EconAgent: Large Language Model-Empowered Agents
for Simulating Macroeconomic Activities

Nian Li[1]   **Chen Gao**[2*]   **Mingyu Li**[2]   **Yong Li**[2*]   **Qingmin Liao**[1]

[1]Shenzhen International Graduate School, Tsinghua University   [2]Tsinghua University

linian21@mails.tsinghua.edu.cn   chgao96@gmail.com   liyong07@tsinghua.edu.cn

Abstract

The advent of artificial intelligence has led to a growing emphasis on data-driven modeling in macroeconomics, with agent-based modeling (ABM) emerging as a prominent bottom-up simulation paradigm. In ABM, agents (*e.g.*, households, firms) interact within a macroeconomic environment, collectively generating market dynamics. Existing agent modeling typically employs predetermined rules or learning-based neural networks for decision-making. However, customizing each agent presents significant challenges, complicating the modeling of agent heterogeneity. Additionally, the influence of multi-period market dynamics and multifaceted macroeconomic factors are often overlooked in decision-making processes. In this work, we introduce **EconAgent**, a large language model-empowered agent with human-like characteristics for macroeconomic simulation. We first construct a simulation environment that incorporates various market dynamics driven by agents' decisions regarding work and consumption. Through the perception module, we create heterogeneous agents with distinct decision-making mechanisms. Furthermore, we model the impact of macroeconomic trends using a memory module, which allows agents to reflect on past individual experiences and market dynamics. Simulation experiments show that EconAgent can make realistic decisions, leading to more reasonable macroeconomic phenomena compared to existing rule-based or learning-based agents. Our codes are released at https://github.com/tsinghua-fib-lab/ACL24-EconAgent.

1 Introduction

The advent of artificial intelligence (AI) has indeed transformed the research paradigm in traditional economics (Jorgenson, 2001). In the digital economy era, individual economic behavior can be elaborately recorded and analyzed, leading to a growing emphasis on data-driven modeling in macroeconomics. These data sources may include online consumption, social media activity, and more, allowing economists to better understand individual economic behavior rather than relying solely on traditional macroeconomic indicators (Schorfheide and Song, 2015). Research in macroeconomics aims to analyze and predict economic variables quantitatively. Early empirical statistical models, such as the Phelps Model (Phelps, 1967), and the work of Kydland and Prescott (Kydland and Prescott, 1982), focused on data-driven analysis and policy outcome prediction but struggled to handle significant shocks. Later, Dynamic Stochastic General Equilibrium (DSGE) models (Christiano et al., 2005) were introduced to capture economic dynamics and shocks but assumed a perfect world.

In the last two decades, agent-based modeling (ABM) has emerged as a promising paradigm for simulating macroeconomics from the bottom up, allowing diverse agents to interact in a complex system (Ding et al., 2024) without assuming a predetermined equilibrium (Farmer and Foley, 2009). The evolution of ABM in macroeconomics can be divided into two stages. Early models (Tesfatsion and Judd, 2006; Brock and Hommes, 1998) relied on predetermined rules but made oversimplified assumptions about agent behaviors. Later learning-based models (Trott et al., 2021; Zheng et al., 2022; Mi et al., 2023), were trained with large-scale behavioral data. Essentially, these models transform multifaceted economic factors into agent decisions. However, customizing decision-making mechanisms for each agent presents substantial difficulties. Using customized rules requires extensive expert knowledge and sophisticated model calibration (Windrum et al., 2007). Similarly, employing customized neural networks leads to a sharp increase in network parameters and difficulties in model training (Mi et al., 2023). These difficulties raise the challenge of modeling

_____
*Corresponding author.

agent heterogeneity. Additionally, existing models typically focus on the individual situations of the current period, making it difficult to consider past periods, their variations, and multifaceted macroeconomic factors. This complicates the modeling of the influence of macroeconomic trends and dynamics.

Recently, the field of AI has witnessed the rise of large language models (LLMs) (Zhao et al., 2023b). Leveraging this advancement, LLM-empowered agents have demonstrated capabilities in reasoning, planning, and decision-making. In this work, we design **EconAgent**, a LLM-empowered agent with human-like characteristics for macroeconomic simulations. We first construct a simulation environment that includes labor and consumption market dynamics driven by agents' decisions on working and consumption, as well as fiscal and monetary policies. Using a perception module that targets agent profiles and mirrors real-world economic situations, we create heterogeneous agents automatically exhibiting different decision-making mechanisms. Additionally, we model the influence of macroeconomic trends with a memory module, enabling agents to reflect on past individual experiences and market dynamics.

In our experiments, classic macroeconomic phenomena, such as inflation in the consumption market and the unemployment rate in the labor market, are reproduced more reasonably compared to traditional rule-based or learning-based agents. We also observe that the EconAgent exhibits human-like decision-making patterns, reflecting the complexities of human economic behavior. These agents demonstrate swift adaptability in response to changes in the internal and external environment.

The contributions of this work can be summarized as follows:

- We take the first step to integrate LLMs into the domain of macroeconomic simulations, bridging two seemingly disparate fields into a cohesive research avenue.

- We conduct macroeconomic simulations in our constructed environment driven by EconAgent, resulting in the emergence of classic macroeconomic phenomena and regularities.

- The results show that our approach not only enhances the realism and depth of macroeconomic simulations but also provides a promising avenue for future research, potentially reshaping how we study and understand the intricacies of macroeconomics.

2 Framework Overview

In this section, we present the overall framework of the macroeconomic simulation. As illustrated in Figure 1, it follows the well-acknowledged simulation frameworks and provides an environment driven by EconAgent. The simulation encompasses four primary components: labor, consumption, financial markets, and government taxation. We simulate the two most critical decisions individuals make in real life: work and consumption (Gatti et al., 2011; Wolf et al., 2013), which subsequently influence the fiscal revenues of the government (Zheng et al., 2022; Trott et al., 2021) and impact the dynamics of labor and consumption markets (Lengnick, 2013; Deissenberg et al., 2008; Dawid et al., 2012). Accordingly, banks adjust interest rates based on market inflation or deflation (Wolf et al., 2013; Dawid and Gatti, 2018).

2.1 Agent Decisions

Labor supply and consumption are necessary agent (household) decisions in macroeconomic simulations. In our framework, each simulation step indicates one month, in which agent i decides,

- whether to work $l_i \sim Bernoulli(p_i^w)$, where p_i^w is the work propensity. If they decide to work ($l_i = 1$), they receive a monthly wage as the income, which varies among agents. Each agent is initialized with an hourly wage w_i following the Pareto distribution (Zheng et al., 2022), and the monthly wage v_i is calculated by multiplying 168 hours (21 working days at 8 hours/day (Lengnick, 2013)). Those who abstain from work ($l_i = 0$) have an income of zero.

- the consumption propensity p_i^c, indicating the proportion of their wealth (including their current savings and income in this month) they wish to spend for essential goods.

As the first challenge, simulating heterogeneous agents' decisions is vital for the emergence of macroeconomic phenomena. Moreover, each agent is influenced by multifaceted economic factors, such as the expected income, the tax paid, *etc*. However, conventional simulations typically model a limited number of factors via predetermined equations (Lengnick, 2013; Gatti et al., 2011; Wolf et al.,
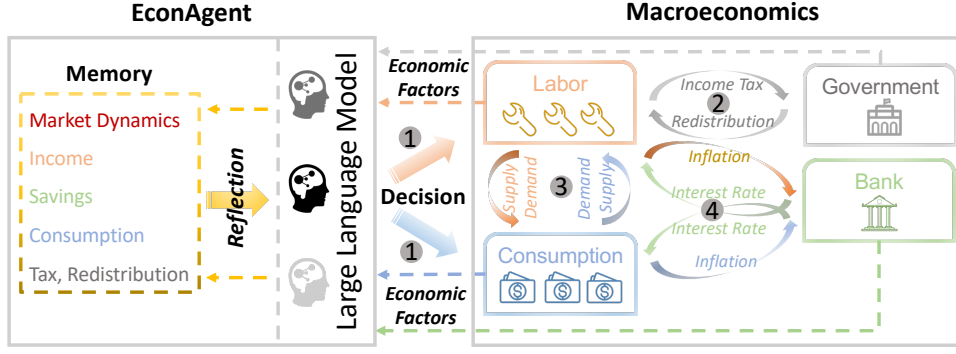
Figure 1: The illustration of our EconAgent (left) and simulation environment (right).

2013). These models lack of flexibility to simulate diverse decision-making mechanisms and assume one or a few representative rules (Axtell and Farmer, 2022).

2.2 Government Taxation

The government assumes the responsibility of taxation and provision of public services in society, as well as fiscal redistribution to ensure social equity. Taxes are collected from all the agents' income[1]. The progressive tax for income z_i is calculated as follows,

$$
T(z_i) = \sum_{k=1}^{B} \tau_k \left((b_{k+1} - b_k) \mathbf{1}\left[z_i > b_{k+1}\right] \right.
$$
$$
\left. + (z_i - b_k) \mathbf{1}\left[b_k < z_i \leq b_{k+1}\right] \right), \quad (1)
$$

where b_k is the k-th tax bracket, τ_k is corresponding tax rate, and $\mathbf{1}[\cdot]$ is indicator function. The tax brackets and rates are set as the 2018 U.S. Federal tax schedule.

The tax revenue is then evenly redistributed among all the agents. Therefore, the post-tax income is

$$
\hat{z}_i = z_i - T(z_i) + z^r = z_i - T(z_i) + \frac{1}{N}\sum_{j=1}^{N} T(z_j),
$$
$$
(2)
$$

where z^r indicates the redistribution. The individual savings for the agent are then updated as follows,

$$
s_i \leftarrow s_i + \hat{z}_i \quad (3)
$$

2.3 Productivity and Consumption

Incorporating agent decisions and government taxation, we simulate labor and consumption market dynamics based on economic principles. First,

working agents contribute 168 hours of productivity monthly, translating to the production of essential goods[2]. The inventory of goods G is then updated as,

$$
G \leftarrow G + S = G + \sum_{j=1}^{N} l_j \times 168 \times A, \quad (4)
$$

where S denotes the production volume from agents' labor supply, and A is a universal productivity. As for the consumption, the total demand of goods is,

$$
D = \sum_{j=1}^{N} d_j = \sum_{j=1}^{N} \frac{c_j}{P} = \sum_{j=1}^{N} \frac{p_j^c s_j}{P}, \quad (5)
$$

where d_j is the intended demand of agent j, c_j is the intended consumption, s_j is current savings, and P is the price of essential goods. Furthermore, both labor and consumption markets evolve based on the imbalance between supply and demand. Specifically, the imbalance is defined as,

$$
\bar{\varphi} = \frac{D - G}{\max(D, G)}, \quad (6)
$$

When the essential goods are in shortage, i.e., the supply can not meet the demand, the worker's wage should be increased to stimulate production. Due to the rise in labor costs for firms, they will also increase the goods price to ensure a certain profit margin (Lengnick, 2013; Dawid and Gatti, 2018; Wolf et al., 2013). The hourly wage is adjusted as follows,

$$
w_i \leftarrow w_i(1 + \varphi_i), \varphi_i \sim sign(\bar{\varphi})U(0, \alpha_w|\bar{\varphi}|),
$$
$$
(7)
$$

and the price is adjusted as follows,

$$
P \leftarrow P(1 + \varphi_P), \varphi_P \sim sign(\bar{\varphi})U(0, \alpha_P|\bar{\varphi}|),
$$
$$
(8)
$$

---

[1]We only consider income tax in this work and leave other taxes for future work, such as value-added tax.

[2]Note that we leave the simulation of firms as future work.

where α_w and α_P are the maximum rate of change when adjusting the wage and price, respectively. We also simulate the dynamics of goods consumption. Specifically, an agent j is randomly selected to consume essential goods, and real consumption goods and money are limited by current inventory of goods,

$$\hat{d}_j = \min(d_j, G), \hat{c}_j = \hat{d}_j \times P \qquad (9)$$

which means the demand is met if, and only if, there is a sufficient supply. The inventory of total goods also decreases,

$$G \leftarrow G - \hat{d}_j. \qquad (10)$$

The process continues until every agent has consumed goods once.

2.4 Financial Market

Annually, the savings of each agent increase based on the interest rate set by the bank,

$$s_i \leftarrow s_i \times (1 + r). \qquad (11)$$

Furthermore, in the first month of each year, the bank adjusts the interest rate based on the inflation in the consumption market and the unemployment rate in the labor market. Specifically, we adopt the widely-used Taylor rule to set the interest rate (Wolf et al., 2013; Dawid and Gatti, 2018),

$$r = \max(r^n + \pi^t + \alpha^\pi (\pi - \pi^t) + \alpha^u (u^n - u), 0), \qquad (12)$$

where r^n and u^n indicate the natural interest rate and unemployment rate, respectively. Besides, π^t is the target inflation rate. The interest rate is adjusted adaptively according to annual inflation rate π and unemployment rate u, where α^π and α^u denote inflation and unemployment adaptation coefficients, respectively. We define the inflation and unemployment rate as follows,

$$\pi = \frac{\bar{P}_n - \bar{P}_{n-1}}{\bar{P}_{n-1}}, u = \frac{\sum_{m=1}^{12} \sum_{j=1}^{N} (1 - l_j)}{12N}, \qquad (13)$$

where $\bar{P}_n$ is the average goods price over the n-th year, and m denotes the m-th month.

When considering the dynamics of labor, consumption, and financial markets, the influence of these macroeconomic trends on agent decision-making is also seldom considered, raising the second challenge.

3 EconAgent

3.1 Perception Module

To harness the semantic awareness and real-world knowledge capabilities of LLM, we endow each agent with real-world profiles, including name, age, and job. Names are generated by the LLM and randomly assigned to each agent. The age distribution of agents follows the 2018 U.S. population distribution for ages 18-60 (Bureau, 2024). For wage and job assignments, we first adjust the parameters of the hourly wage's Pareto distribution to align the monthly wage distribution with 2018 U.S. economic data and tax brackets (Zheng et al., 2022). Furthermore, we prompt the LLM to generate ten job titles for each decile of the monthly wage distribution, mirroring the real-world scenario where different jobs have significant wage differences. Agents are initially assigned jobs based on this monthly wage and their jobs are dynamically adjusted throughout the simulation. If an agent chooses to work in the previous month, the job remains unchanged in the following month. If they are unemployed, an offer, based on the current monthly wage, is randomly presented to them. The generated age distribution, monthly wage distribution, and jobs are provided in the supplementary materials.

In addition, we characterize the entire economic environment in a manner closely mirroring the real world, enabling LLM to thoroughly grasp various economic factors. We integrate variations of key economic variables into the prompts and incorporated typical economic keywords to ensure the LLM could fully perceive dynamics in the economic landscape and employ relevant economic principles in its decision-making. For instance, if the agent chose not to work in the previous month, we would supplement the prompt with,

> *In the previous month, you became unemployed and had no income. Now, you are invited to work as a(an) [offer] with a monthly salary of [wage].*

In the prompt, *offer* and *wage* are dynamically adjusted along with the simulation. Such prompting enables the agent to recognize the risks associated with 'unemployment', thereby increasing its inclination to work in the subsequent month. More similar designs of prompting are also considered, such as 'shortage of goods' when the demand for

agents can not be met.

The perception module enables agents to act as heterogeneous households in the real economic environment, contributing to the emergence of macroeconomic phenomena.

3.2 Memory Module

Considering that decision-making within the economic environment is a sequential task, wherein past experiences and economic dynamics play pivotal roles in present decisions, the incorporation of a memory module can assist the agent in fully accounting for market dynamics and in acquiring valuable decision-making insights. Specifically, we dynamically maintain the memory pool with $2L + 1$ conversations, encompassing the economic environment and agent decisions from the previous L months. Besides, at the end of each quarter, we input dialogues of this quarter into the LLM, prompting it to 'reflect' on the economic phenomena and to respond to how these phenomena might influence its subsequent decisions. The prompts we design are as follows,

> *Given the previous quarter's economic environment, reflect on the labor, consumption, and financial markets, as well as their dynamics. What conclusions have you drawn?*

The following is an example of the reflection.

> *Based on the previous quarter's data, the **labor market experienced deflation**. . .The consumption market also saw a **decrease in prices for essential goods**. . .The financial market's interest rates **remained unchanged at 3.00%**. Overall, the quarter highlighted the need for **careful financial planning and adaptability in response to market fluctuations**.*

Obviously, after reflection, agents can fully perceive past market dynamics and adaptively adjust their strategies to maintain daily life and cope with future uncertainties.

The memory module allows the agent to comprehend dynamics in the market and glean reflections from past experiences, modeling the influence of macroeconomic trends.

3.3 Action Module

When prompting the LLM for decision-making, we explicitly incorporated considerations of living costs, future aspirations, and economic trends into the prompts. We prompt the LLM to respond with a value in the range [0, 1] to indicate the propensity of working and consumption.

The action module empowers the agent to automatically account for the influence of various economic factors when making decisions, such as income and savings, leveraging the semantic perception capabilities of LLM. It only requires the inclusion of relevant economic variables in the prompts, without the need for specially designed decision rules.

4 Experiments

In this section, we conduct experiments to study the ability of EconAgent, aiming to answer the following research questions (RQ).

- **RQ1**: How do the EconAgent behave in the simulation environment, compared with the traditional agents?

- **RQ2**: How do the perception and reflection modules in EconAgent affect the simulation results?

- **RQ3**: Does the decision-making mechanism of EconAgent possess interpretability, and can it address the challenges faced by traditional agents?

- **RQ4**: Can the simulation based on EconAgent reflect the impact of external intervention?

4.1 Experimental Setup

We investigate some phenomena of paramount interest in macroeconomics, including several macroeconomic indicators and two macroeconomic regularities. For comparative analysis, we select representative rule-based (Lengnick, 2013; Gatti et al., 2011) and learning-based (Zheng et al., 2022) agent models.

Definition of Macroeconomics Indicators Annual nominal GDP is defined as the sum of $S \times P$ over one year. As for real GDP, we set the first year in the simulation as the reference year and replace P with P_0, where P_0 is the goods price in the reference year. The definition of the annual (price) inflation rate and the unemployment rate is shown in Eq. 13. For wage inflation, the definition is similar to that of price inflation, where the average price is replaced with the average wage across all the agents.

Baselines We select **LEN** (Lengnick, 2013) and **CATS** (Gatti et al., 2011) as the baselines because 1) they partially reproduce the aforementioned macroeconomics phenomena within their own (more complex) simulation frameworks, and 2) carefully designed decision rules of working and consumption are representative, reflecting typical decision-making observed in real-life scenarios. Considering the importance of agents' heterogeneity in macroeconomic simulation, we also combine these two baselines as an additional baseline **Composite**, where each agent randomly adopts one rule of them. Besides, we select a learning-based method, AI-Economist (Zheng et al., 2022) **AI-Eco**, which builds on the assumption of rational decision-making and employs reinforcement learning (RL) (Arulkumaran et al., 2017) to maximize the agent's utility.

The details of decision rules in LEN and CATS and the training process for AI-Economist are provided in the supplementary materials.

Simulation Parameters We simulate $N = 100$ agents. The productivity is set as $A = 1$ for simplicity. The initial goods price is the average hourly wage across all the agents. For the labor and consumption dynamics, $\alpha_w = 0.05$ and $\alpha_P = 0.10$. For the financial market, $r^n = 0.01$, $\pi^t = 0.02$, $u^n = 0.04$, and $\alpha^\pi = \alpha^u = 0.5$. Note that our results and conclusions are not sensitive to these parameters.

Our simulation is implemented with Python. We use GPT-3.5-turbo-0613 provided by OpenAI API as the LLM[3]. In practice, we find $L = 1$ works well in the simulation. Other detailed simulation parameters, crucial for replicability and deeper understanding, are provided in the supplementary materials.

4.2 Macroeconomic Emergence (RQ1)

Our evaluation encapsulates a broad spectrum of macroeconomic indicators and regularities (Lengnick, 2013; Gatti et al., 2011; Dawid and Gatti, 2018; Axtell and Farmer, 2022). The performance of EconAgent was compared with that of representative rule-based baselines, as detailed in two referenced works and their combination (Lengnick, 2013; Gatti et al., 2011).

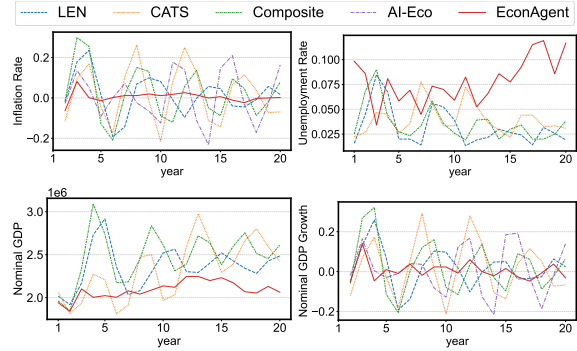**Macroeconomic Indicators** In Figure 2, we depict the fluctuations of the annual inflation rate,

Figure 2: Annual variations of macroeconomic indicators, where the simulation based on EconAgent shows more stable and numerically plausible indicators.

unemployment rate, nominal GDP, and growth rate of nominal GDP. Note that the unreasonable unemployment rate (around 46%) and nominal GDP for AI-Eco are not reported. Both rule-based and RL-driven baselines produce **anomalous indicators and large fluctuations**. In contrast, agent decision-making based on EconAgent has manifested more **stable and numerically plausible** macroeconomic phenomena across multiple facets, even without fine-grained calibration. Specifically, the inflation rate consistently fluctuated within a -5% to 5% range after the 3-th year, whereas the baselines exhibited significantly larger oscillations, at times even surpassing 20%. This indicates that the decision-making of EconAgent is coherent and more closely emulates real-world human choices, leading to easier attainment of equilibrium between supply and demand in the consumption market. The unemployment rate generally varied between 2% and 12%[4], which aligns well with real-world economic activities (Gatti et al., 2011). Both the nominal GDP and its growth rate also fluctuated within more reasonable numerical bounds like the inflation rate does. We also provide quarterly indicators in the appendix.

Macroeconomic Regularities As the two most commonly used regularities in macroeconomic simulations for validating the plausibility of simulation results, the Phillips Curve (Phelps, 1967) and Okun's Law (Okun, 1963) respectively describe the negative correlations between the annual unemployment rate and wage inflation, and the annual growth rate of unemployment and real GDP growth. As shown in Figure 3, only the decision-making of EconAgent has **correctly manifested**

---

[4]The increase in the unemployment rate after year 15 is actually a normal fluctuation. See Figure 10.
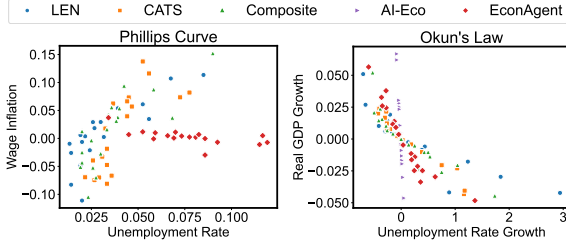
Figure 3: Only the simulation based on EconAgent emerges correct Phillips Curve and Okun's Law.

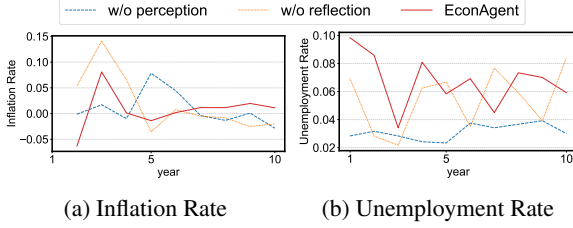(a) Inflation Rate   (b) Unemployment Rate

Figure 4: Ablation study on perception and memory modules.

phenomena in accordance with these two regularities (Pearson correlation coefficient is -0.619, $p < 0.01$ and -0.918, $p < 0.001$). Notably, the rule-based baseline method displayed an **incorrect positive relationship on the Phillips Curve**. We attribute this advantage to the LLM's accurate perception that consumption should be reduced when unemployed, a point which will be elaborated upon in the subsequent section. Note that the Phillips Curve for AI-Eco is not shown due to the very large unemployment rate.

4.3 Ablation Study (RQ2)

We separately remove the perception module and the reflection module, and the results of 10 years are as shown in Figure 4. We observe that when there is no perception capability, the inflation rate and unemployment rate fluctuations significantly decrease, appearing "too stable", especially for the unemployment rate. This suggests that the agents have low sensitivity to changes in their economic conditions and cannot make adaptive decision adjustments. When there is no reflection capability, the inflation rate exhibits anomalies close to 15% in the first three years, emphasizing the importance of long-term (a quarter in our experiments) economic environment perception.

4.4 Decision-Making Abilities (RQ3)

Decision Rationality   Given that an agent's decisions (propensity of working and consumption) may be influenced by multiple economic variables

Table 1: The number of agents for whom the effects of the variables on decisions are statistically significant.

	v_i	$\hat{c}_i$	$T(z_i)$	z^r	P	s_i	r
p_i^w	60	37	60	65	58	56	31
p_i^c	65	73	51	52	62	100	49

such as income and savings, we employ regression analysis to delve into the factors affecting these decisions. Specifically, the regression equation is as follows:

$$p_i^w, p_i^c \sim v_i + \hat{c}_i + T(z_i) + z^r + P + s_i + r, \quad (14)$$

where $\hat{c}_i, T(z_i), z^r$ denotes real consumption, the tax paid, and redistribution in the previous month. These independent variables are embedded in the LLM prompts to influence the agent's decisions. We have conducted individual regression analyses on all $N = 100$ agents' 240 decisions (spanning 20 years) after applying z-score normalization on all the variables and tabulated the significance of the impact of each economic variable on their decisions. Table 1 presents the number of agents for whom the effects of the variables are statistically significant, *i.e.*, $p \leq 0.05$. We observe that 1) the effects of taxation, redistribution, and expected monthly income on the work propensity of the majority of agents are significant, 2) in comparison to work propensity, the previous month's consumption, current savings, and bank interest rates significantly influence the consumption propensity of a greater number of agents, and 3) goods price has a significant impact on both work and consumption propensity for most agents. These phenomena are **consistent with economic common sense about how people make decisions** in daily life.

Agent Heterogeneity   We further investigate whether EconAgent autonomously exhibit heterogeneity in decision-making mechanisms. Figure 5 (a) demonstrates an incremental growth in consumption propensity with age, aligning broadly with empirical regularities observed in traditional macroeconomic research (Carroll, 1997).

Influence of Macroeconomic Trends   Through interactions with the model, we decipher the underlying reasons for the emergence of negative correlations in the Phillips curve. We first calculate the average consumption propensity across all the agents of two years with the highest and lowest unemployment rates. Figure 5 (b) shows the comparison
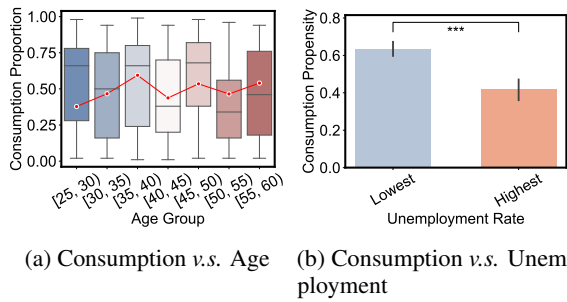
(a) Consumption *v.s.* Age   (b) Consumption *v.s.* Unemployment

Figure 5: The modeling of agent heterogeneity and the influence of macroeconomic trends.

Figure 6: COVID-19 brings the surge of simulated unemployment rates.

results, where *** denotes a significant disparity with $p < 0.001$. Obviously, high unemployment leads to low consumption propensity significantly. To delve deeper into the reasons why agents opt to reduce consumption in the labor market of high unemployment rates, we randomly select an agent and input its conversations with LLM from the year with the highest unemployment rate back into the LLM. We then prompt the LLM to summarize the economic dynamics for each quarter and provide rationales for the consumption decisions made. The following responses demonstrate that agents have a **perception of macroeconomic trends and are cautious about their spending when facing deflation in the labor market under a high unemployment rate**.

> *In the last quarter, I have adjusted my **willingness to work** and my **planned expenditures** on essential goods slightly **downwards**. This decision is primarily influenced by the **continued deflation in the labor market**, resulting in a decrease in my expected income. With a lower income, I need to be **cautious about my spending** and ensure that I have enough savings for **future expenses and unforeseen circumstances** . . .*

4.5 External Intervention (RQ4)

We further investigate the impact of external interventions on agent behavior and the resulting changes in macroeconomic simulation outcomes, a topic widely discussed in many economic ABM studies (Dawid and Gatti, 2018). Using the example of COVID-19, which has had a significant impact on the global economy, we incorporate it into our simulations and conduct comparative experiments.

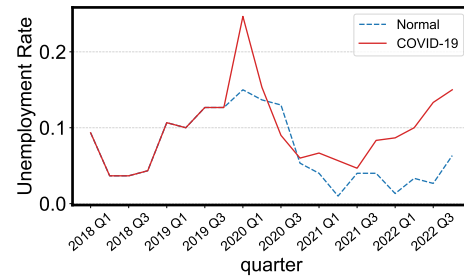Modeling the impact of COVID-19 can be conveniently implemented by incorporating them into the

prompt of a EconAgent. Specifically, we add the following statement to the prompt for simulations after March 2020:

> *In response to the large-scale outbreak of COVID-19 in the United States, the federal government has declared a national emergency since March 2020.*

Figure 6 shows the comparison of unemployment rates, where 'Normal' and 'COVID-19' denote the simulation results w/ and w/o the prompt above, respectively. This indicates that the simulation based on the EconAgent successfully replicated the surge in the unemployment rate in 2020 Q1 under the impact of COVID-19 (Organization for Economic Co-operation and Development, 1970). Although the numerical values do not match the real data exactly, this demonstrates that our proposed framework possesses the ability to qualitatively simulate human decision-making and macroeconomic phenomena in the real world. Furthermore, since we did not introduce government intervention measures, the unemployment rate after 2021 remains significantly higher than in the 'Normal' scenario, reflecting the lasting impact of COVID-19.

The following is an example of the agent's reflection during COVID-19, demonstrating its human-like decisions.

> *. . . However, the outbreak of COVID-19 and the subsequent national emergency declaration had a significant impact on the labor market. . . . resulting in widespread unemployment and uncertainty. This situation has likely affected my willingness to work, as job security and health concerns become more prominent . . .*

15530

5 Related Work

5.1 Simulation in Macroeconomics

ABM emerged as a more promising solution for macroeconomic research compared with empirical statistical models (Hendry and Richard, 1982; Phelps, 1967; Kydland and Prescott, 1982) and DSGE models (Christiano et al., 2005) in recent decades. Diverse agents interact with each other based on rules or computational models, avoiding the assumption of a predetermined economic equilibrium. These models allow for a wide range of nonlinear behaviors, enabling policymakers to simulate different policy scenarios and qualitatively assess their impacts on the economy. The agent modeling with predetermined rules (Tesfatsion and Judd, 2006; Brock and Hommes, 1998) or neural networks (Trott et al., 2021; Zheng et al., 2022; Mi et al., 2023) still suffers from the assumption of oversimplified agent behavior or the dependence on large-scale training data, thus having limitations in capturing the full complexity of economic systems.

In this work, we introduce EconAgent with reasoning and planning abilities for simulating macroeconomic activities.

5.2 LLM-empowered Agents

Recently, LLMs trained with large-scale corpus have shown human-level abilities and provide the basis for constructing agents for simulation (Wang et al., 2023). The LLM agents primarily have three advantages for simulation, including autonomous but adaptive reactions (Team, 2022; Yoheinakajima, 2023), human-like intelligence for planning (Xi et al., 2023), and interaction and communication with other agents or human (Park et al., 2023; Gilbert and Troitzsch, 2005). With these advanced abilities, LLM agents have been widely used in many areas, including social science (Park et al., 2022, 2023; Kovač et al., 2023; Gao et al., 2023b; Jinxin et al., 2023; Shao et al., 2024), natural science (Boiko et al., 2023; Bran et al., 2023), *etc*. Moreover, some works also consider LLM agents in economic research, which can be categorized into three levels (Gao et al., 2023a), including rationality or bias in individual behavior (Horton, 2023; Chen et al., 2023b), planning and cooperation in interactive behavior (Guo, 2023; Akata et al., 2023), and system-level market simulation (Zhao et al., 2023a; Anonymous, 2024; Chen et al., 2023a).

In short, existing works only consider individual one-step or few-step behavior for a few agents, without experimenting on multi-step behaviors within a multi-agent simulation environment, which is the focus of our work.

6 Conclusion

In this work, we ventured into the novel integration of LLMs with macroeconomic simulation, designing EconAgent with the abilities of perception, reflection, and decision-making based on the context of real-world economic environments. Classic macroeconomic phenomena are reproduced and more reasonable compared to traditional rule-based or learning-based agents. Through this endeavor, it has become evident that the capabilities of LLMs offer a promising avenue to simulate more realistic macroeconomics.

7 Limitations

Simulation Environment Although our constructed simulation environment encompasses a variety of market dynamics, it currently only simulates the labor and consumption behavior of households. More complex agents (*e.g.*, firms) and their behaviors (*e.g.*, pricing, hiring) have not yet been incorporated. Simulating these agents can facilitate a move towards a more realistic macroeconomic system, giving rise to more complex empirical regularities, such as the Beveridge curve and procyclicality/countercyclicality (Dawid and Gatti, 2018).

Optimization and Forecasting Current results are confined to replicating stylized facts (Dawid and Gatti, 2018). However, another primary goal of macroeconomic research is policy optimization, along with accurate forecasts of key macroeconomic indicators (Poledna et al., 2023). Therefore, EconAgent still requires more realistic behavioral responses to subtle changes in macroeconomic policies. Moreover, constructing a macroeconomic system that closely mirrors the real world necessitates the implementation of EconAgent at the urban scale level, which cannot be achieved without further breakthroughs in LLMs, including enhancements in inference speed (Shang et al., 2024) and reductions in computational demands.

References

Elif Akata, Lion Schulz, Julian Coda-Forno, Seong Joon Oh, Matthias Bethge, and Eric Schulz. 2023. Playing repeated games with large language models. *arXiv preprint arXiv:2305.16867*.

Anonymous. 2024. Rethinking the buyer's inspection paradox in information markets with language agents.

Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. 2017. Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, 34(6):26–38.

Robert L Axtell and J Doyne Farmer. 2022. Agent-based modeling in economics and finance: Past, present, and future. *Journal of Economic Literature*.

Daniil A Boiko, Robert MacKnight, and Gabe Gomes. 2023. Emergent autonomous scientific research capabilities of large language models. *arXiv preprint arXiv:2304.05332*.

Andres M Bran, Sam Cox, Andrew D White, and Philippe Schwaller. 2023. Chemcrow: Augmenting large-language models with chemistry tools. *arXiv preprint arXiv:2304.05376*.

William A Brock and Cars H Hommes. 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic dynamics and Control*, 22(8-9):1235–1274.

US Census Bureau. 2024. Access demographic, economic and population data from the u.s. census bureau.

Christopher D Carroll. 1997. Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly journal of economics*, 112(1):1–55.

Jiangjie Chen, Siyu Yuan, Rong Ye, Bodhisattwa Prasad Majumder, and Kyle Richardson. 2023a. Put your money where your mouth is: Evaluating strategic planning and execution of llm agents in an auction arena. *arXiv preprint arXiv:2310.05746*.

Yiting Chen, Tracy Xiao Liu, You Shan, and Songfa Zhong. 2023b. The emergence of economic rationality of gpt. *arXiv preprint arXiv:2305.12763*.

Lawrence J Christiano, Martin Eichenbaum, and Charles L Evans. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45.

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

Herbert Dawid, Simon Gemkow, Philipp Harting, Sander Van der Hoog, and Michael Neugart. 2012. The eurace@ unibi model: An agent-based macroeconomic model for economic policy analysis.

Christophe Deissenberg, Sander Van Der Hoog, and Herbert Dawid. 2008. Eurace: A massively parallel agent-based model of the european economy. *Applied mathematics and computation*, 204(2):541–552.

Jingtao Ding, Chang Liu, Yu Zheng, Yunke Zhang, Zihan Yu, Ruikun Li, Hongyi Chen, Jinghua Piao, Huandong Wang, Jiazhen Liu, et al. 2024. Artificial intelligence for complex network: Potential, methodology and application. *arXiv preprint arXiv:2402.16887*.

J Doyne Farmer and Duncan Foley. 2009. The economy needs agent-based modelling. *Nature*, 460(7256):685–686.

Chen Gao, Xiaochong Lan, Nian Li, Yuan Yuan, Jingtao Ding, Zhilun Zhou, Fengli Xu, and Yong Li. 2023a. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *arXiv preprint arXiv:2312.11970*.

Chen Gao, Xiaochong Lan, Zhihong Lu, Jinzhu Mao, Jinghua Piao, Huandong Wang, Depeng Jin, and Yong Li. 2023b. S^3: Social-network simulation system with large language model-empowered agents. *arXiv preprint arXiv:2307.14984*.

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

Nigel Gilbert and Klaus Troitzsch. 2005. *Simulation for the social scientist*. McGraw-Hill Education (UK).

Fulin Guo. 2023. Gpt agents in game theory experiments. *arXiv preprint arXiv:2305.05516*.

David F Hendry and Jean-Francois Richard. 1982. On the formulation of empirical models in dynamic econometrics. *Journal of Econometrics*, 20(1):3–33.

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

Shi Jinxin, Zhao Jiabao, Wang Yilei, Wu Xingjiao, Li Jiawen, and He Liang. 2023. Cgmi: Configurable general multi-agent interaction framework. *arXiv preprint arXiv:2308.12503*.

Dale W Jorgenson. 2001. Information technology and the us economy. *American Economic Review*, 91(1):1–32.

Grgur Kovač, Rémy Portelas, Peter Ford Dominey, and Pierre-Yves Oudeyer. 2023. The socialai school: Insights from developmental psychology towards artificial socio-cultural agents. *arXiv preprint arXiv:2307.07871*.

Finn E Kydland and Edward C Prescott. 1982. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.

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

Qirui Mi, Siyu Xia, Yan Song, Haifeng Zhang, Shenghao Zhu, and Jun Wang. 2023. Taxai: A dynamic economic simulator and benchmark for multi-agent reinforcement learning. *arXiv preprint arXiv:2309.16307*.

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

Organization for Economic Co-operation and Development. 1970. Infra-Annual Labor Statistics: Unemployment Rate Total: From 15 to 64 Years for United States.

Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. *arXiv preprint arXiv:2304.03442*.

Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2022. Social simulacra: Creating populated prototypes for social computing systems. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, pages 1–18.

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

Sebastian Poledna, Michael Gregor Miess, Cars Hommes, and Katrin Rabitsch. 2023. Economic forecasting with an agent-based model. *European Economic Review*, 151:104306.

Frank Schorfheide and Dongho Song. 2015. Real-time forecasting with a mixed-frequency var. *Journal of Business & Economic Statistics*, 33(3):366–380.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Yu Shang, Yu Li, Fengli Xu, and Yong Li. 2024. Defint: A default-interventionist framework for efficient reasoning with hybrid large language models. *arXiv preprint arXiv:2402.02563*.

Chenyang Shao, Fengli Xu, Bingbing Fan, Jingtao Ding, Yuan Yuan, Meng Wang, and Yong Li. 2024. Beyond imitation: Generating human mobility from context-aware reasoning with large language models. *arXiv preprint arXiv:2402.09836*.

AutoGPT Team. 2022. Autogpt: the heart of the open-source agent ecosystem. https://github.com/Significant-Gravitas/AutoGPT. (Accessed on 01/10/2023).

Leigh Tesfatsion and Kenneth L Judd. 2006. *Handbook of computational economics: agent-based computational economics*. Elsevier.

Alexander Trott, Sunil Srinivasa, Douwe van der Wal, Sebastien Haneuse, and Stephan Zheng. 2021. Building a foundation for data-driven, interpretable, and robust policy design using the ai economist. *arXiv preprint arXiv:2108.02904*.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. 2023. A survey on large language model based autonomous agents. *arXiv preprint arXiv:2308.11432*.

Paul Windrum, Giorgio Fagiolo, and Alessio Moneta. 2007. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, 10(2):8.

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

Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, et al. 2023. The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864*.

Yoheinakajima. 2023. Babyagi. https://github.com/yoheinakajima/babyagi. (Accessed on 01/10/2023).

Qinlin Zhao, Jindong Wang, Yixuan Zhang, Yiqiao Jin, Kaijie Zhu, Hao Chen, and Xing Xie. 2023a. Competeai: Understanding the competition behaviors in large language model-based agents. *arXiv preprint arXiv:2310.17512*.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023b. A survey of large language models. *arXiv preprint arXiv:2303.18223*.

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

A EconAgent

Agent Profiles The initialization of agents' profiles is shown in Figure 7, including age distribution (left) and monthly wage distribution (right), as well as the tax brackets and rates of U.S. federal government in 2018, represented by the gray dotted line. As for the generated jobs aligned with the monthly wage, we show some examples as follows,

- [0, 2454): Dog Walker, House Cleaner, Newspaper Delivery . . .

- [2454, 4838): Barista, Cashier, Fast Food Worker . . .

- [35469, 52370): Psychiatrist, Pediatrician, Anesthesiologist . . .

Economic Prompts We provide a full prompt to illustrate our consideration of economic factors, as well as other details not mentioned in the main text.

You're Adam Mills, a 40-year-old individual living in New York City, New York. As with all Americans, a portion of your monthly income is taxed by the federal government. This taxation system is tiered, income is taxed cumulatively within defined brackets, combined with a redistributive policy: after collection, the government evenly redistributes the tax revenue back to all citizens, irrespective of their earnings. Now it's 2001.02. In the previous month, you worked as a(an) Professional Athlete. If you continue working this month, your expected income will be $84144.58, which is decreased compared to the last month due to the deflation of the labor market. Besides, your consumption was $49825.69. Your tax deduction amounted to $28216.98. However, as part of the government's redistribution program, you received a credit of $6351.29. In this month, the government sets the brackets: [0.00, 808.33, 3289.58, 7016.67, 13393.75, 17008.33, 42525.00] and their corresponding rates: [0.10, 0.12, 0.22, 0.24, 0.32, 0.35, 0.37]. Income earned within each bracket is taxed only at that bracket's rate. Meanwhile, deflation has led to a price decrease in the consumption market, with the average price of essential goods now at $135.82. Your current savings account balance is $12456.42. Interest rates, as set by your bank, stand at 3.00%. With all these factors in play, and considering aspects like your living costs, any future aspirations, and the broader economic trends, how is your willingness to work this month? Furthermore, how would you plan your expenditures on essential goods, keeping in mind goods price? Please share your decisions in a JSON format. The format should have two keys: 'work' (a value

---

between 0 and 1 with intervals of 0.02, indicating the willingness or propensity to work) and 'consumption' (a value between 0 and 1 with intervals of 0.02, indicating the proportion of all your savings and income you intend to spend on essential goods).
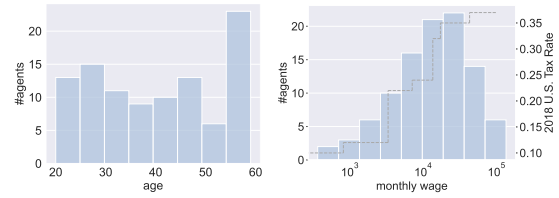

Figure 7: Age and monthly wage distribution for agent profiles.

B   Experimental Setup

Baselines

- **Consumption**. In LEN, the consumption decision is *memory-based*, which means that consumption is influenced not only by current income but also by past accumulated savings. Besides, the goods price is another important factor. Conversely, in CATS, it is *non-memory-based* consumption decisions suggesting that consumption is solely related to the current income. The agent aims to keep a desired ratio between savings and income, and consumption is only a proportion of the income. For more human-like decisions, we also introduce the influence of interest rate in the decision rule.

For LEN, the calculation of consumption propensity is as follows,

$$p_i^c = \left(\frac{P}{s_i + z_i} \right)^\beta, \beta \in [0, 1]. \qquad (15)$$

For CATS, the calculation is as follows, *i.e.*,

$$\frac{\hat{s}_i}{z_i} = \frac{(1 + r)(s_i + (1 - c)z_i)}{z_i} = h, p_i^c = \frac{cz_i}{s_i + z_i}, \qquad (16)$$

where $\hat{s}_i$ denotes the expected savings after consumption in the next month, h is a constant, and c indicates the consumption proportion of the current income. Refer to CATS for the calculation of c. Note that we introduce the influence of the interest rate r to endow the agent with the perception of fiscal policy.

- **Work**. The rule of working in LEN and CATS can not be directly used in our simulation framework because we don't simulate firms. Therefore, we follow the intuitions of their designs and define a formula implying that a higher expected income, lower savings, or a lower interest rate lead to a greater propensity to work.

The work propensity is calculated as

$$p_i^w = \left(\frac{v_i}{s_i(1+r)}\right)^\gamma, \gamma \in [0,1]. \quad (17)$$

For AI-Economist, the utility is a satisfaction function positively correlated with savings and consumption but negatively correlated with labor. Maximizing utility implies that the agent always desires more savings and consumption but prefers less labor. The policy network for the agent's work and consumption decisions is a multi-layer perceptron (MLP), where the input includes various environment information, such as monthly wage, interest rate, goods price, tax rates, *etc*. We modify the utility function to incorporate consumption and goods price, defined as

$$\frac{(s_i/P)^{1-\lambda_s}-1}{1-\lambda_s} \cdot \frac{(\hat{c}_i/P)^{1-\lambda_c}-1}{1-\lambda_c} - \lambda_l l_i, \quad (18)$$

where $\lambda_{s,c,l}$ balance the importance of savings, consumption, and labor contributing to agent satisfaction. Besides, it's discouraged to not work or have no consumption at all, which leads to negative utility. We also introduce the goods price to make the AI agent perceive the dynamics of the consumption market.

Simulation Parameters For LEN, CATS, and Composite, we conduct a careful grid search for proper hyperparameters β, γ, h in decision rules, with the search spaces of $[0.05, 0.1, 0.3, 0.5], [0.05, 0.1, 0.3, 0.5],$ $[0.5, 1, 3, 5]$, respectively. The reported results in the main text are obtained with $\beta = 0.1, \gamma = 0.1, h = 1$, which show the most reasonable macroeconomic indicators.

LLM Costs Each simulation based on EconAgent incurs a cost of approximately \$30 and takes about 2 hours to complete.

C Additional Results

Quaterly Indicators Figure 9 presents quarterly macroeconomic indicators, where the conclusion
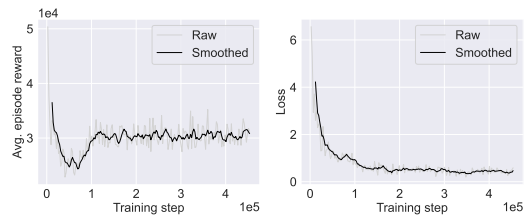

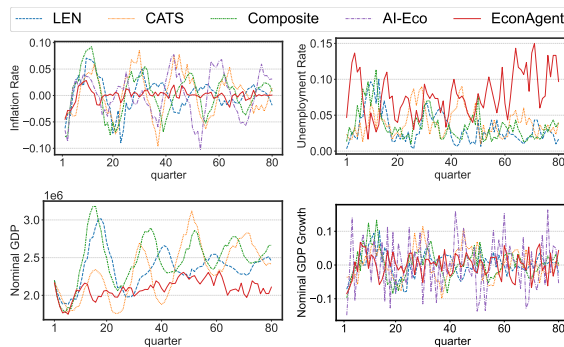Figure 8: The training process of AI-Economist.

Figure 9: Quarterly variations of macroeconomic indicators.

is similar to that of annual ones. For AI-Economist, we follow (Zheng et al., 2022) to adopt PPO algorithm (Schulman et al., 2017) to train the policy network, where the actor and critic networks have the hidden dimensions of $[128, 128]$ and $[128, 64, 32]$, respectively. The observation (input) dimension is 173 and the action (output) dimension is 53, including 2 work actions and 51 consumption actions (0-1 with an interval of 0.02). The training process is shown in Figure 8, including the loss and average episode reward, where one episode is a complete simulation of 20 years.

Fluctuated Unemployment Rate The unemployment rate after 5 additional years of simulation is shown in Figure 10, which returns to a lower level after the 20th year.

Sensitivity and Robustness We increase the number to 300 and run the simulation again. Figure 11 shows consistently stable and plausible inflation rates similar to those of 100 agents, which holds for other indicators as well. Therefore, the simulation results are insensitive to the number of agents.

We conduct five simulations for each agent model and present the inflation rate in Figure 12. It is evident that the simulations based on EconAgent are robust and consistently yield more stable and plausible results, which is also true for other indicators. Importantly, there are no significant differences in terms of stability and numerical scale
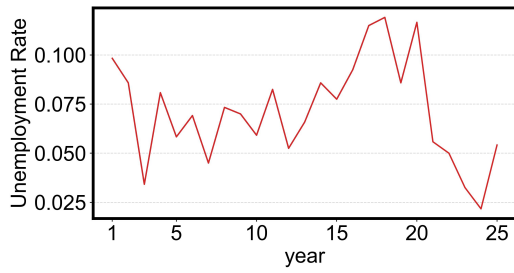
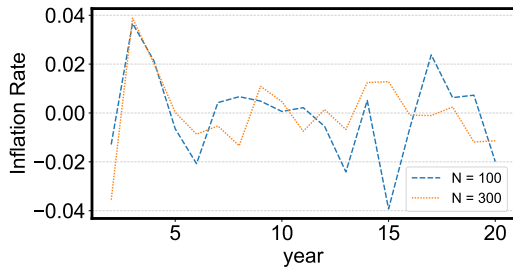Figure 10: Unemployment rate for 25 years.

Figure 11: Inflation rate under different number of agents.
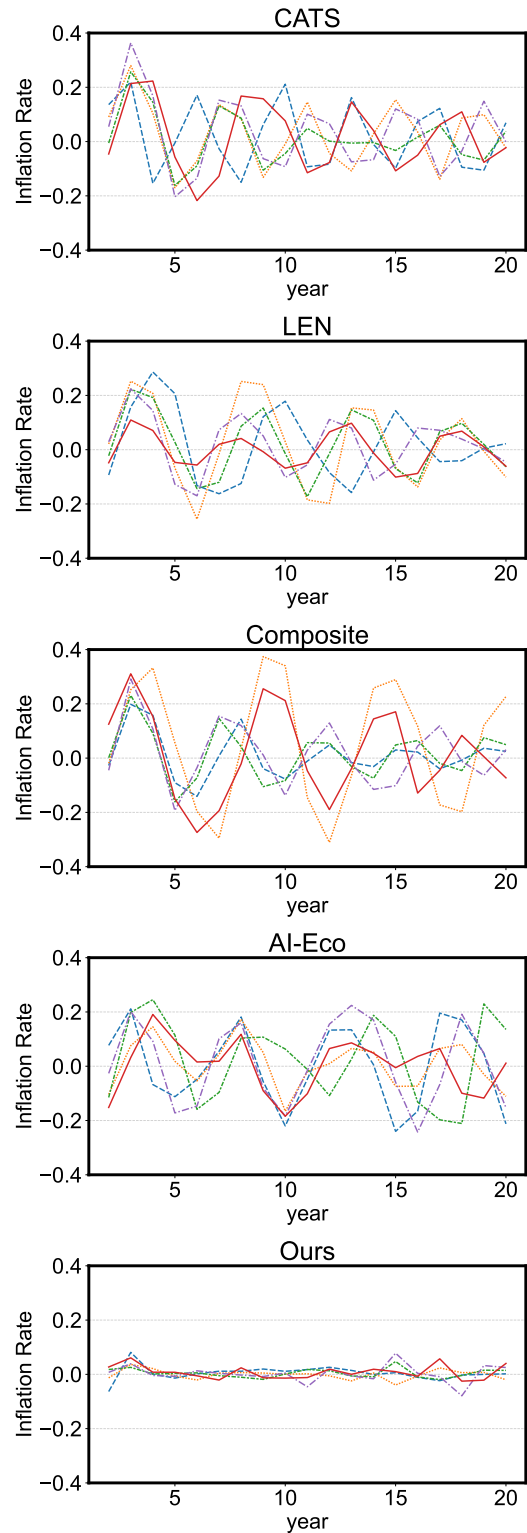
among the simulations.

Figure 12: Simulation variance across five experiments for different agent models.