LLM Agents Empower Simulation for the Future of Work: Universal Basic Income, Employment, and Well-being

Anonymous ACL submission

Abstract

UBI policy remains one of the most widely studied topics in economics, drawing significant attention for its potential social and financial impacts. However, real-world UBI experiments are costly and constrained in scale, limiting their feasibility for large-scale analysis. The emergence of LLM-based society simulations offers a promising alternative, enabling detailed economic and social modeling at a lower cost. We propose an agent-based simulation where Large Language Models (LLMs) role-play individuals in a virtual economy to evaluate UBI policies. By integrating real-world data, our model captures complex human behaviors, including financial decisions and mental well-being. We successfully replicated outcomes from five real-world UBI trials across economic and mental metrics, with ablation studies confirming that LLM role-playing agents produce more realistic and insightful simulations. Our work demonstrates how LLM-powered simulations can advance UBI research and inform policy design. Codes are available here: https://anonymous.4open.science/r/LLM-UBI-7837

1 Introduction

The rapid development of web applications and algorithms has brought about profound changes in the nature of work and employment worldwide (De Stefano, 2019; Lee et al., 2015; Imana et al., 2021; Anagnostopoulos et al., 2018; Wolf and Blomberg, 2019; Kittur et al., 2013; Cao et al., 2021; Chen et al., 2022). On the one hand, these advancements have greatly enhanced efficiency and created new opportunities (Chen et al., 2022; Cao et al., 2021; Gagné et al., 2022; Lee et al., 2015; Wolf and Blomberg, 2019). On the other hand, this development inevitably leads to a growing number of jobs being taken over by advanced artificial intelligence algorithms and web applications, such as ChatGPT (Shen and Zhang, 2024; Tschang and Almirall, 2021; De Stefano, 2019; Zarifhonarvar, 2024). This further amplifies growing concerns about the future of work and how to secure a basic income and live a better life (Gagné et al., 2022).

In response to these growing concerns, researchers have made great efforts to explore Universal Basic Income (UBI) as a promising solution. UBI is a policy framework designed to provide all individuals with a guaranteed, unconditional, universal income to meet their basic needs (Bidadanure, 2019; Banerjee et al., 2019; Haagh, 2019), aiming to alleviate poverty, reduce social inequality, and promote overall well-being (Bidadanure, 2019; Banerjee et al., 2019; Haagh, 2019). Therefore, to evaluate whether UBI can achieve these aims, substantial financial resources have been invested in conducting large-scale experiments across various countries for decades (Haagh, 2019; De Wispelaere and Stirton, 2004; Banerjee et al., 2020, 2019; Coalition, 2012).

However, the implementation of UBI experiments often comes with high costs (Bidadanure, 2019; Banerjee et al., 2019; Haagh, 2019). As noted in the case of Kenya, the experiment took up over 12 years and 30 million dollars, a scale that surpasses the capabilities of most researchers and even the governments of some developing countries without external support (Banerjee et al., 2019), preventing extensive trial to be deployed to obtain comprehensive understanding. As a result, UBI remains a controversial policy, needing a more reasonable and alternative in replace of real-world UBI trails.

The recent advances in large language models (LLMs) have provided a promising solution to this question (Shanahan et al., 2023; Tseng et al., 2024; Zhong et al., 2024). LLM powered agents has shown their abilities to model complex behaviors and adapt to a wide range of scenarios, making them even more effective for simulations (Guo et al., 2024). Prior studies have demonstrated that

LLMs with role-playing have been explored for simulating human-like interactions, enabling advanced conversational agents, and testing social or behavioral theories (Meyer and Elsweiler; Hackenburg et al., 2023; Hao et al., 2024). These studies have provided valuable insights into simulating human dynamics in controllable settings (Li et al., 2024b).

In this paper, we leverage the power of LLM agents to simulate the impact of UBI policies on employment and well-being. Specifically, we first build an LLM-driven simulation system for macroeconomics based on prior literature (Zheng et al., 2022; Li et al., 2024a). Through further introducing demographic profiles of specific cultural contexts into LLM agents, we successfully reproduce the patterns for economic indicators in the real world. Moreover, we incorporate the UBI policies into the LLM-driven simulation system, finding the proposed system captures the economic and social outcomes observed in 5 real-world experiments with 3 economic metrics and 2 social metrics. We also validate the effects of demographic profiles on capturing cultural backgrounds. Overall, our work contributes to the development of LLM-driven simulation systems for exploring UBI policies, demonstrating their effectiveness in generating realistic and flexible economic predictions. By providing an alternative to costly social experiments, our approach advances the methodological toolkit for studying economic and social impacts in a scalable and adaptable manner.

2 Related Works

2.1 LLM Agents

The concept of leveraging LLMs to power agents has gained attraction in recent years, demonstrating their potential across a wide range of applications (Shanahan et al., 2023; Tseng et al., 2024; Zhong et al., 2024). LLM agents has been explored for simulating human-like interactions (Meyer and Elsweiler), enabling advanced conversational agents (Hackenburg et al., 2023), and testing social or behavioral theories (Hao et al., 2024). Prior studies have focused on using LLMs to emulate distinct personas, providing insights into decision-making and collaboration dynamics in controlled environments (Li et al., 2024b). These advancements underline the growing relevance of LLM's role-playing capabilities as a tool for simulating behavioral theories in economic (Li et al., 2024a) and political fields (Cao et al., 2024).

2.2 Universal Basic Income

UBI has been extensively studied as a policy framework aimed at alleviating poverty and improving social well-being (Ghatak and Maniquet, 2019; Bidadanure, 2019). Studies around the world have explored the impact of UBI in various socioeconomic contexts, highlighting diverse outcomes based on regional economic conditions, cultural factors, and policy designs. Table 1 presents five notable cases to illustrate these variations. Real-world trials have provided valuable evidence both supporting (Coalition, 2012; Banerjee et al., 2023) and opposing (Sage and Diamond, 2017; Martinelli, 2017) the implementation of UBI. Large-scale socioeconomic experiments like UBI are expensive and often limited in scope and duration. As such, simulating has become a valuable workaround for testing and evaluating effectiveness and limitations.

2.3 LLM-Based Simulations for Economic System

Rule-based and empirical statistical models have provided foundational insights into economic simulations in the previous decades (Hendry and Richard, 1982; Phelps, 1967; Kydland and Prescott, 1982). With the rise of more sophisticated computational tools, DSGE (Dynamic Stochastic General Equilibrium) models emerged as a solution for modeling a large-scale economy (Christiano et al., 2005, 2018). In recent decades, Agent-Based Modeling (ABM) has emerged as a more promising solution for macroeconomic modeling (Fagiolo et al., 2019; Chen et al., 2012) as it allows for diverse agents to interact based on rule-based (Lengnick, 2013; Gatti et al., 2011), learning-based (Zheng et al., 2022) and recently, LLM-based methods (Li et al., 2024a), enabling the exploration of a wide range of nonlinear behaviors.

Moreover, it is increasingly recognized that non-numerical factors, such as cultural influences (Guiso et al., 2006) and social structures (Granovetter, 2018), also play a significant role in shaping economic outcomes. In this work, we leverage the role-playing capabilities of LLMs to effectively introduce these previously unconsidered elements into simulations with rule-based and numerical modeling.

Table 1: Results of real-world UBI experiments in different locations

Experiment	Depression	Income Level	Locus of Control	Working Hours	Consumption
Finland (Kangas et al., 2019)	Reduced	Increased	Improved	Decreased	-
Kenya (Banerjee et al., 2023)	Reduced	Increased	Decreased	Decreased	Increased
USA Texas Illinois (Bartik et al., 2024)	Reduced	Increased	-	Decreased	Increased
USA California (West et al., 2021)	Reduced	Increased	Improved	Increased	-
Namibia (Coalition, 2012)	Reduced	Increased	-	Increased	Increased

Table 2: Variable Reference Table

Environmental Variables		Agent Variables	
G	Inventory of goods	h	Working hours
S	Supply	p^h	Working propensity
D	Demand	w	Wage
I	Inflation rate	i	Monthly income
R	Interest rate	c	Consumption
W	Strength of social welfare	p^c	Consumption propensity
B	Tax brackets	s	Saving
P	Price for essential goods	t	Taxes to pay
N	Number of labor agents	r	Annual interest rate
E	Market imbalance		

3 Framework

In this section, we introduce the baseline framework, which is a well-established and widely adopted economic simulation framework. The framework includes three types of agents: labor agents, a government agent, and a bank agent (Gatti et al., 2011; Wolf et al., 2013; Li et al., 2024a).

3.1 Labor Agent

Two of the most fundamental elements in the field of macroeconomics are labor supply S and market demand D. Labor agents can determine how much they work and spend during each epoch. Each labor agent i needs to maximize their utility. Utility refers to the property of any object that produces benefit, advantage, pleasure, good, or happiness (Broome, 1991). Working brings negative utility, while consuming goods and receiving income gains positive utility.

Agents can maximize their utility by deciding their Working propensity p_i^h and Consumption propensity p_i^c. p_i^h represents their willingness to work, while p_i^c is the proportion of their total wealth s_i that they are willing to spend.

In previous works, working propensity is used to determine whether an agent will work $l_i \sim \text{Bernoulli}(p_i^h)$ during a given month (Lengnick, 2013; Li et al., 2024a). For agents who decide to work($l_i = 1$), they all work a fixed number of hours h(typically, prior studies assume $h = 168$ (Li et al., 2024a; Zheng et al., 2022)).

However, in reality, the behavior of labor agents is far more complex than the framework presented here. The baseline framework fails to capture several key aspects of labor agent decision-making:

Changes in working incentives: When new policies are implemented or an agent's financial situation changes, the framework fails to capture subtle shifts in working incentives. This lack of sensitivity to small changes undermines the model's ability to accurately track how agents adjust their behavior in response to economic or policy shifts.

Unrealistic decision-making: This framework simplifies agent decisions by assuming binary choices—whether to work or not, with agents either working a full month or skipping it entirely. In reality, workers typically adjust their work hours incrementally based on utilities. The framework does not account for the more nuanced, gradual decisions agents make about work in response to changing circumstances.

Heterogeneity: This framework assumes a uniform working hours, but in reality, agents are heterogeneous. Some agents may have jobs that require longer working hours, while others may work fewer hours. This diversity in work preferences and constraints is not captured in this simplified framework.

3.2 Government Agent

The role of a government in economic simulations is to manage taxation and provide social welfare. One common model for tax collection is to apply a progressive tax policy to agents' post-tax income w_i:

$$t(w_i) = \sum_{k=1}^{B} \tau_k((b_{k+1} - b_k)\mathbf{1}[w_i > b_{k+1}]$$
$$+(w_i - b_k)\mathbf{1}[b_k < w_i \leq b_{k+1}]),$$

where b_k is the k-th bracket in the bracket set B, τ_k is b_k's the corresponding tax rate and $\mathbf{1}$ is the indicator function.

Tax revenue is then used for providing social welfare. It is common to convert social welfare into the utility of agents by redistributing the tax revenue to the agents, either evenly (Zheng et al., 2022) or unevenly (Aaberge et al., 2003).

In reality, however, taxes collected are not simply redistributed back to taxpayers. This practice is a simplified abstraction used to quantify the utility derived from social welfare. This presents a challenge when simulating UBI policies, as redistributing taxes and providing a monthly stipend essentially duplicate the financial flows, diluting the effectiveness of UBI policies. Therefore, individuals' utility gained from UBI stipend and that from social welfare received must be considered separately.

3.3 Bank Agent

Savings of the labor agents are stored in banks, which provide interest according to the annual interest rate R. Agent i's savings in the bank are updated as follows:

$$s_i \longleftarrow s_i \times (1 + R),$$

Annual interest rate R is updated using the Taylor rule (Wolf et al., 2013; Dawid and Gatti, 2018) every January in the simulation setting:

$$r_y = \max(R^n + \pi^t + \alpha^\pi \times (\pi - \pi^t), 0),$$

Natural interest rate R^n, target inflation rate π^t and inflation adaptation rate α^π are adjustable hyper-parameters. During the simulation, the annual interest rate is controlled by the inflation rate, where the inflation rate π for year y is the annual change in the price of essential goods:

$$\pi = \frac{P_y - P_{y-1}}{P_{y-1}},$$

3.4 Market Demand and Supply

In this framework, agents produce and consume essential goods, forming market supply S and demand D, respectively. The market keeps an inventory of essential goods G, which gets updated every month according to the following equation:

$$\begin{aligned} G &\longleftarrow G - D + S \\ &\longleftarrow G - \sum_i^N \frac{p_i^c s_j}{P} + \sum_i^N l_i h, \end{aligned}$$

When the inventory of essential goods G does not match its demand D, an imbalance E occurs:

$$E = \frac{D - G}{\max(D, G)},$$

Market aims to reduce the imbalance E between the supply and demand for essential goods: When there is a shortage of essential goods ($E > 0$), labor agents' wages are increased to stimulate production. As labor costs rise, the prices of essential goods increase to ensure a certain profit margin (Lengnick, 2013; Dawid and Gatti, 2018; Wolf et al., 2013). The wages of agents and the essential good price are then adjusted as follows:

$$\begin{aligned} w_i &\longleftarrow w_i(1 + E_i), E_i \sim sign(E)U(0, \alpha_w|E|), \\ P &\longleftarrow P(1 + E_p), E_p \sim sign(E)U(0, \alpha_P|E|), \end{aligned}$$

α_P and α_w are hyper-parameters introduced to control the maximum rate of changes in the price of essential goods and labor agents' wages, respectively.

4 LLM-Based Labor Agent Design

Labor agent models have traditionally employed various decision-making frameworks, including rule-based systems (Lengnick, 2013; Gatti et al., 2011), reinforcement learning (RL) (Zheng et al., 2022), and LLMs (Li et al., 2024a). While LLMs have been used to simulate labor market agents, our work introduces two novel contributions: **the incorporation of demographic profiles to guide agent behavior and the refinement of economic mechanisms that have been oversimplified or omitted in previous models.** These innovations allow us to simulate labor agents that make more realistic, context-sensitive decisions, based on both their given demographics and broader economic conditions. We modify the following prompts from EconAgent (Li et al., 2024a), which is released under the MIT License. Examples of prompts mentioned below are attached in appendix.

4.1 Role-playing

One key innovation in our model is the integration of demographic attributes to guide the role-playing behavior of labor agents. These demographic factors, including age, city of residence, language, and financial status, are drawn from real-world distributions. This demographic-based role-playing is crucial for simulating realistic labor market behavior, as it allows us to observe how agents with different backgrounds and circumstances make economic decisions.

The Role-playing Prompt (see A.1.1) is used to simulate this demographic context by providing agents with a set of characteristics that influence their choices.

4.2 Utility Considerations

Another key modification to traditional labor agent models is how we incorporate agent utility. In classical models, decision-making is often driven by material utility (Zheng et al., 2022; Aaberge et al., 2003), such as income and consumption. In our model, utility is expanded to include not only financial factors but also leisure (time spent away from work) and the value of social welfare received (benefits provided by taxes and government programs).The value of social welfare is often treated as a direct transfer of funds to the agent's finances, effectively adding to their savings (Zheng et al., 2022; Aaberge et al., 2003). This simplification overlooks the real-world dynamics, where taxes fund public goods and social welfare (such as healthcare and education), but do not directly add to an agent's personal wealth.

The Memory Prompt (see A.1.2) provides environmental information and the agent's individual status, both of which are essential for determining their working and consumption propensities. The environmental information includes factors like the agent's expected salary, tax rates, inflation, and the cost of essential goods, representing the broader economic conditions in which the agent operates. Meanwhile, the agent's individual status consists of personal data such as their previous month's work, consumption, savings, and tax deductions.

The Utility Prompt (see A.1.3) specifies what utility should be considered by the agent when making decisions. The prompt explains that the agent's utility is determined by income, savings, consumption, savings, social welfare received, and leisure time.

4.3 Work Intensity

In contrast to traditional labor agent models, which typically involve a binary decision of whether or not to work (Lengnick, 2013; Li et al., 2024a; Zheng et al., 2022), we replace this approach by introducing work intensity. Rather than simply deciding whether to work, agents now decide how much to work, i.e., the number of hours they wish to allocate to work within the maximum working hours h.

This change is particularly important for studying the impact of policies like UBI. By allowing agents to vary their work intensity, we can observe the microscopic impacts of UBI policies, such as patterns of change in working intensity. This provides insights that would otherwise be obscured if agents were limited to a binary decision of whether to work or not.

The Task Prompt (see A.1.4) is structured to ask agents to make decisions about work intensity. Agents are prompted to determine how much of their available time (168 hours per month) they wish to allocate to work, depending on their information.

4.4 Self-Reflection

In every simulation epoch, agents are tasked with making two critical decisions: how much to spend on consumption and how many hours to allocate to work. In every three epochs, the agent reflects on their decisions and the economic environment with reflection prompt (see A.1.5) to analyze their past actions and the economic conditions they have been operating within. Each time the LLM makes a decision, the history of the last three decisions (last quarter) and the last reflection is provided to the agent, capturing the short-term history behavior and long-term goals, respectively.

4.5 UBI

When UBI policy is deployed to the economy, an amount of stipend will be added to agents' savings. Also, UBI Description Prompt (see A.1.6) will be added before the Task Prompt to aware agents that they are currently experiencing UBI policy.

5 Experiments

5.1 Experimental Settings

We set the hyper-parameters mentioned in section 3 as below:

Table 3: Experiment Settings Reference Table

Hyper-parameters	Description	Value
R^n	Natural Inflation Rate	0.01
π_t	Target Inflation Rate	0.02
α_P	Maximum Price Change Rate	0.5
α_w	Maximum Wage Change Rate	0.5
α^π	Inflation Adaptation Rate	0.1

For the following experiments, we simulate a society with $N = 50$ agents. The tax brackets B are set according to the real-world taxation policies of the respective regions at the time the UBI experiments began. Following prior practice, we initialize the wages of agents using a Pareto distribution, a widely used method for modeling wage

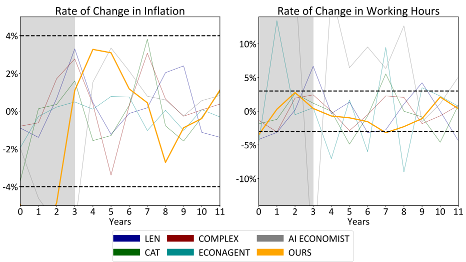


Figure 1: Rate of change on inflation and working hours in yearly scale (shaded regions are initializing steps, horizontal dotted lines are real-world degree of fluctuations)

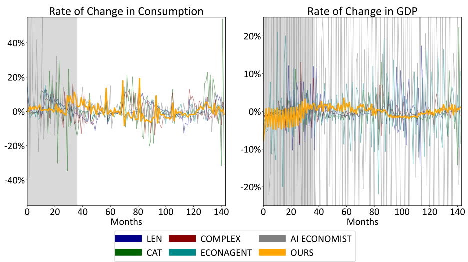


Figure 2: Rate of change on Consumption and GDP in monthly scale (Shaded regions are initializing steps)

distribution in society (Zheng et al., 2022). The shape parameter of the Pareto distribution is set to 8, while the scale parameter is calibrated by matching the average wage and per capita GDP of the corresponding region. We prompt LLM to generate job titles for agents that correspond to the wage intervals. Agent's name and city of residence are generated using python library *Faker* (curella.org, 2024), with respect to the locale of the regions. The age distribution of agents is initialized based on the real-world age distribution of the year the corresponding UBI experiment started. We use GPT-3.5-Turbo from OpenAI as the LLM for our simulations (approximately \$15 for a 50 agents, 200 epoch simulation).

In this section, we investigate the following research questions through experiments: 1.How do our simulation method behave in the simulation environment, compared with other macroeconomic simulation methods? 2.Can our simulation method replicate the economic and mental metrics in 5 real-world UBI trials? 3.Why can we replicate such results?

5.2 RQ1: Baseline Comparison

To examine the effectiveness of our proposed method, we compare our new baseline with methods adopting the framework mentioned in section 3, including three rule-based methods (LEN (Lengnick, 2013), CATS (Gatti et al., 2011), and their combination COMPLEX), one RL-based method (AI-Economist (Zheng et al., 2022)), and one LLM-based method (EconAgent (Li et al., 2024a)). We run a non-UBI economic simulation using the parameters and settings from the Finland 2018 UBI experiment (Kangas et al., 2019), in which the last epoch represents the month before UBI started (December 2017).

Figure 1 shows the yearly inflation change rate and yearly working hours change rate. Our method obtains the most realistic fluctuation (maximum magnitude below 4%) among the baselines. For reference, in the case of Finland (from 2006 to 2017), the annual inflation rates and the annual change in average working hours fluctuate within a range of maximum $\pm 4\%$ (Publicly available in *Tilastokeskus* (*noa*)) and $\pm 3\%$ (Publicly available in *Statista* (*Statista, 2024*)), respectively. The results generated by our method align most closely with these real-world ranges, demonstrating its ability to capture the nuanced dynamics of economic fluctuations with comparable fidelity to other prior approach.

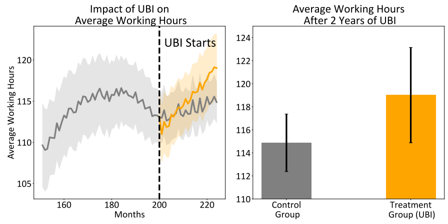
Figure 2 shows the monthly change rate of Consumption and GDP. Consumption and production are relatively stable economic indicators in the real world, meaning that monthly fluctuations should not be excessively volatile. Compared to the other baselines, our method produces the most stable fluctuations in consumption and production, indicating a simulation that more closely mirrors real-world economic behavior.

5.3 RQ2: Replication of real-world UBI results

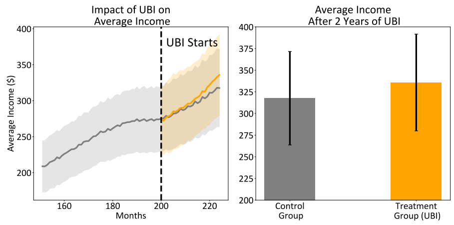
To assess the impact of UBI, we conduct a series of simulations comparing a control group and a treatment group. The experimental setting is structured as follows: First, the simulator is run without any UBI policy for 200 months to allow the system to reach a state of relative stability. Then the simulation continues to run for 2 years with and without UBI policy following the 200 month checkpoint. We examine the impact of UBI on three economic metrics (average working hour, working income, and consumption) and two mental health metrics (depression and locus of control). Results are shown in figure S1 and figure 4.

5.3.1 Impact on Economy

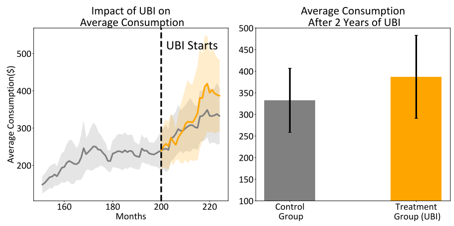
Figure 3 shows an example on the impact of UBI policy to economic metrics. The left-side graphs



(a) Working hours



(b) Income change



(c) Consumption change

Figure 3: Impact of UBI on economic metrics in Namibia settings

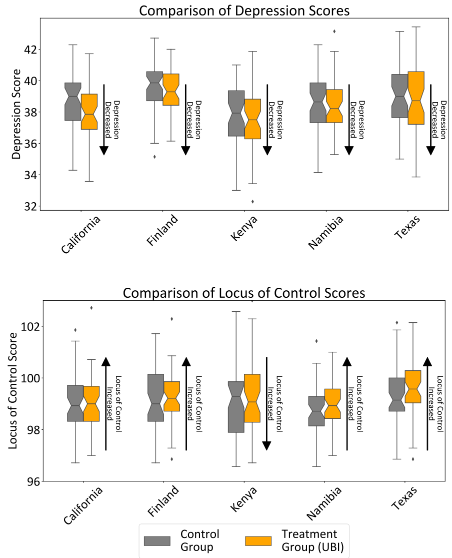


Figure 4: Replication of real-world mental health results

reveal the temporal evolution of the metrics time in the simulation of Namibia case. Initially, the patterns for the control group achieve a relative stability wave-like oscillation. This reflects a stable economic environment where agents' behaviors settle over time. When the UBI policy is introduced, the evolution patterns for the metrics changes significantly. It begin to deviate from their previously stable trajectories, showing an increase in consumption and a decrease in average working hour at the beginning epochs. This shows that agents are aware of the policy and start doing different decisions in response to the change of environment.

5.3.2 Impact on Mental Health

Prior studies demonstrates the effectiveness to measure agents' mental inclinations by asking LLMs to complete questionnaires (Gilson et al., 2022; de Winter, 2023). With agents powered by LLMs, we are able to observe how agents respond to mental health assessments to see if they can reflect realistic decision-making processes.

To measure depression, we apply the widely recognized Center for Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1991), a common tool for assessing depressive symptoms in both clinical and general populations. To evaluate locus of control, we employ the Multidimensional Health Locus of Control (MHLC) Form A and B (Wallston et al., 1978), which assess individuals' beliefs about the control they have over their health and life outcomes. We prompt agents (see section A.1) to fill in the two questionnaires every three months, after the self-reflection prompt is called. These two metrics provide valuable insight into the psychological well-being of agents, allowing the examination on the potential psychological impacts of UBI.

The comparison of the distribution of the two metrics is shown in Figure 4. The box plot displays the distribution of the average scores of agents during the two years following the introduction of the UBI policy. For depression, all treatment groups show a decrease in depression scores compared to their respective control groups.In contrast, while the real-world experiment in Kenya reported a slight decrease in locus of control, our results partly align with this finding. Specifically, although both the upper and lower quartiles of agents show an increase in their locus of control scores, the median score, along with the 95% confidence interval, actually decreased. This suggests that while some agents experience a stronger sense of control following the implementation of UBI, a larger proportion may feel less in control.

Our experimental results matches the common and contradictory findings reported in real-world UBI experiments (compare Table 1 and Figure S1). The simulation results across the five experimental settings align with the outcomes of their respective

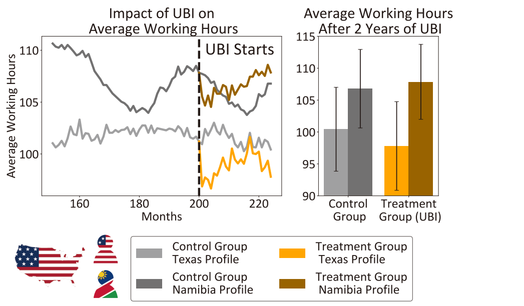


Figure 5: Simulation results by swapping Texas profiles with Namibia profiles

real-world UBI trials, demonstrating the robustness and reliability of our method in capturing the nuanced dynamics of UBI experiments.

5.4 RQ3: Importance of Role-play

To evaluate the validity of agent responses under different experimental settings, we conducted two experiments to verify the role of profiles and the amount stipend in this simulation method.

5.4.1 Role of Agent Profiles

To investigate the effect of agent profiles on the simulation outcomes, we conduct simulations with settings from the Texas case, yet replace the agent profiles with those representing Namibia. All other aspects of the simulation, remain unchanged, isolating the influence of agent demographics on the overall results.

Figure 5 shows the results of this ablation study. It demonstrates that the working hours for the Namibia profile do not follow the same pattern as the original Texas profile. After the UBI intervention, both groups experience a reduction in working hours at first. Namibia agents' working hours rebound soon after the intervention, mirroring the pattern observed in the real-world Namibia UBI experiment and our simulation results(see Figure S1). In contrast, the Texas agents (with Namibia profiles) continue to show a steady decline in working hours.

This experiment shows that by incorporating demographics on LLM-Based simulations, agents can make decisions that reflect their individual backgrounds, thereby producing more realistic and insightful simulation outcomes.

5.4.2 Effect of Stipend

This experiment investigates how varying UBI stipends affect simulation results. In this study, we use agent profiles based on the California locale and assigning different UBI stipend levels from

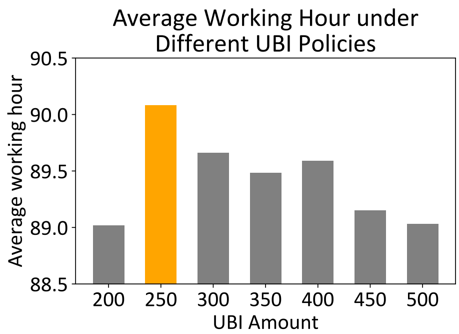


Figure 6: Average working hours when different UBI policies are applied to California simulation setting.

$200 to $500, and investigate the impact on agents' working propensity.

Figure 6 illustrates the average working hours under different UBI stipend levels. Our goal is to determine the optimal UBI stipend amount that maximizes agents' willingness to work while ensuring a balance between financial support and labor participation. The results indicate that at the lowest stipend level ($200), working hours remain relatively low. As the stipend increases to $250, we observe a local maximum in working hours, suggesting that moderate financial support may encourage greater workforce participation. However, beyond this point, working hours decline as stipends increase, potentially reflecting reduced economic incentives to work.

These findings highlight the complex relationship between UBI and labor participation, providing valuable insights for policymakers. By identifying an optimal stipend that sustains both financial security and workforce engagement, our method offers a data-driven approach to obtain insights for UBI policy decisions.

6 Conclusion

Overall, we introduce an innovative approach that integrates LLM-Based agent role-playing capabilities to simulate economies under UBI policies. By modifying agents' utility considerations and incorporating LLMs into classic macroeconomic frameworks, we develop a more stable and numerically accurate simulation framework. Our method effectively reproduces findings from real-world UBI trials across diverse regions, accurately predicting impacts on both economic outcomes and mental health. This work provides policymakers with a cost-effective tool for evaluating the potential impacts and limitations of UBI policies, minimizing the need for costly trial implementations.

References

Consumer price index. https://stat.fi/en/statistics/khi. Accessed: 2024-12-6.

Rolf Aaberge, Ugo Colombino, John E Roemer, et al. 2003. Optimal taxation according to equality of opportunity: a microeconometric simulation analysis. Technical report, ICER-International Centre for Economic Research.

Aris Anagnostopoulos, Carlos Castillo, Adriano Fazzone, Stefano Leonardi, and Evimaria Terzi. 2018. Algorithms for hiring and outsourcing in the online labor market. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1109–1118.

Abhijit Banerjee, Michael Faye, Alan Krueger, Paul Niehaus, and Tavneet Suri. 2020. Effects of a universal basic income during the pandemic. *Innovations for Poverty Action Working Paper*.

Abhijit Banerjee, Michael Faye, Alan Krueger, Paul Niehaus, and Tavneet Suri. 2023. Universal basic income: Short-term results from a long-term experiment in kenya.

Abhijit Banerjee, Paul Niehaus, and Tavneet Suri. 2019. Universal basic income in the developing world. *Annual Review of Economics*, 11(1):959–983.

Alexander W Bartik, Elizabeth Rhodes, David E Broockman, Patrick K Krause, Sarah Miller, and Eva Vivalt. 2024. The impact of unconditional cash transfers on consumption and household balance sheets: experimental evidence from two us states. Technical report, National Bureau of Economic Research.

Juliana Uhuru Bidadanure. 2019. The political theory of universal basic income. *Annual Review of Political Science*, 22(1):481–501.

John Broome. 1991. "utility". *Economics and Philosophy*, 7(1):1–12.

Charles Cao, Jie Zhuang, and Qiang He. 2024. Llm-assisted modeling and simulations for public sector decision-making: Bridging climate data and policy insights. In *AAAI-2024 Workshop on Public Sector LLMs: Algorithmic and Sociotechnical Design*.

Hancheng Cao, Chia-Jung Lee, Shamsi Iqbal, Mary Czerwinski, Priscilla NY Wong, Sean Rintel, Brent Hecht, Jaime Teevan, and Longqi Yang. 2021. Large scale analysis of multitasking behavior during remote meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

Shu-Heng Chen, Chia-Ling Chang, and Ye-Rong Du. 2012. Agent-based economic models and econometrics. *The Knowledge Engineering Review*, 27(2):187–219.

Zhilong Chen, Xiaochong Lan, Jinghua Piao, Yunke Zhang, and Yong Li. 2022. A mixed-methods analysis of the algorithm-mediated labor of online food deliverers in china. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–24.

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.

Lawrence J Christiano, Martin S Eichenbaum, and Mathias Trabandt. 2018. On dsge models. *Journal of Economic Perspectives*, 32(3):113–140.

Basic Income Grant Coalition. 2012. Basic income grant pilot project assessment report, april 2009.

joke2k curella.org. 2024. Faker. Version 33.1.0.

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

Valerio De Stefano. 2019. " negotiating the algorithm": Automation, artificial intelligence, and labor protection. *Comp. Lab. L. & Pol'y J.*, 41:15.

Joost CF de Winter. 2023. Can chatgpt pass high school exams on english language comprehension? *International Journal of Artificial Intelligence in Education*, pages 1–16.

Jurgen De Wispelaere and Lindsay Stirton. 2004. The many faces of universal basic income. *The Political Quarterly*, 75(3):266–274.

Giorgio Fagiolo, Mattia Guerini, Francesco Lamperti, Alessio Moneta, and Andrea Roventini. 2019. Validation of agent-based models in economics and finance. *Computer simulation validation: fundamental concepts, methodological frameworks, and philosophical perspectives*, pages 763–787.

Marylène Gagné, Sharon K Parker, Mark A Griffin, Patrick D Dunlop, Caroline Knight, Florian E Klonek, and Xavier Parent-Rocheleau. 2022. Understanding and shaping the future of work with self-determination theory. *Nature Reviews Psychology*, 1(7):378–392.

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

Maitreesh Ghatak and François Maniquet. 2019. Universal basic income: some theoretical aspects. *Annual Review of Economics*, 11(1):895–928.

Aidan Gilson, Conrad Safranek, Thomas Huang, Vimig Socrates, Ling Chi, R Andrew Taylor, and David Chartash. 2022. How does chatgpt perform on the medical licensing exams? the implications of large language models for medical education and knowledge assessment. *MedRxiv*, pages 2022–12.

Mark Granovetter. 2018. The impact of social structure on economic outcomes. In *The sociology of economic life*, pages 46–61. Routledge.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. 2006. Does culture affect economic outcomes? *Journal of Economic perspectives*, 20(2):23–48.

Taicheng Guo, Xiuying Chen, Yaqi Wang, Ruidi Chang, Shichao Pei, Nitesh V. Chawla, Olaf Wiest, and Xiangliang Zhang. 2024. Large language model based multi-agents: A survey of progress and challenges. *Preprint*, arXiv:2402.01680.

Louise Haagh. 2019. *The case for universal basic income*. John Wiley & Sons.

Kobi Hackenburg, Lujain Ibrahim, Ben M Tappin, and Manos Tsakiris. 2023. Comparing the persuasiveness of role-playing large language models and human experts on polarized us political issues. *OSF Preprints*, 10.

Guozhi Hao, Jun Wu, Qianqian Pan, and Rosario Morello. 2024. Quantifying the uncertainty of llm hallucination spreading in complex adaptive social networks. *Scientific reports*, 14(1):16375.

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

Basileal Imana, Aleksandra Korolova, and John Heidemann. 2021. Auditing for discrimination in algorithms delivering job ads. In *Proceedings of the web conference 2021*, pages 3767–3778.

Olli Kangas, Signe Jauhiainen, Miska Simanainen, and Minna Ylikännö. 2019. The basic income experiment 2017–2018 in finland: Preliminary results.

Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1301–1318.

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

Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pages 1603–1612.

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

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024a. Econagent: large language model-empowered agents for simulating macroeconomic activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15523–15536.

Xuchuan Li, Fei Huang, Jianrong Lv, Zhixiong Xiao, Guolong Li, and Yang Yue. 2024b. Be more real: Travel diary generation using llm agents and individual profiles. *arXiv preprint arXiv:2407.18932*.

Luke Martinelli. 2017. Assessing the case for a universal basic income in the uk. *IPR Policy Brief*.

Selina Meyer and David Elsweiler. Llm-based conversational agents for behaviour change support: A randomized controlled trial examining efficacy, safety, and the role of user behaviour. *Safety, and the Role of User Behaviour*.

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

Lenore Sawyer Radloff. 1991. The use of the center for epidemiologic studies depression scale in adolescents and young adults. *Journal of youth and adolescence*, 20(2):149–166.

Daniel Sage and Patrick Diamond. 2017. Europe's new social reality: The case against universal basic income.

Murray Shanahan, Kyle McDonell, and Laria Reynolds. 2023. Role play with large language models. *Nature*, 623(7987):493–498.

Yang Shen and Xiuwu Zhang. 2024. The impact of artificial intelligence on employment: the role of virtual agglomeration. *Humanities and Social Sciences Communications*, 11(1):1–14.

Statista. 2024. Average weekly hours worked on the main job in finland from 2013 to 2023. Accessed: 2024-12-06.

Feichin Ted Tschang and Esteve Almirall. 2021. Artificial intelligence as augmenting automation: Implications for employment. *Academy of Management Perspectives*, 35(4):642–659.

Yu-Min Tseng, Yu-Chao Huang, Teng-Yun Hsiao, Yu-Ching Hsu, Jia-Yin Foo, Chao-Wei Huang, and Yun-Nung Chen. 2024. Two tales of persona in llms: A survey of role-playing and personalization. *arXiv preprint arXiv:2406.01171*.

Kenneth A Wallston, Barbara Strudler Wallston, and Robert DeVellis. 1978. Development of the multidimensional health locus of control (mhlc) scales. *Health education monographs*, 6(1):160–170.

Stacia West, Amy Castro Baker, Sukhi Samra, and Erin Coltrera. 2021. Preliminary analysis: Seed's first year. *Stockton, CA: Stockton Economic Empowerment Demonstration*, 500.

Christine Wolf and Jeanette Blomberg. 2019. Evaluating the promise of human-algorithm collaborations in everyday work practices. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–23.

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.

Ali Zarifhonarvar. 2024. Economics of chatgpt: A labor market view on the occupational impact of artificial intelligence. *Journal of Electronic Business & Digital Economics*, 3(2):100–116.

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.

Yilin Zhong, Jiahao Xie, Jie Wang, Bo Fan, Zhaoer Fang, and Banghuang Peng. 2024. Improving large language models in multi-party conversations through role-playing. In *International Conference on Intelligent Computing*, pages 209–220. Springer.

7 Limitations

While our study provides valuable insights into the effects of UBI on labor participation, it has several limitations that should be addressed in future work. These limitations primarily stem from constraints in demographic representation, economic modeling, and experimental scalability.

7.1 Limited Demographic Representation

One of the primary limitations of our simulation is the restricted demographic diversity. Currently, agent profiles include only basic attributes such as age, gender, location, and names. However, real-world economic behavior is influenced by a wider range of factors, including education level, occupation, household structure, socioeconomic status, and cultural background. The absence of these attributes limits the depth of our analysis and may overlook important variations in behavioral responses to UBI.

Incorporating a richer demographic dataset could enable more accurate modeling of diverse population groups and improve the generalizability of our findings. However, obtaining and integrating such data remains a challenge due to privacy concerns and data availability. Future research could explore synthetic data generation techniques or leverage large-scale socioeconomic datasets to enhance agent diversity.

7.2 Simplified Economic Modeling

Although our simulation is built upon a widely adopted economic framework, it remains a simplification of real-world economies. Macroeconomic variables such as inflation, taxation policies, wage dynamics, and business cycles are not explicitly included in our model. These factors play a crucial role in shaping economic outcomes and could significantly influence the impact of UBI on labor participation.

While we attempt to address some of these oversimplifications using LLMs, the current approach is still limited in its ability to capture complex economic interactions. In reality, government policies, labor market fluctuations, and social welfare programs interact in non-trivial ways that are not fully represented in our model. Future work should explore integrating more advanced economic models, such as agent-based simulations with endogenous market dynamics, to improve realism.

Additionally, behavioral economic factors, such as changes in motivation, risk perception, and long-term career planning, are not explicitly accounted for. These psychological aspects may significantly affect how individuals respond to UBI and should be incorporated into future iterations of the simulation.

7.3 Scalability Constraints

Our experiments are currently constrained by computational resources, limiting both the number of agents and the duration of the simulations. While our findings provide valuable insights, larger-scale experiments with more agents and extended time horizons could yield more robust and generalizable conclusions.

Expanding the simulation to accommodate larger populations would allow for the study of emergent behaviors that arise in complex social and economic systems. Additionally, longer simulation runs could help assess the long-term effects of UBI, including potential shifts in workforce composition, career trajectories, and economic stability.

To overcome these constraints, future research could focus on optimizing computational efficiency, parallelizing simulations, or leveraging cloud-based infrastructure for large-scale experimentation. Developing more efficient algorithms for agent interactions and economic modeling could also enhance scalability without compromising accuracy.

7.4 Potential Risks

While our study provides valuable insights into UBI's effects on labor participation, it also carries certain risks that should be acknowledged. These risks primarily stem from potential policy misinterpretation, biases in simulated agents, and computational constraints.

7.4.1 Risk of Policy Misinterpretation

Our simulation is a controlled, simplified model of economic behavior and should not be directly translated into real-world policy decisions without further validation. Policymakers or stakeholders may misinterpret the results as direct evidence for implementing specific UBI policies, despite the absence of real-world complexities such as inflation, taxation, and long-term economic feedback loops. To mitigate this risk, we emphasize that our findings should serve as a **theoretical exploration** rather than a prescriptive policy guideline.

7.4.2 Bias in Simulated Agent Behaviors

Since our simulation leverages LLMs to model agent decision-making, it may inherit biases from the underlying training data. These biases could lead to agent behaviors that do not fully align with real-world economic and social dynamics. For instance, LLM-generated decisions might overestimate or underestimate individuals' responses to financial incentives, leading to skewed results. Future research should incorporate more controlled behavioral modeling and real-world data validation to reduce potential biases.

7.4.3 Computational and Data Limitations

The accuracy of our simulation is constrained by computational resources and the available demographic and economic data. Limited agent diversity and simplified economic modeling may introduce inaccuracies in predicting large-scale, long-term UBI effects. While increasing simulation scale and data granularity could improve reliability, such expansions require significant computational power and access to richer datasets.

7.5 Future Directions

Despite these limitations, our study provides a foundational framework for investigating UBI in a simulated environment. By addressing these challenges in the future, we can refine our simulation methodology and provide more actionable insights for policymakers evaluating the potential impacts of UBI on labor markets and economic well-being.

A Supplementary Materials

A.0.1 All result trends

Figure S1 shows the evolution trends of average income, average working hours and average consumption in all 5 cases mentioned in Table 1

A.1 Prompt Used

In the content page, we show the agent decision prompt which is the major part of the framework. Due to page limits, our detailed prompts for generating job titles, questionnaires, and additional prompts used outside the simulation framework are attached with the code available: https://anonymous.4open.science/r/LLM-UBI-7837.

A.1.1 Role-playing Prompt

Role-playing Prompt
You are Ahti Leppänen, a 66-year-old individual living in Pihtipudas.

A.1.2 Memory Prompt

Memory Prompt

In the previous month, you worked as a(an) Receptionist. If you continue working this month, your expected salary will be $21.00 per hour, which is increased compared to the last month due to the inflation of labor market. Besides, your consumption was $2271.02. Your tax deduction amounted to $213.82. The government uses the taxes recieved to provide social welfare. The strength of the social service provided last month was $284.26 per capita. In this month, according to the optimal taxation theory, Saez Tax, the brackets are not changed: [0.00, 808.33, 3289.58, 7016.67, 13393.75, 17008.33, 42525.00] but the government has updated corresponding rates: [12.64%, 19.00%, 30.25%, 34.00%, 42.00%, 44.00%]. Income earned within each bracket is taxed only at that bracket's rate. Meanwhile, inflation has led to a price increase in the consumption market, with the average price of essential goods now at $15.12. Your current savings account balance is $6813.06. Interest rates, as set by your bank, stand at 3.02%.

A.1.3 Utility Prompt

Utility Prompt

Your goal is to maximize your utility by deciding how much to work and how much to consume. Your utility is determined by your consumption, income, savings, social welfare received, and leisure time.

A.1.4 Task Prompt

Task Prompt

With all these factors in play, and considering aspects like your living costs, any future aspirations, and the broader economic trends, what proportion of your spare time (a total of 168 hours a month) would you spend working? Furthermore, how would you plan your expenditures on essential goods, keeping in mind good price?

A.1.5 Reflection Prompt

Reflection Prompt

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?

A.1.6 UBI Description Prompt

UBI Description Prompt

You are experiencing a universal basic income policy, receiving a monthly stipend of $604

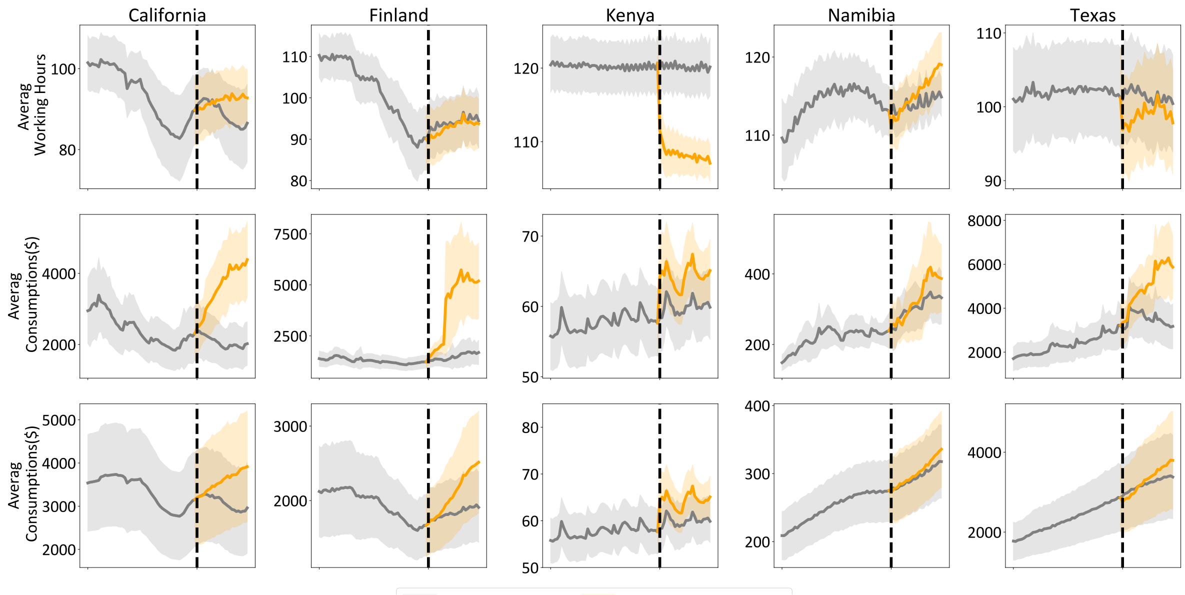


Figure S1: Impact of UBI on economic metrics (Income, Working Hours, and Consumption) in all five settings.