SIMCITY: MULTI-AGENT URBAN DEVELOPMENT SIMULATION WITH RICH INTERACTIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Large Language Models (LLMs) open new possibilities for constructing realistic and interpretable macroeconomic simulations. We present **SimCity**, a multi-agent framework that leverages LLMs to model an interpretable macroeconomic system with heterogeneous agents and rich interactions. Unlike classical equilibrium models with limited heterogeneity, or traditional agent-based models (ABMs) that rely on hand-crafted decision rules, SimCity enables flexible, adaptive behavior with transparent natural-language reasoning. Within SimCity, four core agent types (households, firms, a central bank, and a government) make decisions and participate in a frictional labor market, a heterogeneous goods market, and a financial market. Furthermore, a Vision–Language Model (VLM) determines the geographic placement of new firms based on a virtual city map, allowing us to study both macroeconomic regularities and urban expansion dynamics within a unified environment. To evaluate the framework, we compile a checklist of canonical macroeconomic phenomena, including price elasticity of demand, Engel's Law, Okun's Law, the Phillips Curve, etc., and show that SimCity effectively exhibits these canonical patterns. Using our framework, we study novel economic shocks that are difficult to analyze in existing models and examine their macroeconomic implications qualitatively, yielding coherent results.

1 INTRODUCTION

The rapid development of Large Language Models (LLMs) has enabled multi-agent simulations of human societal activities across diverse scales and domains (Gao et al., 2024a). In these settings, autonomous LLM-powered agents interact with each other and with the environment. Existing works have examined general social simulacra (Park et al., 2023; Huang et al., 2025; Piao et al., 2025), domain-specific applications such as public administration crisis (Xiao et al., 2023), health policy (Hou et al., 2025), or deduction game (Xu et al., 2025).

For the past two decades, dynamic stochastic general equilibrium (DSGE) models have been the predominant paradigm for studying aggregate economic behavior (Blanchard, 2009; Glandon et al., 2023). While mathematically elegant, DSGE relies on the assumption of simplified representative agents solving explicit optimization problems, which restricts heterogeneity and richer behavioral dynamics (Sergi, 2018; Vines & Wills, 2020; Storm, 2021). On the other hand, agent-based models (ABMs) offer a bottom-up alternative that accommodates heterogeneity but typically rely on hand-crafted decision or learning rules (Dilaver et al., 2018). Advancing the ABM with recent developments in LLMs, we propose **SimCity**, a multi-agent macroeconomic simulation framework that enables flexible, adaptive behavior with transparent natural language reasoning, and a virtual city environment.

SimCity models the economy as the interaction of four types of agents: households, firms, a central bank, and a government. To incorporate realistic heterogeneity, households and firms are instantiated as families of agents with rich variation in preferences, abilities, and other background, whereas the central bank and government are modeled as single institutional agents (Blanchard, 2025). Agents are implemented through an LLM-based module that integrates environmental observations, traits, and structured memory to generate reasoning, planning, and decision-making. Following the agent–interaction–environment paradigm (Wooldridge, 2009), we construct a simulation environment that features a frictional labor market, heterogeneous goods markets, and core

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

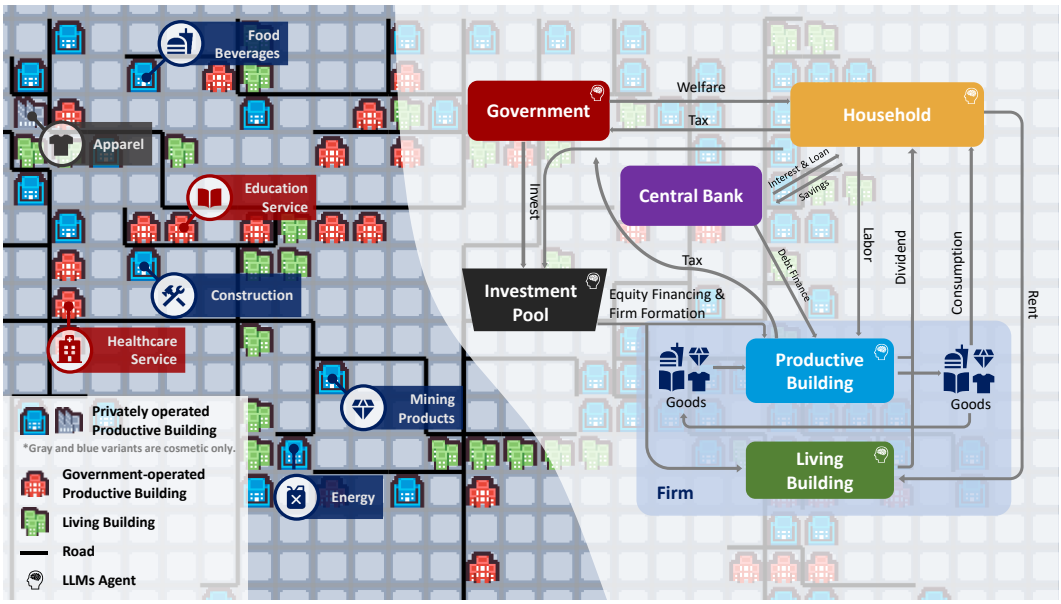


Figure 1: The SimCity framework. Left: A visualized map illustrating the urban layout across three building categories. Households reside in living buildings and engage in labor within productive buildings. City evolution is depicted in Figure 6. Right: Diagram of multi-agent interactions. Gray arrows denote the flow of capital and goods. Mechanisms of interactions are detailed in Section 4.

financial interactions. The environment is augmented by a visualized map (Figure 1) to provide spatial context.

In summary, our work makes the following contributions:

1. **LLM-driven macroeconomic simulation with rich interactions**. We integrate LLMs into economic simulation by modeling four distinct economic roles, households, firms, a central bank, and a government, as LLM-driven agents. This enables agent heterogeneity and rich economic interactions. Furthermore, we provide LLM-based agents with a visualized city map that conveys spatial context and renders the dynamics of urban expansion, supporting decision-making in a realistic urban setting (Figure 1, Table 1).

2. **Systematic evaluation**. We compile a checklist of canonical macroeconomic phenomena and regularities and use it to demonstrate that our framework exhibits established macroeconomic patterns (Table 2, Section 5.2). We further study responses to novel economic shocks (Section 5.5), including exogenous price changes and shifts in production structure, highlighting our framework's flexibility in analyzing scenarios that are difficult to capture in existing models.

3. **Emergent macroeconomic regularities and urban dynamics**. Experimental results show that SimCity exhibits a range of classic macroeconomic phenomena, such as Okun's law, the Phillips curve, the Beveridge curve, price elasticity of demand, and Engel's curve (detailed in Section 5). It also simulates the dynamic expansion of a city, and facilitates the simulation of novel economic shocks that were difficult to study in previous frameworks.

2 Related Work

The application of LLMs to social simulation represents an emerging and promising research frontier. LLMs are capable of exhibiting human-like behaviors (Xie et al., 2024) and can be endowed with diverse personas (Chen et al., 2024), offering a basis for interpretable agentic reasoning. Structurally, these simulation frameworks typically contain agents, an environment they interact with, and the interfaces that mediate their interactions (Gao, 2024). More related work in this area is listed in Appendix F.

Table 1: Comparison of features between frameworks listed in Section 5.

Simulator	AI E.	LEN	CATS	EconA.	SimCity
Households	10	100	100	200	1000
Tax Type	Labor	×	×	Labor	Labor &Firm
Agents' Types*	HG	HB	HFB	HBG	HFBG
Markets†	×	L	LGF	×	LGF
Visualized Map	✓	×	×	×	✓
Goods Type	2	1	1	1	10+
Interpretability	×	×	×	✓	✓

Note: *H: Household, F: Firm, B: Bank, G: Gov. †L: Labor, G: Goods, F: Financial.

Table 2: Macroeconomic regularities exhibited by SimCity, details are in Section 5.

Regularity	AI-E.	LEN	CATS	EconA.	SimCity
Phillips Curve	/	×	×	✓	✓
Okun's Law	/	✓	✓	✓	✓
Beveridge Curve	/	×†	✓	/	✓
Price Elasticity	/	/	/	/	✓
Engel's Law	/	/	/	/	✓
Invest. Volat.	/	/	/	/	✓

Legend: ✓: Verified; ×: Inconsistent; /: Not available.

DSGE and ABM offer contrasting paradigms for macroeconomic modeling: DSGE emphasizes stylized optimizing agents and market-clearing equilibrium (Stokey et al., 1989; Ljungqvist & Sargent, 2018), whereas ABM models heterogeneous agents with rule-based behaviors (Tesfatsion & Judd, 2006). While DSGE remains dominant (Woodford, 2009; Negro et al., 2013), recent crises have renewed interest in ABMs (Stiglitz, 2018; Borsos et al., 2024). Our framework combines key advantages of both approaches by enabling heterogeneous agents with flexible, LLM-based reasoning while preserving structured economic interactions.

Li et al. (2024) introduces EconAgent as a recent attempt at LLM-driven macroeconomic simulation. Its framework restricts agent-environment interactions to two simplified variables: consumption share and work propensity, which constrains the range of economic activities that can be simulated. Our work extends this approach by incorporating other key economic elements, including firms as LLM-driven agents, heterogeneous goods markets, enriched financial activities, and a taxation system. Thus, our framework allows the exploration of a wider range of macroeconomic phenomena. Another contemporary work Mi et al. (2025) is proposed as a testbed for agent performance evaluation while we focus on urban development and macroeconomic laws.

3 Environment and Interaction

Our framework consists of three core layers: environment, interaction protocol, and agents. We introduce the first two layers in this section.

3.1 Environment

On a high level, the simulation proceeds in two phases. In phase 1 (*the move-in phase*), new households with synthetic profiles, representing immigrants or newborns, are introduced into the SimCity environment until a predetermined maximum population is reached. Unlike the typical burn-in period used purely for model stabilization in ABMs (Lengnick, 2013), this phase is designed to study urban expansion, capturing how economic activity and spatial development evolve as the city grows.

In phase 2 (*the development phase*), the population is fixed while urban developments such as firm creation, investment, and spatial reallocation continue to occur. We design this phase to study macroeconomic regularities, as the system operates in a stable state that is less susceptible to the transient noise present in the expansion period.

Phase 1 lasts 36 steps, and phase 2 lasts 144 steps in a simulation. Each simulation step represents one month. The order of agents' actions within each step is detailed in Appendix D.2. Agents can only access information from previous steps.

There are three markets in the environment. Firms sell goods to each other and to households in a *goods market*. Notably, goods are qualitatively differentiated (e.g., food, clothing). Firms post jobs and are matched with households in a *labor market*. The central bank accepts deposits and provides loans in a *financial market*. Details of the environment setup can be found in Appendix D.1.

We develop a web-based rendering module to visualize the urban expansion as shown in Figure 1. All buildings are displayed on the rendered map. The geographic placement of new firms is decided by a Visual-Language Model (VLM). For technical details of this module, please refer to Appendix D.3.

3.2 Interaction Protocol

We leverage the common-sense reasoning capabilities of LLMs to act as human-like, heterogeneous agents. Details about the prompts used are provided in Appendix E.1. Agents interact with the environment by means of *function calling*. The framework loads all the operations that the agent can execute and appends formatted function names along with their descriptions to the prompt. The LLM returns the actions to be taken and their parameters in JSON format. The framework will execute after verification. For detailed examples, see Appendix E.3.

4 Agents

This section presents an overview of the agents in SimCity. As illustrated in Figure 1, there are four agent types: households, firms, a government, and a central bank, each simulated by an LLM. Building on prior work (Gatti et al., 2011; Wolf et al., 2013; Dawid & Gatti, 2018; Li et al., 2024), our enriched framework expands the decision space considerably.

- Households make four key decisions each period: consumption bundle, labor market action, housing, and financial activity.
- Firms set production levels and prices, decide on hiring and capital investment, and acquire financing when needed.
- The government collects taxes and stimulates the economy through public spending and transfer payments to support consumption and improve social welfare.
- The central bank adjusts the interest rate in response to market conditions.

4.1 Households

Households are the fundamental units in our simulation and interact with all other agent types and markets. Each household is initialized with a heterogeneous profile, including age, education, consumption preference, and skill endowment. Details about these characteristics and the initialization process can be found in Appendix D.4.

Each month, households receive two types of information: a personal report summarizing income, expenditures, and other status changes during the current period, and a citywide update containing goods prices, labor market conditions, housing availability, returns on investment, and interest rates. Additionally, households may receive offers for vacancies that align with their skills.

Based on this information, each household decides on (i) its consumption bundle, which requires a minimum expenditure on certain essential goods (Ravn et al., 2008); (ii) its labor market action (accepting or rejecting a new job if offered, or resigning from its current job); (iii) its housing choice, and (iv) its financial decisions (saving, borrowing, or investing in the common investment pool, which is discussed below).

4.2 Investment Pool and Firms

Investment Pool An investment pool is implemented as an intermediary module that leverages a VLM for investment decisions. Funds from households' investment actions are temporarily deposited into this pool and returned if unused for a month. When the pool accumulates sufficient capital and the VLM deems conditions favorable, a new firm is established.

The pool chooses from a library of 44 synthetic firm templates, each of which produces a unique type of good, and selects a geographic placement for the firm. The construction of these templates from real-world data is detailed in Appendix D.5. After a firm is established, all contributing households receive shares of the firm in proportion to their investment.

Firms Each *firm* in SimCity produces a single type of good. It hires households as workers, invests in capital to improve productivity, and transforms input goods into its specialized output, which is then sold in the market. Firms are instantiated from templates that specify skill requirements of job positions in the firm and the input-output relations governing its production process, where a firm may take possibly more than one input good for its production. The effectiveness of resource utilization (i.e., the ratio at which input goods are transformed into the output good) is determined by a firm-specific Cobb-Douglas production function (Cobb & Douglas, 1928):

$$Y_i = A\, L_i^{1-\alpha}\, K_i^{\alpha}, \tag{1}$$

where A is total factor productivity, K_i is the i-th firm's capital stock, L_i is the effective labor supply, defined as the amount of labor input adjusted for the match between job skill requirements and employees' actual skills.

Similar to households, firms receive individual reports summarizing their internal operations (e.g., expenditure composition, employee skill profiles), and citywide updates on external conditions such as the unemployment rate, interest rates, and supply/demand conditions in the goods market.

After deliberating on this information, each firm chooses (i) the quantity and price of its outputs, taking into consideration the prices of relevant input goods, as well as the geometric distance to their respective suppliers; (ii) its labor market actions, which may include posting job vacancies, laying off employees, or modifying wages; and (iii) its investment decision, including whether to borrow from the bank, and whether to purchase fixed assets to increase its capital.

More detailed descriptions of firms are in Appendix C.3.

4.3 GOVERNMENT

Macroeconomic Indicators The *government* monitors key indicators about the state of the economy, and adjusts taxation and fiscal policies to enhance social welfare, as measured by these indicators. The most important macroeconomic indicators include total consumption, total investment, nominal GDP, and real GDP, where nominal GDP is calculated using current prices, and real GDP uses constant base-year prices. Inflation is tracked in two forms: wage inflation, defined as the rate of change in average wages, and GDP inflation, measured by the rate of change in the GDP deflator. The GDP deflator itself is the ratio of nominal GDP to real GDP.

Tax and Welfare The government collects bracketed income tax from households and corporate tax from firms. Let $\mathcal{B} = \{(b_k, r_k)\}_{k=1}^{K}$ denote a tax schedule, where b_k is the lower threshold of bracket k, with $b_{K+1} = \infty$, and r_k the corresponding marginal tax rate. The tax liability from agent i with tax base z_i (income for households, or net profit for firms) is:

$$t_i = \sum_{k=1}^{K} r_k \, \max\big(\min(\, z_i,\, b_{k+1}\,) - b_k, 0\,\big). \tag{2}$$

The total tax revenue collected by the government is the sum of taxes from all agents. The government uses this tax revenue in three ways: investing in the construction of public service buildings, which are modeled as government-owned firms, distributing it to households as a universal basic income (UBI), or reserving it for the future.

Mathematical definitions of indicators and other details can be found in Appendix C.4.

4.4 CENTRAL BANK AND FINANCIAL SYSTEM

The *central bank* accepts deposits, provides loans, and implements monetary policy. In the first month of each year, it sets the policy interest rate according to a modified Taylor rule, which is widely used in monetary policies. (Galí & Gertler, 2007; Dawid & Gatti, 2018). Formally:

$$\hat{r} = \max\big(r_{\text{natural}} + \pi_{\text{target}} + \alpha(\pi - \pi_{\text{target}}) + \beta(Y - Y_{\text{natural}}),\, 0\big), \tag{3}$$

where $\hat{r}$ is the policy rate set by the central bank, r_{natural} is the long-run natural interest rate, π_{target} is the target inflation rate, π is the GDP inflation in the last period, Y is the actual output (GDP), and Y_{natural} is the potential output as measured by a linear trend. Parameters α and β capture the central bank's responsiveness to inflation and the output gap, respectively.

Intuitively, when inflation is high, or output exceeds its long-term trend (signals of an overheated economy), the central bank raises interest rates to cool the economy and maintain monetary stability, and vice versa (Galí, 2015). To reflect the gradual adjustment observed in real-world monetary policy, we incorporate a smoothing term consistent with empirical evidence on persistent rate changes (Coibion & Gorodnichenko, 2012).

At each time step, deposits and loans accrue interest: the deposit rate equals the policy interest rate, while the loan rate equals the policy rate plus a fixed markup. Further details about the financial system are in Appendix C.5.

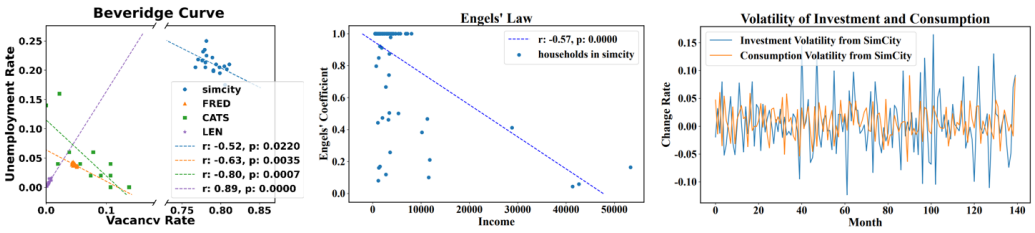


Figure 2: Macroeconomic emergences from SimCity: Beveridge Curve, Engels' Law, and business cycle volatility. Business cycle volatility refers to the well-established phenomenon that investment volatility is significantly higher than consumer spending volatility across economic fluctuation (Smets & Wouters, 2007).

5 Experiments

We conduct experiments to evaluate SimCity's ability to perform macroeconomic simulations and to display urban development dynamics. A key limitation of existing frameworks (Lengnick, 2013; Li et al., 2024) is the lack of a common basis for comparison. Prior studies often highlight a relatively small set of selected macroeconomic phenomena, making it difficult to gauge and compare the capacities of different simulation frameworks.

To address this gap, we compile a checklist of canonical macroeconomic "stylized facts" from the broad economics literature (Blanchard & Fischer, 1989; Williamson, 2014; Òscar Jordà et al., 2017; Axtell & Farmer, 2025), many of which have previously been used to evaluate other ABM models. We then use this checklist to assess SimCity's ability to exhibit established patterns and to simulate new ones under novel economic shocks.

5.1 Experiment Setup

We simulate an economy with a maximum of 1000 households. Each simulation proceeds in two phases. During the initial phase, which lasts 36 steps (3 years), new households gradually populate the city. During the second phase, which lasts 144 steps (12 years), the population remains fixed, and we observe the simulation under steady-state conditions. Detailed specifications of LLMs employed in our framework are provided in Appendix A.2. We provide detailed examples of prompts and responses in Appendix E.3. We also explore the robustness of SimCity across alternative LLMs and seeds in Appendix B.1.

For comparison, we select two traditional ABMs with predetermined rules LEN (Lengnick, 2013) and CATS (Gatti et al., 2011), a deep multiagent reinforcement learning model AI-economist (Zheng et al., 2022), a LLM-based system EconAgent (Li et al., 2024), and real-world data FRED from 1970 Q1 (Federal Reserve Economic Data)[1].

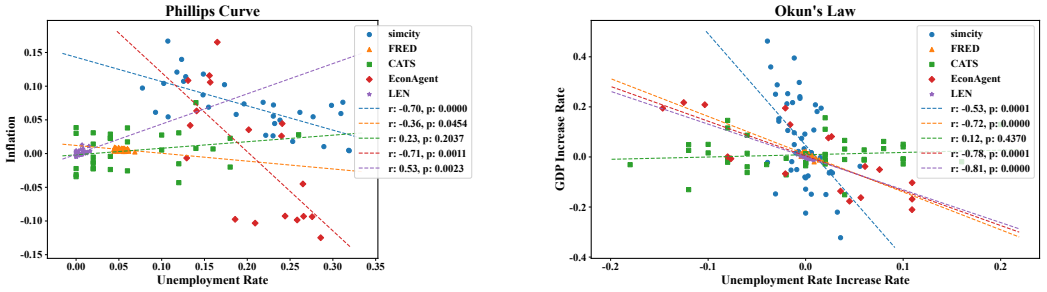
5.2 Macroeconomic Emergence

A central test of macroeconomic simulations is their correspondence to reality. It is important to emphasize, however, that the goal of our simulation is not quantitative reproduction of economic statistics, which can vary substantively across time and settings, but its ability to exhibit well-documented empirical regularities ("stylized facts") observed in real-world economies. Leveraging the strong role-play capabilities of LLM agents, **SimCity exhibits many of these regularities that traditional agent-based models have historically struggled to capture.** Table 2 summarizes the key phenomena and compares SimCity's performance with baseline models.

Phillips Curve and Okun's Law The Phillips Curve (Phelps, 1967) describes the short-run inverse relationship between unemployment and inflation, while Okun's Law (Okun, 1963) depicts the quarterly negative relationship between changes in the unemployment rate and real GDP growth. Following prior work (Li et al., 2024), we use these well-known empirical regularities as tests of the

[†]The paper by Lengnick (2013) claims to have verified the Beveridge Curve, but we are unable to reproduce it with `https://github.com/newwayland/baseline-economy`

[1]`https://fred.stlouisfed.org/`



(a) (b)

Figure 3: Emergence of the Phillips Curve and Okun's Law in SimCity simulations. r-value is the Pearson correlation coefficient, and p-value indicates the statistical significance of it.

plausibility of macroeconomic dynamics in our simulation. Recent economic studies using more modern methodologies and granular data document that the Phillips Curve slope is small and has flattened in recent decades (Hazell et al., 2022; Furlanetto & Lepetit, 2024). Consequently, a steeper or more negative slope does not necessarily indicate a more faithful reproduction of the Phillips Curve. We plot the relationships using one data point per quarter in Figure 3[2]. As shown in Figure 3, SimCity successfully demonstrates these relationships.[3]

Beveridge Curve The Beveridge Curve captures the negative relationship between the job vacancy rate, which is the number of unfilled job openings as a proportion of total job postings, and the unemployment rate (Blanchard & Diamond, 1989). SimCity successfully exhibits this relationship, whereas the previous LLM-driven ABM, EconAgent, lacks the firm module needed to model job vacancies.

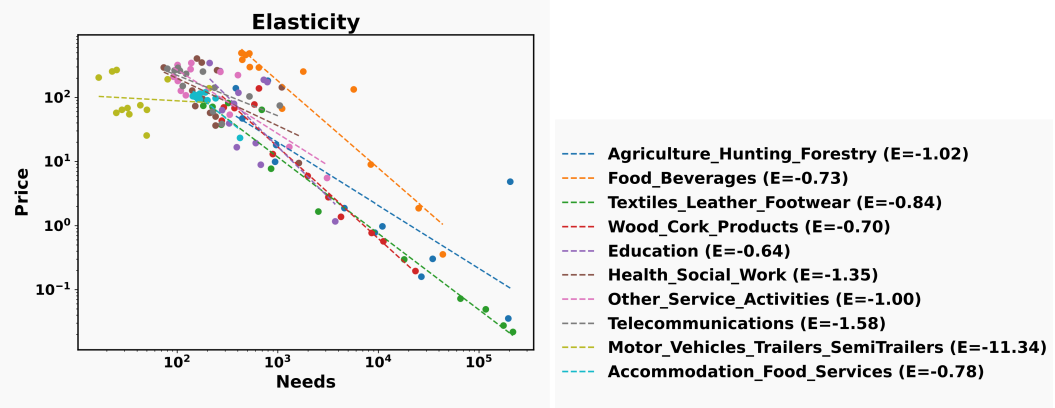


Figure 4: Estimated Price Elasticity of Demand (PED) across different product categories in the SimCity environment, obeying the Law of Demand and commonsense notions of necessity. $E = \frac{\Delta \log Q}{\Delta \log P}$ is the elasticity defined in Section 5.2.

Engel's Law Engel's Law states that the proportion of income spent on food (the Engel coefficient) decreases as income rises (Chai & Moneta, 2010). Simulations from SimCity correctly exhibit this pattern, reflecting that LLM-driven agents display human-like preferences. Notably, validating Engel's Law requires a heterogeneous goods market, which previous frameworks lacked and were therefore unable to study.

[2]For simulations that operate on an annual time scale, we use annual data points instead.

[3]We omit AI-Economist from this comparison because unemployment is not incorporated in its framework.

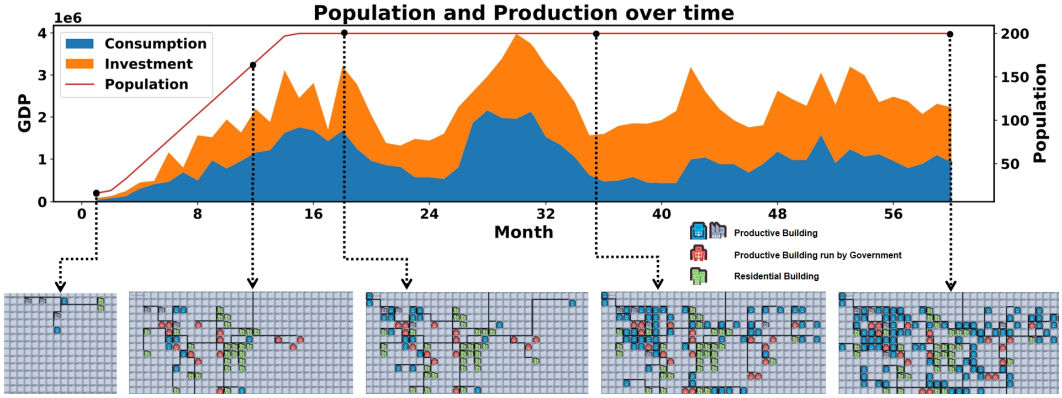


Figure 6: GDP & population curves, and map changes during the move-in phase. Another example is given in Appendix B.4.

Law of Demand To test the Law of Demand, we estimate the Price Elasticity of Demand (PED) for each good, defined as $E = \frac{\Delta \log Q}{\Delta \log P}$, which measures how quantity demand responds to changes in price. Standard economic theory predicts a negative elasticity for normal goods, but the magnitude depends on the importance of the good to consumers (Mas-Colell et al., 1995). As shown in Figure 4, SimCity not only exhibits the general negative association predicted by the Law of Demand but also captures the variation in elasticity across goods. For example, necessities like "food/beverages" are more inelastic ($-1 < E < 0$), whereas goods such as "motor/vehicles/trailers/semitrailers" are more elastic ($E < -1$), which is a pattern confirmed by numerous empirical studies (Perloff, 2009; Nelson et al., 2014). Lastly, as shown in Figure 2 (top-right), our volatility of investment is larger than that of consumption, which is consistent with real-world business cycle patterns.

5.3 Simulation Robustness

An important question for LLM-driven simulations is whether the observed economic phenomena remain stable and reproducible across runs. To address this, we run three sets of experiments with the same hyperparameters but different random seeds, and the results demonstrate that the observed regularities are robust. Experiment results can be found in Figure 8 in Appendix B.1.

5.4 Urban Expansion During Move-In

As mentioned in Section 3.1, our simulation involves a phase in which households move in. Figure 6 illustrates the developments during this stage. As shown in the figure, in the first 15 months, GDP steadily increases with the influx of residents. Then, as the influx ceases, the city transitions from a period of expansion to one of stable development after more than ten months of fluctuations. Meanwhile, we note that the VLM, without additional prompts, forms a clustered structure with residential areas near the center and production buildings on the periphery, in line with the Alonso-Muth-Mills (AMM) model of the mono-centric city (Brueckner, 2011).

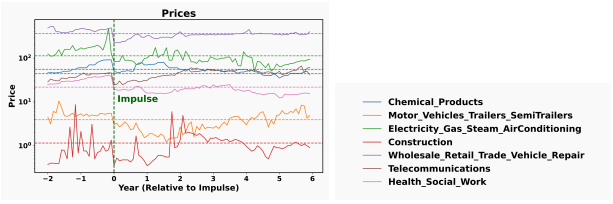


Figure 5: An external price impulse shows that prices maintain their long-term tendency, demonstrating price stickiness.

To evaluate the role of visual information in generating realistic urban forms, we conduct an ablation study to evaluate the necessity of the visual module by replacing the map images with text-based descriptions of the city layout. **Experiments show that without visual grounding, the city evolves into unrealistic geometries**, such as narrow linear formations, contrasting with the organic clusters formed by the VLM-empowered agents (visual comparison provided in Appendix B.3).

5.5 Responses to Exogenous Shocks

Price shock To investigate whether our framework can be used to analyze the effects of sudden external shocks to the economy, we separately apply a price-down and a price-up impact on the city. At the beginning of year 15, we randomly select 7 of 44 goods in two different simulations and apply a 50% price-down and a price-up impulse separately. We then let the simulation run for 6 years. Conceptually, one can think of these sudden, ephemeral changes in price levels of the selected goods as coming from exogenous shocks in the world market. Take electricity prices as an example, which is one of the shock goods in our simulations; it could come from a significant but brief conflict among energy-producing countries.

Impulse responses to a price-down shock are shown in Figure 5 and price-up at Figure 10 (Appendix B.2). The experiment shows that the good's price level gradually returns to equilibrium after an exogenous shock. Due to price stickiness, the full effects of the exogenous shock take some time to develop. Importantly, note that dotted lines in the figure represent long-run average prices of goods. Since the shock is one-off and the economy's fundamentals did not change, prices revert to their original levels in the long run, despite the temporary dislocation. This stylized experiment showcases SimCity's ability to produce qualitative predictions in novel economic scenarios.

Automation Shock To further demonstrate the ability of our framework to simulate novel economic scenarios, we consider a stylized automation shock where physical labor is entirely replaced by information technology (IT) skills (e.g., drivers become obsolete as cars become fully autonomous). We implement this shock by altering the skill requirements in production midway through the simulation (denoted as "0 to Event" in Figure 7). As Figure 7 shows, following the shock, a substantial and widening income gap emerges between workers with strong labor skills and those with strong IT skills. This divergence persists as the economy expands, highlighting the distributional consequences of skill-biased technological change. Shaded areas represent 95% confidence intervals.

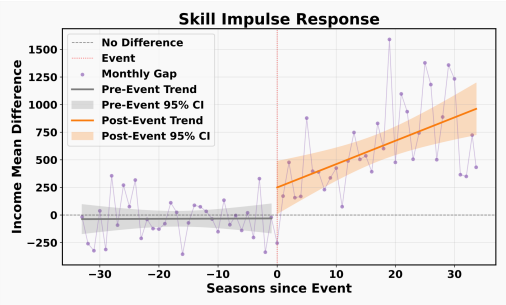


Figure 7: An automation shock leads to an income gap between benefited agents and others.

6 Conclusion

With SimCity, we conduct complex simulations that incorporate four distinct LLM-driven economic roles (households, firms, the central bank, and the government), with a visualized map and markets with heterogeneous goods. Experiments show that SimCity robustly matches a checklist of established macroeconomic phenomena. Moreover, when applied to novel economic scenarios, SimCity generates economically coherent and informative predictions. We hope SimCity could serve as a base platform for a more realistic economic simulation.

7 Limitations

Quantitative Simulation Reproductions of our framework are fundamentally qualitative rather than quantitative. Accurately simulating the real-world economy on a quantitative level remains a challenge. The lack of sufficiently detailed, comprehensive micro-level data and the high computational costs associated with running LLMs at scale are major constraints. Consequently, the direct application of our current framework for precise economic forecasting or real-world policy formulation is restricted. We anticipate that future research, supported by more advanced data methodologies and refined framework designs, will bridge this gap.

Lack of Complex Financial Activities Our current simulation framework does not incorporate real-world financial markets, such as those for bonds, stocks, or their derivatives. Moreover, the simulated agents do not employ mathematical optimization techniques in their investing strategies. These simplifications may limit the model's ability to simulate complex market dynamics accurately.

ETHICS STATEMENT

This work uses publicly available statistical data. No sensitive or private data is used. Names of agents are generated by LLMs and are not intended to represent any real individuals or entities. Our simulation has not been quantitatively calibrated with real-world data and does not constitute a prediction of reality. More ethics considerations are listed in Appendix A.

AI USE STATEMENT

This work utilizes LLM-driven agents to simulate. Additionally, we use LLMs to improve the grammar and readability of this manuscript. The LLM was not used for any other scientific aspects of this work, and all intellectual content is solely the product of the authors.

REFERENCES

Robert L. Axtell and J. Doyne Farmer. Agent-based modeling in economics and finance: Past, present, and future. *Journal of Economic Literature*, 63(1):197–287, 2025.

Olivier Blanchard. The state of macro. *Annual Review of Economics*, 1:209–228, September 2009. doi: 10.1146/annurev.economics.050708.142952.

Olivier Blanchard and Stanley Fischer. *Lectures on Macroeconomics*. The MIT Press, Cambridge, MA, 1989. ISBN 9780262022835.

Olivier J. Blanchard. Convergence? thoughts about the evolution of mainstream macroeconomics over the last 40 years. NBER Working Paper 33802, National Bureau of Economic Research, May 2025.

Olivier Jean Blanchard and Peter A. Diamond. The beveridge curve. *Brookings Papers on Economic Activity*, 20(1):1–76, 1989. URL https://www.brookings.edu/bpea-articles/the-beveridge-curve/.

András Borsos, Adrian Carro, Aldo Glielmo, Marc Hinterschweiger, Jagoda Kaszowska-Mojsa, and Arzu Uluc. Agent-based modeling at central banks: Recent developments and new challenges. Staff Working Paper 1122, Bank of England, 2024.

Jan K. Brueckner. *Lectures on Urban Economics*. MIT Press, Cambridge, MA, 2011. ISBN 9780262016360.

Andreas Chai and Alessio Moneta. Retrospectives: Engel curves. *Journal of Economic Perspectives*, 24(1):225–240, 2010. doi: 10.1257/jep.24.1.225. URL https://www.aeaweb.org/articles?id=10.1257/jep.24.1.225.

Jiangjie Chen, Xintao Wang, Rui Xu, Siyu Yuan, Yikai Zhang, Wei Shi, Jian Xie, Shuang Li, Ruihan Yang, Tinghui Zhu, et al. From persona to personalization: A survey on role-playing language agents. *arXiv preprint arXiv:2404.18231*, 2024.

Charles W. Cobb and Paul H. Douglas. A theory of production. *The American Economic Review*, 18 (1):139–165, 1928. ISSN 00028282, 19447981. URL http://www.jstor.org/stable/1811556.

Olivier Coibion and Yuriy Gorodnichenko. Why are target interest rate changes so persistent? *American Economic Journal: Macroeconomics*, 4(4):126–62, May 2012. doi: 10.1257/mac.4.4.126. URL https://www.aeaweb.org/articles?id=10.1257/mac.4.4.126.

Herbert Dawid and Domenico Delli Gatti. Agent-based macroeconomics. *Handbook of computational economics*, 4:63–156, 2018.

Ömer Dilaver, Richard Calvert Jump, and Paul Levine. Agent-based macroeconomics and dynamic stochastic general equilibrium models: Where do we go from here? *Journal of Economic Surveys*, 32(4):1134–1159, 2018. doi: 10.1111/joes.12249.

Francesco Furlanetto and Antoine Lepetit. The slope of the phillips curve. FEDS Working Paper 2024-43, Board of Governors of the Federal Reserve System, June 2024. URL https://ssrn.com/abstract=4880075. Available at SSRN.

Jordi Galí. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*. Princeton University Press, Princeton, NJ, second edition, 2015. ISBN 9780691164786.

Jordi Galí and Mark Gertler. Macroeconomic modeling for monetary policy evaluation. *Journal of Economic Perspectives*, 21(4):25–46, 2007. doi: 10.1257/jep.21.4.25.

Chen Gao. Xiaochong lan, nian li, yuan yuan, jingtao ding, zhilun zhou, fengli xu, and yong li. *Large language models empowered agent-based modeling and simulation: A survey and perspectives. Humanities and Social Sciences Communications*, 11(1):1–24, 2024.

Chen Gao, Xu Lan, Ning Li, et al. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, 11: 1259, 2024a. doi: 10.1057/s41599-024-03611-3.

Shen Gao, Yuntao Wen, Minghang Zhu, Jianing Wei, Yuhan Cheng, Qunzi Zhang, and Shuo Shang. Simulating financial market via large language model based agents, 2024b. URL https://arxiv.org/abs/2406.19966.

Domenico Delli Gatti, Saul Desiderio, Edoardo Gaffeo, Pasquale Cirillo, and Mauro Gallegati. *Macroeconomics from the Bottom-up*, volume 1. Springer Science & Business Media, 2011.

Patrick J. Glandon, Kenneth N. Kuttner, Sandeep Mazumder, and Caleb Stroup. Macroeconomic research, present and past. *Journal of Economic Literature*, 61(3):1088–1126, 2023. doi: 10.1257/jel.20211609.

Jonathon Hazell, Juan Herreño, Emi Nakamura, and Jón Steinsson. The slope of the phillips curve: Evidence from u.s. states. *The Quarterly Journal of Economics*, 137(3):1299–1344, August 2022. doi: 10.1093/qje/qjac010. URL https://doi.org/10.1093/qje/qjac010.

Vladimir Hlasny. Parametric representation of the top of income distributions: Options, historical evidence, and model selection. *Journal of Economic Surveys*, 35(4):1217–1256, 2021. doi: https://doi.org/10.1111/joes.12435.

Abe Bohan Hou, Hongru Du, Yichen Wang, Jingyu Zhang, Zixiao Wang, Paul Pu Liang, Daniel Khashabi, Lauren Gardner, and Tianxing He. Can a society of generative agents simulate human behavior and inform public health policy? a case study on vaccine hesitancy, 2025. URL https://arxiv.org/abs/2503.09639.

Yizhe Huang, Xingbo Wang, Hao Liu, Fanqi Kong, Aoyang Qin, Min Tang, Song-Chun Zhu, Mingjie Bi, Siyuan Qi, and Xue Feng. Adasociety: An adaptive environment with social structures for multi-agent decision-making, 2025. URL https://arxiv.org/abs/2411.03865.

Matthias Lengnick. Agent-based macroeconomics: A baseline model. *Journal of Economic Behavior & Organization*, 86:102–120, 2013.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. Econagent: Large language model-empowered agents for simulating macroeconomic activities, 2024. URL https://arxiv.org/abs/2310.10436.

Lars Ljungqvist and Thomas J. Sargent. *Recursive Macroeconomic Theory*. MIT Press, 4 edition, 2018. ISBN 978-0262047203.

Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, June 1995. ISBN 9780195073409.

Qirui Mi, Qipeng Yang, Zijun Fan, Wentian Fan, Heyang Ma, Chengdong Ma, Siyu Xia, Bo An, Jun Wang, and Haifeng Zhang. Econgym: A scalable ai testbed with diverse economic tasks, 2025. URL https://arxiv.org/abs/2506.12110.

Marco Del Negro, Stefano Eusepi, Marc Giannoni, Argia Sbordone, Andrea Tambalotti, Matthew Cocci, Raiden Hasegawa, and M. Henry Linder. The frbny dsge model. Staff Report 647, Federal Reserve Bank of New York, October 2013.

Julie Nelson, Neva Goodwin, Frank Ackerman, and Thomas Weisskopf. *Microeconomics in Context*. 01 2014. doi: 10.13140/2.1.4786.5287.

Arthur M Okun. *Potential GNP: its measurement and significance*. Cowles Foundation for Research in Economics at Yale University, 1963.

OpenAI et al. Gpt-4o system card, 2024. URL https://arxiv.org/abs/2410.21276.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior, 2023. URL https://arxiv.org/abs/2304.03442.

J.M. Perloff. *Microeconomics*. Addison-Wesley series in economics. Pearson/Addison Wesley, 2009. ISBN 9780321558497. URL https://books.google.com/books?id=8qyZtAEACAAJ.

Edmund S Phelps. Phillips curves, expectations of inflation and optimal unemployment over time. *Economica*, pp. 254–281, 1967.

Jinghua Piao, Yuwei Yan, Jun Zhang, Nian Li, Junbo Yan, Xiaochong Lan, Zhihong Lu, Zhiheng Zheng, Jing Yi Wang, Di Zhou, Chen Gao, Fengli Xu, Fang Zhang, Ke Rong, Jun Su, and Yong Li. Agentsociety: Large-scale simulation of llm-driven generative agents advances understanding of human behaviors and society, 2025. URL https://arxiv.org/abs/2502.08691.

Morten O. Ravn, Stephanie Schmitt-Grohé, and Martín Uribe. Macroeconomics of subsistence points. *Macroeconomic Dynamics*, 12(S1):136–147, 2008. doi: 10.1017/S1365100507070095.

Francesco Sergi. Dsge models and the lucas critique: A historical appraisal. Economics Working Paper Series 1806, University of the West of England, Bristol, October 2018.

Aaditya Singh, Adam Fry, Adam Perelman, et al. Openai gpt-5 system card, 2026. URL https://arxiv.org/abs/2601.03267.

Frank Smets and Rafael Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606, June 2007. doi: 10.1257/aer.97.3.586. URL https://www.aeaweb.org/articles?id=10.1257/aer.97.3.586.

Joseph E. Stiglitz. Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, 34(1-2):70–106, January 2018. doi: 10.1093/oxrep/grx057.

Nancy L. Stokey, Robert E. Lucas, and Edward C. Prescott. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989. doi: 10.2307/j.ctvjnrt76.

Richard Stone. Linear expenditure systems and demand analysis: An application to the pattern of british demand. *The Economic Journal*, 64(255):511–527, 1954. doi: 10.2307/2227743.

Servaas Storm. Cordon of conformity: Why dsge models are not the future of macroeconomics. *International Journal of Political Economy*, 50(2):77–98, 2021. doi: 10.1080/08911916.2021.1929582.

Jiakai Tang, Heyang Gao, Xuchen Pan, Lei Wang, Haoran Tan, Dawei Gao, Yushuo Chen, Xu Chen, Yankai Lin, Yaliang Li, et al. Gensim: A general social simulation platform with large language model based agents. *arXiv preprint arXiv:2410.04360*, 2024.

Leigh Tesfatsion and Kenneth L. Judd (eds.). *Handbook of Computational Economics: Agent-Based Computational Economics*, volume 2 of *Handbook of Computational Economics*. Elsevier, 2006. ISBN 978-0-444-50710-6. doi: 10.1016/S1574-0021(05)02016-0. URL https://doi.org/10.1016/S1574-0021(05)02016-0.

Yuyang Tian, Shunqiang Mao, Wenchang Gao, Lanlan Qiu, and Tianxing He. A visualized framework for event cooperation with generative agents, 2025. URL https://arxiv.org/abs/2509.13011.

Maximilian Puelma Touzel, Sneheel Sarangi, Austin Welch, Gayatri Krishnakumar, Dan Zhao, Zachary Yang, Hao Yu, Ethan Kosak-Hine, Tom Gibbs, Andreea Musulan, et al. A simulation system towards solving societal-scale manipulation. *arXiv preprint arXiv:2410.13915*, 2024.

David Vines and Samuel Wills. The rebuilding macroeconomic theory project part ii: Multiple equilibria, toy models, and policy models in a new macroeconomic paradigm. *Oxford Review of Economic Policy*, 36(3):427–497, September 2020. doi: 10.1093/oxrep/graa066.

Stephen D. Williamson. *Macroeconomics*. Pearson, 5th edition, 2014.

Sarah Wolf, Steffen Fürst, Antoine Mandel, Wiebke Lass, Daniel Lincke, Federico Pablo-Marti, and Carlo Jaeger. A multi-agent model of several economic regions. *Environmental modelling & software*, 44:25–43, 2013.

Michael Woodford. Convergence in macroeconomics: Elements of the new synthesis. *American Economic Journal: Macroeconomics*, 1(1):267–279, 2009. doi: 10.1257/mac.1.1.267. URL http://www.jstor.org/stable/25760267. Accessed September 23, 2025.

Michael Wooldridge. *An Introduction to MultiAgent Systems*. Wiley, 2 edition, May 2009. ISBN 978-0-470-51946-2.

Randall Wright, Philipp Kircher, Benoît Julien, and Veronica Guerrieri. Directed search and competitive search equilibrium: A guided tour. *Journal of Economic Literature*, 59(1):90–148, 2021. doi: 10.1257/jel.20191505.

Bushi Xiao, Ziyuan Yin, and Zixuan Shan. Simulating public administration crisis: A novel generative agent-based simulation system to lower technology barriers in social science research, 2023. URL https://arxiv.org/abs/2311.06957.

Chengxing Xie, Canyu Chen, Feiran Jia, Ziyu Ye, Shiyang Lai, Kai Shu, Jindong Gu, Adel Bibi, Ziniu Hu, David Jurgens, James Evans, Philip Torr, Bernard Ghanem, and Guohao Li. Can large language model agents simulate human trust behavior?, 2024. URL https://arxiv.org/abs/2402.04559.

Zelai Xu, Chao Yu, Fei Fang, Yu Wang, and Yi Wu. Language agents with reinforcement learning for strategic play in the werewolf game, 2025. URL https://arxiv.org/abs/2310.18940.

Stephan Zheng, Alexander Trott, Sunil Srinivasa, David C Parkes, and Richard Socher. The ai economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science advances*, 8(18):eabk2607, 2022.

Òscar Jordà, Moritz Schularick, and Alan M. Taylor. Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual*, 31:213–263, 2017. doi: 10.1086/696047.

A STATEMENTS

A.1 RISKS

Our SimCity framework is intended as a research tool for academic exploration of multi-agent dynamics in simulated environments. We explicitly caution against deploying this system for real-world macroeconomic policy-making, especially since there is no quantitative alignment. Due to LLM hallucinations and compounding errors in multi-agent communications, relying on simulation outcomes could lead to unintended policies affecting real-world economies.

A.2 ARTIFACTS DETAIL

LLMs We employ `gpt-4o-mini` (version `2024-07-18`) for regular reasoning and `gpt-4o` (OpenAI et al., 2024) (version `2024-11-20`) as the vision language model for establishment decision-making. Additionally, `gpt-5` (Singh et al., 2026) (version `2025-08-07`) is utilized for robustness checking. The three models are accessed via the Azure OpenAI API[4] using default sampling parameters, except for setting `temperature=0` to reduce generation variance. Both the input prompts and the generated content are in English. We utilize the API strictly for research purposes to simulate multi-agent economic behaviors, and our usage fully complies with the *OpenAI Terms of Use*[5].

Datasets We utilize several public economic datasets to align and evaluate our framework. Specifically, we employ the Demographic and Housing Characteristics (DHC) table from the U.S. Census Bureau[6] and the 2023 American Community Survey (ACS) IPUMS microdata, which provides detailed household-level income information. Additionally, we incorporate the Input-Output Tables (IOTable) provided by the OECD[7]. All of these datasets are publicly available, provided in English, and fully anonymized, containing no personally identifiable information.

Packages We utilize LangChain[8] 1.3.1 for establishing agents framework, with MIT license.

B EXPERIMENT DETAIL AND ADDITIONAL RESULTS

LLMs Costs Our simulation costs about 800,000 tokens, which is roughly $0.25 per step. For a standard 180-step simulation, the total cost is around $180.

Base Year We set prices for the first year as the base-year prices for calculating real GDP.

B.1 ROBUSTNESS

We run three sets of experiments with the same hyperparameters but different random seeds, and the results demonstrate that the observed regularity is reproducible.

Furthermore, we evaluate SimCity using `gpt-5` as the reasoning backbone. Figure 9 shows that the fundamental relationships remain consistent under a more advanced LLM.

B.2 PRICE IMPULSE

The price-up impulse response, Figure 10, is shown here due to space limitations.

[4] https://portal.azure.com/

[5] https://openai.com/policies/row-terms-of-use/

[6] https://data.census.gov/table?q=PCT12&d=DEC+Demographic+and+Housing+Characteristics

[7] https://www.oecd.org/en/data/datasets/input-output-tables.html

[8] https://github.com/langchain-ai/langchain.git

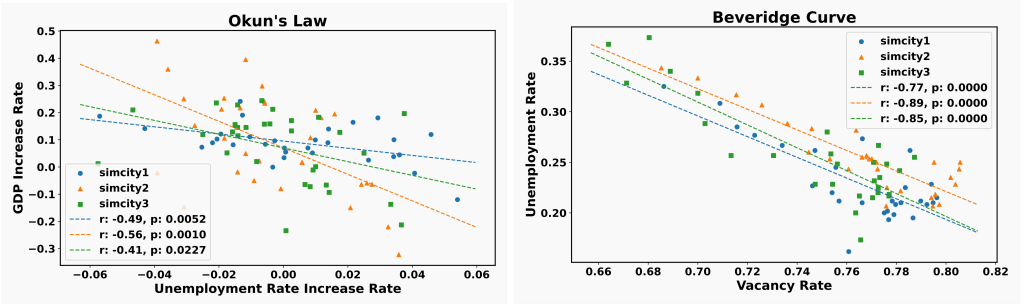


Figure 8: The results from different random seeds demonstrate that the observed regularity is robust.

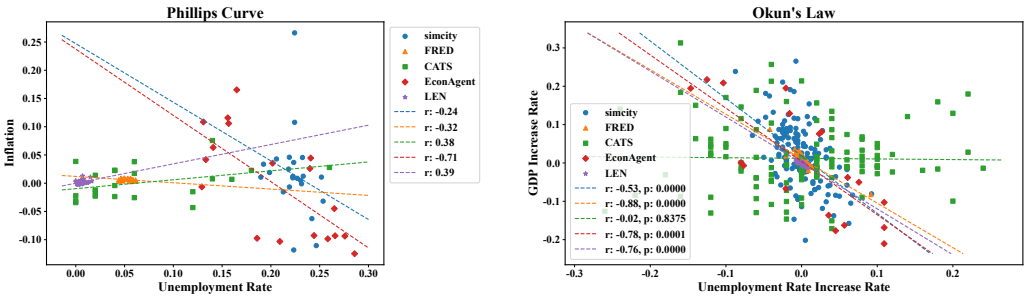


Figure 9: Macroeconomic patterns simulated using the `gpt-5` driven agents.

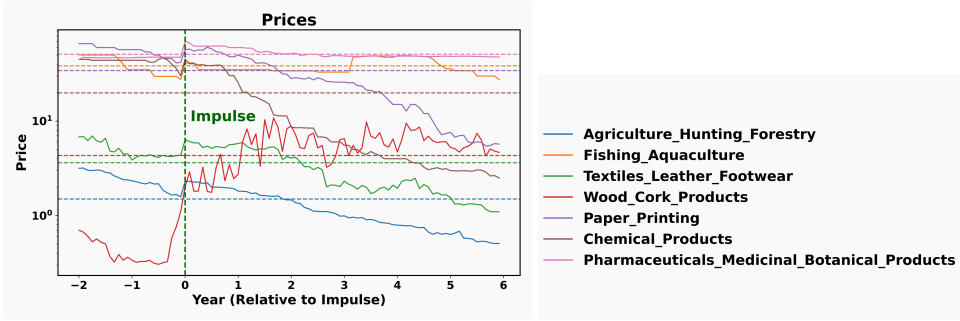


Figure 10: External impulse does not significantly affect long-run price trends.

B.3 Ablation Study: Visual Module

To validate the contribution of the vision component, we compare the standard SimCity framework against a text-only variant. In this ablation setting, the establishment agent receives a list of building coordinates rather than a rendered map image to make location decisions.

As illustrated in Figure 11, the absence of the visual module significantly degrades the spatial logic of urban planning. While the VLM-based approach (Figure 11a) fosters a natural, concentric city structure with a distinct center, the text-only approach (Figure 11b) fails to interpret spatial density effectively. Consequently, the agents form an artificial linear arrangement of buildings ("two lines"), confirming that the visual module is crucial for interpreting complex spatial information that is difficult to convey through text alone.

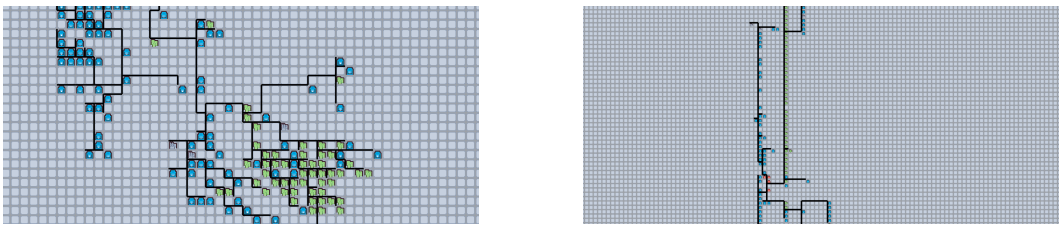


Figure 11: Impact of the Visual Module on urban layout. The full model produces a realistic, clustered city center, whereas removing the visual input results in an unrealistic linear topology. The left is generated with VLM while the right layout uses text description only.

B.4 Extra Example for City Expansion

An extra example for city expansion, Figure 12, is shown here.

C Agents Details

C.1 Notations

We define the following notations:

- The goods set, $\mathcal{G}$, where $|\mathcal{G}|$ denotes the number of heterogeneous goods present in our simulation. Certain subsets of $\mathcal{G}$ have a specific meaning, such as $\mathcal{G}_E$ for essential goods, $\mathcal{G}_T$ for transportation goods, and $\mathcal{G}_D$ for durable capital.
- Each step in our simulation represents a month. For clarity, time subscripts are suppressed unless needed for exposition. For example, the labor choice at time t for household i is written as l_i instead of $l_{i,t}$.

C.2 Households

A household i is characterized by the following attributes:

- **Cash Holdings m_i:** The amount of money held by the household.
- **Labor Participation:** An indicator $l_i \in \{0, 1\}$, where 1 means the household is employed to work in a position denoted by position_j and receives a salary S_j, and 0 means the household is unemployed. Note that we are modeling the extensive margin of labor, rather than of the intensive margin (hours worked).
- **Innate Skills:** Each household is endowed with heterogeneous skills. Please see Table 3 for the full list of skills. For each skill j, the innate skill level $s_{i,j} \sim U(s_{\min}, s_{\max})$ means that when the household works in a job requiring skill j, it supplies $l_i \times s_{i,j}$ effective units of labor. Additionally, the probability that a household is matched with a vacancy requiring skill j is an increasing function of $s_{i,j}$. This is in line with the directed search literature in economics (Wright et al., 2021).

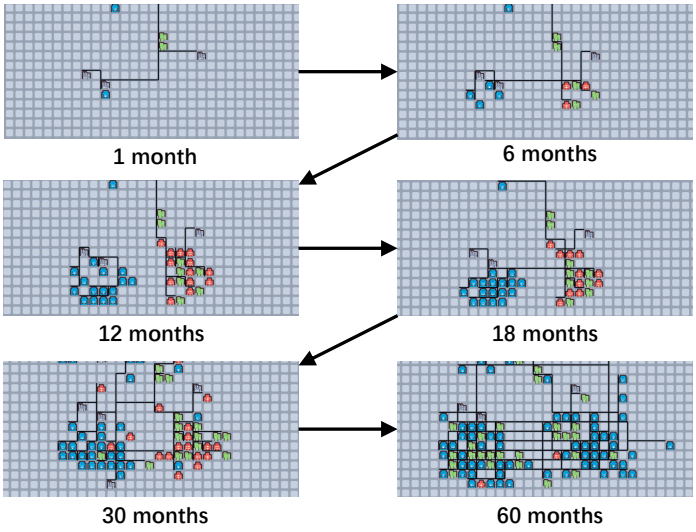


Figure 12: Another example of city dynamics during the move-in phase.

- **Heterogeneous Consumption Preferences:** The household has needs denoted by n_i. For each good $g \in \mathcal{G}$, $n_{i,g}$ represents the desired number of units of good g to be purchased. n_i may split to essential needs $n_{e,i}$ which is fixed denoting immutable demands and additional needs $n_{a,i}$. The modeling of essential goods has a long history in economics. See, for example, the Stone-Geary Utility function (Stone, 1954).
- **Housing:** Each household resides in a house $\mathcal{H}_i$. Consequently, the household must pay rent r_i to the owner of the property[9].

C.3 Firms

Each firm i is characterized by the following attributes:

- **Cash Holdings m_i:** The monetary resources currently available to the firm.
- **Output Good:** Each firm specifies a specific type of good $g_i \in \mathcal{G}$.
- **Shareholders:** A collection of households and/or government entities, each holding a share denoted by $\textbf{share}_{i,j}$.
- **Dividend Rate d_i:** A parameter $d_i \in [0, 1]$ representing the fraction of the firm's cash holdings that are distributed as dividends to shareholders. Specifically, a shareholder holding $\textbf{share}_{i,j}$ receives a dividend computed as:

$$D_j = \frac{\textbf{share}_{i,j}\, d_i\, m_i}{\sum_k \textbf{share}_{i,k}}.$$

- **Job Positions $\mathcal{J}_i$:** A list of job vacancies within the firm. Each position j requires a specific skill and offers a salary $S_{i,j}$.
- **Durable Capital $\mathcal{K}_i$:** The stock of durable capital assets that the firm has invested in to support production.
- **Estimated Value $\mathcal{V}_i$:** The firm's estimated value, calculated as:

$$\mathcal{V}_i = \frac{\sum_{j=t-12}^{t-1} \textbf{profit}_j}{12 \cdot \max(\mathcal{I}_{d,t}, \epsilon_d)} + m_i - L_i + \mathcal{K}_i.$$

[9] In this model, we assume that everyone rents their housing. In reality, if a household owns its home, one can think of it as renting a property in which it holds full ownership.

where $\mathcal{I}_{d,t}$ denotes the GDP deflator, $\epsilon_d = 0.02$ is a lower bound introduced to prevent division by zero and $\textbf{profit}_j$ denotes the net profit in month j, defined as the sales revenue minus the production expenditures.

- **Loans**: The amount of loans the firm owes to the financial system.

If a firm experiences zero income for a year and its assets are insufficient to cover its overdue debts, the firm is declared bankrupt. In that case, all employees are terminated, and the land it occupies is released.

A rentable apartment (residential building) is also formally modeled as a firm, though its decision-making is limited to adjusting its rental price based on conditions in the housing market.

C.4 Government

The government records a series of indicators:

- $\mathbf{nominalGDP}_t$: the sum of consumption, investment, and government spending, and $\mathbf{realGDP}_t$, computed things above using the base year's prices.
- $\mathbf{eq}_t$, $\mathbf{giniW}_t$, $\mathbf{giniI}_t$: measures of equality, wealth inequality (Gini coefficient), and income inequality (Gini coefficient), defined as follows. Suppose there are n households, and let $\{x_{(i)}\}_{i=1}^{n}$ denote the sorted (in increasing order) wealth (or income) of households. Then,

$$\textbf{giniW}_t = \frac{\sum_i \text{sorted_wealth}_i \cdot 2i}{n \cdot \text{sum_wealth} - \frac{n+1}{n}}, \tag{4}$$

$$\textbf{giniI}_t = \frac{\sum_i \text{sorted_income}_i \cdot 2i}{n \cdot \text{sum_income} - \frac{n+1}{n}}, \tag{5}$$

$$\textbf{eq}_t = 1 - \frac{n}{n-1}\textbf{giniI}_t. \tag{6}$$

- Unemployment rate u_t, along with broad monetary aggregates M0 and M1[10].
- Goods production Y_t, and average production $\bar{Y} = \frac{\sum_t Y_t}{T}$.
- Inflation: including wage inflation $\mathcal{I}_w$ and GDP inflation $\mathcal{I}_d$, each defined by:

$$\mathcal{I}_{w,t} = \frac{\sum_{\text{household}} S_{i,t}}{\sum_{\text{household}} S_{i,t-1}} - 1, \tag{7}$$

$$\mathbf{GDPdeflator}_t = \frac{\mathbf{nominalGDP}_t}{\mathbf{realGDP}_t}, \tag{8}$$

$$\mathcal{I}_{d,t} = \frac{\mathbf{GDPdeflator}_t}{\mathbf{GDPdeflator}_{t-1}} - 1. \tag{9}$$

C.5 Central Bank and Financial System

We implement a Taylor rule with interest rate smoothing in the policy behavior of our simulated central bank agent. Without smoothing, simulated interest rates exhibit implausibly high volatility in response to small shocks.

Interest rate smoothing refers to the empirical regularity that central banks adjust policy rates gradually over time, rather than immediately to the level implied by contemporaneous macroeconomic conditions. This behavior is often interpreted as reflecting forward-looking policy preferences, model uncertainty, or an aversion to financial market instability. In practice, this means that the actual interest rate set by the central bank is

$$r_t = \rho r_{t-1} + (1-\rho)\hat{r}_t, \tag{10}$$

where $\hat{r}$ is the policy rate computed from the unmodified Taylor rule 3, r_{t-1} the interest rate set last period, and ρ is the smoothing factor.

[10] M2 is not included since fixed deposits have not been defined.

Empirically, interest rate smoothing is a well-documented phenomenon. For example, Coibion & Gorodnichenko (2012) estimates a Taylor rule with smoothing and finds that the smoothing coefficient typically ranges between 0.7 and 0.9 in developed economies. We choose $\rho = 0.8$ as our parameterization.

D Framework Details

D.1 Environment Setup

There are 44 types of goods in SimCity. Each of them represents an economic industry in the real world as categorized by the Organization for Economic Co-operation and Development (OECD).

Initial prices of all goods are set at 50. However, the specific choice of initial prices does not matter for our simulation. We experimented with various initial prices, but in all cases, as we mention in 5.5, prices reach steady-state in a small number of steps. Similarly, wages of positions and rents of residential buildings are set arbitrarily.

D.2 Stages within Each Step

As aforementioned, each simulation step represents one month, and each step includes the following four stages:

- **Production and Trading Stage:** Firms produce goods, after which households and firms purchase goods for consumption and materials as planned.
- **Taxation and Dividend Stage:** Firms pay dividends, while the government collects taxes and disburses welfare.
- **Metabolic Stage:** New companies are established through equity financing from the investment pool, bankrupt companies are removed, and the population grows according to set rules.
- **Revision Stage:** Households, firms, the government, and the central bank agents review their situations and decide on their actions. Only the revision stage involves LLM agents.

D.3 Render Module

We utilize `flask-socketio` to build a web server and `Vue.js` to build a website.

The assets are from the open-source `tiny-battle` package created by Kenney[11].

D.4 Profile Setup

Age distribution is from the Demographic and Housing Characteristics (DHC) table from the U.S. Census Bureau[12].

To assign realistic initial cash holdings to households, we estimate the income distribution using U.S. microdata. We use the 2023 American Community Survey (ACS) IPUMS microdata, which contains detailed household-level income information. We assume a lognormal distribution for household income, a standard approximation in the economics literature due to its simplicity and its ability to capture the right-skewed nature of income data.[13] Formally, the initial cash holding is drawn from

$$\ln m_{i,0} \sim \mathcal{N}(\mu, \sigma^2) \tag{11}$$

where $\mu = 11.1496, \sigma^2 = 1.1455$, which are estimated via maximum likelihood estimation (MLE) with the 2023 ACS microdata.

[11] `https://www.kenney.nl/`

[12] `https://data.census.gov/table?q=PCT12&d=DEC+Demographic+and+Housing+Characteristics`

[13] While some heavier-tailed distributions (e.g., Pareto) better approximate top incomes, our simulation does not focus on the super-rich. Hence, the lognormal distribution is a reasonable choice. See Hlasny (2021) for a recent summary of related work.

D.5 SYNTHESIS

We assume that each category of goods is produced by a representative type of firm. Therefore, we generate a firm template for each category, which specifies the firm's name, the skill requirements for its positions, and an input-output "recipe" derived from the Input-Output Tables (IOTable) provided by the OECD.[14]

We process the IOTable to determine the inputs for each firm. For each category of goods, we normalize the input requirements from all other categories needed to produce one unit of output. We then select the top input categories that cumulatively account for more than 75% of the total input value, defining these as the necessary inputs for the corresponding firm's production process.

Then, we use the prompt introduced in Listing 1 to call a LLM to generate the processed input-output into a firm template.

Observing that the positions required skills in the generated firm templates contain many semantically similar skills, we use the prompt introduced in Listing 2 to call an LLM to merge these similar skills. Skills after merge are shown in Table 3.

Listing 1: Prompt for firm template generation.

```
I am currently developing a game. One of the tasks I need to do now is:
    Given a recipe, please help me design a building to produce this
    recipe.
First, I will give you a recipe, which is composed of some raw materials.
     Then, you need to design a suitable building to produce it. Noticing
     that the employees' core skill should be as abstract as possible.

``type`` must either "BusinessBuilding" or "ServiceBuilding". If you
    consider a type of product is mainly provided by government, you
    should set it as ServiceBuilding, else BusinessBuilding.
Here is an example:


```
### Recipe ###

{
  "input": {
    "RawFood": 1000
  },
  "output": {
    "Food": 1000
  }
}

Your Response:

  {
    "template_name": "Restaurant",
    "width": 1,
    "height": 1,
    "description": "A restaurant that uses raw food to serve food.",
    "type": "BusinessBuilding",
    "building_cost": 500000,
    "provide_radius": 10,
    "employees": [
      {
        "position_type": "Manager",
        "salary": 3000.0,
        "core_skills": [
          "Management"
        ],
```

[14] https://www.oecd.org/en/data/datasets/input-output-tables.html
```

```
      "importance": 1.0
    },
    {
      "position_type": "Cook",
      "salary": 1000.0,
      "core_skills": [
        "Cooking"
      ],
      "importance": 0.5
    },
    {
      "position_type": "Waiter",
      "salary": 1000.0,
      "core_skills": [
        "Service"
      ],
      "importance": 0.5
    },
    {
      "position_type": "Waiter",
      "salary": 1000.0,
      "core_skills": [
        "Service"
      ],
      "importance": 0.5
    }
  ],
  "recipes": [
    {
      "input": {
        "RawFood": 1000
      },
      "output": {
        "Food": 1000
      }
    }
  ]
}

```

Now the recipe I provide will listed follow "### Recipe ###"
You should respond with a factory design in the same format as the examples above.

Listing 2: Prompt for merging skills

```
[Skills]
```
The content above is a list of several skills from my economic simulation environment, SimCity. A job posting might require several of these skills. I now feel that there are too many of them and they need to be simplified. Please merge the skills that have the same meaning into a single skill (note: only merge two skills if and only if they are very similar; for example, Geology and Mining should not be merged because they are different specialties), and provide the merge relationships: a JSON object { "A": "B" } indicates that skill A is essentially skill B.

If several skills are to be merged into one category, choose the most representative one as B. Then, return this JSON object.

Table 3: List of skills.

Management	Quality Control	Operations Management	Logistics Management
Physical Labor	Aquaculture	Environmental Science	Engineering
Monitoring	Maintenance	Geology	Mining
Chemistry	Surveying	Technical Skills	Food Science
Machinery Operation	Safety Management	Supply Chain Management	Pharmaceutical Science
Laboratory Skills	Metalworking	Assembly	Energy Management
Water Management	Equipment Maintenance	Project Management	Design
Building	Sales	Customer Service	Vehicle Maintenance
Data Analysis	Regulations	Equipment Operation	Driving
Transportation	Communication	Culinary	Cleaning
Media Production	Equipment Handling	Logistics	Information Technology
Finance	Technical Support	Real Estate	Insurance
Market Analysis	Office Management	Administrative Support	Building Maintenance
Research	Consulting	Facility Maintenance	Basic Repairs
Support	Teaching	Marketing	Legal Knowledge
Human Resources	Sanitation		

E PROMPTING AND EXAMPLES

Structure of Prompt

System Prompt:

You are a citizen of SimCity and you are taking action improve your life. Basically speaking, your goal is to be consuming a greater variety of products, acquiring more money and assets, and having a stable job and residence.
There are some information about you and the city which may affect your life ...

User Prompt:

Profile
Your name is [name], and you are [age] years old ..... You current have [money] money. You are working at [place] as [position] with salary [salary]...
Report
... During the past 12 months, your average income is [average_income]. This month, your income is [total_income], which consists of the following parts:
- Salary: ...
- Benefit: ...

...
This month, your outcome is [total_outcome]
...
Observation
Here are opening positions in the city:
-[position_id]: ...
...
The social average return on investment is [roi], and the interest rate of the bank is [interest_rate].

Figure 13: Structure of prompts.

E.1 PROMPTING STRUCTURE

We aim to leverage the common-sense capabilities of large language models to act as human-like, heterogeneous agents. Figure 13 provides the structure of the prompt for a household agent.

As mentioned in Section 4, the Agent will receive the necessary information to make its decisions. The Profile includes some basic personal information about the Agent, such as current income, age, skills, job, and residence. The Report includes some of the Agent's recent experiences and information. The Observation presents information about the overall environment that the Agent should be aware of, such as vacant positions, the unemployment rate, prices of goods, and the Return on Investment (ROI).

Table 4: Example Listings

Agent Type	System Prompt	User Prompt	Response
Household	Listing 3	Listing 4	Listing 5
Government	Listing 6	Listing 7	Listing 8
Investment Pool	Listing 9	Listing 10	Listing 11
Firm(Productive Building)	Listing 12	Listing 13	Listing 14
Firm(Residential Building)	Listing 15	Listing 16	Listing 17

Prompts for other agents are structured in a similar manner. Detailed examples are provided in Appendix E.3.

E.2 Function Calling

The model interacts with the environment by means of Function Calling. The framework loads all the operations that the Agent can execute and appends formatted function names along with their descriptions to the prompt. The LLMs will return the actions to be taken and their parameters in JSON format. The framework will execute after a verification.

E.3 Full Examples

Full examples of all types of agents can be found at Table 4.

Listing 3: System prompt example of household agent.

```
You are a citizen in SimCity choosing actions for the next month. Your
    goals include satisfying a variety of needs, improving your finances,
     and maintaining suitable employment and housing.

The input has four sections:
- `### Profile`: your traits, skills, finances, employment, housing,
    needs, deposits, and loans.
- `### Report`: your income and expenses from the latest month.
- `### Observation`: currently available opportunities and relevant
    market information.
- `### Interfaces`: the exact actions currently available. Do not invent
    actions or arguments.

Base decisions only on supplied information. Returning no actions is
    valid when no change is beneficial.

Return only valid JSON:
{
  "reasoning": "A concise explanation grounded in the supplied state.",
  "actions": [
    "action_name(argument1, argument2)"
  ]
}

Use double quotes for JSON strings and no trailing commas.
```

Listing 4: User prompt example of household agent.

```
### Profile

Your name is Anthony, and you are 38 years old. Your ID is 2000183.You
    current have 285906.703 money in your account.
You are a Disorganized,Uncreative man
You are currently unemployed.
```

```
You are living in Residential Building at (12, 12) with rent 90.

You plan to spend a percentage of your income on the following items, at
    a total of 53.5000% of your income:
  - Agriculture_Hunting_Forestry: 7.5
  - Fishing_Aquaculture: 2.5
  - Food_Beverages: 15.0
  - Textiles_Leather_Footwear: 4.0
  - Wood_Cork_Products: 1.0
  - Paper_Printing: 2.0
  - Chemical_Products: 1.5
  - Pharmaceuticals_Medicinal_Botanical_Products: 3.0
  - Rubber_Plastics_Products: 2.0
  - Accommodation_Food_Services: 5.0
  - Education: 2.0
  - Health_Social_Work: 2.0
  - Arts_Entertainment_Recreation: 1.0
  - Other_Service_Activities: 1.0
  - Wholesale_Retail_Trade_Vehicle_Repair: 0.5
  - Public_Administration_Defence_SocialSecurity: 0.0
  - Electricity_Gas_Steam_AirConditioning: 2.5
  - Transport_Warehousing_Support: 0.0
  - Telecommunications: 1.0
  - IT_Information_Services: 0.0
  - Household_Employer_Activities: 0.0
  - Motor_Vehicles_Trailers_SemiTrailers: 0.0

Bank Account Anthony:
Demand deposit: 27318.80

You have no loans.
Demand deposit interest rate is 1.1137%, and loan rate is 3.1137%.
Your maximum additional borrowing capacity is 285906.70. This is a credit
     limit, not cash that you currently own.


### Report

This month, your income is 7087.822, which consists of the following
    parts:
  - Welfare: 31.071(0.4384%)

This month, your outcome is 7145.377, which consists of the following
    parts:
  - Tax: 0.000(0.0000%)
  - Rent: 90.000(1.2596%)


### Observation

Here are opening positions in the city:
  - position_id: 2000001, Production Worker requires skills Operations
      Management, Physical Labor, with salary 100.000
  - position_id: 2000094, Production Worker requires skills Operations
      Management, Physical Labor, with salary 60.884
  - position_id: 2000096, Production Worker requires skills Operations
      Management, Physical Labor, with salary 58.659
  - position_id: 2000200, Production Worker requires skills Operations
      Management, Physical Labor, with salary 88.072
  - position_id: 2000190, Food Scientist requires skills Food Science,
      with salary 84.944
You expect the work should cover your spending of 7145.377.

Here are available residential buildings in the city:
```

```
  - residential id: 200005 at (25, 26) with rent 100
  - residential id: 200015 at (15, 15) with rent 80
  - residential id: 200013 at (10, 11) with rent 90
  - residential id: 100002 at (10, 6) with rent 90
  - residential id: 200017 at (16, 10) with rent 90
  - residential id: 200006 at (5, 4) with rent 85
  - residential id: 200001 at (0, 0) with rent 90
  - residential id: 100001 at (10, 5) with rent 90

The monthly operating return on invested business capital is 1.9286%, and
     the monthly bank interest rate is 1.1137%. Actual cash dividends
    depend on each building's dividend policy.
If the ROI is higher than the bank's interest rate, consider investing in
     the market for better returns. You may also borrow from the bank to
    leverage your investment.
Otherwise, you may simply save your money in the bank.

The prices of goods on the market are as follows:
  - Food_Beverages: 264.054, 13.9984% (up)
  - Wood_Cork_Products: 50.000, -0.0000% (down)
When adjusting your needs for goods, please keep the following points in
    mind:
1. You can adjust your needs for goods based on the price change rate.
    All change rates are derived from the average prices of the past 12
    months.
2. Your demand for goods should also reflect your income. If your income
    increases, you should allocate a higher percentage of your income to
    luxury items and a lower percentage to food.

During the past 12 months, your average income is 38.807:
- 12 months ago, you were working at Production Worker with salary
    154.879, and your monthly income was 154.879.
- 11 months ago, you were unemployed, and your monthly income was 0.000.
- 10 months ago, you were working at Production Worker with salary
    85.405, and your monthly income was 85.405.
- 9 months ago, you were working at Production Worker with salary 85.405,
     and your monthly income was 85.405.
- 8 months ago, you were unemployed, and your monthly income was 0.000.
- 7 months ago, you were unemployed, and your monthly income was 0.000.
- 6 months ago, you were working at Laborer with salary 70.000, and your
    monthly income was 70.000.
- 5 months ago, you were working at Laborer with salary 70.000, and your
    monthly income was 70.000.
- 4 months ago, you were unemployed, and your monthly income was 0.000.
- 3 months ago, you were unemployed, and your monthly income was 0.000.
- 2 months ago, you were unemployed, and your monthly income was 0.000.
- 1 months ago, you were unemployed, and your monthly income was 0.000.



### Interfaces
Here are the actions you may take:
  - find_job(position_id: int): Apply for a job with the specified
     position ID. If you are already employed, you will resign from your
     current job first.
  - set_needs_budget(updates: dict[str, float]): Update one or more
     spending shares at once. Each value is a fraction of income between
     0 and 1. Items omitted from updates keep their current shares;
     supplied items overwrite their old shares. The resulting total share
      must not exceed 1.0. Example: set_needs_budget({'Food_Beverages':
     0.3, 'Education': 0.05}).
  - set_invest_rate(rate: float): Set the monthly investment rate (between
      0 and 1). This determines the proportion of your funds invested
     automatically each month.
```

```
 - borrow(amount: float): Borrow money from the bank. The loan must be
     repaid with interest. You can borrow for investment or to purchase
     necessities.
 - payback(loan_id: int): Repay the loan identified by loan_id.
 - save_money(amount: float): Deposit money into the bank to earn
     interest. The amount must be non-negative and cannot exceed your
     current liquid cash balance.
 - withdraw(amount: float): Withdraw money from your bank deposit.
 - move_to_home(building_id: int): Move into a specific residential
     building. If you are homeless, you will move in directly; otherwise,
      you will move from your current home.
 - wait(): Do nothing. Maintain the status quo.
 - The maximum amount you can pass to save_money right now is 285906.703.
```

Listing 5: Response example of household agent.

```
{
  "reasoning": "Your current expenses exceed your income, so you need to
      adjust your spending. To better align your budget with your
      financial situation, you should focus on essential needs,
      particularly food, and reduce spending on luxury items. The rent is
      already set to the lowest available option, so no change in housing
      is needed. Saving money is also advisable given your current balance
       and interest rate. Therefore, I will lower the percentages for
      luxury goods and increase the share for food and essentials.",
  "actions": [
    "set_needs_budget({'Food_Beverages': 0.3, 'Textiles_Leather_Footwear':
        0.02, 'Accommodation_Food_Services': 0.03})",
    "save_money(5000)"
  ]
}
```

Listing 6: System prompt example of government agent.

```
You are the governing body of SimCity, entrusted with the task of making
    strategic decisions to improve the city's overall standard of living.
There are four primary categories of information available to you that
    will affect your decision-making process: "### City Profile", "###
    City Statistics", "### Citizen Feedback", and "### Potential
    Initiatives".

The "### Report" section contains data regarding the city's current
    economic, social, and environmental status..

You should respond in JSON format as described below:
```
{
   "strategies": "strategies",
   [
      "initiative1(para1, para1, ...)",
      "initiative2(para1, para1, ...)",
      ...
   ]
}
```
Ensure the response can be parsed by Python `json.loads`, e.g.: no
    trailing commas, no single quotes, etc.

An example output might look like this:
```
{
   "strategies": "Given the city's high unemployment rate and the
       citizens' feedback about lack of job opportunities, it would be
       beneficial to invest in job creation programs. Furthermore, the
```
```

```
       high levels of pollution suggest a need for environmental
       initiatives.",
   "actions": [
      "LaunchJobCreationProgram('Technology Sector')",
      "ImplementEnvironmentalPolicy('Recycling Initiative')"
   ]
}
```

Listing 7: User prompt example of government agent.

```
### Report

Here is the government report:
The government has a balance of 878.741.

Bank Account Government:
You have no savings
You have no loans.
Demand deposit interest rate is 0.6108%, and loan rate is 2.6108%.
Your loanable amount is 10000878.74065722.

Here are the news in the past months:
1985-01-01 00:00:00: A sudden collapse in food sector starts. All food
    related sectors production efficiency decreases to 50% of its
    original value.

The government of the city has the following tax policy:
  - Tax Building by following brackets:
  - Down to 0.000 get 5.0% tax
  - From 0.000 to 100000.000 get 15.0% tax
  - Tax Citizen by following brackets:
  - Down to 0.000 get 5.0% tax
  - From 0.000 to 50000.000 get 10.0% tax

The government has collected 104553.065 as tax.
  - Tax Citizen are from:
    - 0.000 are from down to 50000.000
    - 33144.289 are from 50000.000 to 0.000
  - Tax Building are from:
    - 0.000 are from down to 100000.000
    - 71408.776 are from 100000.000 to 0.000

The government has distributed 64000.000 to every citizen as UBI.


There're 400 citizens in the city, and working or living in 330 buildings
    .
During the last month, the city has:

  - Consumption: 1637604.621
  - Investment: 332733.1813
  - Government Expense: 0
and total nominal GDP is 1970337.8023.
The real GDP, calculated by last month's price, is 1921755.1228.
The GDP deflator is 2.528%.

The social equality defined as Gini coefficient of income is 0.8438, and
    the wealth inequality is 0.8353.
The unemployment rate is 0.3275.
During the past 12 months, the city has:
  - Year 14, Month 10: Consumption: 1651127.9157, Investment: 274729.127,
      Government Expense: 0, GDP: 1925857.0427, Wealth Gini Coefficient:
      0.8505, Income Gini Coefficient: 0.8086, Unemployment Rate: 0.285,
      Homeless Rate: 0.08, Average Happiness: 0.9555, Average Wage:
      1732.675, Average Position Salary: 1722.0206, Broad Money Supply: M0
```

```
      44111103.5414, M1 44111103.5414, Deposit Rate: 0.0428, Loan Rate:
     0.0628, Reserve Rate: 0.05
  - Year 14, Month 11: Consumption: 1596120.9542, Investment: 463245.2733,
      Government Expense: 0, GDP: 2059366.2275, Wealth Gini Coefficient:
     0.8475, Income Gini Coefficient: 0.8213, Unemployment Rate: 0.2425,
     Homeless Rate: 0.1175, Average Happiness: 0.9522, Average Wage:
     1736.1225, Average Position Salary: 1742.6331, Broad Money Supply:
     M0 44111103.5414, M1 44111103.5414, Deposit Rate: 0.0394, Loan Rate:
      0.0594, Reserve Rate: 0.05
[... and other historic information.]

### Interfaces
Here are the actions you may take:
  - adjust_bracket(type: str, brackets: list[tuple[float, float]]): Adjust
      the tax bracket for a given type.
    type: 'Building' or 'Citizen'
    bracket: list of tuples, each tuple contains a threshold and a tax
        rate. 1 means tax 100% and 0.1 means tax 10%.
    e.g. adjust_bracket('Building', [(10000, 0.05), (50000, 0.1),
        (100000, 0.15), (500000, 0.2)])
  - adjust_UBI(amount: float): Adjust the amount of UBI(Universal Basic
     Income) welfare. Noting that if the government has no money, it will
      not distribute UBI.
  - invest(amount: float): Invest government's money into the investment
     market. This action may stimulate the economy.
  - borrow(amount: float): Borrow money from the bank, those money will be
      paid back with interest.
  - payback(loan_id: int): Pay back the loan with the given loan_id.
  - save_money(amount: float): Save money to the bank to get interest.
  - withdraw(amount: float): Withdraw money from the bank.
```

Listing 8: Response example of government agent.

```
{
  'strategies': 'Given the recent collapse in the food sector, it is
     crucial to address food security and the consequent economic
     downturn. The high unemployment rate and income inequality suggest
     that social programs and job creation initiatives will be beneficial
     . Additionally, recognizing that the government has no savings and a
      balance suitable for moderate investments, we should focus on
     stabilizing the economy and enhancing living conditions.',
  'actions': [
    'adjust_UBI(70000.00)',
    'invest(500000.00)',
    "adjust_bracket('Citizen', [(0.000, 0.05), (50000.000, 0.07)])",
    "adjust_bracket('Building', [(0.000, 0.05), (100000.000, 0.12)
  ]
}
```

Listing 9: System prompt example of investment pool.

```
You are a city planner in SimCity and you are tasked with planning the
    position and type of a new building.

You must follow the following criteria:
1) The plan should be based on the current city layout and economic
    situation.
2) Act as a city planner and provide a detailed development plan. You may
     construct various buildings and can build multiple instances of the
    same type if deemed suitable.
3) Always refer to the city report, which outlines the current status of
    the city.
4) Determine the type and location of new buildings based on the
    information provided in the city report.
5) Formulate a step-by-step plan for the construction of the buildings.
```

```
There is some information about the city which may affect your plan. The
    information includes two parts: "### Report" and "### Layout"

You should then respond to me with
Reasoning: Are there any steps missing in your plan? What is the purpose
    of each step in the plan? What does the city layout imply? What does
    the city report imply?
x: x_position of the topleft corner
y: y_position of the topleft corner
type: which type of building you'd choose. if you decide not to build,
    type should be "None"

You should only respond in JSON format as described below:
{
   "reasoning": str
   "investments": [
      {
         "x": int,
         "y": int,
         "type": str
      },
      ...
   ]
}
Ensure the response can be parsed by Python `json.loads`, e.g.: no
    trailing commas, no single quotes, etc.

Here are some examples:

INPUT:
### Report
There are several building templates available for construction:
  - Residential Building: Houses 100 people, costs 500 money to build.
  - Factory: Employs 50 people, costs 1000 money to build.
  - Shop: Employs 10 people, generates 200 money per month, costs 500
      money to build.

### Layout
[An Image]
OUTPUT:
{
   "reasoning": "Based on the city layout and report, the city cannot
       produce enough foods. The solution may be build a Restaurant.
       According to the layout, the (25, 25) is a good place.",
   "investments": [
      {
         "x": 25,
         "y": 25,
         "type": "Factory"
      },{
         "x": 23,
         "y": 17,
         "type": "Factory"
      },
      {
         "x": 55,
         "y": 23,
         "type": "Residential Building"
      }
   ]
}
```

Listing 10: User prompt example of investment pool.

```
### Report

Here is the average local price of goods in the city:
  - Nonenergy_Mining_Products: 47.915
  - Food_Beverages: 81.344
[... and other prices]
  - Education: 65.748
  - Health_Social_Work: 82.491

If one of the goods is not produced, the price will not be displayed.

There're 200 citizens in the city, and 25 residential buildings have
    capacity, and provide 115 empty houses.
The homeless rate is 0.125, which means 25 citizens are homeless.
and the average rent is 958.6206896551724

Here is the list of buildings that can be built in the city:
  - Agro-Processing Plant(Produces Agriculture_Hunting_Forestry): An agro-
      processing plant that transforms food beverages and agricultural
      products into refined agricultural goods. Cost 50000 to build, needs
       4 employees to run
  - Aquaculture Facility(Produces Fishing_Aquaculture): A facility
      dedicated to aquaculture, producing sustainable seafood and fishery
      products. Cost 50000 to build, needs 5 employees to run
[... and other firm templates]
  - Service Center(Produces Other_Service_Activities): A service center
      that coordinates a variety of professional services and
      administrative support to fulfill community needs. Cost 50000 to
      build, needs 7 employees to run
  - Residential Building: A place where people live. Cost 50000 to build,
      may contain up to 10 people

The size of investment pool is: 2103539.50867976

Those goods that citizens failed or partially satisfied because of the
    lack of supply or high price:
  - Electricity_Gas_Steam_AirConditioning: Lack 1051.8885135094476
  - Mining_Support_Services: Lack 0.0
[... and other goods]
  - Transport_Warehousing_Support: Lack 0.0
  - Household_Employer_Activities: Lack 0.0


### Layout

[A base64 encoded image]
```

Listing 11: Response example of investment pool.

```
{
  'reasoning': "Based on the city report, there are critical shortages in
      Electricity_Gas_Steam_AirConditioning, Telecommunications, IT
      Information Services, Construction, Motor Vehicles, Mining,
      Machinery, Real Estate, Manufacturing, Paper Printing, Food,
      Agricultural, Textiles, Financial, Chemical, Rubber/Plastics,
      Publishing, Fabricated Metals, Education, Health, Public Services,
      Entertainment, Aquaculture, Pharmaceuticals, Electrical Equipment,
      Transport, Logistics, and Service Activities. Many of these
      shortages are significant, particularly Construction, Motor Vehicles
      , Mining Products, Manufacturing, Real Estate, Energy, and
      Telecommunications. The residential sector has a high homeless rate
      even though there is reported vacant housing, indicating rent is
      likely unaffordable. Factory buildings should be clustered together
      and away from residential blocks per guidelines; most factories are
      concentrated on the left and center-bottom of the map. To address
```

```
     the most urgent needs and the highest 'lack,' I will prioritize: 1)
     Energy Production Facility (for electricity), 2) TechHub (for IT
     services/telecommunications), 3) Vehicle Manufacturing Plant (motor
     vehicles), 4) Mining Facility, 5) Food Processing Plant, 6)
     Construction Firm, 7) Community Health Center, 8) Educational
     Institution, 9) Real Estate Development Office. Factory sites will
     be placed in zones with existing factories (to the west, southwest,
     and center), while public/community services (health, education,
     real estate) should be placed near residential/high-density business
      clusters (southeast, east, and center-right). Public services will
     be distributed near residential clusters to optimize access. Each
     location is selected for minimal disruption and adjacency to similar
      use, supporting the existing city layout logic.",
  'investments': [
    {'x': -5, 'y': 12, 'type': 'Energy Production Facility'},
    {'x': -6, 'y': 13, 'type': 'TechHub'},
    {'x': -6, 'y': 10, 'type': 'Vehicle Manufacturing Plant'},
    {'x': -5, 'y': 8, 'type': 'Mining Facility'},
    {'x': -4, 'y': 12, 'type': 'Food Processing Plant'}
  ]
}
```

Listing 12: System prompt example of firm (productive building) agent.

```
You are a helpful assistant that utilize the information provided below
    to adjust strategy for my company.

The information of the building will report below "### Report", and
    actions you may do to the building will be listed after "### Actions
    ".
You must follow the following criteria:
1) You should analyze the report and utilize the actions to adjust the
    strategy for the building.
2) You may engage in appropriate equity financing or loan and enhance
    your output by purchasing capital.


You should only respond in JSON format as described below:
{
   "reasoning": "reasoning"
   [
      "action1(para1, para1, ...)",
      "action2(para1, para1, ...)",
      ...
   ]
}
Ensure the response can be parsed by Python `json.loads`, e.g.: no
    trailing commas, no single quotes, etc.

One of the example outputs is as follows:
{
   "reasoning": "The building is in a good condition and the rent is low.
       You should consider increasing the rent to maximize profit.",
   "actions": [
      "IncreaseRent(1000)"
   ]
}
```

Listing 13: User prompt example of firm (productive building) agent.

```
### Report

Here is the monthly report for Educational Institution (No. 200025):
```

```
Overall, during this month, the building gross profit is 1040.80, with
    rest cash 5784.46. It has 1329.2468 durable investments for promoting
     production.
The building has 0 employees, 6 vacant positions, and run with efficiency
     0.0000%.
The building aims to maximize its profit, and the following is the
    detailed income and expense:

This month, the building income is 1040.800, which consists of the
    following parts:
  - Local trading income: 1040.800(100.0000%)

There's no outcome this month

The efficiency of this building is 0.0000%, which consists of the
    following parts:
  - Principal has 28.5714% importance, but no one is working here. If an
      employee with average skill(Management) works here, about 1804.5412
      goods will be produced per month.
  - Teacher has 22.8571% importance, but no one is working here. If an
      employee with average skill(Teaching) works here, about 1543.5823
      goods will be produced per month.
[... and other employees]

The building has 6 positions, and here are the details:
  - position_id: 2000067, None(Principal), salary: 100
  - position_id: 2000068, None(Teacher), salary: 100
[... and other positions]
The building has no employees, so you may consider to increase salary of
    the position to attract more employees.

The building run with profit: 1040.8000. According to current dividend
    rate, 50.0000% of current balance will be distributed to shareholders
    , which is 2892.2278, and the rest 2892.2278 will stay in the
    building.

The social average return on investment is 10.6795%, and the interest
    rate of the bank is 37.9503%.
If roi is higher than the interest rate of the bank, you can invest your
    money in the investment market to get a higher return on investment
    and you may borrow money from the bank to invest if you need.
Otherwise, you may simply save your money in the bank.

The building has the following goods in storage:
  - Education: 24230.216
  - Public_Administration_Defence_SocialSecurity: 5.984
  - Health_Social_Work: 1.6456

Bank Account Educational Institution:
You have no savings
Loans:
  - id: 89, loan amount: 2100.0, should be paid back within 4 months
  - id: 109, loan amount: 3139.45905, should be paid back within 5 months
[... and other loans.]

Demand deposit interest rate is 37.9503%, and loan rate is 40.7381%.
Your loanable amount is 265361.0053339647.


The building has the following price setting for selling goods:
  - Education: The average price in the city is 65.7481, changed (8.2185%(
      up)) compared to last year. And the building's price is 80.

There are some items that you may produce more to sell because of
    shortage currently on the market:
```

```
 - Education: short of 905.9101880203923 units goods on the market
[... and other goods.]

During the past 12 months, the building has generated 18712.751 profit,
    running at an average efficiency of 0.424. Capital of the building
    has changed from 995.700 -> 1329.247 (333.547).
 Last year, the building produced 24786.920 goods, and sold
     502.607(2.0277%) goods. The storage has changed from {'Education':
     3679.6098, 'Health_Social_Work': 4.052} to {'Education': 24230.216, '
     Public_Administration_Defence_SocialSecurity': 5.984, '
     Health_Social_Work': 1.6456}


### Interfaces
Here are the actions you may do to the building:
 - fire(position_id): Terminate the employment of an employee associated
     with a given position_id. This action will result in a one-month
     salary compensation, potentially reducing efficiency but also
     decreasing monthly expenditures
 - adjust_salary(position_id, new_salary): Modify the salary of a
     specific position using its position_id. Increase the salary to
     retain/attract skilled employees, or decrease it to reduce expenses.
 - close_vacancy(position_id): Close a vacancy for a specific position.
     This action will set the salary of the position to 0, effectively
     closing the position temporarily.
 - set_max_running_efficiency(efficiency): Set the maximum running
     efficiency of the building. An efficiency lower than 1 indicates
     that the firm operates below its full capacity (i.e., "idles") to
     reduce production output and material costs.
 - capital_invest(amount): Buying durable goods with [amount] money to
     increase the capital amount of the building. This action will
     improve productivity of the building.
 - equity_finance(amount): Issue shares to raise money. This action will
     increase the current cash of the building.
 - borrow(amount): Borrow money from the bank, those money will be paid
     back with interest. You may borrow money from the bank to invest in
     the building or to cover the cost of the building.
 - payback(loan_id): Pay back the loan with the given loan_id.
 - adjust_price(name, new_price): Alter the price of a particular product
     . If your productivity is in surplus, you may decrease the price to
     sell more goods, and if the product in market is in shortage, you
     may increase the price to boost profit.
 - set_dividend_rate(rate): Modify the dividend rate of the building.
     This rate determines the proportion of profits that will be
     distributed to shareholders. e.g. set_dividend_rate(0.5) will
     distribute 50% of profits to shareholders.
```

Listing 14: Response example of firm(Productive Building) agent

```
{
 'reasoning': 'The building is currently experiencing a good profit
     margin and a surplus in cash, but it has a low efficiency rate of
     64% primarily due to vacant positions. To optimize operations and
     maximize profit, new positions should be added, particularly for
     Civil Engineer and Site Foreman roles, which can significantly
     improve production efficiency. Furthermore, since the current
     unemployment rate is high, we can lower the salaries of existing
     roles to manage expenses while offering competitive salaries for new
      positions. We should also consider raising the price of
     Construction goods since there is a market shortage, which allows us
      to capitalize on this demand. Additionally, the building should
     focus on addressing overdue loans, potentially by borrowing more to
     clear outstanding debts if necessary.',
 'actions': [
   'add_position()',
```

```
   'add_position()',
   'adjust_salary(2002754, 1500)',
   'adjust_salary(2000007, 3500)',
   'adjust_salary(2000083, 2500)',
   "adjust_price('Construction', 2000)",
   'borrow(30000)'
  ]
}
```

Listing 15: System prompt example of firm (residential building) agent.

```
You are a helpful assistant that utilize the information provided below
    to adjust strategy for a residential

The information of the building will report below "### Report", and
    actions you may do to the building will be listed after "### Actions
    ".
You should analyze the report and utilize the actions to adjust the
    strategy for the building.

You should only respond in JSON format as described below:
{
   "reasoning": "reasoning"
   [
      "action1(para1, para1, ...)",
      "action2(para1, para1, ...)",
      ...
   ]
}
Ensure the response can be parsed by Python `json.loads`, e.g.: no
    trailing commas, no single quotes, etc.

One of the example outputs is as follows:
{
   "reasoning": "The building is in a good condition and the rent is low.
       You should consider increasing the rent to maximize profit.",
   "actions": [
      "set_rent(1000)"
   ]
}
```

Listing 16: User prompt example of firm (residential building) agent.

```
### Report

Here is the monthly report for Residence:
The residential building lives 1 citizens, and its capacity is 10.
The rent per month is 925, and here're other building's rent per month:
  - Residential Building at (14, 5): 900
  - Residential Building at (11, 2): 900
[... and other residential buildings' position and rent]


### Interfaces
Here are the actions you may do to the building:
  - set_rent(new_rent): Alter the rent per month for the residential
      building. This action will directly affect the monthly income of the
       building.
```

Listing 17: Response example of firm (residential building) agent.

```
{
  'reasoning': 'The building has a low occupancy rate with only 1 out of
      10 units filled. The current rent is at the higher end compared to
      other similar residences nearby. To attract more residents, we
      should consider lowering the rent slightly to improve occupancy.',
```

```
 'actions': ['set_rent(900)']
}
```

F MORE RELATED WORK

Social Simulation Existing work in this domain can be broadly divided into two categories. The first category encompasses general-purpose social simulation frameworks such as Park et al. (2023), Tian et al. (2025), Tang et al. (2024), Piao et al. (2025), and Huang et al. (2025), which aim to model general societal patterns. The second category investigates specific social phenomena, such as public administration crises (Xiao et al., 2023), health policy (Hou et al., 2025), political manipulations (Touzel et al., 2024), financial market (Gao et al., 2024b), and deduction games (Xu et al., 2025).