafted rules. Taken together, the experiments suggest that the LLM Economist can serve as tractable test beds for policy design long before real-world deployment, providing a bridge between modern generative AI and classical economic theory. While our approach assumes static skills and fixed population, and the framework could be misused to craft biased policies, it offers a controlled environment for exploring economic mechanisms before real-world deployment.

References

[1] A. AL, A. Ahn, N. Becker, S. Carroll, N. Christie, M. Cortes, A. Demirci, M. Du, F. Li, S. Luo, et al. Project sid: Many-agent simulations toward ai civilization. *arXiv preprint arXiv:2411.00114*, 2024. 1

[2] K. J. Arrow et al. *Essays in the theory of risk-bearing*, volume 121. North-Holland Amsterdam, 1974.

[3] Y. Bai, C. Jin, H. Wang, and C. Xiong. Sample-efficient learning of stackelberg equilibria in general-sum games. *Advances in Neural Information Processing Systems*, 34:25799–25811, 2021.

[4] G. Brero, A. Eden, D. Chakrabarti, M. Gerstgrasser, A. Greenwald, V. Li, and D. C. Parkes. Stackelberg pomdp: A reinforcement learning approach for economic design. *arXiv preprint arXiv:2210.03852*, 2022.

[5] G. Brero, E. Mibuari, N. Lepore, and D. C. Parkes. Learning to mitigate ai collusion on economic platforms. *Advances in Neural Information Processing Systems*, 35:37892–37904, 2022.

[6] T. B. Brown. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*, 2020.

[7] R. Chetty. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, 1(1):451–488, 2009.

[8] A. Chopra. Large population models. *arXiv preprint arXiv:2507.09901*, 2025. 1

[9] A. Chopra, S. Kumar, N. Giray-Kuru, R. Raskar, and A. Quera-Bofarull. On the limits of agency in agent-based models. *arXiv preprint arXiv:2409.10568*, 2024.

[10] J. J. Chung. Money as simulacrum: The legal nature and reality of money. *Hastings Bus. LJ*, 5:109, 2009.

[11] X. Deng, Y. Gu, B. Zheng, S. Chen, S. Stevens, B. Wang, H. Sun, and Y. Su. Mind2web: Towards a generalist agent for the web. *Advances in Neural Information Processing Systems*, 36, 2024.

[12] P. A. Diamond and J. A. Mirrlees. Optimal taxation and public production i: Production efficiency. *The American economic review*, 61(1):8–27, 1971.

[13] Y. Du, L. Han, M. Fang, J. Liu, T. Dai, and D. Tao. Liir: Learning individual intrinsic reward in multi-agent reinforcement learning. *Advances in Neural Information Processing Systems*, 32, 2019.

[14] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

[15] P. Duetting, V. Mirrokni, R. Paes Leme, H. Xu, and S. Zuo. Mechanism design for large language models. In *Proceedings of the ACM on Web Conference 2024*, pages 144–155, 2024.

[16] M. F. A. R. D. T. (FAIR)†, A. Bakhtin, N. Brown, E. Dinan, G. Farina, C. Flaherty, D. Fried, A. Goff, J. Gray, H. Hu, A. P. Jacob, M. Komeili, K. Konath, M. Kwon, A. Lerer, M. Lewis, A. H. Miller, S. Mitts, A. Renduchintala, S. Roller, D. Rowe, W. Shi, J. Spisak, A. Wei, D. Wu, H. Zhang, and M. Zijlstra. Human-level play in the game of <i>diplomacy</i> by combining language models with strategic reasoning. *Science*, 378(6624):1067–1074, 2022. doi: 10.1126/science.ade9097. URL `https://www.science.org/doi/abs/10.1126/science.ade9097`.

[17] M. F. A. R. D. T. (FAIR)†, A. Bakhtin, N. Brown, E. Dinan, G. Farina, C. Flaherty, D. Fried, A. Goff, J. Gray, H. Hu, et al. Human-level play in the game of diplomacy by combining language models with strategic reasoning. *Science*, 378(6624):1067–1074, 2022.

[18] E. Farhi. Capital taxation and ownership when markets are incomplete. *Journal of Political Economy*, 118(5):908–948, 2010.

[19] X. Feng, Z. Wan, M. Wen, Y. Wen, W. Zhang, and J. Wang. Alphazero-like tree-search can guide large language model decoding and training. *arXiv preprint arXiv:2309.17179*, 2023.

[20] X. Feng, Z. Wan, H. Fu, B. Liu, M. Yang, G. A. Koushik, Z. Hu, Y. Wen, and J. Wang. Natural language reinforcement learning. *arXiv preprint arXiv:2411.14251*, 2024.

[21] M. Fleurbaey. Normative economics and economic justice. 2004.

[22] X. Gabaix. A behavioral new keynesian model. *American Economic Review*, 110(8): 2271–2327, 2020.

[23] S. Garg, D. Tsipras, P. S. Liang, and G. Valiant. What can transformers learn in-context? a case study of simple function classes. *Advances in Neural Information Processing Systems*, 35:30583–30598, 2022.

[24] S. Hao, Y. Gu, H. Ma, J. J. Hong, Z. Wang, D. Z. Wang, and Z. Hu. Reasoning with language model is planning with world model. *arXiv preprint arXiv:2305.14992*, 2023.

[25] S. Hoderlein. Nonparametric demand systems and a heterogeneous population. Technical report, Working Paper, Uni Mannheim, 2004.

[26] M. Hong, H. Wai, Z. Wang, and Z. Yang. A two-timescale framework for bilevel optimization: Complexity analysis and application to actor-critic, dec. 20. *arXiv preprint arXiv:2007.05170*, 2020.

[27] J. J. Horton. Large language models as simulated economic agents: What can we learn from homo silicus? Technical report, National Bureau of Economic Research, 2023.

[28] S. Hu, T. Huang, F. Ilhan, S. Tekin, G. Liu, R. Kompella, and L. Liu. A survey on large language model-based game agents. *arXiv preprint arXiv:2404.02039*, 2024.

[29] C. Ilut and R. Valchev. Economic agents as imperfect problem solvers. *The Quarterly Journal of Economics*, 138(1):313–362, 2023.

[30] D. Jeurissen, D. Perez-Liebana, J. Gow, D. Cakmak, and J. Kwan. Playing nethack with llms: Potential & limitations as zero-shot agents. *arXiv preprint arXiv:2403.00690*, 2024.

[31] S. Karten, A. L. Nguyen, and C. Jin. Pokéchamp: an expert-level minimax language agent. *arXiv preprint arXiv:2503.04094*, 2025.

[32] M. Klissarov, P. D'Oro, S. Sodhani, R. Raileanu, P.-L. Bacon, P. Vincent, A. Zhang, and M. Henaff. Motif: Intrinsic motivation from artificial intelligence feedback. *arXiv preprint arXiv:2310.00166*, 2023.

[33] J. Y. Koh, R. Lo, L. Jang, V. Duvvur, M. C. Lim, P.-Y. Huang, G. Neubig, S. Zhou, R. Salakhutdinov, and D. Fried. Visualwebarena: Evaluating multimodal agents on realistic visual web tasks. *arXiv preprint arXiv:2401.13649*, 2024.

[34] A. Korinek. Generative ai for economic research: Llms learn to collaborate and reason. Technical report, National Bureau of Economic Research, 2024.

[35] M. Laskin, L. Wang, J. Oh, E. Parisotto, S. Spencer, R. Steigerwald, D. Strouse, S. Hansen, A. Filos, E. Brooks, et al. In-context reinforcement learning with algorithm distillation. *arXiv preprint arXiv:2210.14215*, 2022. 1

[36] J. Z. Leibo, V. Zambaldi, M. Lanctot, J. Marecki, and T. Graepel. Multi-agent reinforcement learning in sequential social dilemmas. *arXiv preprint arXiv:1702.03037*, 2017.

[37] J. Z. Leibo, A. S. Vezhnevets, W. A. Cunningham, S. Krier, M. Diaz, and S. Osindero. Societal and technological progress as sewing an ever-growing, ever-changing, patchy, and polychrome quilt. *arXiv preprint arXiv:2505.05197*, 2025.

[38] Y. Leng and Y. Yuan. Do llm agents exhibit social behavior? *arXiv preprint arXiv:2312.15198*, 2023.

[39] N. Li, C. Gao, M. Li, Y. Li, and Q. Liao. Econagent: large language model-empowered agents for simulating macroeconomic activities. *arXiv preprint arXiv:2310.10436*, 2023. 1

[40] J. Liu, D. Shen, Y. Zhang, B. Dolan, L. Carin, and W. Chen. What makes good in-context examples for gpt-3? *arXiv preprint arXiv:2101.06804*, 2021.

[41] X. Liu, H. Yu, H. Zhang, Y. Xu, X. Lei, H. Lai, Y. Gu, H. Ding, K. Men, K. Yang, et al. Agentbench: Evaluating llms as agents. *arXiv preprint arXiv:2308.03688*, 2023.

[42] Y. Lu, M. Bartolo, A. Moore, S. Riedel, and P. Stenetorp. Fantastically ordered prompts and where to find them: Overcoming few-shot prompt order sensitivity. *arXiv preprint arXiv:2104.08786*, 2021.

[43] R. E. Lucas Jr. Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. North-Holland, 1976. 1

[44] R. D. Luce et al. *Individual choice behavior*, volume 4. Wiley New York, 1959. 1

[45] K. Ma, H. Zhang, H. Wang, X. Pan, W. Yu, and D. Yu. Laser: Llm agent with state-space exploration for web navigation. *arXiv preprint arXiv:2309.08172*, 2023.

[46] W. Ma, Q. Mi, X. Yan, Y. Wu, R. Lin, H. Zhang, and J. Wang. Large language models play starcraft ii: Benchmarks and a chain of summarization approach. *arXiv preprint arXiv:2312.11865*, 2023.

[47] L. Maliar and S. Maliar. The representative consumer in the neoclassical growth model with idiosyncratic shocks. *Review of Economic Dynamics*, 6(2):362–380, 2003.

[48] N. G. Mankiw, M. Weinzierl, and D. Yagan. Optimal taxation in theory and practice. *Journal of Economic Perspectives*, 23(4):147–174, 2009.

[49] C. F. Manski. What is the general welfare? welfare economic perspectives. Technical report, National Bureau of Economic Research, 2025.

[50] R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995. 1

[51] J. A. Mirrlees. An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208, 1971.

[52] J. A. Mirrlees. Optimal tax theory: A synthesis. *Journal of public Economics*, 6(4): 327–358, 1976.

[53] A. Moeini, J. Wang, J. Beck, E. Blaser, S. Whiteson, R. Chandra, and S. Zhang. A survey of in-context reinforcement learning. *arXiv preprint arXiv:2502.07978*, 2025. 1

[54] G. Monea, A. Bosselut, K. Brantley, and Y. Artzi. LLMs are in-context reinforcement learners, 2024. URL `https://openreview.net/forum?id=YW79lAHBUF`. 1

[55] G. H. Orcutt. Simulation of economic systems. *The American Economic Review*, 50(5): 893–907, 1960.

[56] D. Paglieri, B. Cupiał, S. Coward, U. Piterbarg, M. Wolczyk, A. Khan, E. Pignatelli, Ł. Kuciński, L. Pinto, R. Fergus, et al. Balrog: Benchmarking agentic llm and vlm reasoning on games. *arXiv preprint arXiv:2411.13543*, 2024.

[57] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages 1–22, 2023. 1

[58] J. S. Park, C. Q. Zou, A. Shaw, B. M. Hill, C. Cai, M. R. Morris, R. Willer, P. Liang, and M. S. Bernstein. Generative agent simulations of 1,000 people. *arXiv preprint arXiv:2411.10109*, 2024. 1

[59] A. Rees-Jones and D. Taubinsky. Taxing humans: Pitfalls of the mechanism design approach and potential resolutions. *Tax Policy and the Economy*, 32(1):107–133, 2018.

[60] E. Saez. Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229, 2001. 1

[61] E. Saez and S. Stantcheva. Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(01):24–45, 2016. 1

[62] N. E. Sanders, A. Ulinich, and B. Schneier. Demonstrations of the potential of ai-based political issue polling. *arXiv preprint arXiv:2307.04781*, 2023.

[63] P. Srivastava, S. Golechha, A. Deshpande, and A. Sharma. Nice: To optimize in-context examples or not? *arXiv preprint arXiv:2402.06733*, 2024.

[64] O. Topsakal and J. B. Harper. Benchmarking large language model (llm) performance for game playing via tic-tac-toe. *Electronics*, 13(8):1532, 2024.

[65] A. Trott, S. Srinivasa, D. van der Wal, S. Haneuse, and S. Zheng. Building a foundation for data-driven, interpretable, and robust policy design using the ai economist. *arXiv preprint arXiv:2108.02904*, 2021. 1

[66] U.S. Census Bureau. American community survey, 2023 public-use microdata sample (pums). https://www.census.gov/programs-surveys/acs, 2023. Accessed May 14, 2025. 2, 3

[67] H. Von Stackelberg. *Market structure and equilibrium.* Springer Science & Business Media, 2010.

[68] Y. Wang, Q. Liu, Y. Bai, and C. Jin. Breaking the curse of multiagency: Provably efficient decentralized multi-agent rl with function approximation. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 2793–2848. PMLR, 2023.

[69] R. Willis, Y. Du, J. Z. Leibo, and M. Luck. Will systems of llm agents cooperate: An investigation into a social dilemma. *arXiv preprint arXiv:2501.16173*, 2025.

[70] C. Yang, X. Wang, Y. Lu, H. Liu, Q. V. Le, D. Zhou, and X. Chen. Large language models as optimizers, 2024. URL https://arxiv.org/abs/2309.03409.

[71] J. C. Yang, M. Korecki, D. Dailisan, C. I. Hausladen, and D. Helbing. Llm voting: Human choices and ai collective decision making. *arXiv preprint arXiv:2402.01766*, 2024.

[72] Z. Yang, Z. Zhang, Z. Zheng, Y. Jiang, Z. Gan, Z. Wang, Z. Ling, J. Chen, M. Ma, B. Dong, et al. Oasis: Open agents social interaction simulations on one million agents. *arXiv preprint arXiv:2411.11581*, 2024. 1

[73] W.-B. Zhang. A discrete heterogeneous-group economic growth model with endogenous leisure time. *Discrete Dynamics in Nature and Society*, 2009(1):670560, 2009.

[74] Y. Zhang, S. Mao, T. Ge, X. Wang, A. de Wynter, Y. Xia, W. Wu, T. Song, M. Lan, and F. Wei. Llm as a mastermind: A survey of strategic reasoning with large language models. *arXiv preprint arXiv:2404.01230*, 2024. 1

[75] S. Zheng, A. Trott, S. Srinivasa, N. Naik, M. Gruesbeck, D. C. Parkes, and R. Socher. The ai economist: Improving equality and productivity with ai-driven tax policies. *arXiv preprint arXiv:2004.13332*, 2020. 1

[76] S. Zheng, A. Trott, S. Srinivasa, D. C. Parkes, and R. Socher. The ai economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science advances*, 8(18):eabk2607, 2022. 1

[77] A. Zhou, K. Yan, M. Shlapentokh-Rothman, H. Wang, and Y.-X. Wang. Language agent tree search unifies reasoning acting and planning in language models. *arXiv preprint arXiv:2310.04406*, 2023.

[78] R. Zhou, S. S. Du, and B. Li. Reflect-rl: Two-player online rl fine-tuning for lms. *arXiv preprint arXiv:2402.12621*, 2024.