How does AI affect society in a human-AI mixed environment?

Peiyuan Zhang *

[0.5ex]

Date: 2024.4.8

Abstract

Nowadays, with the advance of web and internet technology, Artificial Intelligence (AI) has significantly improved productivity in production and become an integral part of labor market. Motivated by this, our study empirically examines the impact of AI on the economy. We construct a multi-agent economic environment that incorporates both human and AI agents. A key distinction between AI and human agents lies in the seamless sharing of skills among AI agents, while humans continue to develop their own skills. Furthermore, we thoroughly discussed the implications of AI ownership on social welfare. We employ the Proximal Policy Optimization (PPO) algorithm to train all agents in the environment. Through our experiments, we uncover several important findings: Firstly, the introduction of AI results in a substantial increase in overall productivity. The presence of AI leads to a crowding-out effect, replacing humans in the production Secondly, AI's value creation heavily relies on the attention it garners, and an open AI market/technology notably boosts its progress and worth. Companies with diverse products and AI tech display greater influence in this trend. Furthermore, we highlight the significance that implementing a proper tax policy enables the effective utilization of the productivity advantages brought by AI while mitigating its negative impact. Overall, our work provides new insights on the complex interactions between humans and AI in economic activities.

Keywords: AI ; Optimal Taxation; Welfare Analysis; Multi-Agent Simulation

1 Introduction

The development and widespread adoption of web and internet technology revitalized numerous fields, infusing them with new energy. As we enter the post-internet era, the pervasive

*School

influence of web and internet technology remains unabated, continuously shaping and impacting other emerging domains. Nowadays, as a natural extension of the internet era, Artificial Intelligence (AI) has seamlessly integrated into society, revolutionizing productivity processes and significantly enhancing productivity. AI technologies have demonstrated their prowess in various domains, such as online displaying advertisement (Gharibshah et al., 2020), manufacturing (Arinez et al., 2020), recommendation systems (Zhang et al., 2019), healthcare (Secinaro et al., 2021), and automated driving (Grigorescu et al., 2020). In addition to its positive impact, AI has also brought negative effects on the labor market, saying substituting labor. The advent of AI leads to higher unemployment and job polarization (Autor and Dorn, 2013). Frey and Osborne (2017) and Bowles (2014) estimate that up to 47% and 54% of jobs in the US and EU could be automated within 10 to 20 years. According to Goyal and Aneja (2020), AI will displace about 20% of the global workforce, and there is a growing concern about job loss. Moreover, not only labor market is affected, but also the resulting productivity and inequality balance is shifted. AI substitutes both skilled and unskilled labor, leading to heterogeneous income change among human agents. This brings higher output and welfare but at the cost of increased income and wealth inequality (Korinek and Stiglitz, 2018).

Therefore, it is crucial to examine the impact of AI technology on the economy. While a limited body of literature has started exploring this subject, there remains a need for further investigation. For example, Acemoglu and Restrepo (2017) develop a model to understand how AI replaces labor in the production process. Aghion et al. (2018) identify constraints on economic growth arising from Baumol's cost-disease effects in the context of AI. Sachs (2018) acknowledge the potential income redistribution from workers to owners of "business capital" as AI advances. Bessen (2018) emphasizes the importance of income elasticities of price and demand when assessing the implications of AI. Gries and Naudé (2018) extend the existing literature by incorporating AI-facilitated automation into standard product variety models and accounting for demand-side constraints on outcomes. In our study, we contribute to this literature by incorporating learning and skill-sharing AI in a multi-agent reinforcement learning (MARL) environment, further exploring the dynamics of AI's impact on the economy.

In addition to exploring the impact of AI on individuals, as global competition among major powers intensifies in the technological sphere, technology blockades and technological protectionism are gradually becoming the new normal in international competition. In this environment, major technology companies have to face various challenges posed by technology blockades in order to maintain their competitive edge. The issue of technology "ownership" is particularly important in this context. It not only concerns the core competitiveness of enterprises but also involves strategic planning for national security and economic development. However, unfortunately, this critical issue has not received sufficient attention and research in traditional economic fields. This is mainly because existing economic theories and models struggle to deal with complex scientific issues involving ownership, especially when different advanced technologies are in the hands of different companies, increasing complexity exponentially.

Currently, academic research primarily focuses on the specific applications of AI-generated technologies, such as the ownership of AI-generated images (Aziz, 2023). Although these studies have their value, they do not address the core issue of technology ownership. We know that it is crucial for companies and countries aspiring to make breakthroughs in technology to clarify and grasp technology ownership. This not only concerns economic interests but also affects the direction of technological development and the future social landscape.

In this paper, we study how AI affects human workers in a Human-AI hybrid environment. The key difference between AI and human agents is that we assume AI agents can share skills frictionlessly. The rationale behind such an assumption is that we consider the case where all AI agents share their parameters with each other. This assumption is supported by the rapid progress of federated learning (Li et al., 2020, Zhang et al., 2021) and web and internet technology (Korinek and Stiglitz, 2018). In federated learning, different AI agents can collect data and update their local parameters in a decentralized manner, and then aggregate all the information from the AI agents to update a centralized set of model parameters. Additionally, advancements in web and internet technology have made information transmission more efficient (Korinek and Stiglitz, 2018), enabling AI agents to share skills effectively. These techniques enable AI agents to efficiently share skills and continuously stay at the forefront of technological advancements. In contrast, humans gradually improve their skills through practice, experience, and minor innovations, resulting in a significant cost of learning.

We expanded upon the work of Zheng et al. (2021, 2022), constructing a multi-agent environment that introduces two types of agents (AI and human agents) to study how AI affects the economy. The specific settings of the environment and agents enable us to study how AI affects the labor market in detail and the resulting aggregate impacts. Specifically, our environment has three objects, World, Agents, and the Planner. The World consists of resource distributions and information of active agents. There are two types of Agents (AI and human agents), whose main difference is whether they can share skills frictionlessly. The objective of agents is to choose actions to maximize their lifetime utility. The Planner is to redistribute the income of all agents through taxation, with a trade-off between productivity and equality. The ownership setting is as follows: If a human owns the ownership of an AI, then all products/incomes produced by this AI agent will be freely obtained by this human agent. To solve this model, we apply the PPO algorithm Schulman et al. (2017) to solve the above multi-agent environment and achieve equilibrium.

Firstly, We find that the application of AI agents has various impacts on the economy from the micro to macro level, which can be explicitly described as follows:

1. The presence of AI agents, which possess the capability to share knowledge with each other, enhance overall productivity in the economy.
2. The application of AI leads to a crowding-out effect, particularly impacting individuals with middle incomes, while having a minimal effect on those with high or low incomes.
3. Optimal taxation, implemented through income redistribution, can foster productivity and

equality within the Human-AI hybrid society.

Then, the ownership research found that:

1. If all products/incomes produced by AI agents are simply attributed directly to a human, AI agent will choose to give up all labor, as they will not have any utility.
2. In order to stimulate the labor enthusiasm of AI agents, we envision that social planner can allow AI agents to retain a portion of their income as investment. This initiative greatly increases the enthusiasm of AI agents to engage in work.
3. When exploring the setting of skill sharing, we found that the more open the technology community is, or the more technology the AI owner have, the faster his tech will develop.

2 Model

This section provides an overview of the environment setting. We construct a simulated multi-agent economic environment that mirrors a real-world environment, which is akin to a Gather-and-Build game, yet its core mechanism centers on interactions among agents and economic behaviors. By continuously optimizing the behavioral strategies of various agents within the environment, we aspire to ultimately uncover the motivations and patterns behind agents actions. The model primarily consists of three aspects: the World, Agents, and Planner. In the following subsections, we will delve into these three aspects, exploring the constituent elements of each part and their economic implications.

2.1 World

The World is composed of a two-dimensional grid (each $(i, j) \in N \times N$) that encompasses a variety of resources. The world serves as the core of agent activities and resource distribution, forming the fundamental platform for various economic behaviors.

The world is composed of several layers, each containing unique resources and settings. At the bottom lies a detailed world map featuring grassland and water areas. Agents can move freely on the grassland, but they cannot cross the water bodies. In addition, there is a resource layer primarily consisting of two renewable resources: wood and stone. These resources have a probability of $\rho_{i,j}$ to regenerate at each grid point. Following this is the construction layer, agents have the option to consume wood and stone to build houses. Once a house is built, it permanently stands at that location, and different houses cannot occupy the same grid. The topmost layer is dedicated to recording agent activities, precisely marking their positions in the world and clearly defining their visible range and activity areas. The specific attributes of each resource are shown in table 1.

The initial World generates a random distribution of resources, which can be visually represented in fig. 1. All agents actively engage with the world by taking actions and interacting

Table 1: World resource attributes Table

	Collectability	Renewability	Private	Passability	Tradeable
Wood	Yes	Yes	No	Yes	Yes
Stone	Yes	Yes	No	Yes	Yes
House	No	No	Yes	Yes	No
Water	No	No	No	No	No

with it. Certain actions have the potential to alter the distribution of resources or modify agent information within the world.

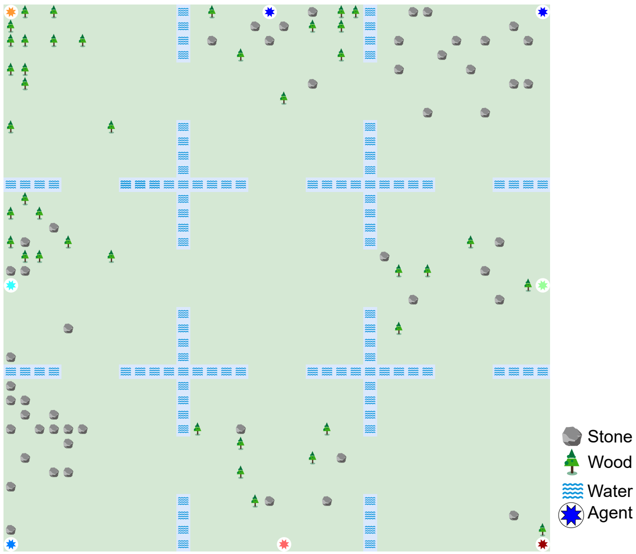

Figure 1: An example of a world map includes resources such as wood, stone, and the basic outline of the map, along with the initial positions of the agents.

2.2 Agent

In our benchmark, we consider two types of agents: AI agents and human agents. Within each agent type, there is a degree of heterogeneity. Agents have distinct attributes, namely their inventory $x_{i,t}$ and skills $s_{i,t}$. The inventory attribute reflects the number of collectible resources an agent possesses, including wood ($x^w_{i,t}$), stone ($x^s_{i,t}$), and coins ($x^c_{i,t}$). Skills, on the other hand, represent an agent's proficiency in various tasks, particularly labor actions.

2.2.1 Action

Agents can choose four actions: Move, Gather, Build, and Trade.

- **Move.** Agents have the ability to move within the World in four cardinal directions: up, down, left, and right. However, there are strict constraints that govern their movement. Agents are not allowed to cross the boundaries of the World or traverse impassable grids – water. Each movement from one grid to another incurs a labor cost (l^M), and once the Move action is completed, the agent's position within the World is adjusted accordingly.

- **Gather.** When an agent reaches a grid containing collectible resources, they can choose to perform a Gather action. The act of Gathering requires the expenditure of labor (l^G). After the completion of the Gather action, the resources within the grid are depleted, and the agent's inventory is simultaneously augmented with an increased quantity of the collected resources.
- **Build.** In a scenario where the agent's inventory comprises one unit of wood and one unit of stone, and the current grid the agent occupies is vacant, the agent has the option to select a Build action. The act of Building requires the expenditure of labor (l^B) and resources (1 wood and 1 stone). Once the Build action is completed, a house is constructed in the corresponding grid, and the agent receives a coin income as a result.
- **Trade.** The Trade action in our environment is similar to the continuous double auction mechanism described in Zheng et al. (2022). The central market emphasizes that buyers and sellers can submit their trading requests to the market, which then matches transactions based on these requests. In the model, a central market is established within the economic world map, where economic participants can submit requests to buy or sell ($\{ask, bid\}$). Specifically, they can propose to buy or sell a quantity of h of item $i \in \{W, S\}$ (x^i) and give a price of p^i for each item ($\{ask/bid, h, x_i, p^i\}$). If the bid is not less than the ask, the two requests will be successfully matched, and the transaction will be established, resulting in an exchange of resources. Subsequently, buyers and sellers proceed to the Central Market to collect their respective purchases or coins based on the trade's outcome.

2.2.2 Learning

Agents can improve their productivity through practice, self-perfection, and minor innovations.Such behavior is called *learning by doing* (Arrow, 1971). In our model, agents accumulate working skills while taking actions: Move, Gather, and Build ($\{s^M, s^G, s^B\}$), which affects their corresponding productivity. The skill accumulation is affected by labor expenditure l^a and learning rate ρ^a, which follows $s^a_{i,t} = s^a_{i,t-1} + l^a * \rho_a \; a \in \{M, G, B\}$.

The main difference between AI and human agents lies in their approach to acquiring working skills. Human agents primarily rely on the process of “learning by doing" to enhance their skills over time. They gradually improve their abilities through practice, experience, and minor innovations. In contrast, AI agents possess the unique capability to effortlessly share working skills with each other in real-time. This enables them to quickly adapt and benefit from the collective knowledge and advancements within the AI network. Prior to each period, all AI agents synchronize their working skills with the skill frontier, aligning themselves with the most up-to-date and effective approaches. This dynamic skill-sharing mechanism allows AI agents to continuously improve and excel in their tasks, surpassing the limitations of individual learning.

2.2.3 Skill

Agents possess skills in three specific actions: Move, Gather, and Build, denoted as s^M, s^G, s^B, respectively. These skills have distinct effects on the agents' performance in each action.

- **Gather**: In the Gather action, an agent's skill level influences the gather probability. A more skilled agent has a higher probability ($\min(1, \frac{1}{1-e^{-(s^G-2)}})$) of successfully gathering an extra unit of resource. The gather probability increases exponentially with the agent's skill level, allowing highly skilled agents to gather additional resources more frequently.
- **Move**: In the Move action, an agent's skill level determines their efficiency in traversing the world. More skilled agents can cover greater distances with reduced labor consumption. The labor cost for moving is calculated as ($l^M = l_0^M \times (1 - \max(0.5, \frac{1}{1-e^{-(s^M-10)}}))$), where l_0^M denotes the baseline labor cost for movement. Skilled agents can move more efficiently, conserving labor resources.
- **Build**: The action of Build is positively related to the coin income, which can be explained as more skilled agents tend to have more income after the Build action is finished.($\max(pay_{max}, s^B)$, where pay_{max} denotes maximum Build income multiplier.)

2.2.4 Utiltiy

Agents derive utility from their coin holdings while incurring a negative utility from engaging in work. The utility function can be expressed as follows:

$$u_i(x_{i,t}^c, l_{i,t}) = \frac{{x_{i,t}^c}^{1-\sigma} - 1}{1-\sigma} - l_{i,t}, \quad l_{i,t} = \sum_a^A l_{i,t}^a \; A = \{M, G, B\} \tag{1}$$

Where x is the stock of agents coins, l represents the labor expenditure of individuals, and σ represents the risk aversion index of agents. In other words, the larger stock of coins, the smaller the labor expenditure, the agents will have a higher utility.

2.2.5 Optimization

We have provided the utility function for agents. However, it's important to clarify that agents action within a simulated environment, utilizing the given utility function to adjust and optimize their actions in pursuit of maximizing their lifetime utility function. Taking tree gathering as an example, untrained agents may wander randomly in the environment and occasionally choose to gather when they stumble upon a tree. In contrast, well-trained agents can quickly locate the nearest tree and head directly there for gathering based on all available information.

To further optimize the entire model, we have extended the simulation to $T = 1000$ periods. This means that agents will continuously operate and optimize their behaviors in this simulated environment for up to 1000 periods. After each period, we randomly reset the environment and continue the optimization process until the agent's lifetime utility function stabilizes. Therefore, the agent's lifetime utility function can be expressed as:

$$\max_{\pi_i} U_{a\sim\pi,s'\sim\Gamma,o'\sim\Gamma,\tau}\left[u_i(x^c_{i,0},l_{i,0})+\sum_{t=1}^{T}\beta^t(u_i(x^c_{i,t},l_{i,t})-u_i(x^c_{i,t-1},l_{i,t-1}))\right] \tag{2}$$

Here, β is the discount factor in periods, π_i represents the action policy of agent i, cause our model contains many agents, the decisions of different agents may affect each other, so that each agents takes into account the decisions of other agents when making decisions. s denotes the current state of the agent (including inventory, skills and locations), and o represents the information set that the agent can currently observe (social state - tax rate, $n \times n$ around world). In our model, the measure of utility is based on the stock of coins. Thus, the utility gained by an agent in each period can be defined as $r_{i,t} = u_i(x^c_{i,t}, l_{i,t}) - u_i(x^c_{i,t-1}, l_{i,t-1})$. This equation implies that agents need to determine their action policy π_i to maximize their total lifetime rewards, which is represented by the summation $\sum \beta^t r_{i,t}$, given the taxation τ, the agent's state s, and the action policies of other agents π_{-i}.

Therefore, the challenge mentioned above can be described as an optimization problem within a multi-agent environment that encompasses various taxation systems.

2.3 Planner

The planner assumes a role similar to decision-making bodies such as governments, with the goal of maximizing social welfare. In the aforementioned model, variations in agents' birthplaces and resource endowments give rise to income disparities. Taking this into account, and based on the characteristics of the model, we set the planner's objective as maximizing the overall benefit of social total productivity and human equality (without AI agents). To achieve this goal, the planner possesses the authority to levy income taxes, collecting them from all agents in the economy (including AI agents) at a tax rate denoted as τ. Subsequently, these taxes are redistributed evenly among human agents, a process we refer to as "redistribution." The planner's objective can be formulated as a maximization problem, expressed as follows:

$$\max_{\tau}\ \mathrm{Productivity}(\mathrm{All},\tau)\times\mathrm{Equality}(\mathrm{Human},\tau)_{o_p\sim\Gamma} \tag{3}$$

Productivity is the total social coin output in period T, Equality is the Gini coefficient, and o_p denotes the information set that is currently available, enabling the planner to observe and assess relevant factors influencing the system's dynamics and outcomes. Then the tax rate determined through the above maximization process is referred to as the **optimal taxation**.

$$
\begin{aligned}
\text{Productivity}(\text{All}, \tau) &= \sum_{i}^{N} x_{i,T}^{c} \\
\text{Equality}(\text{Human}, \tau) &= 1 - \text{Gini}(x_{i,T}^{c}) \ \forall \ i \in \text{Human agent}
\end{aligned}
\tag{4}
$$

Meanwhile, based on the specific settings of the model, we have decided to implement a taxation strategy throughout the simulated T periods of individual activities. It's important to note that the definition of a "period" here differs slightly from that in traditional economic models. In our simulated environment, a "period" represents a single action decision made by an agent. Some agent's actions may not generate income, but they still consume a "period" opportunity. Therefore, we have chosen to conduct taxation every $T_{tax} = 100$ periods. In other words, after every T_{tax} periods of simulate/training, we will calculate the income generated by all the agent's activities during those periods and proceed with taxation and redistribution accordingly. Specifically,

$$
\begin{aligned}
Tax_{i,nT_{tax}} &= \tau * \sum_{t=(n-1)T_{tax}}^{nT_{tax}} (x_{i,t}^{c} - x_{i,t-1}^{c}) \\
Lump_{i,nT_{tax}} &= \frac{\sum_i Tax_{i,nT_{tax}}}{n_{Human\ Agent}}
\end{aligned}
\tag{5}
$$

the above equation shows the tax for agent i in period nT_{tax}, while $\tau$is tax rate, $x_{i,t}^{c} - x_{i,t-1}^{c}$ is net income in each t. And the below one is the lump transfer for each human agent.

2.4 Reinforcement Learning Methods

The Planner's problem is inherently challenging as it requires considering the endogenous responses of all agents when making decisions. The complex dynamics and strategic interactions resemble a Stackelberg Game or Ramsey Problem. In such scenarios, the Planner acts as the leader, making decisions that take into account the reactions and behavior of other agents in the system. This interdependence adds an additional layer of complexity to the optimization problem.

In our environment, agents take their best strategy based on the planner's policy, while the planner selects the optimal taxation policy based on the agents' actions. To solve the strategies of agents and the planner, a dual-loop reinforcement learning approach is employed, consisting of inner and outer loops. The inner loop focuses on agents finding optimal actions, while the outer loop involves the Planner choosing the optimal tax policy. Iterative iterations of these loops allow for the model to be solved and refined using the dual-loop reinforcement learning framework. Such dual-loop reinforcement algorithm is illustrated in fig. 2.

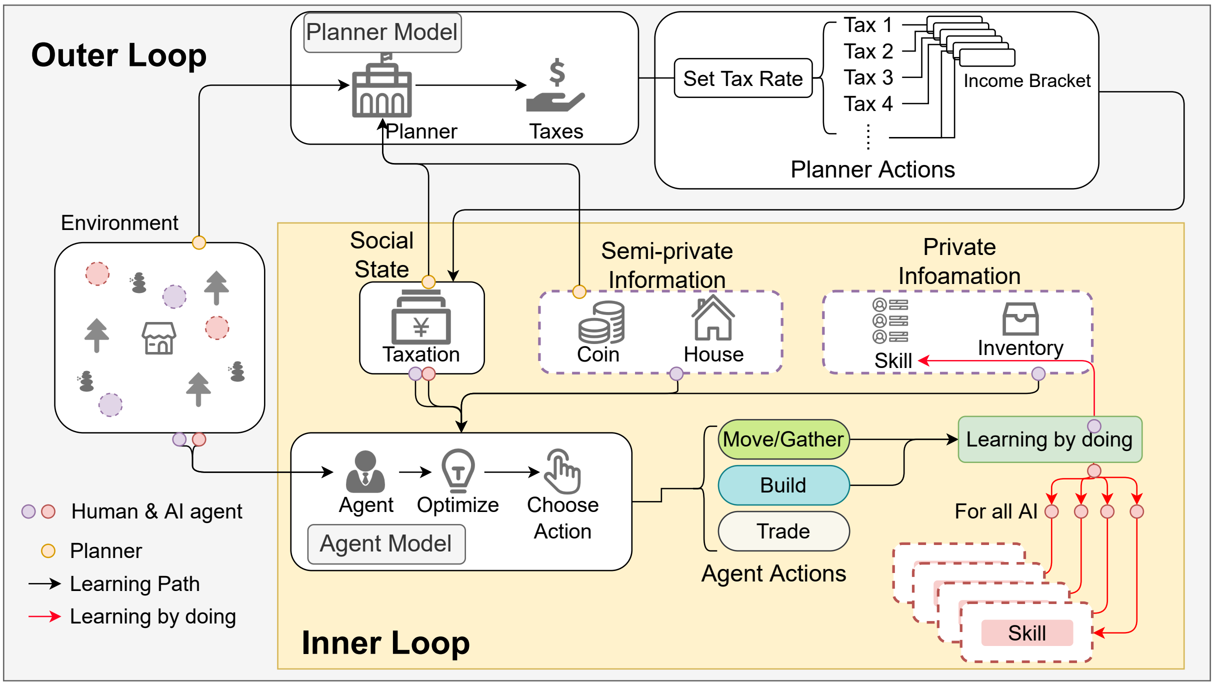

Figure 2: Training flow chart, the yellow part is an inner loop, and the gray part is an outer loop. For the **inner loop** (agents), it optimizes the agents' actions within the environment. By observing the environment and private information, the agent must choose appropriate actions to maximize their lifetime utility. As for the **outer loop** (planner), the planner maximizes social welfare by optimizing taxation. This is achieved through observing the environment and each agent's semi-private information.

Inner Loop In the inner loop, each agent observes her own information from the environment. Such information includes the current state of the agent i s_i (skill level, inventory level, etc.), the information the agent i can observe about the world o_i and the current taxation τ. On top of that, they take action based on their policy

$$a_i \sim \pi(\cdot|s_i, o_i, \tau; \Phi), \tag{6}$$

where Φ is the parameters of the agents' policy networks and $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ denotes the joint action of all the n agents. Notice that the current taxation τ is determined by the planner's policy. Importantly, the planner does not actively engage in the inner loop, resulting in a fixed taxation policy for each iteration of the inner loop.

Outer Loop In the outer loop, the focus is solely on the planner. It is responsible for iterative refining and optimizing the tax policy based on the observed income dynamics of the agents. Specifically, as illustrated in Algorithm 1, the planner updates the tax policy for every T_{Tax} iterations as follows:

$$\tau \sim \pi(\cdot|s', o'_p; \Phi_p), \tag{7}$$

where Φ_p is the parameters of the planner's policy network. Through careful analysis of the agents' income patterns and performance, the planner makes adjustments to the tax policy to

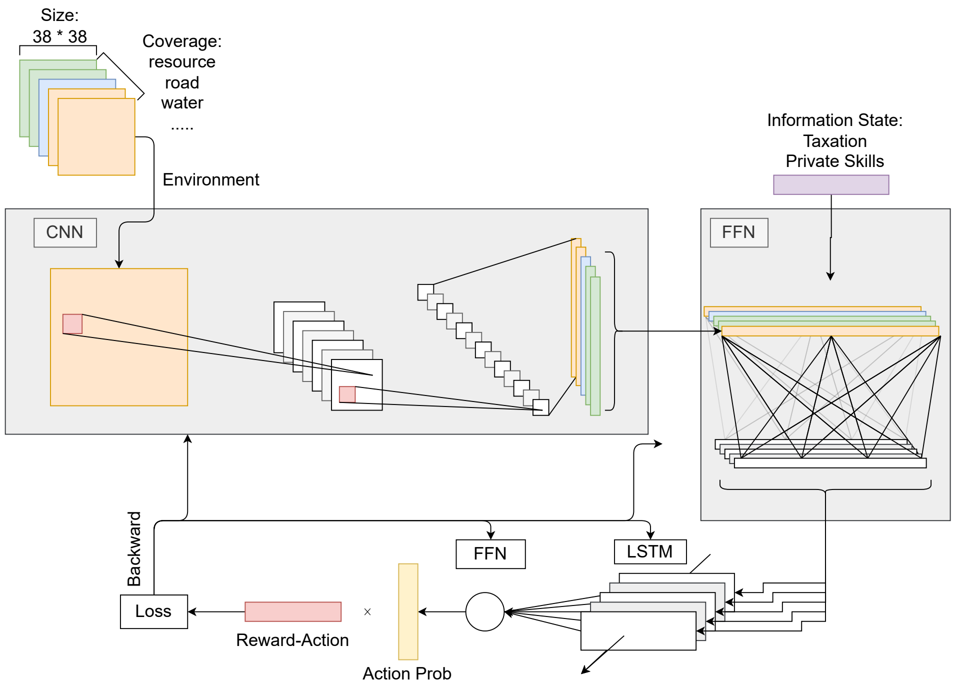

Figure 3: A schematic view of Agent/Planner Policy Network, which takes the map layer, resources layer, and agents layers as inputs. Firstly, we extract feature information from each layer of the map using CNN networks. Subsequently, we refine the extracted features from each layer using FFN networks. Afterward, we reduce the dimension of the refined content from each layer using LSTM based on the map layer. Then, we utilize FFN networks to generate a probability vector for the agents' action set. Finally, we construct the loss function based on the maximization objective and optimize the network structures.

maximize overall social welfare and achieve desired objectives. The outer loop allows the planner to continuously adapt and improve the tax policy based on the observed outcomes from the inner loop.

Dual Loop Finally, after we collect the transition data of one episode through the inner loop and the outer loop, we apply PPO (Schulman et al., 2017), a classic deep reinforcement learning algorithm, to update the policy networks of all the agents and the planner.

Training Network The network used by agents and planner to make actions is shown in the fig. 3. Taking agents as an example, they observe the world state and social state and then decide their actions to maximize the current reward. The process can be described as follows: first, observe the overall environment, then extract features from the environment by CNN, and then merge the social state and obtain the action matrix (probability for each action) of the agent by FFN-LSTM-FFN. Then calculate each action's reward, weigh it as the loss of the current network output, and then backward to optimize all the above networks.

Pesudo Code We also provide a pseudo-code in 1.

Algorithm 1: Dual-Loop Reinforcement Learning

Input: Episode Length T, Tax Frequence T_{tax}
Output: Planner's policy $\pi(\cdot; \Phi_p)$, Agent's policy $\pi(\cdot; \Phi)$

Initialize $\Phi_p \leftarrow \Phi_{p,0}, \Phi \leftarrow \Phi_0$
for *Each Training Episode* **do**
  World$(s, o, o_p) \leftarrow$ World$(s_0, o_0, o_{p,0})$ `/* Reset World */`
  $\tau \sim \pi(\cdot|s, o_p; \Phi_p)$ `/* Sample Planner Taxation */`
  `// Sampling`
  for *timestep* $= 1$ **to** T **do**
    `// Agent Inner Loop`
    $a_i \sim \pi(\cdot|s, o; \Phi)$ `/* Sample agent actions */`
    $s', o', o'_p, r', r'_p \leftarrow World(\mathbf{a}, \tau|s, o, o_p).step$ `/* Next step and get reward */`

    `// Planner Outer Loop`
    if $t \mod T_{tax} == T_{tax} - 1$ **then**
      $s', o', o'_p, r', r'_p \leftarrow World(\tau|s', o_p).taxation$ `/* Apply taxes */`
    end
    if $t \mod T_{tax} == 0$ **then**
      $\tau \sim \pi(\cdot|s', o'_p; \Phi_p)$ `/* Update Planner Taxation */`
    end

    `// Save sample data`
    $D \leftarrow D \bigcup \{(s, o, a, r, o', s')\}$
    $D_p \leftarrow D_p \bigcup \{(s, o_p, \tau, r_p, s', o'_p)\}$
  end

  `// Training`
  Update Φ_p, Φ using the PPO algorithm based on data D, D_p.
end

3 Experiments - Impact on inclucing AIs

In this paper, we explore five different environments, categorized as fixed taxation and optimal taxation (without/with a planner). Within each category, we examine environments with AI and without AI. The experiments are labeled as follows to indicate the specific conditions under investigation:

- **All human (8H)**: This environment shares similarities with the one described in Zheng et al. (2022), where all agents are human agents. The taxation setting in this environment is based on the US Federal Taxation system (Scarboro, 2018), utilizing Bracketed Tax Schedules. The primary objective for the planner in this setting is to learn the optimal tax schedule that maximizes the social objective.
- **AI and Human (4H4A)**: The experiment is almost identical to **All Human (8H)** except that half of the agents are AI agent. And the planner only cares about the inequality of

human agents.

- **AI and Human with free market (4H4A_free))**: This experiment abstracts from **AI and Human (4H4A)** by removing the planner, where there is no taxation or redistribution.
- **AI and Human with US taxation (4H4A_US)**: This experiment modifies **AI and Human (4H4A)** by removing optimal taxation. Thus, the model maintains the initially given taxation. The initial taxation is provided by US Federal Income Tax Rates (Scarboro, 2018).
- **AI and Human with the optimal taxation in 8H (4H4A_8Htax)**: Same as **AI and Human with US taxation (4H4A_US)**, but the fixed taxation is given by convergent optimal taxation result in **All human (8H)**.

Table 2: Main settings for the aforementioned environments and introduction of key differences. The naming convention for the model abbreviations follows the pattern **a**H**b**A_**c**, where **a** represents the number of human agents, **b** represents the number of AI agents, and **c** indicates the corresponding taxation rule if present, or indicates the need to optimize for the optimal taxation if absent. For instance, **4H4A_US** represents a model with 4 human agents, 4 AI agents, and taxation based on the U.S. taxation.

Model Name	Agents	Fixed Taxation?	(Initial) Taxation
8H	(8 human) agents	no	Trained Taxation
4H4A	(4 human + 4 AI) agents	no	Trained Taxation
4H4A_free	(4 human + 4 AI) agents	yes	Free Market
4H4A_US	(4 human + 4 AI) agents	yes	US Federal Income Taxation
4H4A_8Htax	(4 human + 4 AI) agents	yes	Trained Taxation in 8H

3.1 Model Training

3.1.1 Optimal Taxation

The process, known as optimal taxation, involves continuous optimization of taxation by planners, aiming to maximize social welfare based on agents' information and behavioral decisions within the environment. We conducted our experiments using the RLlib framework (Liang et al., 2018) and employed the proximal policy optimization algorithms (Schulman et al., 2017) with the Adam optimizer (Kingma and Ba, 2015) to solve the multi-agent problem described above. The training process consisted of two phases. In the first phase, agents were pre-trained without any taxation, allowing them to learn and optimize their actions in the absence of tax-related constraints. This phase aimed to establish a baseline understanding of agent behavior and performance. The second phase involved training with the inner-outer loops, where the planner was introduced. This phase focused on training agents to adapt their actions based on the taxation policy implemented by the planner. To collect training samples efficiently, we ran 15 parallel replicas of the environment. In total, we collected 45,000 episodes, with 5,000 episodes in the first phase and 40,000 episodes in the second phase. Each episode consisted of 1,000 timesteps.

The training rewards result is shown in Figure 4.

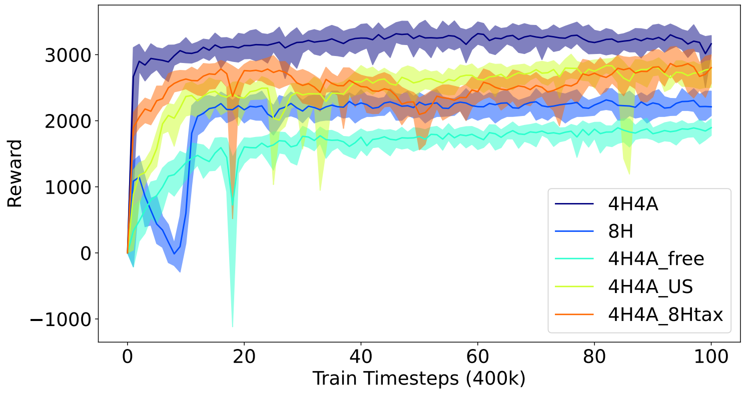

Figure 4: Empirical training progress for all models. The **4H4A**(Dark Blue) achieves significantly better social welfare than other models. As the training iterations increase, all models have converged.

From the functional trend, the model gradually converges to a steady state. We then calculate the average tax rate for the last 50 optimizations as the training result. Then substitute into the fixed taxation model, and train the model 5000 episodes to get the optimal agent performance under that taxation.

3.1.2 Fixed Taxation

In contrast to the Optimal Taxation model, the Fixed Taxation model does not allow the planner to change the taxation. As a result, only the inner loop is present in this model. Consequently, we only utilize 5000 episodes during the second phase of training, as there is no need for the planner to refine the tax policy in this scenario.

3.2 Learning by doing

The skill accumulation between AI (agents 1, 3, 4, 6) and human (agents 0, 2, 5, 7) agents after the introduction of Learning by doing and AI sharing is shown in Figure 5:

The horizontal axis represents the number of timesteps, while the vertical axis represents the skill value. The skill accumulation trends of AI agents are depicted as dashed lines. The figure consists of six subplots, each representing a different action skill and model. Subplots A1 to A3 demonstrates the skill accumulation trends of the three skills when only human agents (8H) are present, while subplots B1 to B3 showcase the skill accumulation trends when both AI and human agents (4H4A) coexist.

The skill level of AI agents demonstrates a remarkable superiority over human agents, exhibiting a higher growth rate. Notably, AI agents working longer hours exhibit a significantly faster rate of skill improvement. These findings underscore the potential of AI to gradually replace human agents, particularly in low-skilled work environments. Moreover, they highlight AI's capacity for large-scale learning and ongoing advancements facilitated by knowledge sharing through advanced Internet technology.

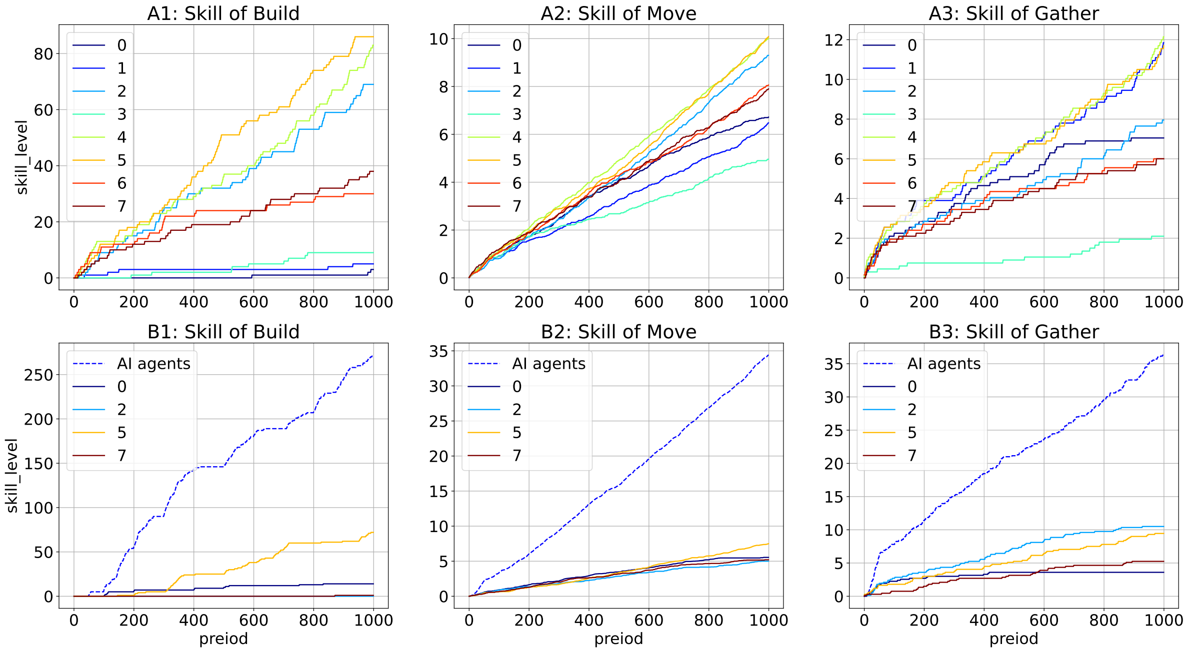

Figure 5: Skill level change without AI(A1-3,**8H**) agents and with AI(B1-3,**4H4A**) agents. The skill level and growth rate of AI agents are significantly higher than those of human agents.

3.3 AI's Impact

The essence of our model is a Stackelberg Game, characterized by a hierarchical structure involving two primary entities: agents and a planner. The actions or policies made by these entities are dynamic and mutually influential. This enables us to observe the operational logic of the economy and the impact of AI from the unique perspectives of decision-making by both entities - the agents and the overarching planner.

3.3.1 The agents' decision

Income/Coin Figure 6 compares the economic performance with and without the presence of AI. A1-3 represent scenarios with only human agents (**8H**), while B1-3 represents scenarios with the introduction of AI agents (**4H4A**, dashed line). Plot 1 (A1&B1) displays the aggregated economic information for each agent throughout the **entire periods**, including net income, taxes, and transfers. The pre-tax income is the sum of net income and taxes, while the after-tax income is the sum of net income and transfers. Plot 2 shows the pre-tax coin state of the agents **in each period**, excluding taxes and transfers. Plot 3 illustrates the after-tax coin state of the agents, considering taxes and transfers. In both Plot 2 and Plot 3, the horizontal axis represents the number of periods, and the vertical axis represents the number of coins.

In all humans (**8H**) model, notable disparities in income among agents emerge, influenced by factors such as resource distribution and skill endowment. The ranking of pre-tax income, from highest to lowest, is as follows (agents' index): $5, 4, 2, 7, 6, 1, 0, 3$. In subplot A2, the pre-tax income differences and trends become more pronounced. In A3, the impact of taxation

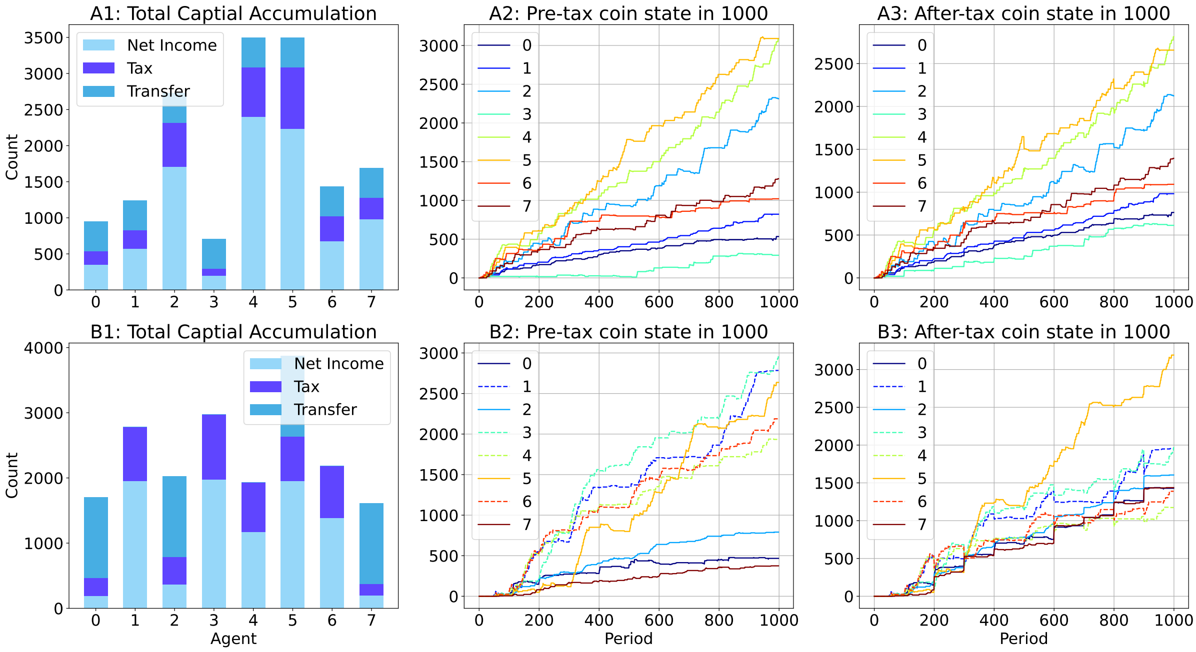

Figure 6: Wealth information without AI(A1-3,**8H**) agents and with AI(B1-3,**4H4A**) agents. There is a decrease in pre-tax income for all human agents, which is also referred to as the crowding-out effect. However, the magnitude of this effect varies depending on the strength of human endowments, with individuals who have stronger initial endowments experiencing a relatively smaller impact. Additionally, the optimal taxation implemented by the planner, which only considers human agents, can result in more transfers for human agents. This ensures that their after-tax income (with taxes and transfers) remains similar to or even better than before the introduction of AI agents.

on inequality is small, and the income trends remain consistent with those observed in A2.

The introduction of AI agents leads to several notable changes in the system:

1. The pre-tax income of all human agents experiences a decrease, affecting both those with originally higher and lower pre-tax incomes. In contrast, AI agents exhibit a higher level of pre-tax income, as observed in the comparison between subplots A2 and B2.
2. This effect varies across different agent types. High-income and lower-income agents in **8H** (index: $0, 2, 5$) can maintain their per-tax income at the original level, while middle-income agents (index: 7) experience a more pronounced decline.
3. Furthermore, the planner implements redistribution through taxation, levying taxes on all agents and subsequently distributing the collected tax equally among agents. As previously mentioned, since the welfare of AI agents is not taken into consideration by the planner, no transfer payments are made to them. Instead, all tax transfers are allocated to human agents, resulting in even higher after-tax income for human agents compared to the model with only human agents (**8H**).

Working Hours Figure 7 depicts the influence of introducing AI agents on the total working hours of human agents. The graph provides a comparison of working hours before (**8H**) and after (**4H4A**) the introduction of AI agents, excluding the labor saved through skill acquisition.

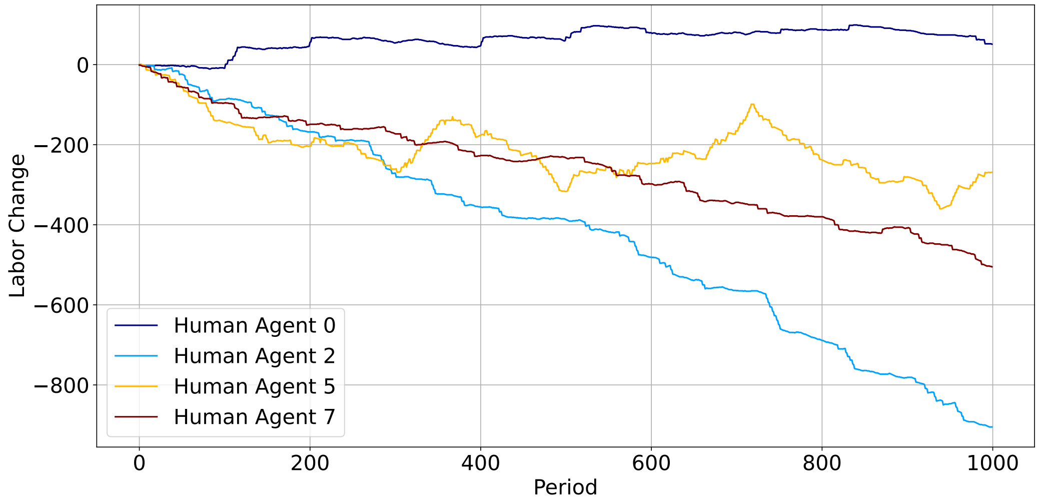

Figure 7: The change in labor for human agents without AI (**8H**) agents and with AI (**4H4A**) agents. It is evident that most human agents opt to reduce their working hours. This represents a displacement of human labor.

Each line represents a distinct human agent, with the horizontal axis representing the period and the vertical axis indicating the change in working hours.

From this figure, it is evident that most human agents opt to reduce their working hours, while there is a slight increase in working hours for human agent 0.This can be attributed to the fact that human agent 0, who is initially in a disadvantaged productivity position and faced competition from surrounding agents (as observed in Figure 6 A2), continues to face competition from AI agents, resulting in small changes in working hours. However, for other human agents, the impact of competition from AI agents is more pronounced, leading to a noticeable reduction in their working hours.

Working hours are closely tied to the action choices made by agents. Therefore, with the introduction of AI agents, we observe a decrease in the overall working hours of human agents. This indicates that AI agents begin to displace the labor of certain agents, ultimately leading to a reduction in their pre-tax income.

Division of labor Figure 8 illustrates the changes in different forms of social division of labor – **Trade** (trade income, Plot 1), **Build** (house count, Plot 2), and **Gather** (resource collection, Plot 3)) before (**8H**) and after (**4H4A**) the introduction of AI agents. The plots are represented as bar charts, with two columns representing the corresponding data (trade income, house count, resource collection) before and after the introduction of AI agents. Agents generate income through building and trading activities, with the building being a high-income activity due to its higher revenue per unit compared to trade income from building resources (1 Stone + 1 Wood).

From Plot 2, it is evident that the introduction of AI agents leads to a decrease in the number of houses for most human agents. This effect is particularly pronounced for agents who initially

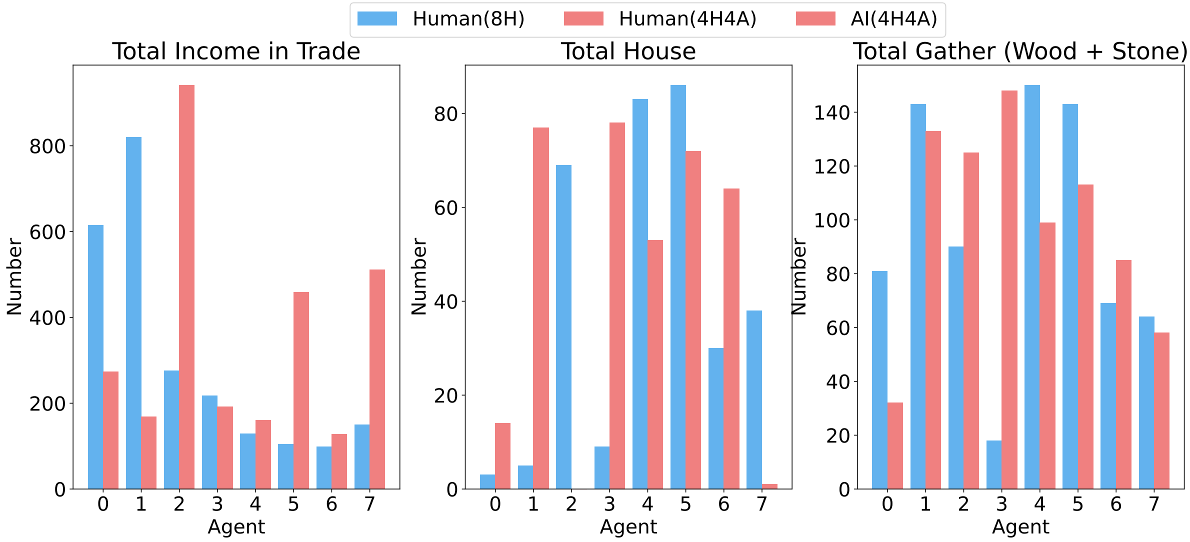

Figure 8: Trade, Trade income and Gather information of agents, before (**8H**) and after (**4H4A**) the introduction of AI agents. It is evident that the introduction of AI agents leads to a decrease in the number of houses for most human agents. Additionally, there is a significant increase in resource trade income for human agents (Plot 2), indicating their evolving role as suppliers of resources for the construction of AI agents' houses.

had a medium or high number of houses (2,7). In contrast, the number of houses built by AI agents experiences a significant increase. Additionally, due to the strong crowding-out effect, agents with higher incomes build fewer houses after transitioning to AI agents.

In terms of resources (Plot 3), the overall change in the number of resources collected is negligible. However, there is a clear trend of decreasing resource numbers for human agents and increasing numbers for AI agents. Conversely, there is a substantial increase in resource trade income for human agents (Plot 2), indicating their gradual transition into resource providers for AI agents' house construction. The introduction of AI agents significantly displaced human agents from high-income activities (such as building), leading to a shift in their behavioral choices towards becoming resource providers.

In summary The introduction of AI agents has a significant crowding-out effect on human agents, causing a reduction in their per-tax income.

1. **Substitution of high-skilled jobs.** AI agents crowd out human agents across all activities (Build and Gather), particularly in high-income activities (Build). Thus, AI agents firstly crowd out human agents from high income and then from other activities, and finally force a shift in the agents' main income methods (become resource providers). This phenomenon contrasts with the traditional utility of machines, which primarily replaces repetitive human labor, often associated with lower wages. This is one of the reasons why AI technologies are regarded as transformative for the future, as they possess the capability to replace jobs that demand higher levels of skills.
2. **Crowding out effect on different human agents.** For different human agents, the

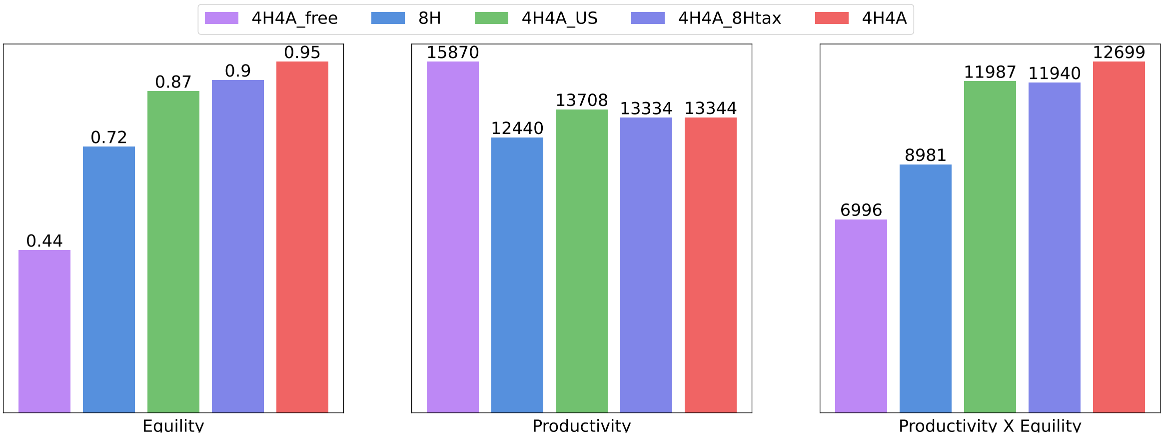

Figure 9: "Equality", "Productivity" and "Productivity $\times$ Equality" in all Models. If a planner chooses proper taxation, it enables the effective utilization of the productivity advantages brought by AI while mitigating its negative impact, such as crowding-out effects. Among all taxation policies considered, the optimal taxation learned in our model training maximizes social welfare by achieving a balance between equality and productivity.

original high-income agents tend to have little impact, as they can rely on their abilities and endowments to maintain a higher pre-tax income; For low-income agents, they are already crowded out by other human agents, and the crowding out after the introduction of AI will have little impact on them; The middle-income agents are the most affected, from skills to working hours, and are heavily dominated by AI agents, resulting in a severe crowding-out effect.

3. **Channels of the crowding out effect.** The crowding-out effect mainly has two channels. Firstly, AI agents have a higher skill to dominate low-skilled human agents, thus crowding out their working hours. So, this generates the crowding-out effect due to substitution. Secondly, as mentioned before, after the introduction of AI agents, the planner only considers human welfare and human agents can get more transfers. The after-tax income of human agents increases significantly. For the original human agents with lower pre-tax income, because of higher transfers and the negative utility generated by labor, they have no incentive to labor anymore and choose to stay and give up their labor opportunities to AI agents. This generates the crowding-out effect due to leisure.

3.3.2 The Planner's decision

Welfare Analysis Figure 9 represents the equality, productivity (all agents), and productivity $\times$ equality in five models. Equality is measured among human workers. In the Human-AI hybrid environment, the planner aims to achieve equality through a tax and transfer system. In the presence of the planner, the use of proper taxation can effectively address the negative impacts caused by the introduction of AI. Though AI crowds out some human labor, the taxation can redistribute income from AI to human workers as compensation to the displaced human agents, reducing the negative effects of AI introduction on social welfare (productivity $\times$ equality). In summary, the planner plays a crucial role in optimizing social welfare in a Human-AI hybrid environment. Next, we will analyze specifically from three perspectives illustrated in Figure 9.

1. **Equality** From the left plot in Figure 9, if the planner employs effective taxation (compared to free market **4H4A_free**), the society is more equal when AI presents. This is because planner aims to maximize the welfare of human workers, higher transfers can be provided to human workers by taxing heavily on AI. The substantial transfers can help reduce the level of income inequality to some extent.
2. **Productivity** The middle plot in Figure 9 represents the overall social productivity. The presence of AI agents leads to higher social productivity (See **4H4A**, **4H4A_US**, **4H4A_free**, **4H4A_8Htax**) than before (See **8H**). This is because AI agents possess higher skill levels, which leads to increased productivity. In addition, due to the limited resources in the environment, the substitution of human agents for AI agents allows for more efficient resource utilization. As a result, there is an overall boost in productivity.
3. **Productivity** $\times$ **Equality** The right plot in Figure 9 represents the level of social welfare, measured as a trade-off between productivity and equality. Firstly, if a planner chooses proper taxation, it enables the effective utilization of the productivity advantages brought by AI while mitigating its negative impact, such as crowding-out effects. Secondly, among all taxation policies considered, the optimal taxation learned in our model training maximizes social welfare by achieving a balance between equality and productivity.

Optimal taxation It is well concerned that the crowding-out effect leads to job losses and income reduction. Typically, such an effect would also increase inequality. Fortunately, we can find from the above welfare analysis that, society can achieve an improvement in both equality and productivity under a social planner. This is due to the proper taxation policy taken by the social planner. Such results demonstrate a feasible Pareto optimality. AI agents effectively replace human agents and compensate them through government transfer. Therefore, we can achieve an increase in economic productivity without compromising human equality, leading to a welfare-improving society.

Furthermore, the optimal tax rates in different experiments are shown in Figure 10. The three tax schedules consist of U.S. Federal Taxation (red line), taxation for society with only humans, saying **8H** (green line), and taxation for Human-AI hybrid society, saying **4H4A** (blue line). We can see that, in a society with only humans, the optimal tax is roughly flat across income brackets, since the objective of the planner is both productivity and equality. However, in a Human-AI hybrid society, the optimal tax rate is high tax rates for the super high-income groups and lower tax rates for high-income groups, reflecting a progressive tax schedule for AI. A roughly flat tax schedule is for lower-income groups (0 - 200 as in the figure x-axis), most of which are human workers. The tax schedule seems to be flat or regressive, but the net tax rate (tax - transfer) is very progressive since the welfare transfer is only provided for human workers. This tax schedule is well related to the policy debate, saying providing Universal Basic Income

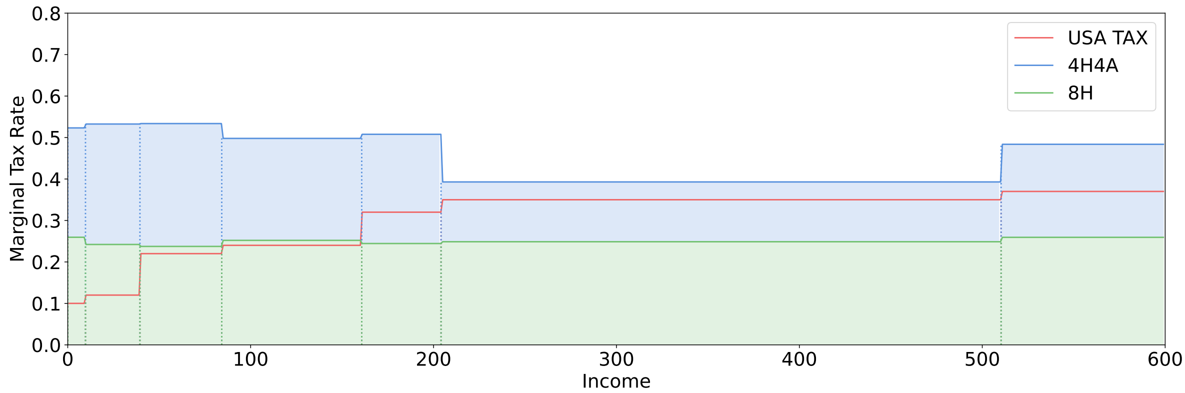

Figure 10: Taxation of **8H, 4H4A** and the real US. When there are only human agents, there is a tendency towards an equalized tax burden. However, with the inclusion of AI, the tax levels are higher for both low-income and high-income segments.

in a world with robots and AI.

4 Experiments - Onwership

4.1 Models

In this section, we primarily discuss two models. One serves as a benchmark, focusing on the key modifications required to introduce ownership. The other model incorporates a keeping-rate setting based on the previous one, aiming to incentivize AI agents to engage in productive activities.

4.1.1 Onwership Models

First, AI agents can possess two types of ownership: ownership by society or ownership by an human agent. Ownership by society implies that the AI agent is publicly owned, so all of its income goes directly into taxation and then goes through transfer payments. In contrast, ownership by an human agent signifies that the income belongs to that specific human agent. Then, we need to make crucial modifications to the **Agent** and **Planner** section of the model. Specifically, we need to clarify the specific timing of when the AI agent transfers income to the owner which is during the taxation period(T_{tax}), but prior to the imposition of taxes, the AI agent transfers all of its income to its owner. which:

$$\begin{aligned} Income_{i,nT_{tax}} = &\sum_{t=(n-1)T_{tax}}^{nT_{tax}} (x^c_{i,t} - x^c_{i,t-1}) + \sum_{t=(n-1)T_{tax}}^{nT_{tax}} (x^c_{i_{belong},t} - x^c_{i_{belong},t-1}) \\ with\ i \in &\ Human\ Agent \\ i_{belong} \in &\ AI\ Agent \end{aligned} \tag{8}$$

Furthermore, during taxation, since the AI agents are now owned by human agents and its productivities are also belong to society/human, the planner will not collect taxes from the AI itself, but only from the human owner. Similarly, transfer payments will only be considered for human agents. Consequently, in this environment, AI agents effectively become mere production tools.

$$\max_{\tau} \text{Productivity}(\text{Human}, \tau) \times \text{Equality}(\text{Human}, \tau)_{o_p \sim \Gamma} \tag{9}$$

4.1.2 Keeping-Rate Models

For the Keeping-Rate model, our primary assumption is that AI agents do not give all of their income to their owners. From an economic intuition, if AI agents have no investments or income of their own, their motivation would be extremely low, rendering the ownership system ineffective. The introduction of a Keeping-Rate(ζ) setting is akin to making an investment, whether it be for the "maintenance" or "innovation" of AI agents, which can generate a certain utility for them. Consequently, the formula for agents income is updated as follows:

$$\begin{aligned}
Income_{i,T_{tax}} &= \sum_{t=(n-1)T_{tax}}^{nT_{tax}} (x^c_{i,t} - x^c_{i,t-1}) + (1-\zeta) \sum_{t=(n-1)T_{tax}}^{nT_{tax}} (x^c_{i_{belong},t} - x^c_{i_{belong},t-1}) \\
Income_{j,T_{tax}} &= \zeta \sum_{t=(n-1)T_{tax}}^{nT_{tax}} (x^c_{j,t} - x^c_{j,t-1}) \\
& \quad with\ i \in Human\ Agent \\
& \quad \quad i_{belong}, j \in AI\ Agent
\end{aligned} \tag{10}$$

Moreover, the question of who determines this Keeping-Rate is worthy of discussion. Firstly, AI agents cannot be allowed to decide their own Keeping-Rate, as the outcome would converge entirely to 1, the AI agents would maximize their own utility. The decision, therefore, falls to the owners or social planners. Currently, we are discussing having social planners determine the Keeping-Rate, with the goal of maximizing social welfare. Hence, social planners need to optimize two objectives simultaneously: optimal taxation and optimal social Keeping-Rate.

$$\max_{\tau} \text{Productivity}(\text{Human}, \tau, \zeta) \times \text{Equality}(\text{Human}, \tau, \zeta)_{o_p \sim \Gamma} \tag{11}$$

Our model becomes more complex at this point. Taxation affects overall social inequality and labor motivation, while the Keeping-Rate impacts the work enthusiasm of AI agents, thereby influencing the overall productivity of society and subsequent transfer payments.

Pesudo Code We also provide a pseudo-code in 2.

Algorithm 2: Keeping-Rate Model Pesudo Code

Input: Episode Length T, Tax Frequence T_{tax}, Keeping-Rate ζ

Output: Planner's policy $\pi_\tau(\cdot;\Phi_p^\tau)$ and $\pi_\zeta(\cdot;\Phi_p^\zeta)$, Agent's policy $\pi(\cdot;\Phi)$

Initialize $\Phi_p^{\tau,\zeta} \leftarrow \Phi_{p,0}^{\tau,\zeta}, \Phi \leftarrow \Phi_0$
for *Each Training Episode* **do**
  World$(s, o, o_p) \leftarrow$ World$(s_0, o_0, o_{p,0})$ `/* Reset World */`
  $\tau \sim \pi(\cdot|s, o_p; \Phi_p^\tau)$ `/* Sample Planner Taxation */`
  $\zeta \sim \pi(\cdot|s, o_p; \Phi_p^\zeta)$ `/* Sample Keeping-Rate */`
  `// Sampling`
  for *timestep* $= 1$ **to** T **do**
    `// Agent Inner Loop`
    $a_i \sim \pi(\cdot|s, o; \Phi)$ `/* Sample agent actions */`
    $s', o', o'_p, r', r'_p \leftarrow World(\mathbf{a}, \tau, \zeta|s, o, o_p).step$ `/* get reward */`

    `// Planner Outer Loop`
    if $t \mod T_{tax} == T_{tax} - 1$ **then**
      $s', o', o'_p, r', r'_p \leftarrow World(\tau, \zeta|s', o_p).taxation$ `/* Apply taxes */`
    end
    if $t \mod T_{tax} == 0$ **then**
      $\tau \sim \pi(\cdot|s', o'_p; \Phi_p^\tau)$ `/* Update Planner Taxation */`
      $\zeta \sim \pi(\cdot|s', o'_p; \Phi_p^\zeta)$ `/* Update Keeping-Rate */`
    end

    `// Save sample data`
    $D \leftarrow D \bigcup \{(s, o, a, r, o', s')\}$
    $D_p \leftarrow D_p \bigcup \{(s, o_p, \tau, r_p, s', o'_p)\}$
  end

  `// Training`
  Update $\Phi_p^{\tau,\zeta}, \Phi$ using the PPO algorithm based on data D, D_p.
end

4.2 Experments

In this part, based on **4H4A** Model, we explore 4 different environments for ownership Model. The experiments are labeled as follows to indicate the specific conditions under investigation:

Ownership Model

- **4B_each**: Each human agent has a AI agent.
- **2B2S**: Two human agent each has a AI agent and another two AI agent is belong to society.
- **4S**: All AI agent is belong to society.
- **4B_bias**: A human agent has two AI agents, two human agent each has a AI agent and one human agent has not AI agent.

Also, we explore 4 different environments for Keeping-Rate Model.

Keeping-Rate Model

- **2B2S_same**: All AI agents have a same Keeping-Rate
- **2B2S_cate**: Different category of AI agents(belong to human/society) have different Keeping-Rate.
- **2B2S_private**: AI agents belonging to society have a common Keeping-Rate and AI agents belonging to humans each have a private Keeping-Rate.
- **4B_private**: All AI agents each have a private Keeping-Rate
- **4S_same**: All AI agents have a same Keeping-Rate

4.3 The Impact of Ownership

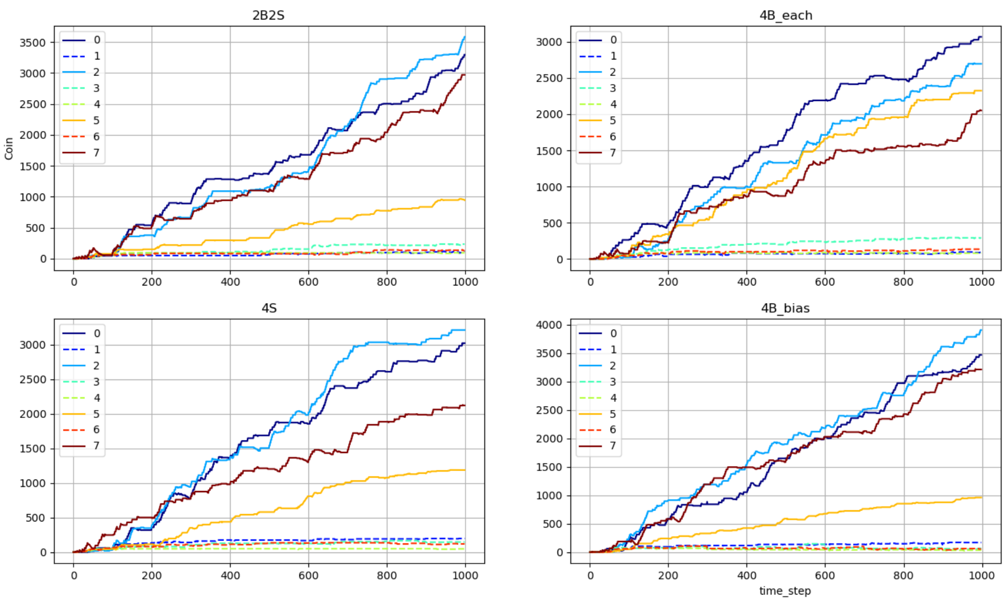

Figure 11: The image represents the cumulative pre-tax income of agents across all ownership models. The horizontal axis indicates the number of periods, while the vertical axis represents the coin stock.

As evident from the Figure11, due to the crude introduction of ownership, the AI agent transfers all of its income to its owner. Consequently, the AI individuals have no incentive to perform any additional actions because they would be giving away all of their income. Therefore, the multi-agent environment with **4A4H** degrades into one involving only 4 human agents, where income disparities among agents depend on their endowments and random parameters. From the results, it is apparent that the graphical outcomes of the four models are almost identical.

Furthermore, as observed from Figures 12 and 13, the net resource purchases of AI agents are close to zero but negative, indicating that they are selling resources, with very small quantities being sold. Therefore, they gradually become a low-level efficiency tool primarily engaged in

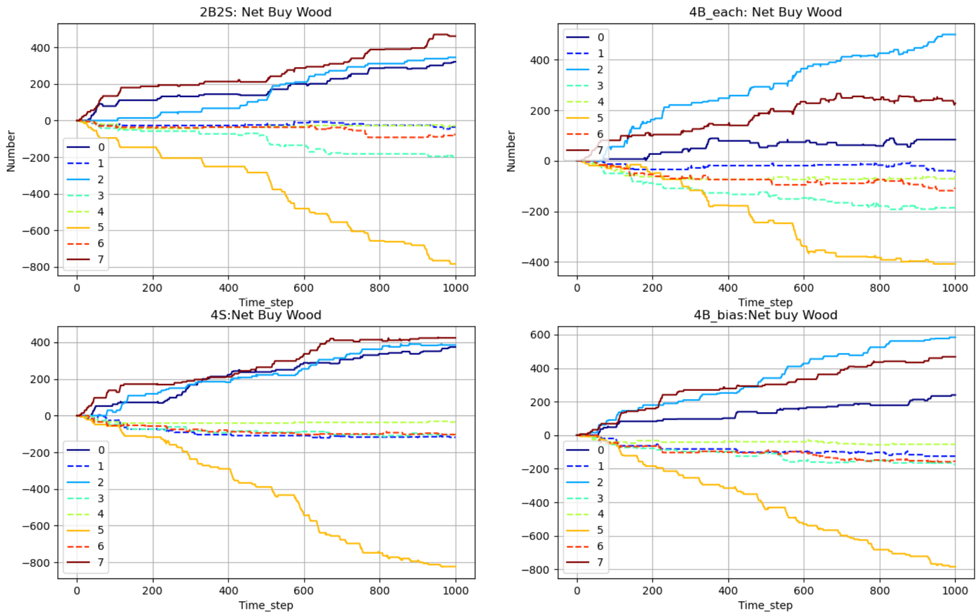

Figure 12: Net sood buy(buy - sell) number

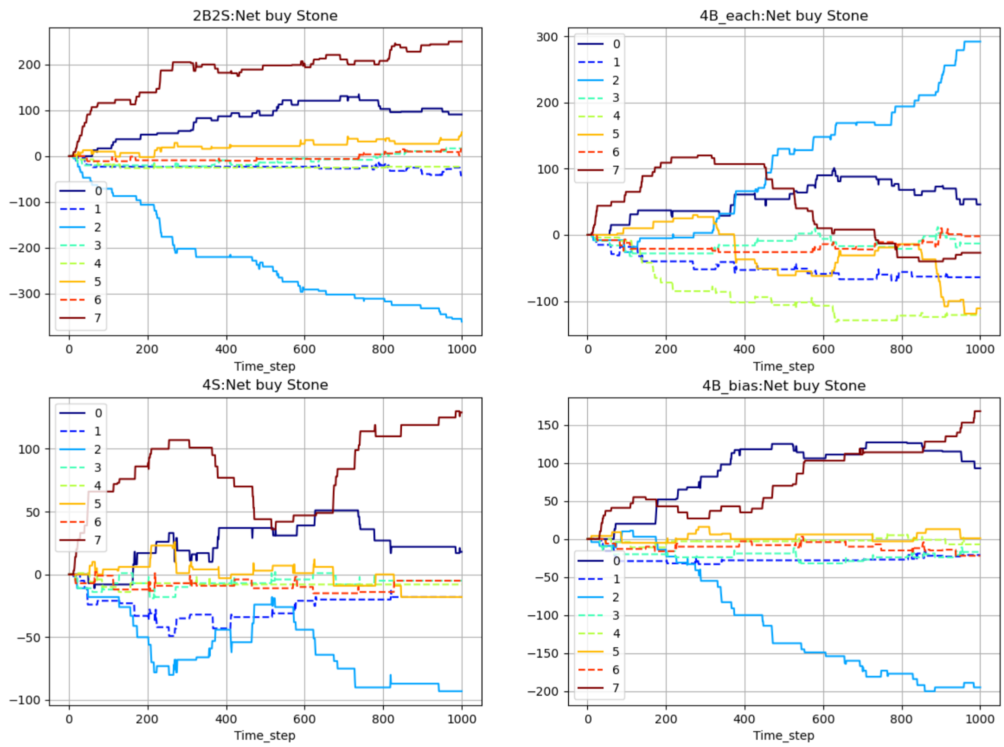

Figure 13: Net stone buy(buy - sell) number

collecting and trading small amounts of resources, and their main source of income is selling the resources they collect. Additionally, due to the low efficiency of AI tools in collecting resources, a specific human agent has emerged to focus on resource collection.

The introduction of an ownership system where AI agents give everything to their owners significantly reduces the labor motivation of AI agents. At this point, AI agents are more like low-level tool substitutes, mainly used for resource collection and gathering. Giving everyone their own AI agents is equivalent to providing them with a resource collector, enabling them to engage in more construction activities and optimizing social welfare (compared to other ownership systems). Realistically speaking, in the initial stages of substitution, there is a tendency to distribute these "tools" to individuals for use rather than for public use, as unclear ownership in public use does not incentivize agents.

4.4 The Impact of Keeping-Rate

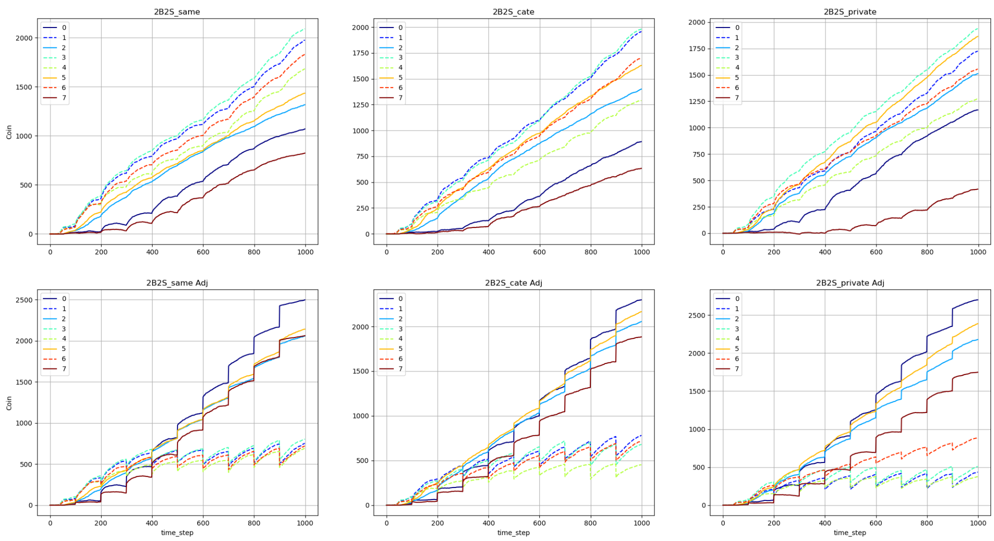

Figure 14: Result for Keeping-Rate Model **2B2S**

From the Figure 14 and 15, it is easy to observe that the Keeping-Rate continuously optimizes as time steps progress. Therefore, due to the significant randomness in single experiments, averaging multiple experiments can effectively reduce randomness and provide a better understanding of the equilibrium trend of the Keeping-Rate. After introducing the Keeping-Rate, AI agents have significantly increased their labor enthusiasm. The decrease in human agents' income is attributed to the crowding-out effect of AI agents and the "lying flat" phenomenon caused by more AI agents' ownership payments. The existence of the Keeping-Rate is also quite interesting. Specifically, in the early stages, the Keeping-Rate is relatively high, allowing AI agents' incomes to continue growing. However, after reaching a certain threshold, planners tend

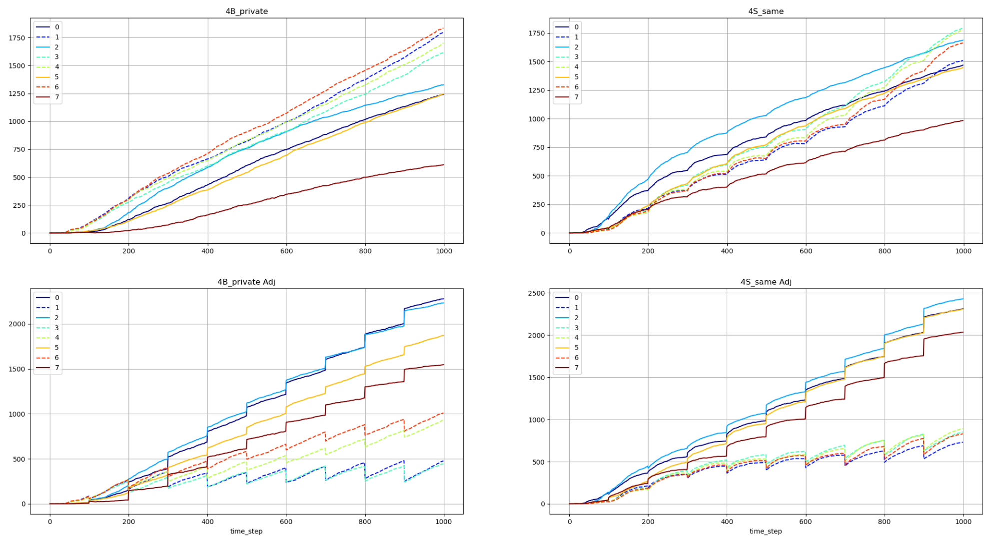

Figure 15: Result for Keeping-Rate Model **4B/4S**

to keep AI agents at a certain wealth level for production and only take their new additions. This approach not only provides AI agents with a certain degree of utility but also maximizes their labor enthusiasm and productivity on the margin.

Table 3: The eq. Keeping-Rate for each Model

Model Name	AI Agents cate	Keeping-Rate
2B2S_same	2 private + 2 society	0.714
2B2S_cate	2 private	0.711
	2 society	0.669
2B2S_private	2 private	[0.514, 0.829]
	2 society	0.581
4B_private	4 private	[0.547, 0.552, 0.782, 0.812]
4S_same	4 society	0.762

Optimal taxation and Social Welfare Based on the Figure 16, it is evident that tax brackets are generally high across almost all income ranges. This is primarily due to the significant increase in agents incomes after the introduction of ownership (regardless of whether Keeping-Rate are implemented or not, as the absence of Keeping-Rate causes the model to revert to an environment with only 4 human agents, resulting in less competition for resources). Consequently, incomes mostly fall within the higher tax brackets, affecting all sectors during training. In this context, taxation serves primarily to ensure social equity rather than motivate agents to engage in productive activities. This is because taxation has minimal impact on the productivity incentives of AI agents.

In this scenario, the introduction of the Keeping-Rate mechanism proves effective. It stimulates AI agents to engage in labor, ultimately enhancing social welfare. Simultaneously,

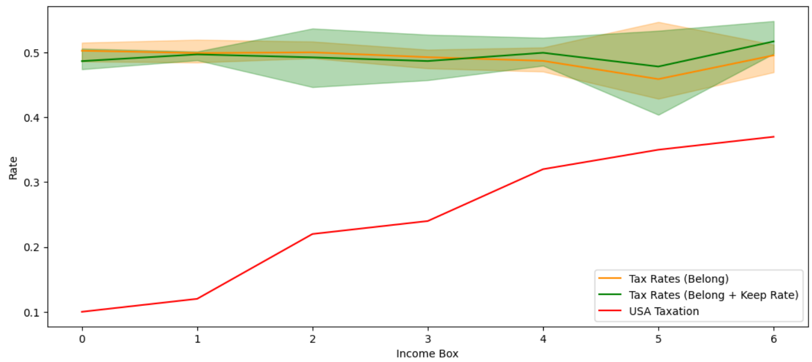

Figure 16: Trained optimized taxation in both ownership and Keeping-Rate Model(**4B**)

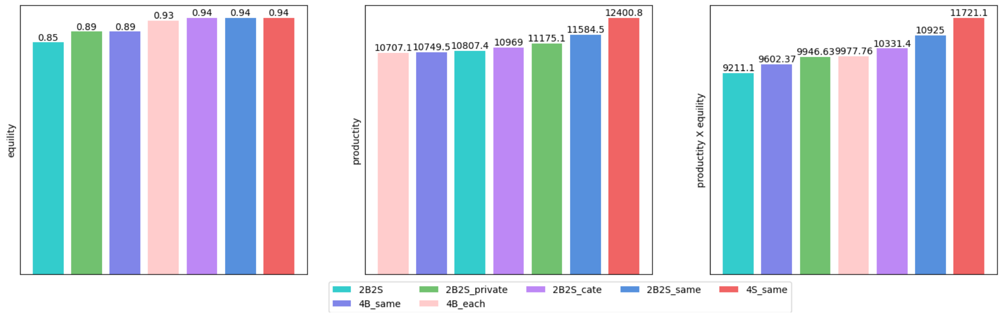

Figure 17: Final Social Welfare for all Models

efficient AI skills can substitute for some human labor without reducing individual utility, allowing people to enjoy more leisure time. This aligns with a future societal goal: to replace most repetitive tasks with artificial intelligence and distribute the value created by AI to everyone through various transfer payments, enabling individuals to devote more time to pursuits of spiritual value.

5 Conclusion

In this paper, we empirically study the economic impacts of introducing AI in a Human-AI hybrid environment we constructed. The environment is built on a multi-agent framework and solved through reinforcement learning. We analyze how the planner should implement policy in this Human-AI hybrid environment to maximize social welfare. The key feature in our environment is that, AI can share skills and knowledge in real-time, but humans accumulate human capital through working experience.

From the agents' perspective, we find that AI increases productivity and brings higher output for the whole society. However, AI agents also crowd out human agents in labor markets. The human worker with higher or lower initial skills experiences a lesser degree of crowding out. In addition, the planner can provide more transfers to human agents. This enables human agents to a propensity to opt for work withdrawal and further intensifies the crowding-out effect.

From the planner's perspective, the presence of AI can improve both equality and productivity. This can primarily be attributed to two factors: 1) The effective management of the Human-AI hybrid environment by the social planner through proper taxation; and 2) The ability of AI agents sharing knowledge through web and internet technology, enabling them to consistently operate at high skill levels. Specifically, the implementation of effective tax systems allows us to leverage the productivity advantages brought by AI while addressing its negative impact, such as the crowding-out effect. Consequently, this leads to enhancements in both equality and productivity. It also reveals the pivotal role of web and internet technology as a driving force for continuous progress and innovation in various other fields.

And we then discuss the Ownership part, here, we find that: The established Keeping-Rate can stimulate the labor enthusiasm of AI agents. At this point, AI agents can utilize their informational advantages to engage in high-skilled labor substitution (such as construction) through learning. For high-skill substitution, adopting a communal approach can enhance its effectiveness, providing AI agents with more freedom for development.

Combining the conclusions from the previous sections, we can draw an insight:

1. When AI substitution technology is in its early stages, planners should tend to grant this technology to individuals or firms. These technologies can replace some low-skilled labor, freeing up individuals to engage in more advanced labor and bringing more social output to enhance social welfare.
2. When AI technology reaches a high level of skill substitution, if it remains within the firm, it can still bring some optimization. However, if it is merged into a communal system and opened up to every individual in society, this technology can significantly replace human labor, liberating people from heavy workloads and allowing them to engage in lighter tasks. Furthermore, through reasonable taxation and economic measures, this communal output can be distributed fairly to every individual, providing more income for all and increasing overall social welfare.

References

Acemoglu, D. and Restrepo, P. (2017). Secular stagnation? the effect of aging on economic growth in the age of automation. *American Economic Review*, 107(5):174–79.

Aghion, P., Jones, B. F., and Jones, C. I. (2018). Artificial intelligence and economic growth. In *The economics of artificial intelligence: An agenda*, pages 237–282. University of Chicago Press.

Arinez, J. F., Chang, Q., Gao, R. X., Xu, C., and Zhang, J. (2020). Artificial intelligence in advanced manufacturing: Current status and future outlook. *Journal of Manufacturing Science and Engineering*, 142(11):110804.

Arrow, K. J. (1971). The economic implications of learning by doing. In *Readings in the Theory of Growth*, pages 131–149. Springer.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, 103(5):1553–1597.

Aziz, A. (2023). Artificial intelligence produced original work: A new approach to copyright protection and ownership. *European Journal of Artificial Intelligence and Machine Learning*, 2(2):9–16.

Bessen, J. (2018). Ai and jobs: The role of demand. Technical report, National Bureau of Economic Research.

Bowles, J. (2014). The computerisation of european jobs. *Bruegel, Brussels*.
Frey, C. B. and Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change*, 114:254–280.
Gharibshah, Z., Zhu, X., Hainline, A., and Conway, M. (2020). Deep learning for user interest and response prediction in online display advertising. *Data Science and Engineering*, 5(1):12–26.
Goyal, A. and Aneja, R. (2020). Artificial intelligence and income inequality: Do technological changes and worker's position matter? *Journal of Public Affairs*, 20(4):e2326.
Gries, T. and Naudé, W. (2018). Artificial intelligence, jobs, inequality and productivity: Does aggregate demand matter? *Available at SSRN 3301777*.
Grigorescu, S., Trasnea, B., Cocias, T., and Macesanu, G. (2020). A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics*, 37(3):362–386.
Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In Bengio, Y. and LeCun, Y., editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.
Korinek, A. and Stiglitz, J. E. (2018). Artificial intelligence and its implications for income distribution and unemployment. In *The economics of artificial intelligence: An agenda*, pages 349–390. University of Chicago Press.
Li, T., Sahu, A. K., Talwalkar, A., and Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60.
Liang, E., Liaw, R., Nishihara, R., Moritz, P., Fox, R., Goldberg, K., Gonzalez, J., Jordan, M., and Stoica, I. (2018). Rllib: Abstractions for distributed reinforcement learning. In *International Conference on Machine Learning*, pages 3053–3062. PMLR.
Sachs, J. D. (2018). R&d, structural transformation, and the distribution of income. In *The Economics of Artificial Intelligence: An Agenda*, pages 329–348. University of Chicago Press.
Scarboro, M. (2018). State individual income tax rates and brackets for 2018. *Tax Foundation*.
Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.
Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., and Biancone, P. (2021). The role of artificial intelligence in healthcare: a structured literature review. *BMC medical informatics and decision making*, 21:1–23.
Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., and Gao, Y. (2021). A survey on federated learning. *Knowledge-Based Systems*, 216:106775.
Zhang, S., Yao, L., Sun, A., and Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM computing surveys (CSUR)*, 52(1):1–38.
Zheng, S., Trott, A., Srinivasa, S., Parkes, D. C., and Socher, R. (2021). The AI economist: Optimal economic policy design via two-level deep reinforcement learning. *arXiv preprint arXiv:2108.02755*.
Zheng, S., Trott, A., Srinivasa, S., Parkes, D. C., and Socher, R. (2022). The AI economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science Advances*, 8(18):eabk2607.