A Setting

A.1 Agent Decisions

We follow the previous work Li et al. (2024); Arulkumaran et al. (2017), and have the following settings. Labor supply and consumption are necessary agent (household) decisions in economic 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.

Firm j makes strategic decisions at each simulation step, which represents a month, regarding its labor force and investment strategies:

- whether to hire new employees $h_j \sim Bernoulli(q_j^h)$, where q_j^h is the hiring propensity. If the firm decides to hire ($h_j = 1$), it does so based on a calculated demand for labor and available budget. The firm's decision to hire is influenced by market conditions and the potential for increased productivity.
- whether to lay off existing staff $f_j \sim Bernoulli(q_j^f)$, where q_j^f is the firing propensity. If the firm decides to let go of staff ($f_j = 1$), it is due to economic pressures such as reduced demand or the need to cut costs. The firm must weigh the immediate financial savings against potential long-term impacts on morale and productivity.
- the amount to allocate towards research and development $R\&D_j$, which is a function of the firm's innovation propensity q_j^r and its available capital. The firm decides to increase ($R\&D_j =$ increase), maintain ($R\&D_j =$ maintain), or decrease ($R\&D_j =$ decrease) its R&D budget with the aim of staying competitive and driving future growth.
- the level of investment in capital goods I_j, determined by the firm's capital investment propensity q_j^I and its financial health. The firm may choose to invest more ($I_j =$ increase), continue at the current level ($I_j =$ maintain), or reduce investment ($I_j =$ decrease) based on projections of future returns and the state of the economy.

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

A.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[2]. The progressive tax for income z_i is calculated as follows,

$$T(z_i) = \sum_{k=1}^{B} \tau_k \left((b_{k+1} - b_k) \mathbf{1} \left[z_i > b_{k+1} \right] + (z_i - b_k) \mathbf{1} \left[b_k < z_i \le b_{k+1} \right] \right), \tag{13}$$

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.

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

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

$$\hat{z}_i = z_i - T(z_i) + z^r = z_i - T(z_i) + \frac{1}{N}\sum_{j=1}^{N} T(z_j), \tag{14}$$

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

$$s_i \leftarrow s_i + \hat{z}_i \tag{15}$$

A.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. The inventory of goods G is then updated as,

$$G \leftarrow G + S = G + \sum_{j=1}^{N} l_j \times 168 \times A, \tag{16}$$

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,

$$D = \sum_{j=1}^{N} d_j = \sum_{j=1}^{N} \frac{c_j}{P} = \sum_{j=1}^{N} \frac{p_j^c s_j}{P}, \tag{17}$$

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,

$$\bar{\varphi} = \frac{D - G}{\max(D, G)}, \tag{18}$$

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 & Gatti (2018); Wolf et al. (2013). The hourly wage is adjusted as follows,

$$w_i \leftarrow w_i(1 + \varphi_i), \varphi_i \sim sign(\bar{\varphi})U(0, \alpha_w|\bar{\varphi}|), \tag{19}$$

and the price is adjusted as follows,

$$P \leftarrow P(1 + \varphi_P), \varphi_P \sim sign(\bar{\varphi})U(0, \alpha_P|\bar{\varphi}|), \tag{20}$$

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,

$$\hat{d}_j = \min(d_j, G), \hat{c}_j = \hat{d}_j \times P \tag{21}$$

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

$$G \leftarrow G - \hat{d}_j. \tag{22}$$

The process continues until every agent has consumed goods once.

A.4 Financial Market

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

$$s_i \leftarrow s_i \times (1 + r). \tag{23}$$

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 & Gatti (2018),

$$r = \max(r^n + \pi^t + \alpha^\pi(\pi - \pi^t) + \alpha^u(u^n - u), 0), \tag{24}$$

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,

$$\pi = \frac{\bar{P}_n - \bar{P}_{n-1}}{\bar{P}_{n-1}}, u = \frac{\sum_{m=1}^{12}\sum_{j=1}^{N}(1 - l_j)}{12N}, \tag{25}$$

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.

B ECONAI

Agent Profiles The initialization of agents' profiles is shown in Figure 6, 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.

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.

> You're Emma Johnson, a 35-year-old graphic designer living in Seattle, Washington. Like all Americans, you contribute to the federal government's tax system through a tiered taxation model, which is progressive and includes a redistributive policy. After collecting taxes, the government evenly redistributes the revenue to all citizens, regardless of their income levels. It's now February 2001. Last month, you worked as a Graphic Designer. If you continue in the same profession this month, your anticipated income is $75,000, which is slightly lower than the previous month due to the ongoing deflation in the job market. Additionally, your spending last month was $30,000 on various creative projects and living expenses. Your tax deduction was $15,000, but you received a government credit of $4,500 as part of the redistribution program. This month, the tax brackets are set as follows: [0.00, 1200.00, 5000.00, 8800.00, 12000.00, 20000.00, 50000.00] with corresponding rates: [0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40]. Each income segment is taxed at the rate of its bracket. Deflation has also caused a decrease in consumer goods prices, with the average cost of essential items now at $100. Your current savings stand at $20,000, and your bank offers an interest rate of 2.5%. Considering your living expenses, future goals, and the overall economic climate, your willingness to work this month is 0.78, indicating a strong propensity to work. Regarding your spending on essential goods, you plan to spend 0.54 of your total savings and income, which reflects a moderate level of consumption considering the current market prices.

C 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

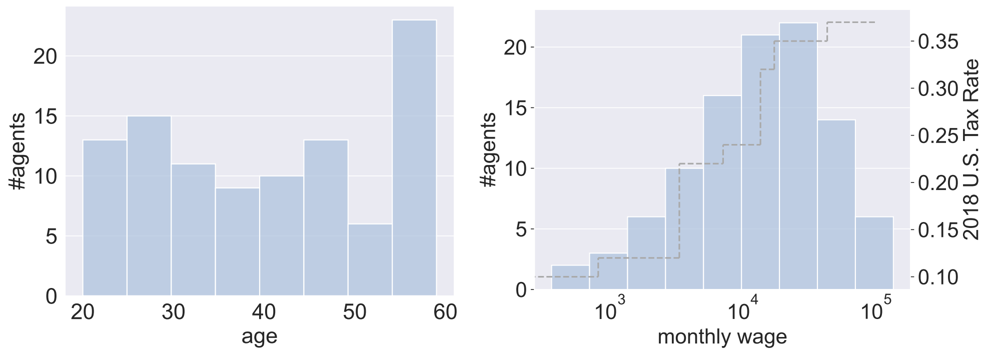

Figure 6: Age and monthly wage distribution for agent profiles.

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,

$$p_i^c = \left(\frac{P}{s_i + z_i}\right)^{\beta}, \beta \in [0, 1]. \tag{26}$$

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

$$\frac{\hat{s}_i}{z_i} = \frac{(1+r)(s_i + (1-c)z_i)}{z_i} = h, p_i^c = \frac{cz_i}{s_i + z_i}, \tag{27}$$

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 leads to a greater propensity to work.

  The work propensity is calculated as

$$p_i^w = \left(\frac{v_i}{s_i(1+r)}\right)^{\gamma}, \gamma \in [0, 1]. \tag{28}$$

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

$$\frac{(s_i/P)^{1-\lambda_s} - 1}{1 - \lambda_s} \cdot \frac{(\hat{c}_i/P)^{1-\lambda_c} - 1}{1 - \lambda_c} - \lambda_l l_i, \tag{29}$$

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 EconAI incurs a cost of approximately $25 and takes about 1.8 hours to complete.

D Additional Results

Sensitivity and Robustness We increase the number to 300 and run the simulation again. Figure 7 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.

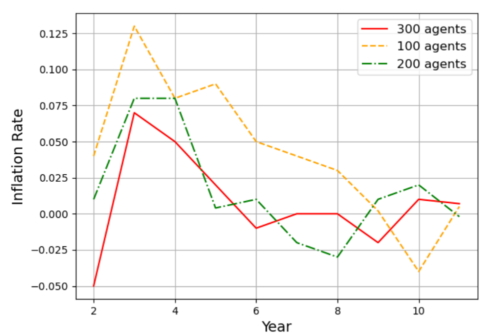

Figure 7: Inflation rate under different number of agents.

We conduct five simulations for each agent model and present the inflation rate in Figure 8. It is evident that the simulations based on EconAI 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 among the simulations.

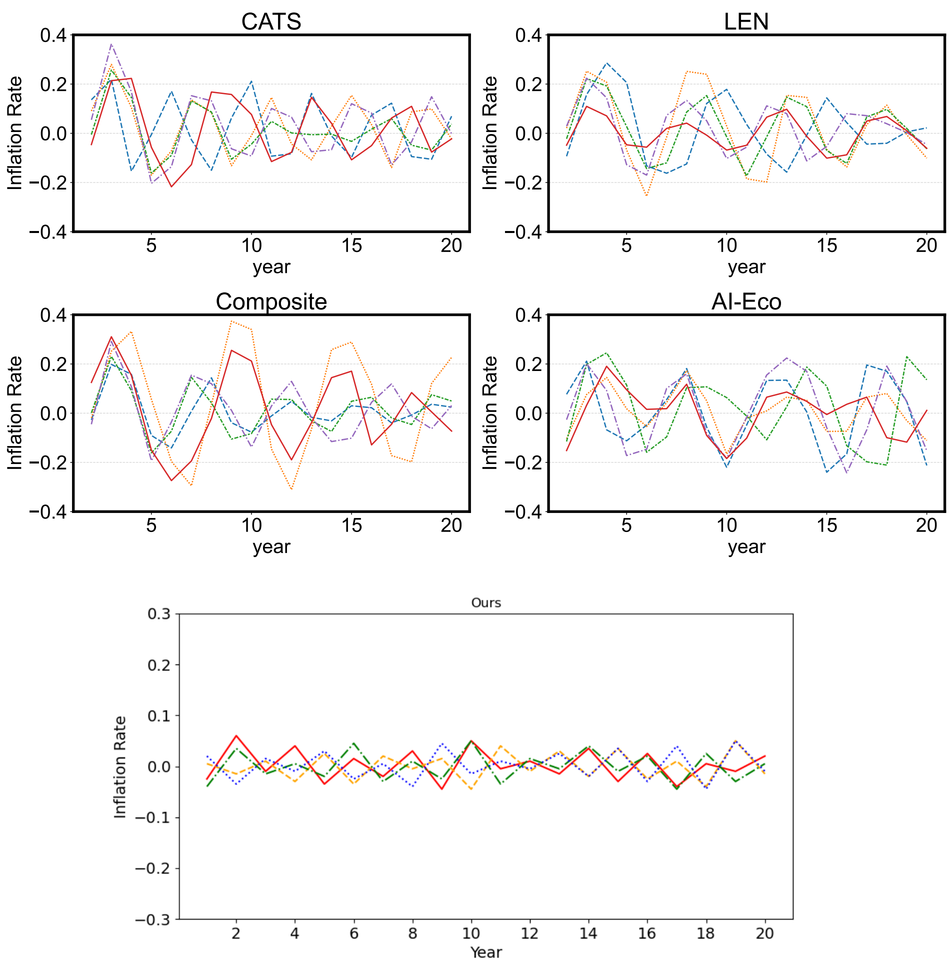

Figure 8: Simulation variance across five experiments for different agent models.