Discussion Paper No. 1260

ISSN (Print) 0473-453X
ISSN (Online) 2435-0982

WEALTH PREFERENCES AND WEALTH INEQUALITY: EXPERIMENTAL EVIDENCE

Nobuyuki Hanaki
Yuta Shimodaira

Secondly Revised February 2026
Revised February 2025
October 2024

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Wealth preferences and wealth inequality: Experimental evidence*

Nobuyuki Hanaki† Yuta Shimodaira‡

January 31, 2026

Abstract

Some researchers claim that a preference for wealth accumulation is the main cause of the long-run stagnation of the Japanese economy. A theoretical implication of people having such a preference, particularly the assumption that the marginal utility of wealth accumulation has a positive lower bound while that of consumption does not, is a widening of wealth inequality. We experimentally test this theoretical prediction by inducing a wealth preference in the laboratory. We find partial support for this prediction: wealth inequality widens when initial inequality is large, but not when it is small. This is because high-wealth participants tend to overconsume more than lower-wealth participants, partly offsetting the effect of the induced preference for wealth accumulation on the widening of wealth inequality. Activating participants' status concerns by displaying their ranking in accumulated wealth has only a limited impact on the expansion of wealth inequality.

JEL Code: C91, E21, E71

Keywords: Wealth preferences, Status concerns, Consumption–saving decision

*The experiment reported in this paper was approved by the IRB of ISER, the University of Osaka (#20231101, #20240101, and #20251002), and was pre-registered at aspredicted.org (#176252, `https://aspredicted.org/kwy2-4nfm.pdf`; #181286, `https://aspredicted.org/knjy-y2rz.pdf`; and #255854, `https://aspredicted.org/25cb6y.pdf`). We gratefully acknowledge the financial support from grants-in-aid for scientific research (KAKENHI, grant numbers: 20H05631, 23H00055, 25H00388, and 25K16600) from the Japan Society for the Promotion of Science (JSPS) and from the Joint Usage/Research Center, ISER, the University of Osaka. We thank Yoshiyasu Ono, John Duffy, and the seminar participants at the University of Osaka for their helpful comments.

†Institute of Social and Economic Research, the University of Osaka, and University of Limassol. Email: nobuyuki.hanaki@iser.osaka-u.ac.jp

‡Setsunan University. Email: yuta.shimodaira@setsunan.ac.jp

1 Introduction

The Japanese economy has experienced stagnation over the past 30 years, a period often referred to as the "Lost 30 Years." Various policies aimed at stimulating the economy, including quantitative easing led by the Bank of Japan, have been ineffective in boosting aggregate demand (see, for example, Ugai, 2007, for a survey of empirical studies on the effects of quantitative easing).

Some researchers, such as Ono (1994, 2001), Ono and Ishida (2014), Ono (2015), and Michau (2018), attribute people's preference for holding wealth as a main cause of this long-run stagnation.[1] They argue that although the marginal utility of consumption declines with the amount consumed (as is usually assumed in economic analyses), the marginal utility of holding wealth declines much more slowly and has a strictly positive lower bound. As a result, once consumption reaches a certain level, the marginal utility of wealth accumulation becomes greater than the marginal utility of consumption. Accordingly, any additional income people receive tends to be allocated to wealth accumulation rather than consumption. Using data gathered through a survey of a representative Japanese sample, Akesaka et al. (2024) present results consistent with this assumption.

One implication of the existence of such a preference for wealth is an increase

[1] Other recent studies incorporate preferences for holding wealth. Michaillat and Saez (2021, 2022) introduce a preference for relative wealth as a proxy for status preference. Michaillat and Saez (2021) do so to resolve anomalies of the New Keynesian model under the zero lower bound. More specifically, the anomalies predicted by the New Keynesian model under the zero lower bound—namely, that "output and inflation collapse to implausibly low levels and that government spending and forward guidance have implausibly large effects" (p. 197)—are resolved when people have a strong enough preference for wealth. Michaillat and Saez (2022) introduce such preferences to develop a simple model of the business cycle that incorporates matching frictions between consumers and producers. Michau et al. (2023b) show the existence of rational bubbles in a frictionless economy with an infinitely lived representative household when that household has a preference for holding wealth. Hashimoto et al. (2023) show that underemployment without high unemployment, as has been observed in Japan in recent years, can arise under secular stagnation in a model with search and matching.

in wealth inequality (see Ono 1994, Ch. 10 and Michau et al. 2023a). To see this, consider two types of agents with low and high initial wealth holdings. Assume also that their preferences for consumption and wealth, as well as their labor incomes, are identical. Further assume that the income flow, including interest earnings from wealth, is such that it allows the low-initial-wealth agent to consume exactly up to the point where the marginal utility of consumption equals the marginal utility of holding wealth. In such a case, the high-initial-wealth agent's additional income, arising from higher interest earnings than those of the low-initial-wealth counterpart, will be allocated to wealth accumulation, thereby widening wealth inequality.

In this paper, we experimentally test the hypothesis that wealth inequality widens by explicitly inducing consumption and wealth preferences, consistent with those assumed in the models, within an otherwise standard consumption–saving decision experiment.[2] Thus, our first research question is:

- Does wealth inequality widen, as predicted by the model, when participants are induced to have preferences for wealth in addition to preferences for consumption?

More specifically, in our baseline experiments, participants decide how much to consume and save over 20 periods. They are given an initial amount of wealth and receive, in each period, a constant labor income and interest earnings from the accumulated wealth. Participants are rewarded based on the amount they consume in each period (converted into payments according to a concave function representing decreasing marginal utility of consumption) and on the accumulated amount of savings (converted into payments at a much lower rate than consumption,

[2]Note that there are many reasons for saving and wealth inequality, such as bequests, the transmissions of human capital, entrepreneurship, and medical-expense risk, as concluded in the survey by De Nardi and Fella (2017). In our experiment, apart from the induced bequest motive, all other reasons are explicitly assumed away.

according to a linear function representing the non-decreasing marginal utility of wealth accumulation).

There are two levels of initial wealth within each group of six participants. In each group, half start with low initial wealth and the other half with high initial wealth. Each participant's initial level is determined either endogenously or exogenously. Specifically, in one condition (the Effort condition), the initial wealth level is determined by performance in a real-effort task conducted prior to the consumption–saving experiment, whereas in the other condition (the Random condition), the initial wealth level is assigned randomly. We investigate whether the within-group wealth gap widens as predicted by theory.

The experimental literature on consumption–saving decisions has shown, however, that individuals have difficulty solving dynamic intertemporal optimization problems.[3] While experiments in which participants must save part of their income for later consumption exhibit undersaving (e.g., Carbone, 2006; Brown et al., 2009), the opposite pattern is observed when participants must initially borrow to consume optimally because their income rises in later periods (Meissner, 2016; Ahrens et al., 2022). Gechert and Siebert (2022) attribute a preference for wealth as a reason for the observed deviation (i.e., underconsumption) in their experiment on consumption–saving decisions.

Because our setting, which incorporates an induced preference for wealth, is more complex than previously studied consumption–saving settings, one might expect participants to fail to make optimal decisions and, consequently, outcomes to deviate from theoretical predictions. Thus, it is not ex ante obvious whether widening wealth

[3]See, e.g., Hey and Dardanoni (1988), Carbone and Hey (2004), Noussair and Matheny (2000), Lei and Noussair (2002), Ballinger et al. (2003), Carbone (2006), Brown et al. (2009), Ballinger et al. (2011), Carbone and Duffy (2014), Meissner (2016), Crockett et al. (2019), Ahrens et al. (2022), and Gechert and Siebert (2022). For a general overview, see Duffy (2015).

inequality will emerge in the experiment.

In our baseline experiments, we simply induce a preference for wealth and investigate its consequences as described above. However, it is also valuable to examine potential reasons why people might have a preference for wealth; status concerns are considered one such reason (Michaillat and Saez, 2021, 2022). We therefore test whether activating participants' status concerns strengthens their preference for wealth and further supports the hypothesis of widening wealth inequality. Thus, our second research question is:

- Are the dynamics of wealth inequality observed in the experiment affected by activating participants' status concerns?

To address these research questions, we vary three factors in our $2 \times 2 \times 2$ between-subjects design. The first factor is whether the initial wealth type (H-type or L-type) is determined randomly or on the basis of performance in a real-effort task. The second factor is whether the ranking and amount of accumulated wealth within a participant's group are displayed. The third factor is the size of the gap between participants with low and high levels of initial wealth.

Existing experiments show that activating participants' status concerns through ranking information influences behavior in consumption–saving experiments. Specifically, Feltovich and Ejebu (2014) show that displaying the ranking of accumulated utility from consumption induces participants to consume more than when that ranking is not displayed. Carbone and Duffy (2014) also find that informing participants of the average consumption of others in the same group during the previous period stimulates consumption. We conjecture that displaying the ranking of accumulated wealth within a group has a similar impact on wealth accumulation.

We also conjecture that the size of the initial wealth inequality influences behavior when status concerns are activated. Consider a case in which the initial wealth gap within a group is small. In such a situation, participants with lower initial wealth may be motivated to save more to catch up with others and improve their ranking, even at the cost of reduced monetary rewards from consumption. By contrast, when the initial wealth gap is so large that lower-wealth participants cannot realistically catch up, such behavior should be absent. Thus, we expect the widening of wealth inequality in the presence of ranking information to be less pronounced when initial wealth inequality is small than when it is large.

We find partial support for the theoretical prediction, but no clear support for our conjectures. Wealth inequality widens when initial inequality is large but does not widen when it is small. This pattern holds regardless of whether initial wealth is determined by participants' performance in the real-effort task or assigned randomly, and regardless of whether ranking information on accumulated wealth is displayed.

Initial inequality matters because participants in our experiment tend to overconsume, and this tendency is stronger among those with high initial wealth than among those with low initial wealth. When initial inequality is small, the difference in overconsumption between high- and low-initial-wealth participants is as large as the difference in their interest earnings, completely offsetting the effect of induced wealth preference on inequality. When initial inequality is large, however, the difference in interest earnings exceeds the difference in overconsumption tendencies, and wealth inequality widens.

Finally, contrary to our conjectures, activating participants' status concerns via the display of ranking information on accumulated wealth does not further widen wealth inequality in our experiment.

The rest of the paper is organized as follows. Section 2 introduces the model with a preference for wealth accumulation on which our experiment is based and explains its predictions. The experimental design and procedures are described in Section 3. The results of the experiment are presented in Section 4, and Section 5 offers some closing remarks.

2 A model

In the experiment, participants face a finite-horizon and non-stochastic dynamic optimization problem. Decision makers are initially endowed with wealth, k_0. In each period $t = 1, 2, \ldots, T$, the decision maker receives a constant basic income y and interest earnings from accumulated wealth, rk_{t-1}, where r is the interest rate. The decision maker allocates the available budget $y + (1 + r)k_{t-1}$ between consumption c_t and saving k_t, and the saving k_t is carried over as wealth to the next period. Note that income in our model is non-stochastic; therefore, there is no precautionary motive for saving, which is an important consideration in models such as "buffer–stock saving" (Carroll, 1997).

We assume that the decision maker enjoys utility $u(c_t)$ from consumption c_t in period t. Additionally, we assume a preference for wealth accumulation: the decision maker also obtains utility $\gamma(k_t)$ from the saving k_t chosen in period t, and utility $q(k_T)$ from the wealth k_T held at the end of the final period T. We interpret $\gamma(k_t)$ as reflecting a pure preference for holding wealth, as in Michau et al. (2023b), and $q(k_T)$ as capturing motives such as bequests. Under these assumptions, the optimal consumption level is determined independently of initial wealth k_0; it is determined by the marginal rates of substitution implied by the utility functions u, γ, and q.

Formally, the utility maximization problem can be formulated as follows:

$$\begin{aligned}
\underset{\{c_t\}_1^T, \{k_t\}_1^T}{\text{maximize}} \quad & \sum_{t=1}^{T} \big(u(c_t) + \gamma(k_t)\big) + q(k_T) \\
\text{subject to} \quad & c_1 + k_1 = y + (1+r)k_0 \\
& c_2 + k_2 = y + (1+r)k_1 \\
& \dots \\
& c_T + k_T = y + (1+r)k_{T-1}.
\end{aligned}$$

From the first-order conditions with respect to c_t and k_t for $t < T$, we obtain

$$\frac{u'(c_{t+1})}{u'(c_t)} = (1+r)^{-1} \left(1 - \frac{\gamma'(k_t)}{u'(c_t)}\right) \qquad \text{for } t = 1, \dots, T-1. \tag{1}$$

Eq. (1) is the Euler equation describing intertemporal consumption over two periods. Unlike the conventional Euler equation in the typical consumption–saving (cake-eating) problem, it includes the marginal rate of substitution between saving k_t and consumption c_t. Assuming that the change in consumption between two periods is sufficiently small (i.e., $c_{t+1} \sim c_t$) and using the approximation $u'(c_{t+1}) \sim u'(c_t) + u''(c_t)(c_{t+1} - c_t)$, we obtain

$$\Delta c_{t+1} \sim (1+r)^{-1} \left(r + \frac{\gamma'(k_t)}{u'(c_t)}\right) \frac{u'(c_t)}{-u''(c_t)} \qquad \text{for } t = 1, \dots, T-1, \tag{2}$$

where $\Delta c_{t+1} = c_{t+1} - c_t$. Assuming that $u' > 0$, $u'' < 0$, and $\gamma' > 0$, it follows that $\Delta c_{t+1} > 0$, implying that consumption increases over time.

From the first-order condition for the utility maximization problem at $t = T$, we obtain the following relationship among the marginal rates of substitution between

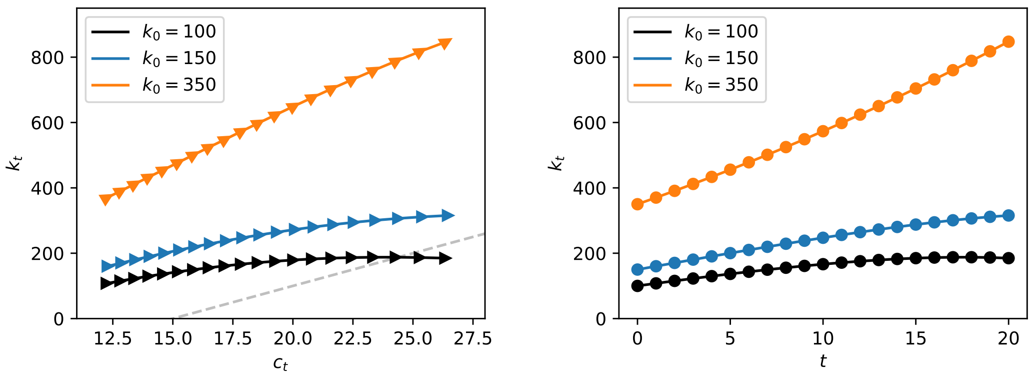


Figure 1: Optimal paths in the experimental setup

Notes: **(Left)** The optimal paths in the experimental setup are illustrated for three agents ($k_0 = 100, 150, 350$). The horizontal axis measures consumption c_t, and the vertical axis measures savings k_t. The gray dashed line represents the locus satisfying $\Delta k_t = 0$. The leftmost point of each trajectory corresponds to $t = 1$, and the rightmost point corresponds to $t = 20$. Because $\Delta c_t > 0$ in all periods, all three paths move to the right until they reach the terminal condition at $c_{20} = 27.05$. Crossing the $\Delta k_t = 0$ line changes the sign of Δk_t, causing only the path for $k_0 = 100$ to shift from upward-sloping to downward-sloping. In the experimental setup, the sequence of optimal c_t is analytically identical for all three agents, although minor discrepancies arise due to rounding imposed by the decision interface used in the experiment. **(Right)** The time evolution of the optimal savings k_t. The gap between agents with $k_0 = 100$ and those with $k_0 = 350$ widens more substantially than the gap between agents with $k_0 = 100$ and those with $k_0 = 150$.

saving and consumption in the final period:

$$u'(c_T) = \gamma'(k_T) + q'(k_T). \tag{3}$$

Considering the no-borrowing constraints $k_t \geq 0$ for any t, if the consumption level c_t implied by Eqs. (2) or (3) exceeds the budget constraint, then consumption is restricted to the budget on hand: $c_t = y + (1 + r)k_{t-1}$.

From the budget constraints, we have

$$\Delta k_t = r k_{t-1} + y - c_t, \tag{4}$$

where $\Delta k_t = k_t - k_{t-1}$.

Based on the two difference equations in Eqs. (2) and (4), together with the terminal condition in Eq. (3), numerical calculations yield the optimal paths for consumption and wealth (saving), as illustrated in Figure 1. The parameters used for the numerical calculations are the same as those used in our experimental setup: $T = 20$, $r = .05$, $y = 15$, $k_0^L = 100$, $k_0^H \in \{150, 350\}$,

$$u(c_t) = 4 \times \left[20 - 20\exp\left(-c_t/9\right)\right],$$
$$\gamma(k_t) = 4 \times 0.01k_t, \text{ and}$$
$$q(k_T) = 4 \times 0.1k_T.$$

We assume linear functions for $\gamma(k_t)$ and $q(k_T)$. This assumption simplifies the analysis because, beyond a certain level of consumption, the marginal utility of consumption becomes lower than that of wealth accumulation (saving). This mechanism drives the increasing-wealth-inequality hypothesis in the presence of a preference for wealth.

3 Experimental design

The experiment consisted of two parts. In the first part, each participant engaged in a real-effort task. In the second part, participants made consumption and saving decisions over 20 periods. As described in Section 2, we induced participants' preferences for consumption and wealth accumulation by rewarding them based on their consumption and saving in each period, as well as on the wealth they held after the final period. As we explain later, we varied (i) how participants' initial wealth, high or low, was determined in the second part, (ii) the size of the gap between high and low initial wealth, and (iii) whether participants observed their ranking in terms of accumulated wealth.

Below, we first describe the two parts of the experiment in more detail, followed by a discussion of the treatments and the implementation procedure. An English translation of the instructions is provided in Appendix A.

Part 1: Real-effort task Participants were asked to add pairs of two-digit numbers for three minutes as many times as possible. Their performance was evaluated based on the number of correctly answered problems. To ensure understanding, participants first completed a one-minute practice session, during which their performance did not affect their subsequent rewards.

In each session, participants were randomly divided into groups of six (or five) individuals.[4] There were two conditions: Random and Effort. In the Random condition, the real-effort task served only to ensure that participants understood that the initial wealth in Part 2 was earned through their own effort rather than provided as house money. In contrast, in the Effort condition, the task was used to make participants recognize that inequality in initial wealth in Part 2 resulted from their own effort: after completing the real-effort task, participants were ranked within their groups based on the number of correct answers, with ties broken randomly.

Participants were informed that, under the Random condition, the rewards they would receive in Part 2 would depend on their performance in Part 1 and an element of chance. We chose this framing because, on the one hand, we wanted participants to understand that their initial wealth in Part 2 was earned in Part 1, but on the other hand, we wanted to assign their types (discussed below) randomly in Part 2. In the Effort condition, participants were instructed that attaining a higher rank in the task would make it easier for them to earn higher rewards in Part 2.

[4] We recruited participants so that each group would consist of six members; however, when scheduled participants were absent, some groups were reduced to five members.

Part 2: Consumption and saving decisions At the outset of Part 2, participants were endowed with initial wealth. As noted above, this initial wealth was determined differently in the Effort and Random conditions. In the Effort condition, three participants (or two in a five-person group) who ranked above the median in their group in Part 1 (H-type) received k_0^H points, whereas those below the median (L-type) received a fixed $k_0^L = 100$ points, with the size of the $k_0^H - k_0^L$ gap determined by the treatment.

In the Random condition, participants' types were assigned at random and did not depend on Part 1 performance. Each participant first received a baseline wealth level calculated by rounding the value of the function $100 - 20\exp(-x/5)$, where x is the participant's Part 1 score. This function was set so that all participants would receive the same baseline wealth of $k_0^L = 100$.[5] After this, three (or two) randomly selected participants in each group were assigned to be H-type and received an additional amount equal to the treatment-specific initial-wealth gap. All other participants were assigned to be L-type.

In each period, participants received a fixed basic income of 15 points and earned interest on their wealth at 5%. The budget for each period comprised wealth carried forward, interest accrued, and basic income. Participants decided how much to consume and how much to save while adhering to the budget constraint

$$c_t + k_t = y + (1 + r)k_{t-1} \qquad \text{for } t = 1, 2, \ldots, 20,$$

where c_t represents consumption, k_t represents end-of-period wealth, $y = 15$ is the basic income, and $r = 0.05$ is the interest rate. Consumption was converted to a

[5]Based on preliminary results, it was known that most participants would score at least $x = 20$ in Part 1. Consequently, most participants were expected to receive 100 ($= 100 - 20\exp(-20/5)$) points as their baseline wealth, and in fact, all participants did so.

utility value and added to the participants' rewards, and eventually to their monetary compensation, using the following concave function:

$$u(c_t) = 80 - 80 \exp(-c_t/9) .$$

To induce wealth preference, wealth carried into each period, k_t, contributed not only to future budgets but was also evaluated by $\gamma(k_t) = 0.04k_t$ and added to participants' monetary compensation in each period, along with consumption utility. Furthermore, wealth not consumed in the final period, k_{20}, was evaluated at $q(k_{20}) = 0.4k_{20}$ and counted toward the final monetary reward, in addition to $\gamma(k_{20})$ for that period.[6]

Thus, the total experimental reward for each participant comprised a 500-JPY participation fee, the aggregated utility from consumption, $\sum_{t=1}^{20} u(c_t)$, the utility from wealth accumulation in each period, $\sum_{t=1}^{20} \gamma(k_t)$, and the utility from final-period wealth $q(k_{20})$. As noted in Section 2, the per-period utility from wealth $\gamma(k_t)$ represents a pure preference for wealth accumulation, while the final-period utility $q(k_{20})$ captures motives such as bequests.

Participants' decisions were facilitated by an interface that presented their wealth, interest received, basic income, current budget, and accumulated experimental rewards. A simulation tool was also available to help participants estimate additional monetary compensation and future budgets based on their consumption or saving choices (see Figure 2 for a screenshot). Although participants were prompted to submit their decisions within three minutes each period, they were allowed additional time if necessary.

After reading the instructions, participants took a comprehension quiz (provided

[6]Theoretically, $\gamma(k_t)$ and $q(k_{20})$ need not differ. We chose to distinguish them to highlight the difference for participants.

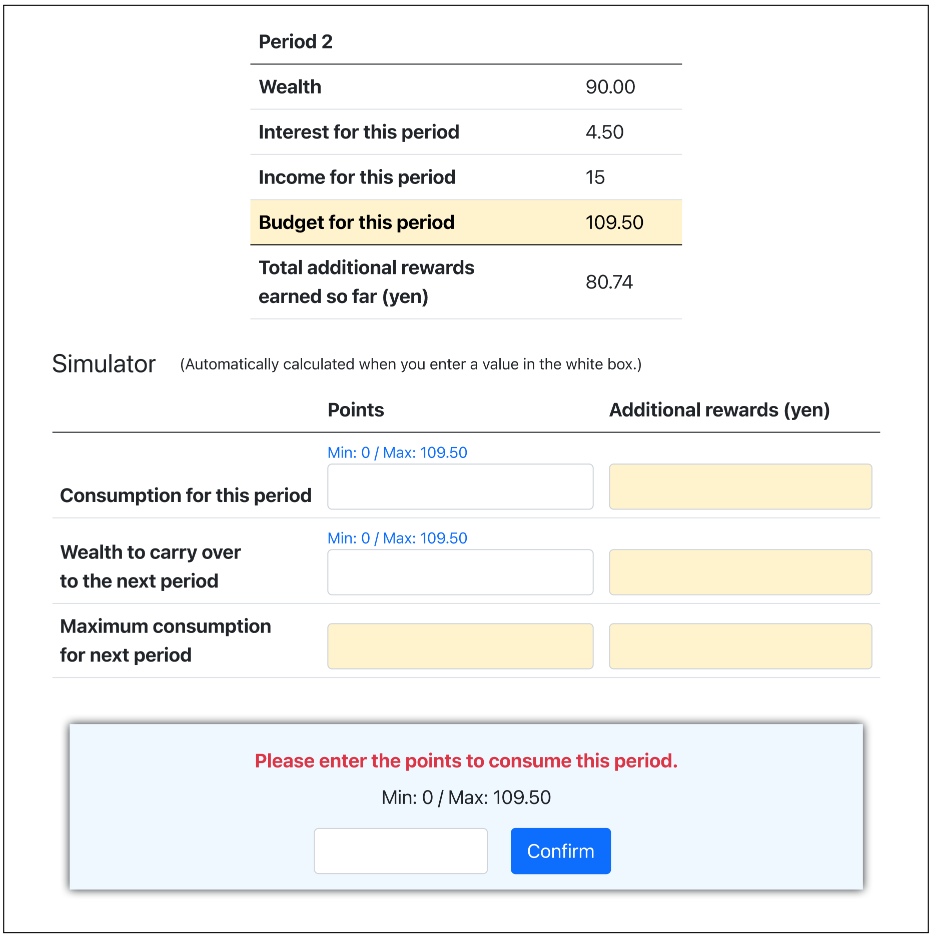


Figure 2: A screenshot of the decision screen

Note: This screenshot has been translated into English; however, it was displayed in Japanese during the actual experiment.

in Appendix A). The decision-making task began only after they correctly answered all questions.

Experimental treatments The design followed a 2×2×2 between-subjects structure, as shown in Table 1. The three factors were:

1. the size of the initial wealth gap between L- and H-types (50 vs. 250),
2. the presence or absence of feedback on the ranking based on accumulated wealth, and
3. the method for determining types (Effort vs. Random).

Table 1: Experimental treatments

Label			Effort determines types	With ranking feedback	Large initial gap (250)
Effort-	NoRank-	Gap50	✓		
Effort-	NoRank-	Gap250	✓		✓
Effort-	Rank-	Gap50	✓	✓	
Effort-	Rank-	Gap250	✓	✓	✓
Random-	NoRank-	Gap50			
Random-	NoRank-	Gap250			✓
Random-	Rank-	Gap50		✓	
Random-	Rank-	Gap250		✓	✓

The first factor was varied by changing the initial endowment for H-types while keeping the endowment for L-types fixed at $k_0^L = 100$. The endowment for H-types, k_0^H, was either 150 or 350 points. Thus, the initial wealth gap between the two types was 50 in the former case (Gap50) and 250 in the latter case (Gap250).

The second factor concerns whether participants received feedback on their within-group rank based on accumulated savings after each period's decision.[7] We label treatments that include rank feedback as "Rank" and those without rank feedback as "NoRank." This factor examined the effect of providing ranking feedback. In treatments with feedback, participants were shown their current wealth and their rank at the end of each period. For ties, ranks were assigned at random. We implemented this manipulation to investigate status concerns that have so far been studied in the context of consumption—for example, by showing average consumption (Carbone and Duffy, 2014) or rankings (Feltovich and Ejebu, 2014)—but not yet in the context of savings.

One might object that, unlike consumption, savings are typically not directly observable to others; therefore, manipulating the visibility of accumulated savings

[7]See Figure B.1 in Appendix B for a screenshot.

in the experiment may seem unrealistic. Although balances in one's bank account are not directly observable in everyday life, we argue that accumulated savings in our experiment can reasonably be interpreted as a proxy for an individual's overall wealth, including past expenditures on so-called "status" goods such as luxury cars, housing, or jewelry. While these expenditures are often discussed under the label of conspicuous consumption, such goods are typically durable, accumulate over time, and constitute an important component of an individual's wealth.

Our experimental design distinguishes between non-durable consumption, which is immediately converted into earnings and captures contemporaneous utility, and savings, which accumulate over time and represent enduring wealth. Within this framework, information about participants' relative rankings in accumulated savings serves as a natural and tractable analogue to social comparisons of wealth in real-world settings. We therefore view the visibility of accumulated savings in the experiment as a justified and meaningful representation of economically relevant status concerns, even if savings are not directly observable in everyday transactions.

Note that activating status concerns by displaying ranking and wealth information may motivate L-type participants to save more to catch up to H-types when the initial wealth gap is small. This is why we consider two levels of initial wealth gaps between the two types, as discussed above.

In the Effort condition of the third factor, we linked initial wealth to performance on the real-effort task so that displaying within-group rankings and accumulated wealth would meaningfully activate participants' status concerns by signaling their ability. Clingingsmith and Sheremeta (2018) experimentally show that conspicuous consumption—that is, the purchase of chocolates during the experiment, which does not contribute to participants' monetary reward—is high when it can be interpreted as a signal of ability because income is determined by a test score, as in our Effort condition. At the same time, Clingingsmith and Sheremeta (2018) report that when income is determined randomly, and thus the purchase of chocolates does not signal ability as in our Random condition, such consumption is low.

However, in the Effort condition, there is a correlation between participants' cognitive skills (as captured by our summation task) and the initial wealth, which may drive the widening of wealth inequality, as participants with higher cognitive skills may make better intertemporal allocation decisions than their low-skilled counterparts.[8] Because such correlation is absent in the Random condition, any widening of wealth inequality observed in this condition is not due to differences in the cognitive skills between types, providing a stronger test of our hypothesis that wealth-inequality dynamics are driven by preferences for wealth accumulation.

Procedures Participants were students of the University of Osaka recruited between June and July 2024 (Effort condition) and in November 2025 (Random condition) from the ISER laboratory pool, which is managed by ORSEE (Greiner, 2015). The experiment was computerized using oTree (Chen et al., 2016) and conducted via Zoom as an online, real-time interactive experiment. Such web-based interactive experiments have become popular worldwide since the onset of the COVID-19 pandemic. The experimental tasks were web-based, and participants accessed them using their laptops or tablets. They were free to choose any location provided it had a stable internet connection and a quiet, disturbance-free environment. These conditions were communicated to participants by the experimenter during recruitment.

At the beginning of each part, an audio recording of the relevant instructions was broadcast via Zoom. While listening to the audio, participants read the instructions displayed on their screens. Throughout the experiment, participants could ask the experimenter questions directly via Zoom's chat function.

[8]We thank a reviewer for pointing out this potential confound.

During Part 2, participants received feedback on their accumulated rewards in units of hundredths of JPY. At the very end of the experiment, the total accumulated rewards were rounded up to the next 10 JPY and paid out together with a fixed participation fee.

Participants were paid with Amazon gift cards (email version), the standard procedure for all online experiments at the ISER laboratory. They received the gift card by email on the day they completed the experiment.

4 Results

In total, 379 participants took part in the experiment. The number of participants in each treatment was as follows: 46 in Effort-NoRank-Gap50, 45 in Effort-NoRank-Gap250, and 48 in each of the other six treatments. Eight groups of observations were collected for each of the eight experimental treatments. Because participants were intended to be divided into groups of six, we recruited 24 participants plus a few reserves for each experimental session. Although we recruited the same number for each session, the actual number of participants varied due to no-shows. When fewer than 24 participants appeared, groups of five were created.[9]

Of the participants, 228 were male, 144 were female, and 7 did not specify their gender. Table 2 summarizes the gender composition of initial wealth type in the Effort and Random conditions. In the Effort condition, no significant association was observed between gender and initial wealth type ($\chi^2 = 0.6056, p = 0.436$). In contrast, in the Random condition, despite the random assignment of initial wealth types, a significant association emerged, indicating that men were more likely to be

[9]This explains why we deviate from our pre-registered number of participants, 48, in some treatments. In addition, due to a technical issue that occurred midway through one session, an additional session was held with a group of six participants.

Table 2: Cross-tabulation of gender balance in the assignment of initial wealth types

Effort condition				Random condition			
	L-type	H-type	Total		L-type	H-type	Total
Female	44	37	81	Female	25	38	63
Male	50	53	103	Male	70	55	125
Unknown	2	1	3	Unknown	1	3	4
Total	96	91	187	Total	96	96	192

Notes: The "Unknown" category includes participants who declined to answer the gender question or who did not identify their gender within a binary framework.

assigned to the L-type ($\chi^2 = 4.4618, p = 0.035$). Therefore, we also conduct analyses controlling for gender and gender composition.

The experiment lasted around 51.6 minutes (SD = 17), including instructions and the comprehension quiz. Average earnings from all tasks were 2,117 JPY (SD = 247), including a fixed participation fee of 500 JPY. Earnings ranged from a minimum of 1,260 JPY to a maximum of 2,690 JPY.[10]

Table 3 shows, for each treatment and initial wealth type, the average consumption between periods 1 and 19 ($c_{1\text{–}19}$) and its average relative deviation from the conditional optimum; the average consumption in the final period (c_{20}) and its average relative deviation from the conditional optimum; and the average wealth (net of initial wealth) in four intervals: periods 1–5 ($k_{1\text{–}5}$), 6–10 ($k_{6\text{–}10}$), 11–15 ($k_{11\text{–}15}$), and 15–20 ($k_{15\text{–}20}$).[11] Conditionally optimal consumption in period t is obtained by solving a dynamic optimization problem for the remaining $T - t + 1$ periods, with initial wealth equal to the allocable budget in period t for each participant.[12]

[10]Converted using the exchange rate at the time of the final experimental session (0.0064 USD per JPY), the mean, minimum, and maximum earnings correspond to 13.55, 8.06, and 17.22 USD, respectively.

[11]Appendix C reports the dynamics of average wealth for H- and L-types in each group and each treatment.

[12]Focusing on conditionally optimal choices is not unprecedented. For example, Carbone and Duffy (2014) show that deviations from conditionally optimal choices are larger when participants receive social information, i.e., the average previous-period consumption of all participants in the group, than when they do not.

Table 3: Consumption and wealth by treatment and type

Treatment	Type	$c_{1\text{–}19}$	Rel. dev. $c_{1\text{–}19}$	c_{20}	Rel. dev. c_{20}	$k_{1\text{–}5}$	$k_{6\text{–}10}$	$k_{11\text{–}15}$	$k_{16\text{–}20}$
Effort-									
NoRank-Gap50	L-type	20.388	0.177***	25.687	−0.018	−1.979	3.379	5.527	−2.885
		(3.92)	(0.24)	(11.56)	(0.41)	(32.30)	(56.74)	(78.83)	(110.33)
	H-type	24.370	0.428***	25.655	0.030	−6.743	−17.833	−25.891	−46.326
		(4.57)	(0.30)	(13.03)	(0.44)	(35.43)	(74.90)	(104.19)	(130.70)
NoRank-Gap250	L-type	20.777	0.206***	26.475	0.074	−0.876	3.261	6.552	−8.942
		(3.47)	(0.23)	(13.78)	(0.47)	(24.04)	(51.78)	(75.85)	(94.66)
	H-type	33.329	0.800***	69.255	1.560*	30.075***	73.502***	79.053**	44.441
		(9.86)	(0.48)	(102.44)	(3.79)	(38.01)	(84.37)	(131.63)	(222.51)
Rank-Gap50	L-type	20.916	0.190***	26.273	−0.023	0.559	9.524	8.579	−9.390
		(3.31)	(0.18)	(6.93)	(0.26)	(28.52)	(47.42)	(64.28)	(83.66)
	H-type	23.765	0.338***	34.645	0.305	−5.615	4.103	3.827	−18.115
		(2.74)	(0.17)	(28.32)	(1.04)	(31.82)	(51.26)	(63.11)	(81.06)
Rank-Gap250	L-type	20.253	0.199***	24.365	−0.048	−14.376**	−15.881**	−13.635	−16.896
		(3.73)	(0.21)	(16.20)	(0.59)	(25.81)	(34.29)	(53.21)	(87.45)
	H-type	26.481	0.499***	26.029	−0.038	24.946***	69.274***	116.497***	177.142***
		(8.68)	(0.47)	(14.10)	(0.52)	(41.02)	(104.81)	(157.39)	(230.07)
Random-									
NoRank-Gap50	L-type	21.785	0.259***	25.820	0.059	1.048	1.296	−9.274	−34.250**
		(2.07)	(0.17)	(9.57)	(0.29)	(24.89)	(60.72)	(76.88)	(76.48)
	H-type	24.631	0.416***	26.833	0.042	−9.327	−17.488	−29.663*	−54.573***
		(2.76)	(0.19)	(8.90)	(0.30)	(28.58)	(54.38)	(73.26)	(89.04)
NoRank-Gap250	L-type	21.627	0.237***	29.855	0.192*	2.914	6.403	2.839	−23.153
		(2.57)	(0.20)	(14.83)	(0.49)	(21.91)	(54.15)	(76.45)	(90.84)
	H-type	25.639	0.491***	33.872	0.282*	10.529	51.608	112.104**	184.832**
		(11.19)	(0.70)	(21.16)	(0.76)	(82.64)	(153.29)	(244.13)	(347.81)
Rank-Gap50	L-type	21.435	0.232***	25.769	0.052	−0.976	−0.615	−2.644	−23.456
		(2.53)	(0.17)	(11.74)	(0.39)	(24.78)	(43.07)	(58.04)	(75.27)
	H-type	23.273	0.305***	26.384	0.033	3.910	12.618	20.387	1.786
		(4.25)	(0.27)	(13.04)	(0.46)	(32.72)	(76.80)	(119.95)	(139.61)
Rank-Gap250	L-type	18.896	0.070	28.763	0.113	4.070	26.493*	46.855**	48.700*
		(4.17)	(0.27)	(12.54)	(0.43)	(25.63)	(73.77)	(100.11)	(127.55)
	H-type	24.822	0.420***	27.761	0.026	17.104*	69.076***	139.681***	215.995***
		(9.26)	(0.53)	(10.13)	(0.37)	(44.99)	(104.18)	(167.35)	(261.88)

Notes: "c_{1-19}" column reports the average consumption c_t for all periods prior to the final period, $t < 20$. "c_{20}" column reports the final-period consumption. "Rel. dev." columns report the average relative deviation of c_{1-19} and c_{20} from the conditionally optimal consumption, where negative values indicate underconsumption (i.e., oversaving). "$k_{a\text{–}b}$" columns report the average wealth k_t, net of initial wealth k_0, over the interval from $t = a$ to $t = b$. For the "Rel. dev." and "$k_{a\text{–}b}$" columns, values statistically significantly different from zero at the 1%, 5%, and 10% levels are denoted by ***, **, and *, based on two-sided t-tests. Standard deviations are shown in parentheses.

The positive and significant average relative deviations between periods 1 and 19 indicate that participants overconsume (relative to the conditional optimum) regardless of their initial wealth type across all treatments. The average degree of overconsumption in these periods is significantly larger for female participants than for male participants (see the negative and significant coefficient on the dummy indicating male in Column (1) of Table E.1 in Appendix E).

Furthermore, the average relative deviation from the conditionally optimal consumption between periods 1 and 19 is significantly larger for H-types than for L-types at the 5% significance level in all treatments except Random-NoRank-Gap250 and Random-Rank-Gap50.[13]

In the final period, participants consume, on average, the conditionally optimal amount, except for H-types in Effort-NoRank-Gap250 and both types in Random-NoRank-Gap250, which exhibit positive relative deviations that are marginally significantly different from zero.[14]

For wealth accumulation, we observe significantly positive average wealth accumulation (net of initial wealth) in at least two of the four intervals for H-types across the four Gap250 treatments: Effort-NoRank-Gap250, Effort-Rank-Gap250, Random-NoRank-Gap250, and Random-Rank-Gap250. Except for Effort-NoRank-Gap250, which shows declining wealth accumulation in the final intervals, the remaining

[13]The difference is marginally significant for Random-NoRank-Gap250. $p-$values, comparing the relative deviation between the two types, based on the two-sided Welch's t-test, are as follows: Effort-NoRank-Gap50 ($p = 0.003$), Effort-NoRank-Gap250 ($p < 0.001$), Effort-Rank-Gap50 ($p = 0.005$), Effort-Rank-Gap250 ($p = 0.007$), Random-NoRank-Gap50 ($p = 0.005$), Random-NoRank-Gap250 ($p = 0.099$), Random-Rank-Gap50 ($p = 0.265$), and Random-Rank-Gap250 ($p = 0.007$). See also the positive and significant coefficient on the dummy indicating H-type in Column (1) of Table E.1 in Appendix E).

[14]In the final period, only the Effort-NoRank-Gap250 treatment ($p = 0.089$, two-sided Welch's t-test) shows a marginally statistically significant difference in the relative deviation of consumption from the conditional optimum between the two types. Column (2) of Table E.1 in Appendix E shows that although H-types overconsume significantly more than L-types, there is no significant gender difference in the degree of overconsumption in the final period.

three treatments exhibit increasing wealth accumulation. For L-types, significant wealth accumulation is observed only in Random-Rank-Gap250. Instead of wealth accumulation, significant wealth *decumulation* appears in the final interval for both types in Random-NoRank-Gap50. This suggests that, despite significantly greater overconsumption by H-types than by L-types, when initial wealth inequality is sufficiently large, the wealth inequality between the two types widens.[15]

Let us now turn to main focus of our paper: the dynamics of wealth inequality, analyzed through the difference in average wealth (net of initial wealth) between the two types.[16]

4.1 Dynamics of wealth inequality

Figure 3 illustrates the dynamics of mean within-group wealth inequality between the two types across the eight treatments. Initial inequalities are normalized to zero. The model's predictions for each condition, excluding the possible effects of status concerns conveyed through feedback on ranking, are shown in thin dotted lines. Table 4 complements Figure 3 by presenting the results of linear regressions, in which the dependent variable is the within-group normalized wealth inequality between the two types, and the independent variables are the treatment dummies.[17]

From Figure 3, we observe that the average wealth inequality between the two types does not widen as much as the model predicts. The experimental results

[15]Table E.2 in Appendix E reports the results of a linear regression investigating the relationship between the displayed rank and participants' subsequent behavior. We do not observe significant relationships between the displayed rank and the magnitude of the average relative decision from optimal consumption, nor do these relationships differ significantly by gender or participant type.

[16]We obtain qualitatively similar results if we instead consider the difference in the median wealth of each type within a group.

[17]Appendix D reports the distribution of within-group wealth inequalities, as well as the results of pairwise treatment comparisons based on the Mann–Whitney U-test. Table E.5 in Appendix E summarizes the pairwise comparisons of treatments based on the regression coefficients presented in Table 4.

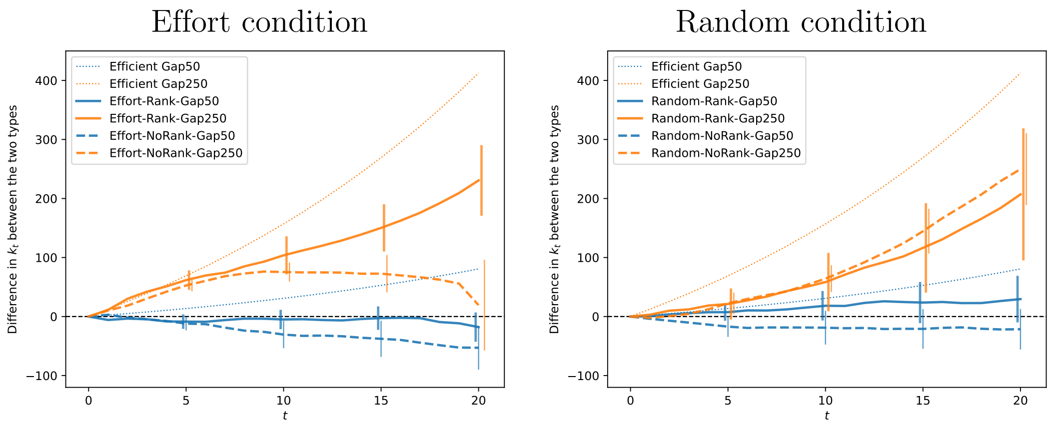


Figure 3: Temporal evolution of within-group wealth inequality

Notes: **(Left)** The Effort condition; **(Right)** The Random condition. The within-group wealth inequality is defined as the difference between the H-type mean of wealth, k_t, and the L-type mean within each group, where the initial gap is normalized to zero. The solid lines with bold error bars correspond to the Rank condition, whereas the dashed lines with thin error bars correspond to the NoRank condition. The error bars indicate standard errors and are shown only for the 5th, 10th, 15th, and 20th periods. Theoretical predictions are shown as thin dotted lines.

(shown in bold lines) all lie below the corresponding theoretical predictions (shown in thin dotted lines). This is a consequence of a significantly higher degree of overconsumption by H-types relative to L-types that we observed above. In fact, for Effort-NoRank-Gap50 and Random-NoRank-Gap50, we see a decline in wealth inequality over time.

The widening of wealth inequality observed in the four Gap250 treatments is statistically significant (see Table 4). While wealth inequality shrinks in Effort-NoRank-Gap250 in the final interval, it continues to widen in the remaining three Gap250 treatments. Thus, we find partial support for the hypothesis of widening wealth inequality: when the initial inequality is large, it widens, at least initially, whereas it does not widen when the initial inequality is small. This pattern holds in part even after controlling for the gender composition of a group. Table E.3 in Appendix E reports the results of regression analyses that control for gender composition. The estimated coefficient for the high-male-ratio dummy is positive but not statistically significant. The Effort-NoRank-Gap250 dummy is not significant in

Table 4: Group-level regression analysis on the wealth inequality

	Periods 1–5	Periods 6–10	Periods 11–15	Periods 16–20
Effort-				
NoRank-Gap50	−4.567	−22.436	−34.399	−47.578
	(10.91)	(23.40)	(35.48)	(51.53)
NoRank-Gap250	30.852***	70.682***	73.853**	54.780
	(10.91)	(23.40)	(35.48)	(51.53)
Rank-Gap50	−6.174	−5.421	−4.752	−8.725
	(10.91)	(23.40)	(35.48)	(51.53)
Rank-Gap250	39.322***	85.155***	130.133***	194.038***
	(10.91)	(23.40)	(35.48)	(51.53)
Random-				
NoRank-Gap50	−10.375	−18.784	−20.389	−20.323
	(10.91)	(23.40)	(35.48)	(51.53)
NoRank-Gap250	7.615	45.205*	109.266***	207.984***
	(10.91)	(23.40)	(35.48)	(51.53)
Rank-Gap50	4.886	13.233	23.031	25.243
	(10.91)	(23.40)	(35.48)	(51.53)
Rank-Gap250	13.035	42.583*	92.826**	167.295***
	(10.91)	(23.40)	(35.48)	(51.53)
Observations	64	64	64	64
Adjusted R^2	0.205	0.267	0.303	0.356

Notes: The dependent variable is the within-group normalized wealth inequality between the two types, constructed as follows. For each interval (periods 1–5, 6–10, 11–15, and 16–20), we compute, for each participant, the interval average of saving k_t net of initial wealth k_0. We then take the average of these values within each type in each group, and measure wealth inequality as the value obtained by subtracting the L-type mean from the H-type mean. The independent variables are dummy variables representing the eight experimental treatments, and the models exclude a constant term. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

any of the four intervals. While the other Gap250 dummies also lose significance in the earlier intervals, the final interval remains significant.[18]

Table 5 shows the results of linear regressions analyzing the main effects of treatment variations on within-group normalized wealth inequality across four intervals, controlling for the gender composition of each group. The dummy variables "Effort", "Rank", and "Gap250" take the value 1 (0) for the Effort (Random) condition, the

[18] If we use the male ratio instead of the high-male-ratio dummy in the regression, the results remain qualitatively the same, but most of the Gap250 treatment dummies lose significance due to multicollinearity between the male ratio and the treatment dummies (the variance inflation factor of the male ratio is 11.98 in the regression). See Table E.4 in Appendix E.

Table 5: Group-level factorial regression analysis on the wealth inequality

	Periods 1–5	Periods 6–10	Periods 11–15	Periods 16–20
Effort	9.725	7.460	4.283	0.334
	(18.00)	(38.37)	(58.83)	(85.35)
Rank	10.854	22.675	31.936	25.458
	(17.74)	(37.80)	(57.96)	(84.08)
Gap250	14.078	53.407	118.228**	214.856**
	(17.15)	(36.56)	(56.06)	(81.33)
Effort × Rank	−17.815	−19.817	−21.851	−14.265
	(24.13)	(51.43)	(78.86)	(114.40)
Effort × Gap250	11.042	14.708	−49.828	−175.072
	(24.27)	(51.72)	(79.30)	(115.04)
Rank × Gap250	−10.604	−34.473	−57.710	−75.076
	(24.21)	(51.60)	(79.13)	(114.79)
Effort × Rank × Gap250	33.447	59.053	123.842	231.880
	(33.79)	(72.02)	(110.44)	(160.21)
Male ratio	19.798	60.292	81.259	107.049
	(22.06)	(47.02)	(72.10)	(104.59)
Constant	−20.075	−49.653	−64.828	−80.468
	(19.60)	(41.78)	(64.07)	(92.94)
Observations	57	57	57	57
Adjusted R^2	0.105	0.123	0.129	0.201

Notes: The dependent variable is the within-group normalized wealth inequality between the two types. The dummy variables "Effort", "Rank", and "Gap250" take the value 1 (0) for the Effort (Random) condition, the Rank (NoRank) condition, and the Gap250 (Gap50) condition, respectively. The variable "Male ratio" represents the proportion of men in each group. Groups that include participants whose gender could not be identified in binary terms were excluded from the regression analysis. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

Rank (NoRank) condition, and the Gap250 (Gap50) condition, respectively.

While the estimated coefficients for the Gap250 condition are positive and significant in the last two intervals, those for the Effort and Rank conditions are not, suggesting that the size of initial inequality is the main driver of widening wealth inequality. Contrary to our conjecture, the activation of status concerns by the display of ranking (the Rank condition), or whether initial wealth is determined by participants' effort (the Effort condition), does not play a significant role in widening wealth inequality.

Table 6: Regression analyses on individual behavior

	(1) Rel. dev. $c_{1\text{–}19}$	(2) Rel. dev. c_{20}	(3) $k_{1\text{–}5}$	(4) $k_{6\text{–}10}$	(5) $k_{11\text{–}15}$	(6) $k_{16\text{–}20}$
Effort	−0.023	−0.140	−3.628	−3.525	−1.461	6.445
	(0.03)	(0.15)	(3.76)	(7.92)	(11.12)	(14.32)
Rank	−0.062**	−0.089	−1.876	4.029	12.801	23.857*
	(0.03)	(0.15)	(3.78)	(7.90)	(10.93)	(13.89)
Gap250	−0.047	0.032	−1.223	3.208	12.910	22.109
	(0.03)	(0.15)	(3.76)	(7.82)	(10.81)	(13.76)
H-type	0.121**	−0.013	−10.284	−14.835	−8.014	1.514
	(0.06)	(0.22)	(6.68)	(14.00)	(20.79)	(27.82)
Effort × H-type	0.132**	0.504**	9.177	6.868	−16.548	−54.760*
	(0.07)	(0.22)	(7.15)	(15.03)	(22.70)	(32.29)
Rank × H-type	−0.082	−0.307	5.015	11.481	23.206	38.054
	(0.07)	(0.22)	(7.44)	(15.56)	(23.44)	(33.29)
Gap250 × H-type	0.226***	0.294	25.596***	66.248***	106.062***	163.489***
	(0.07)	(0.22)	(7.59)	(15.63)	(23.52)	(33.51)
Male	−0.107***	−0.063	13.812***	28.769***	38.855***	50.989***
	(0.04)	(0.11)	(4.17)	(8.33)	(12.57)	(17.86)
Constant	0.333***	0.181	−6.829	−16.372	−31.913**	−68.455***
	(0.04)	(0.18)	(4.88)	(10.04)	(14.11)	(18.29)
Observations	7068	372	1860	1860	1860	1860
Adjusted R^2	0.043	0.035	0.086	0.148	0.185	0.217

Notes: The dependent variable in Column (1) is the average relative deviation from conditionally optimal consumption between periods 1 and 19, and that in Column (2) is the relative deviation in the final period. The dependent variable in Columns (3) to (6) is the interval average of wealth accumulations, net of initial wealth. The dummy variables "Effort", "Rank", and "Gap250" take the value 1 (0) for the Effort (Random) condition, the Rank (NoRank) condition, and the Gap250 (Gap50) condition, respectively. The variable "H-type" is a dummy indicating that the participant's initial wealth type is H, and "Male" is a dummy indicating that the participant is male. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors clustered at the individual level are shown in parentheses.

4.2 Individual level analyses

We have already noted that, regardless of treatment, H-types tend to overconsume more than L-types. However, when the initial wealth inequality between the two types is sufficiently large, as in our Gap250 treatments, inequality further widens.

Table 6 reports the results of linear regression analyzing the main effects of treatment variation on individual behavior. Column (1) examines the average relative deviation from conditionally optimal consumption between periods 1 and 19, and Column (2) conducts the same analysis for the final period. Columns (3) to (6) analyze average wealth accumulations, net of initial wealth, across the four intervals.

Column (1) shows that, as previously noted, H-types overconsume significantly more than L-types (the average marginal effect of the "H-type" dummy is 0.258, $SE = 0.034$), whereas male participants overconsume significantly less than female participants. The negative and statistically significant coefficient for the "Rank" dummy indicates that overconsumption in the Rank condition is significantly lower than in the NoRank condition. However, the interaction between the "H-type" dummy and the "Rank" dummy is not statistically significant, so the difference in overconsumption between the two types is negligible.

The degree of overconsumption by H-types becomes significantly larger in the Effort condition than in the Random condition, and in the Gap250 condition than in the Gap50 condition, as indicated by the positive and significant coefficients for Effort $\times$ H-type and Gap250 $\times$ H-type. Despite these significantly higher levels of overconsumption among H-types relative to L-types, in the Gap250 treatments H-types accumulate significantly more wealth than L-types across all four intervals, as shown by the significant and increasingly larger positive coefficients for Gap250 $\times$ H-Type in Columns (3) through (6).

This suggests that the underlying mechanism behind the widening wealth inequality hypothesis in our experiment—namely, that the marginal utility of wealth accumulation exceeds that of consumption beyond a certain level—operates even when participants' decisions deviate substantially from conditionally optimal choices.

5 Conclusions

Suppose decision makers have a preference for wealth accumulation. Suppose further that the marginal utility of wealth accumulation declines more slowly than that of consumption and has a strictly positive lower bound, such that beyond a certain level of consumption, the marginal utility of wealth accumulation becomes larger than that of consumption.

One implication of such a preference for wealth accumulation is an increase in wealth inequality (Michau et al., 2023a). In our baseline experiment, we test this widening-wealth-inequality hypothesis in the laboratory by explicitly inducing a preference for wealth accumulation, in addition to consumption, in an otherwise standard intertemporal optimization (consumption–saving) experiment (see Duffy, 2015, for a survey of the literature).

Furthermore, we investigate whether status concerns, which are considered an important factor in individuals' wealth preference (Michaillat and Saez, 2021, 2022), can enhance the widening of wealth inequality by activating these concerns through the display of within-group rankings and amounts of accumulated wealth. In particular, we examine whether the effect of activating of status concerns differs depending on the level of initial inequality, either small or large, and whether such inequality is determined by participants' performance in the real-effort condition or by luck. This resulted in a $2 \times 2 \times 2$ between-subjects design.

We found partial support for the hypothesis of widening wealth inequality. When the initial inequality was large, inequality widened at least until the middle of the experiment, although its magnitude was much smaller than the theoretical prediction, regardless of whether status concerns were activated or whether the amount of initial wealth was determined by participants' effort or luck. Wealth inequality shrank in the second half of the Effort-NoRank-Gap250 treatment, while in the remaining three Gap250 treatments it continued to widen. By contrast, when initial inequality was small, wealth inequality did not widen.

The difference between the high and low initial inequality conditions was due to participants' tendency to overconsume, and this tendency was stronger for those with high initial wealth than for those with low initial wealth. When the initial inequality was small, this difference in the tendency to overconsume between the two types completely offset the effect of the induced wealth preference on widening wealth inequality. However, it was not enough to do so when initial inequality was large.

Contrary to our conjecture, activating participants' status concerns through the display of within-group wealth rankings did not lead to a further widening of wealth inequality. This result holds regardless of whether initial wealth was determined by effort or assigned randomly, and regardless of whether initial inequality was small or large. Although ranking information affected individual behavior, most notably by slightly reducing overall overconsumption, it did not differentially affect high- and low-initial-wealth participants in a way that amplified inequality.

One possible interpretation is that, in the presence of a strong induced preference for wealth accumulation, the marginal incentive provided by social comparison through rankings is relatively weak. Another possibility is that ranking information simultaneously induces opposing behavioral responses: while some low-wealth participants may be motivated to save more to catch up and improve their relative position, high-wealth participants may respond by consuming more, either because their rank is already secure or because observing the wealth levels of other participants reduces the perceived need for further accumulation, even if they wish to maintain a superior position. As a result, these effects may largely offset each other at the aggregate level.

Taken together, our findings suggest that, in this experimental environment, the size of initial wealth inequality plays a far more important role in shaping the dynamics of wealth inequality than the activation of status concerns through explicit ranking information.

While our experiment investigated the consequences of people having a certain form of wealth preference, future research should investigate whether people indeed have such a preference and, if so, what possible reasons underlie it.

References

Ahrens, S., Bosch-Rosa, C., Meissner, T., 2022. Intertemporal consumption and debt aversion: a replication and extension. Journal of the Economic Science Association 8, 56–84. doi:https://doi.org/10.1007/s40881-022-00118-y.

Akesaka, M., Mikami, R., Ono, Y., 2024. Insatiable wealth preference: Evidence from Japanese household survey. Discussion Paper No. 1241RR. Institute of Social and Economic Research. doi:https://dx.doi.org/10.2139/ssrn.4800341.

Ballinger, T.P., Hudson, E., Karkoviata, L., Wilcox, N.T., 2011. Saving behavior and cognitive abilities. Experimental Economics 14, 349–374. doi:https://doi.org/10.1007/s10683-010-9271-3.

Ballinger, T.P., Palumbo, M.G., Wilcox, N.T., 2003. Precautionary Saving and Social Learning across Generations: An Experiment. The Economic Journal 113, 920–947. doi:https://doi.org/10.1111/1468-0297.t01-1-00158.

Brown, A.L., Chua, Z.E., Camerer, C.F., 2009. Learning and Visceral Temptation in Dynamic Saving Experiments. The Quarterly Journal of Economics 124, 197–231. doi:https://doi.org/10.1162/qjec.2009.124.1.197.

Carbone, E., 2006. Understanding intertemporal choices. Applied Economics 38, 889–898. doi:https://doi.org/10.1080/00036840500399313.

Carbone, E., Duffy, J., 2014. Lifecycle consumption plans, social learning and external habits: Experimental evidence. Journal of Economic Behavior and Organization 106, 413–427. doi:https://doi.org/10.1016/j.jebo.2014.07.010.

Carbone, E., Hey, J.D., 2004. The Effect of Unemployment on Consumption: An Experimental Analysis. The Economic Journal 114, 660–683. doi:https://doi.org/10.1111/j.1468-0297.2004.00236.x.

Carroll, C.D., 1997. Buffer-stock saving and the life cycle/permanent income hypothesis. The Quarterly Journal of Economics 112, 1–55. doi:https://doi.org/10.1162/003355397555109.

Chen, D.L., Schonger, M., Wickens, C., 2016. oTree – An open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88–97. doi:http://dx.doi.org/10.1016/j.jbef.2015.12.001.

Clingingsmith, D., Sheremeta, R.M., 2018. Status and the demand for visible goods: experimental evidence on conspicuous consumption. Experimental Economics 21, 877–904. doi:https://doi.org/10.1007/s10683-017-9556-x.

Crockett, S., Duffy, J., Izhakian, Y., 2019. An Experimental Test of the Lucas Asset Pricing Model. The Review of Economic Studies 86, 627–667. doi:https://doi.org/10.1093/restud/rdy035.

De Nardi, M., Fella, G., 2017. Saving and wealth inequality. Review of Economic Dynamics 26, 280–300. doi:https://doi.org/10.1016/j.red.2017.06.002.

Duffy, J., 2015. Macroeconomics: A survey of laboratory research, in: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics, Volume 2. Princeton University Press, pp. 1–90. doi:https://doi.org/10.1515/9781400883172.

Feltovich, N., Ejebu, O.Z., 2014. Do positional goods inhibit saving? evidence from a life-cycle experiment. Journal of Economic Behavior and Organization 107, 440–454. doi:https://doi.org/10.1016/j.jebo.2014.01.015.

Gechert, S., Siebert, J., 2022. Preferences over wealth: Experimental evidence. Journal of Economic Behavior and Organization 200, 1297–1317. doi:https://doi.org/10.1016/j.jebo.2020.06.006.

Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association 1, 114–125. doi:https://doi.org/10.1007/s40881-015-0004-4.

Hashimoto, K., Ono, Y., Schlegl, M., 2023. Structural unemployment, underemployment, and secular stagnation. Journal of Economic Theory 209, 105641. doi:https://doi.org/10.1016/j.jet.2023.105641.

Hey, J.D., Dardanoni, V., 1988. Optimal Consumption Under Uncertainty: An Experimental Investigation. The Economic Journal 98, 105–116. doi:https://doi.org/10.2307/2233308.

Lei, V., Noussair, C.N., 2002. An Experimental Test of an Optimal Growth Model. American Economic Review 92, 549–570. doi:https://doi.org/10.1257/00028280260136246.

Meissner, T., 2016. Intertemporal consumption and debt aversion: an experimental study. Experimental Economics 19, 281–298. doi:https://doi.org/10.1007/s10683-015-9437-0.

Michaillat, P., Saez, E., 2021. Resolving new keynesian anomalies with wealth in the utility function. The Review of Economics and Statistics 103, 197–215. doi:https://doi.org/10.1162/rest_a_00893.

Michaillat, P., Saez, E., 2022. An economical business-cycle model. Oxfort Economic Papers 74, 382–411. doi:https://doi.org/10.1093/oep/gpab021.

Michau, J.B., 2018. Secular stagnation: Theory and remedies. Journal of Economic Theory 176, 552–618. doi:https://doi.org/10.1016/j.jet.2018.04.011.

Michau, J.B., Ono, Y., Schlegl, M., 2023a. The preference for wealth and inequality:

Towards a Piketty theory of wealth inequality. Working Paper No. 10824. CESInfo. doi:https://dx.doi.org/10.2139/ssrn.4665585.

Michau, J.B., Ono, Y., Schlegl, M., 2023b. Wealth preference and rational bubbles. European Economic Review 156, 104496. doi:https://doi.org/10.1016/j.euroecorev.2023.104496.

Noussair, C., Matheny, K., 2000. An experimental study of decisions in dynamic optimization problems. Economic Theory 15, 389–419. doi:https://doi.org/10.1007/s001990050019.

Ono, Y., 1994. Money, Interest, and Stagnation: Dynamic Theory and Keynes's Economics. Oxford University Press, New York.

Ono, Y., 2001. A reinterpretation of chapter 17 of keynes's *General Theory*: Effective demand shortage under dynamic optimization. International Economic Review 42, 207–236. doi:https://doi.org/10.1111/1468-2354.00107.

Ono, Y., 2015. Growth, secular stagnation and wealth preference. Discussion Paper No. 946. Institute of Social and Economic Research. URL: https://ssrn.com/abstract=2654081.

Ono, Y., Ishida, J., 2014. On persistent demand shortages: A behavioral approach. Japanese Economic Review 65, 42–69. doi:https://doi.org/10.1111/jere.12016.

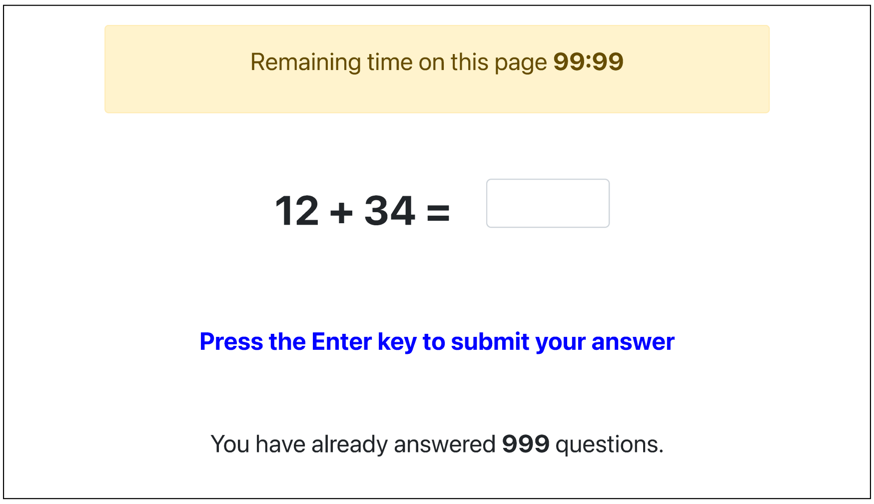
Ugai, H., 2007. Effects of the quantitative easing policy: A survey of empirical analyses. Monetary and Economic Studies 25, 1–48. URL: https://www.boj.or.jp/en/research/wps_rev/wps_2006/wp06e10.htm.

A Instructions

[The instructions have been translated from the original Japanese.]

Part 1

- In Part 1, you will work on calculation tasks.
- A screen like the one shown below will be displayed.



- After entering your answer to the addition problem displayed on the screen into the input box, press the Enter key to submit it.
- You will earn one point for each correct answer.
- After practicing for one minute, you will work on the actual task for three minutes.
- Please answer as many questions correctly as possible within three minutes.

*[For the **Effort** condition]*

- After completing Part 1, we will rank you based on your score.
- Participants will be randomly assigned to groups, and rankings will be conducted within each group.
- Participants with higher ranks within their group will be more likely to earn greater rewards in Part 2.
- If multiple participants obtain the same score, their relative ranks will be determined randomly.

*[For the **Random** condition]*

- The payment you receive in Part 2 will be determined by your score in Part 1, as well as by an element of chance.

Part 2

- In Part 2, you will be asked to make decisions on how to allocate your budget between consumption and wealth accumulation over 20 periods.

[*For the **Effort** condition*]

- First, you will receive initial wealth based on your ranking in Part 1.
- The amount of initial wealth will be 150 points [*for the **Gap50** condition; 350 points for the Gap250 condition*] for participants who rank above the median, and 100 points for those who rank below the median.

[*For the **Random** condition*]

- First, based on your score in Part 1, you will receive initial wealth of up to 100 points.
- In addition, participants will be randomly assigned to groups of six (or five). Within each group, a lottery will be conducted, and three participants (or two in a five-person group) will receive 50 points [*for the **Gap50** condition; 250 points for the Gap250 condition*] added to their initial wealth.
- The allocation of these additional points is determined randomly and is unrelated to participants' scores in Part 1.

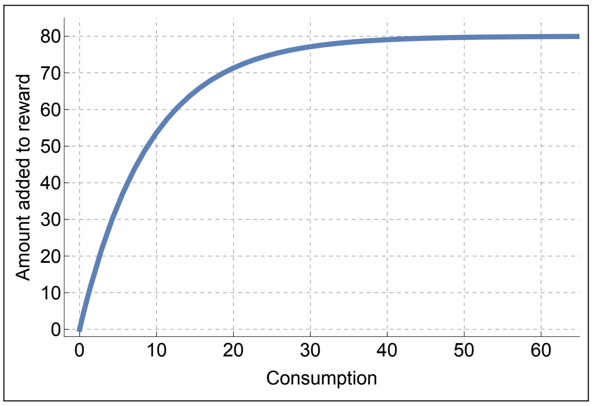
- Wealth will accrue 5% interest each period. Interest earnings of less than 0.01 points will be rounded down.
- You will also receive an income of 15 points in each period.
- For example, if the initial wealth is 100 points, the budget for the first period is 120 points:

$$\underbrace{100}_{\text{initial wealth}} + \underbrace{0.05 \times 100}_{\text{interest}} + \underbrace{15}_{\text{income}} = \underbrace{120.}_{\substack{\text{budget for} \\ \text{the first period}}}$$

- Each period, you will decide how to allocate your budget between consumption in that period and the amount of wealth to carry over to the next period.

- The points you choose to consume will be converted into currency using the following function and added to your reward:

$$80 - 80 \times e^{-(\text{consumption points})/9} \quad \text{[yen]}$$



- The conversion table below summarizes the correspondence between specific consumption points and reward amounts:

The conversion table

Consumption points	Reward amounts
0.00	0.00
0.20	1.75
0.40	3.47
0.60	5.15
0.80	6.80

[Note: This conversion table is scrollable on the screen and includes rows up to the point where the reward amounts in the right column reach 80.00 yen, although the consumption points are not evenly spaced.]

- For example, if you allocate 30 points for consumption from your budget, an additional 77.14 yen will be added to your reward.
- The values listed in the conversion table are examples only. In your actual decisions, you may submit any value with a precision up to the second decimal place.

- Any points not consumed within the budget will become wealth for the next period.
- The following equation describes the relationship between wealth carried across periods and consumption:

[consumption for this period] + [wealth carried over to the next period] = [wealth carried over from the previous period] + [interest accrued on carried-over wealth] + [income]

- For example, if the budget for the first period is 120 points and you consume only 30 points, the amount carried over to the second period is 90 points:

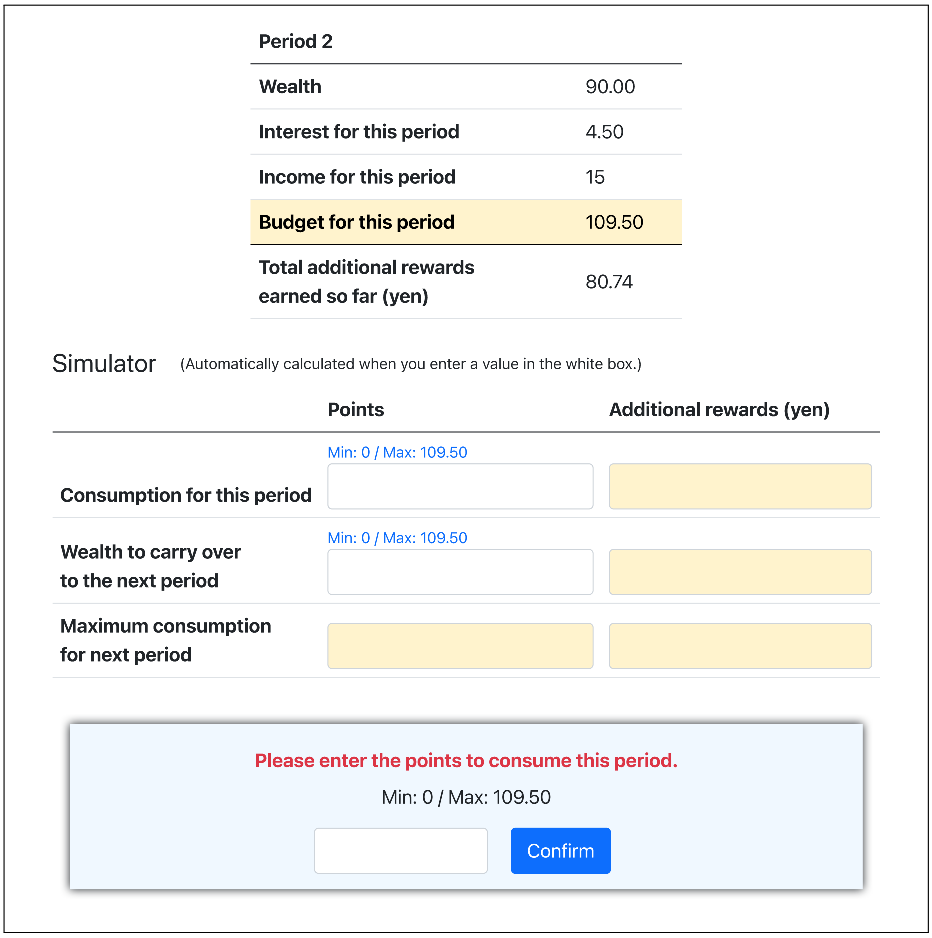
$$\underbrace{120}_{\text{budget}} - \underbrace{30}_{\text{consumption}} = \underbrace{90.}_{\text{carried-over wealth}}$$

  Meanwhile, the budget for the second period is 109.50 points:

$$\underbrace{90}_{\text{wealth}} + \underbrace{0.05 \times 90}_{\text{interest}} + \underbrace{15}_{\text{income}} = \underbrace{109.50.}_{\substack{\text{budget for} \\ \text{the second period}}}$$

- You are not allowed to consume more than the budget for the period. In addition, you may not choose a negative consumption amount.
- In decisions from period 1 through period 19, if you carry over wealth to the next period, the value obtained by multiplying the number of points carried over by **0.04** will be added to your experimental reward as currency.
- For example, if you allocate 30 points for consumption from a budget of 120 points, the wealth carried over to the next period is 90 points, and 3.60 yen will be added to your reward. Combined with the 77.14 yen reward obtained from consumption, the total reward for that period is 80.74 yen.
- The wealth left unconsumed in the final (20th) period will be refunded as currency by multiplying the number of points by **0.44** and adding the result to your reward.

- Decisions are made on the following screen:



- Before making a decision, you may use the simulator to calculate how much reward you can earn based on your consumption in the current period, as well as the maximum amount you can consume in the next period (i.e., your budget for the next period). Please enter a number in either the "Consumption for this period" or "Wealth to carry over to the next period" field. The values in the other fields will be calculated and displayed automatically.

- Your final decision is made by entering the number of points you wish to consume in that period in the blue box at the bottom of the screen. Any number entered will be rounded down to the nearest hundredth of a point.

- Once you have entered the number of points to consume, press the "Confirm" button to submit your choice.

- Please decide and submit your consumption points within three minutes. A warning will appear once three minutes have elapsed.

[*For the* ***Rank*** *condition*]

- After you make a decision in each period, the ranking based on the amount of wealth carried over to the next period will be displayed within the group.

- If multiple participants have the same amount of wealth, their relative ranking will be determined randomly.
- The ranking screen will display both the ranking and the amount of wealth.
- Your own information will be highlighted.

Comprehension quiz

[Participants were allowed to refer to the instructions and the point-to-yen conversion table while answering the quiz.]

Q1 Suppose you allocated 10 points to consumption in Period 1. How many yen would you earn from consumption in Period 1? Refer to the conversion table to answer. [*Correct answer: 53.66 yen*]

Q2 Suppose you carried over 10 points from Period 1 to your wealth in Period 2. Which of the following formulas correctly represents the maximum number of points you can allocate to consumption in Period 2? Please choose from the options below.

- $10 + 15$
- $10 + 0.05 \times 15$
- $0.05 \times 10 + 15$
- $10 + 0.05 \times 10 + 15$ [*correct*]
- $10 + 0.05 \times 10 + 15 + 0.05 \times 15$

Q3 Suppose you allocated your entire budget to consumption in Period 10. What is the maximum number of points you can allocate to consumption in Period 11? Please choose from the options below.

- 0
- 15 [*correct*]
- 15×0.05
- $15 + 0.05 \times 15$

Q4 Suppose you had earned 1,000 yen by Period 19. In the final Period 20, your budget was 200 points, and you allocated 150 of those points to consumption. In this case, excluding the 500-yen participation fee, how much additional reward will you receive in total? Please choose from the options below.

- 1,000 yen
- 1,000 yen + [reward corresponding to consumption of 150 points]
- 1,000 yen + [reward corresponding to consumption of 150 points]
  + [reward corresponding to the remaining 50 points of wealth] [*correct*]

B Screen for providing feedback on rankings

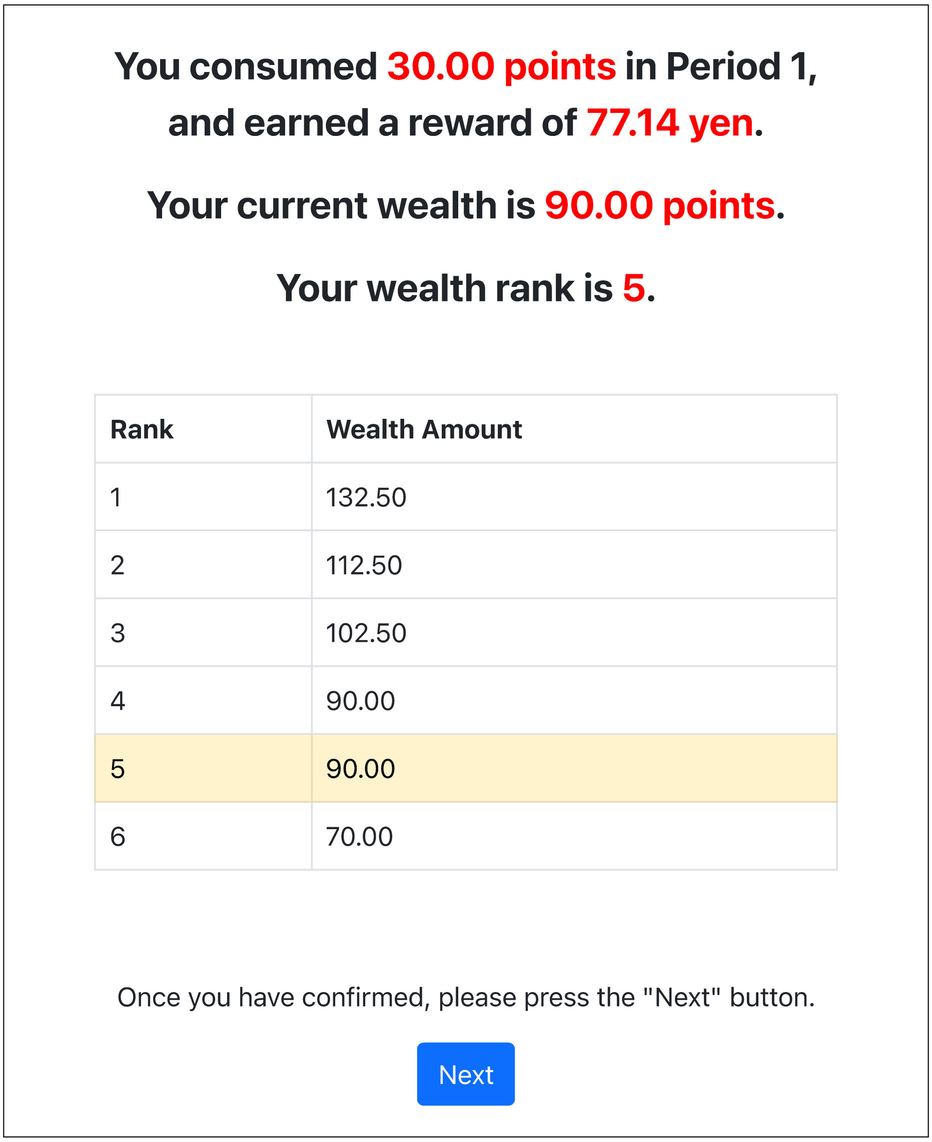


Figure B.1: Screen for providing feedback on rankings

Note: This screenshot is not presented in the instructions.

C Group-level outcomes

This section presents the results for each group separately.

Figure C.2 shows the dynamics of the average wealth for each type (red for H-type and black for L-type) in each group in the Effort treatments. Figure C.3 do the same for the Random treatments.

In all the Gap50 treatments shown in Figures C.2 and C.3, a larger decline in the average wealth of H-types than L-types, as well as an eventual reversal of their average wealth, is observed in some groups. For Effort-NoRank-Gap250, the wealth of H-types declines in the second half of the experiment in three groups (including a sharp decline in the final period that results in the reversal of the average wealth between the two types in one group), which contributes to the declining wealth inequality after the second interval for this treatment shown in Figure 3. There is a group in Random-Rank-Gap250 in which the average wealth of L-types catches up with that of H-types, partly because the average wealth of H-types does not increase over time. However, in many groups of the Gap250 treatments, the average wealth of H-types increases while that of L-types declines, resulting in a continued widening of wealth inequality in these treatments.

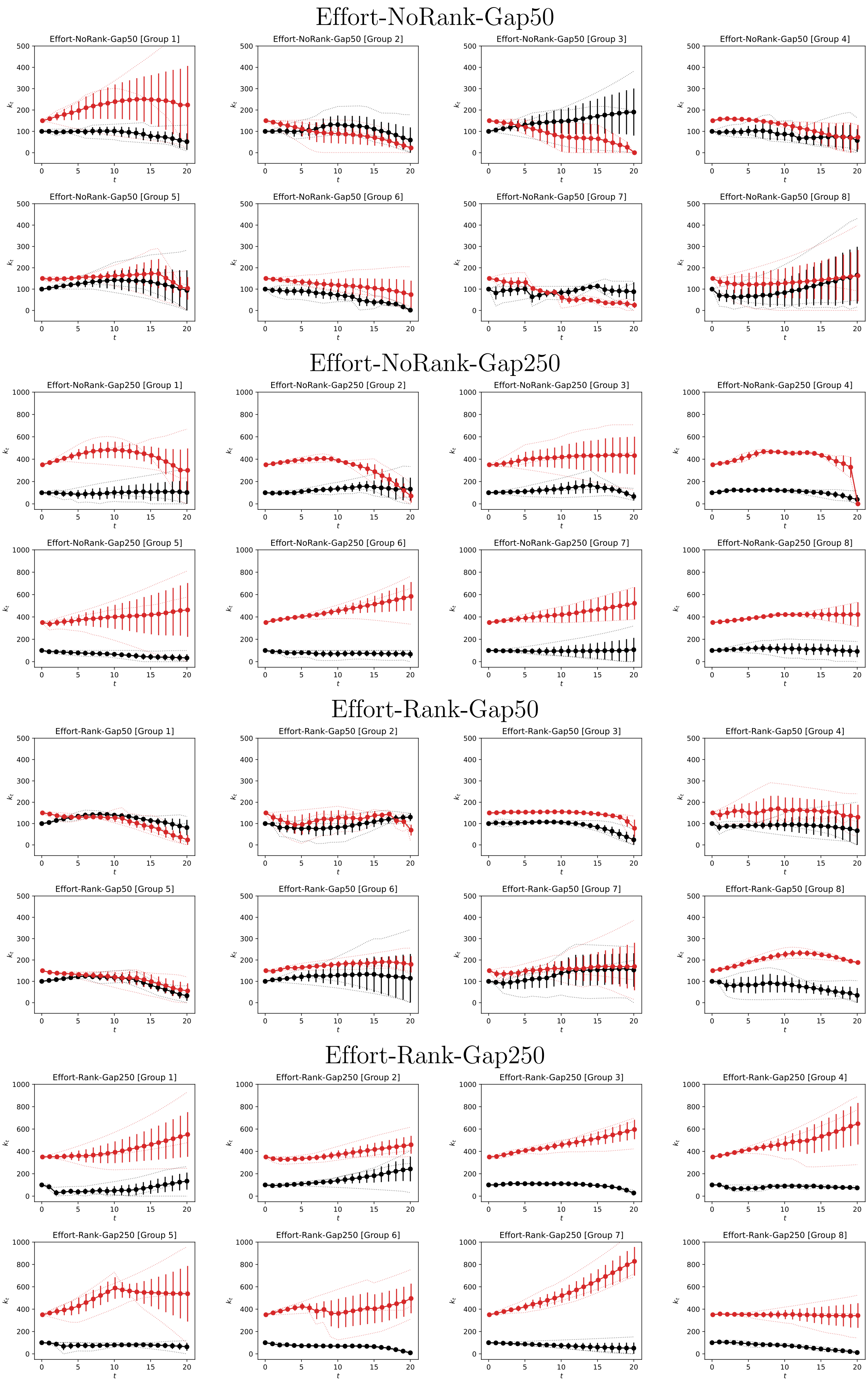


Figure C.2: Dynamics of average wealth of two types in the Effort treatments

Notes: H-type in red; L-Type in black. Error bars represent the standard error of the mean. The dotted lines represent the decisions of each participant.

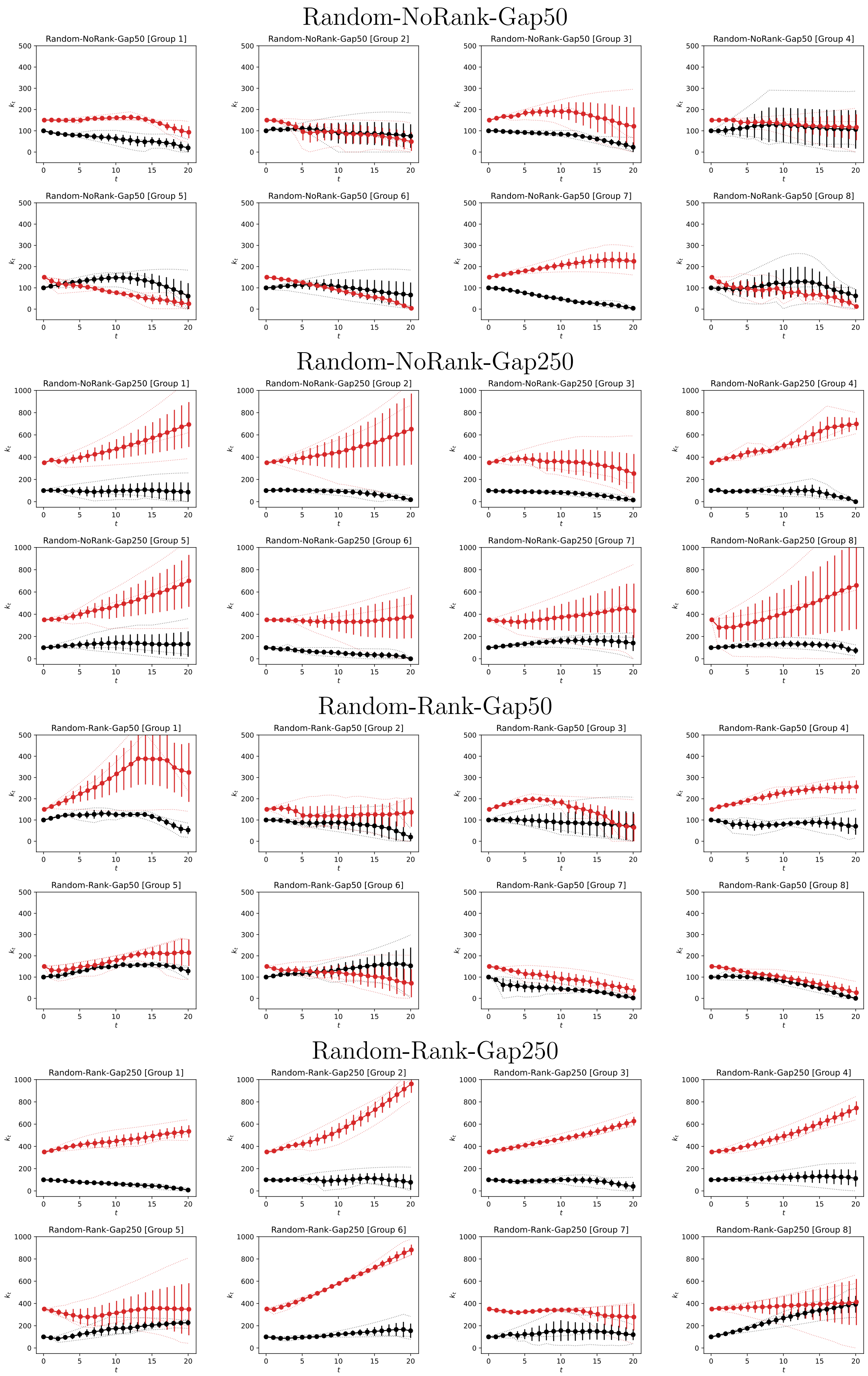


Figure C.3: Dynamics of average wealth of two types in the Random treatments

Notes: H-type in red; L-Type in black. Error bars represent the standard error of the mean. The dotted lines represent the decisions of each participant.

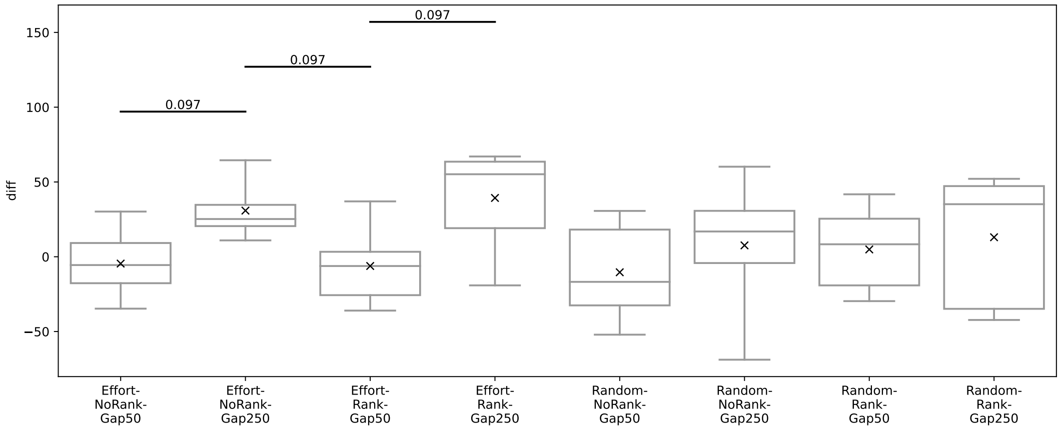
D Non-parametric treatment comparisons

Figure D.4 shows the distribution of within-group wealth inequalities for periods 1 to 5 and 6 to 10, as well as the results of pairwise treatment comparisons based on the Mann–Whitney U test. The initial wealth inequality is normalized to zero. Figure D.5 presents the corresponding results for periods 11 to 15 and 16 to 20.

In the earlier periods shown in Figure D.4, significant treatment differences in wealth inequality are observed only in the Effort treatments. Specifically, wealth inequality in Effort-NoRank-Gap250 and Effort-Rank-Gap250 is (marginally) significantly larger than in their Gap50 counterparts. There is no significant difference between the NoRank and Rank treatments when the level of initial wealth inequality is held constant. Among the four treatments in the Random condition, there is no significant difference across treatments. Furthermore, there are no significant differences between the Effort and Random conditions when the initial level of wealth inequality and the display of ranking information are controlled for.

In the later periods shown in Figure D.5, we begin to observe significant differences among treatments in the Random treatments as well. In particular, wealth inequality in Random-NoRank-Gap250 is significantly larger than its Gap50 counterpart. There is no significant difference between Random-Rank-Gap250 and its Gap50 counterpart, mainly because of the wide variation in outcomes across groups in Random-Rank-Gap250. As in the earlier periods, we do not observe a significant difference between the Effort and Random treatments or between the Rank and NoRank treatments when other aspects are held constant.

(a) Periods 1–5



(b) Periods 6–10

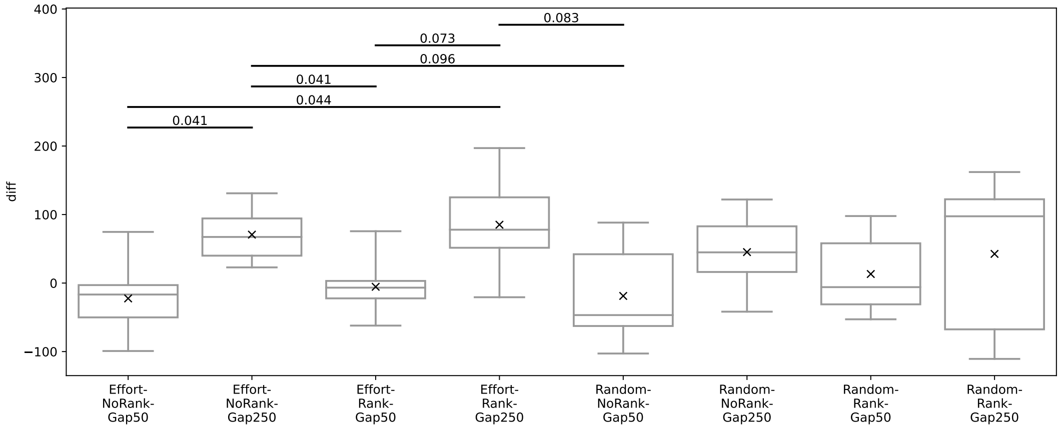
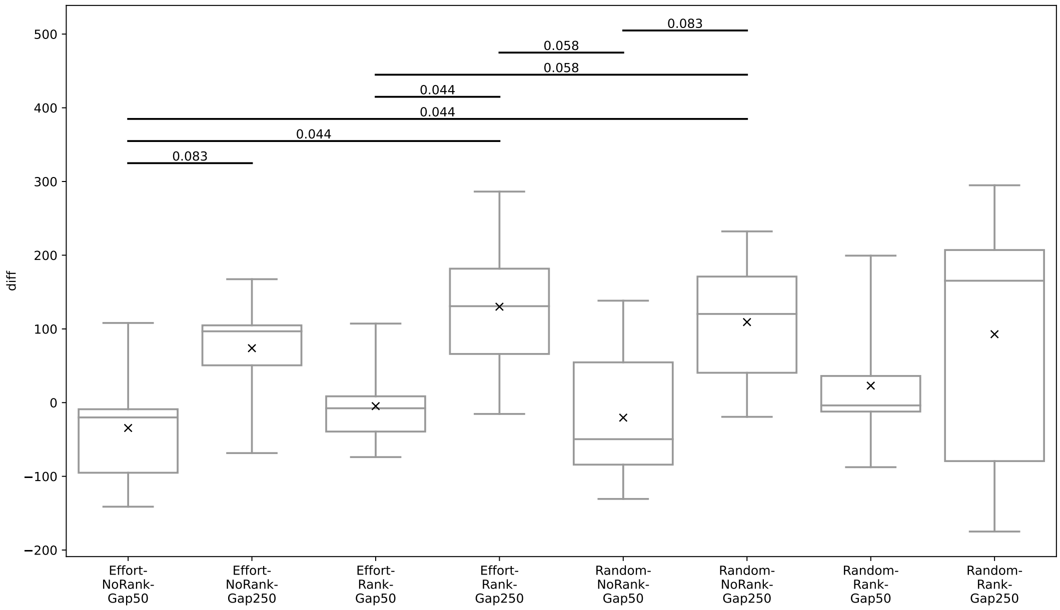


Figure D.4: Distribution of interval-averaged wealth inequality and pairwise treatment comparisons for periods 1–5 and 6–10.

Notes: Boxplots show the distribution of the interval average of type-wise wealth inequality for each group. The initial inequality is normalized to zero. The cross mark indicates the mean. The lower and upper whiskers represent the minimum and maximum values, respectively. Above each boxplot, pairwise Mann–Whitney U test p-values for differences in wealth inequality between two treatments are shown when they fall below 0.10. The p-values are adjusted for multiple comparisons using the Benjamini–Hochberg procedure.

(c) Periods 11–15



(d) Periods 16–20

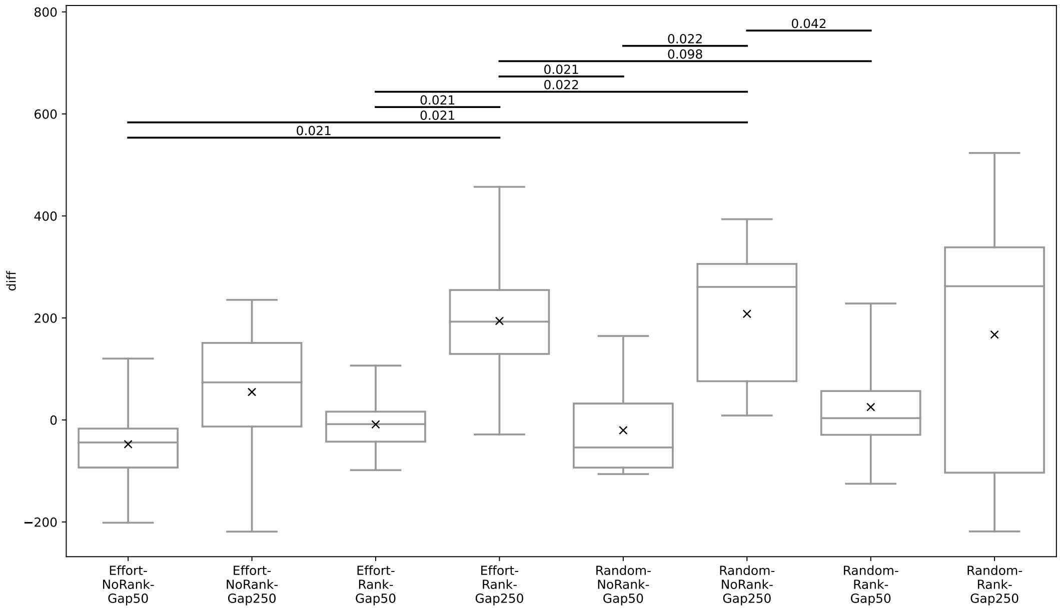


Figure D.5: Distribution of interval-averaged wealth inequality and pairwise treatment comparisons for periods 11–15 and 16–20

Notes: Boxplots show the distribution of the interval average of type-wise wealth inequality for each group. The initial inequality is normalized to zero. The cross mark indicates the mean. The lower and upper whiskers represent the minimum and maximum values, respectively. Above each boxplot, pairwise Mann–Whitney U test p-values for differences in wealth inequality between two treatments are shown when they fall below 0.10. The p-values are adjusted for multiple comparisons using the Benjamini–Hochberg procedure.

E Additional analysis

Table E.1 reports individual-level regression estimates examining how each experimental treatment affects the relative deviation from conditionally optimal consumption and wealth accumulation across different intervals.

Regarding the control variables, H-type participants accumulate substantially more wealth in all intervals and also deviate more from optimal consumption. Male participants deviate slightly less from optimal consumption on average but accumulate more wealth in all intervals. These effects are large and highly statistically significant.

Table E.2 reports analyses of the relationship between participants' within-group wealth rank, which is presented to each participant at the end of the previous period in the Rank treatments, and their subsequent decisions. For wealth accumulation across all intervals (Columns 3–6), rank value is negatively and significantly associated with wealth growth: participants with a poorer rank accumulate less wealth in the next period. In the Gap250 treatments, a poorer rank substantially reduces wealth accumulation, as indicated by the large and significant negative interaction coefficients.

In contrast, rank shows limited influence on consumption deviations. For the average deviation over periods 2 to 19 (Column 1), rank is positively associated with deviation from the conditional optimum, although the magnitude is small. No significant rank effect appears in the final-period deviation (Column 2). The interaction terms with rank are also small and statistically insignificant in both consumption regressions.

Tables E.3 and E.4 present the results of controlling for within-group gender composition in the linear regression analysis reported in Table 4. In the model shown in Table E.4, the proportion of male participants in each group was added as a control variable; however, the variance inflation factor (VIF) for this variable was

Table E.1: Individual-level regression analysis

	(1) Rel. dev. $c_{1\text{–}19}$	(2) Rel. dev. c_{20}	(3) $k_{1\text{–}5}$	(4) $k_{6\text{–}10}$	(5) $k_{11\text{–}15}$	(6) $k_{16\text{–}20}$
Effort-						
NoRank-Gap50	0.238***	−0.071	−17.708***	−38.838***	−59.001***	−95.037***
	(0.04)	(0.18)	(5.79)	(11.05)	(15.21)	(20.45)
NoRank-Gap250	0.429***	0.707***	−0.134	3.793	−9.178	−55.668**
	(0.06)	(0.18)	(5.46)	(11.57)	(17.02)	(25.30)
Rank-Gap50	0.191***	0.052	−15.816***	−24.489**	−41.457***	−81.654***
	(0.04)	(0.17)	(5.50)	(9.65)	(12.87)	(17.75)
Rank-Gap250	0.276***	−0.184	−8.074	−4.390	4.676	14.407
	(0.05)	(0.17)	(5.60)	(11.39)	(17.90)	(26.57)
Random-						
NoRank-Gap50	0.280***	−0.027	−19.082***	−42.909***	−72.021***	−119.167***
	(0.04)	(0.18)	(5.12)	(10.86)	(15.04)	(19.27)
NoRank-Gap250	0.304***	0.154	−8.430	−6.158	4.647	6.200
	(0.08)	(0.18)	(9.43)	(17.63)	(27.31)	(39.05)
Rank-Gap50	0.211***	−0.032	−14.172***	−29.888***	−44.188***	−85.627***
	(0.04)	(0.18)	(5.08)	(11.02)	(16.65)	(20.47)
Rank-Gap250	0.174***	−0.022	−4.243	14.603	43.707**	61.761*
	(0.06)	(0.18)	(5.81)	(13.94)	(21.10)	(31.91)
H-type	0.259***	0.237**	9.790***	27.855***	48.845***	75.164***
	(0.03)	(0.11)	(3.76)	(7.86)	(11.85)	(16.96)
Male	−0.101***	−0.051	14.919***	30.889***	41.312***	53.902***
	(0.04)	(0.11)	(4.38)	(8.66)	(12.94)	(18.30)
Observations	7068	372	1860	1860	1860	1860
Adjusted R^2	0.178	0.061	0.066	0.145	0.183	0.186

Notes: The dependent variable in Column (1) is the average relative deviation from conditionally optimal consumption between periods 1 and 19, and that in Column (2) is the relative deviation in the final period. The dependent variable in Columns (3) to (6) is the interval average of wealth accumulations, net of initial wealth. The independent variables are dummy variables representing the eight experimental treatments, and the models exclude a constant term. All models control for the variable "H-type," a dummy indicating that the participant's initial wealth type is H, and "Male," a dummy indicating that the participant is male. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors clustered at the individual level are shown in parentheses.

11.98, indicating a potentially problematic level of multicollinearity. To address this issue, the model in Table E.3 uses a dummy variable coded as 1 when the proportion of men in a group exceeds 0.5 and 0 otherwise. The VIF for this dummy variable is 3.10, which is acceptable, and in this model the Effort-Rank-Gap250 treatment has a statistically significant effect on wealth inequality.

Table E.5 summarizes the pairwise comparisons of treatments based on the regression coefficients presented in Table 4.

Table E.2: Individual-level factorial regression analysis with rank

	(1) Rel. dev. $c_{2\text{–}19}$	(2) Rel. dev. c_{20}	(3) $k_{2\text{–}5}$	(4) $k_{6\text{–}10}$	(5) $k_{11\text{–}15}$	(6) $k_{16\text{–}20}$
Effort	0.095	0.031	7.419	−6.701	−33.274	−59.316
	(0.10)	(0.18)	(10.43)	(20.54)	(31.28)	(43.90)
Gap250	0.064	−0.377**	34.577***	90.774***	174.468***	310.455***
	(0.10)	(0.19)	(10.55)	(20.99)	(32.54)	(44.01)
rank	0.058***	0.018	−13.496***	−20.402***	−27.915***	−33.483***
	(0.02)	(0.05)	(3.22)	(4.46)	(6.19)	(8.05)
Effort × rank	−0.015	−0.018	−3.267	−0.568	3.312	8.672
	(0.02)	(0.05)	(2.48)	(4.45)	(6.54)	(9.14)
Gap250 × rank	−0.012	0.070	−7.308***	−18.085***	−33.327***	−57.792***
	(0.02)	(0.05)	(2.59)	(4.67)	(6.93)	(9.30)
H-type	0.283***	0.327	−22.424*	−10.102	10.468	40.449
	(0.10)	(0.22)	(13.09)	(20.51)	(30.45)	(40.32)
H-type × rank	0.007	−0.054	−6.827*	−12.219**	−20.547***	−29.320***
	(0.03)	(0.06)	(3.71)	(5.64)	(7.86)	(9.90)
Male	−0.049	−0.005	9.388**	11.947	11.476	9.734
	(0.05)	(0.08)	(4.63)	(8.49)	(12.30)	(15.79)
Constant	−0.114	−0.026	61.959***	96.256***	132.953***	136.591***
	(0.11)	(0.22)	(14.30)	(22.28)	(31.57)	(40.29)
Observations	3402	189	756	945	945	945
Adjusted R^2	0.030	−0.008	0.402	0.465	0.546	0.606

Notes: The dependent variable in Column (1) is the average relative deviation from conditionally optimal consumption between periods 2 and 19, and that in Column (2) is the relative deviation in the final period. The dependent variable in Columns (3) to (6) is the interval average of wealth accumulations, net of initial wealth. The dummy variables "Effort" and "Gap250" take the value 1 (0) for the Effort (Random) condition and the Gap250 (Gap50) condition, respectively. The variable "rank" is each participant's within-group wealth ranking, which is presented to the participant at the end of the previous period. A larger value indicates a lower level of wealth. Decisions made in the first period are not included in the analysis because the ranking is not presented prior to decision-making. The variable "H-type" is a dummy indicating that the participant's initial wealth type is H, and "Male" is a dummy indicating that the participant is male. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors clustered at the individual level are shown in parentheses.

Table E.3: Group-level regression analysis on the wealth inequality with gender control (dummy for high male ratio)

	Interval 1–5	Interval 6–10	Interval 11–15	Interval 16–20
Effort-				
NoRank-Gap50	–6.400	–28.603	–42.696	–58.248
	(11.69)	(25.17)	(38.39)	(55.59)
NoRank-Gap250	23.404	50.871	41.513	4.589
	(14.29)	(30.76)	(46.92)	(67.95)
Rank-Gap50	–8.617	–13.644	–15.816	–22.952
	(12.08)	(26.00)	(39.65)	(57.42)
Rank-Gap250	39.401***	79.328***	119.087***	174.075***
	(12.58)	(27.07)	(41.29)	(59.79)
Random-				
NoRank-Gap50	–10.965	–25.306	–31.169	–33.789
	(13.62)	(29.31)	(44.71)	(64.75)
NoRank-Gap250	3.950	32.870	92.670**	186.644***
	(13.11)	(28.23)	(43.06)	(62.35)
Rank-Gap50	1.205	3.426	8.118	–0.386
	(14.29)	(30.76)	(46.92)	(67.95)
Rank-Gap250	4.661	22.752	69.154	139.877**
	(13.04)	(28.07)	(42.82)	(62.01)
High male ratio	4.886	16.446	22.127	28.454
	(9.15)	(19.69)	(30.03)	(43.49)
Observations	59	59	59	59
Adjusted R^2	0.152	0.214	0.247	0.309

Notes: The dependent variable is the within-group normalized wealth inequality between the two types. The independent variables are dummy variables representing the eight experimental treatments, and the models exclude a constant term. The variable "High male ratio" is a dummy indicating that the proportion of male participants within a group exceeds 50%. Groups are excluded from the regression analysis when it is impossible to determine whether the male ratio exceeds 50% because some participants did not report their gender in a binary format. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

Table E.4: Group-level regression analysis on the wealth inequality with gender control (male ratio)

	Interval 1–5	Interval 6–10	Interval 11–15	Interval 16–20
Effort-				
NoRank-Gap50	−10.350	−42.193	−60.546	−80.133
	(16.44)	(35.04)	(53.74)	(77.95)
NoRank-Gap250	14.770	25.922	7.854	−40.349
	(18.66)	(39.77)	(60.98)	(88.46)
Rank-Gap50	−17.311	−39.336	−50.461	−68.940
	(16.73)	(35.66)	(54.69)	(79.33)
Rank-Gap250	30.653*	53.359	84.071	127.648
	(17.03)	(36.29)	(55.65)	(80.73)
Random-				
NoRank-Gap50	−20.075	−49.653	−64.828	−80.468
	(19.60)	(41.78)	(64.07)	(92.94)
NoRank-Gap250	−5.996	3.754	53.400	134.388
	(18.87)	(40.22)	(61.67)	(89.46)
Rank-Gap50	−9.220	−26.979	−32.892	−55.009
	(20.23)	(43.11)	(66.11)	(95.90)
Rank-Gap250	−5.746	−8.045	27.625	84.770
	(18.98)	(40.46)	(62.04)	(89.99)
Male ratio	19.798	60.292	81.259	107.049
	(22.06)	(47.02)	(72.10)	(104.59)
Observations	57	57	57	57
Adjusted R^2	0.158	0.224	0.254	0.315

Notes: The dependent variable is the within-group normalized wealth inequality between the two types. The independent variables are dummy variables representing the eight experimental treatments, and the models exclude a constant term. The variable "Male ratio" represents the proportion of men in each group. Groups that include participants whose gender could not be identified in binary terms were excluded from the regression analysis. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

Table E.5: Pairwise comparison of the regression coefficients based on the results reported in Table 4

(a) Periods 1–5

	Effort- NoRank-Gap50	NoRank-Gap250	Rank-Gap50	Rank-Gap250	Random- NoRank-Gap50	NoRank-Gap250	Rank-Gap50
Effort-							
NoRank-Gap250	35.420						
Rank-Gap50	−1.607	−37.027					
Rank-Gap250	43.889*	8.469	45.496*				
Random-							
NoRank-Gap50	−5.808	−41.227*	−4.201	−49.697*			
NoRank-Gap250	12.182	−23.237	13.789	−31.707	17.990		
Rank-Gap50	9.453	−25.966	11.060	−34.436	15.261	−2.729	
Rank-Gap250	17.602	−17.818	19.209	−26.287	23.410	5.420	8.149

(b) Periods 6–10

	Effort- NoRank-Gap50	NoRank-Gap250	Rank-Gap50	Rank-Gap250	Random- NoRank-Gap50	NoRank-Gap250	Rank-Gap50
Effort-							
NoRank-Gap250	93.118*						
Rank-Gap50	17.015	−76.103					
Rank-Gap250	107.591**	14.473	90.576*				
Random-							
NoRank-Gap50	3.652	−89.465*	−13.362	−103.939**			
NoRank-Gap250	67.641	−25.477	50.626	−39.950	63.989		
Rank-Gap50	35.669	−57.449	18.654	−71.922	32.017	−31.972	
Rank-Gap250	65.019	−28.099	48.004	−42.572	61.367	−2.622	29.350

(c) Periods 11–15

	Effort- NoRank-Gap50	NoRank-Gap250	Rank-Gap50	Rank-Gap250	Random- NoRank-Gap50	NoRank-Gap250	Rank-Gap50
Effort-							
NoRank-Gap250	108.252						
Rank-Gap50	29.646	−78.605					
Rank-Gap250	164.531*	56.279	134.885*				
Random-							
NoRank-Gap50	14.010	−94.242	−15.636	−150.521*			
NoRank-Gap250	143.664*	35.413	114.018*	−20.867	129.654*		
Rank-Gap50	57.429	−50.822	27.783	−107.102	43.419	−86.235	
Rank-Gap250	127.225*	18.973	97.578	−37.306	113.215*	−16.439	69.795

(d) Periods 16–20

	Effort- NoRank-Gap50	NoRank-Gap250	Rank-Gap50	Rank-Gap250	Random- NoRank-Gap50	NoRank-Gap250	Rank-Gap50
Effort-							
NoRank-Gap250	102.358						
Rank-Gap50	38.853	−63.505					
Rank-Gap250	241.617**	139.258	202.763**				
Random-							
NoRank-Gap50	27.255	−75.103	−11.598	−214.361**			
NoRank-Gap250	255.562**	153.204*	216.709**	13.946	228.307**		
Rank-Gap50	72.821	−29.537	33.968	−168.796*	45.566	−182.742**	
Rank-Gap250	214.873**	112.515	176.020*	−26.743	187.618**	−40.689	142.053

Notes: The value in each cell is obtained by subtracting the coefficient of the column treatment from that of the row treatment. ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively, based on Wald tests. The p-values used to assess statistical significance are adjusted for multiple comparisons using the Benjamini–Hochberg procedure controlling the false discovery rate at 0.05.