ISSN (Print) 0473-453X
Discussion Paper No. 1147
ISSN (Online) 2435-0982

THE IMPACT OF ASSET PURCHASES IN AN EXPERIMENTAL MARKET WITH CONSUMPTION SMOOTHING MOTIVES

Jieyi Duan
Nobuyuki Hanaki

Revised September 2022
November 2021

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

The impact of asset purchases in an experimental market with consumption smoothing motives*

Jieyi Duan† Nobuyuki Hanaki‡

September 19, 2022

Abstract

We investigate the effect of preannounced market intervention on an asset price as well as participants' welfare in an experimental framework where participants have consumption smoothing motives to trade the asset. The results show that, on one hand, the preannounced intervention results in significantly larger overpricing of the asset relative to the rational expectations equilibrium level in periods prior to the intervention compared with the treatment without it. The participants' welfare, measured by the discounted sum of the payoffs at the beginning of the experiment, on the other hand, are not significantly worsened by the intervention.

JEL Code: C90, D84

Keywords: Asset Pricing, Consumption Smoothing, Bounded Rationality, Quantitative Easing

*The experiments reported in this paper has been approved by IRB of the Institute of Social and Economic Research, Osaka University. We gratefully acknowledge financial support from grant-in-aid for scientific research (KAKENHI, grant number: 18K19954, 20H05631) from the Japan Society for the Promotion of Science (JSPS) as well as support from the Joint Usage/Research Center at ISER, Osaka University. We thank Daiki Maeda and participants at the international workshop on lab and field experiments (Osaka), 2021 ESA Global Online Meeting for comments and discussion.

†Faculty of Economics and Management, Kyoto University of Advanced Science. E-mail: duan.jieyi@kuas.ac.jp

‡Institute of Social and Economic Research, Osaka University. E-mail: nobuyuki.hanaki@iser.osaka-u.ac.jp

1 Introduction

The biggest monetary policy experiment in recent years has been quantitative easing (QE). Between 2008 and 2017, seven central banks (Mexico, England, Japan, Europe, US, Switzerland, and Sweden) have employed this policy (Committee on the Global Financial System, 2019). Furthermore, QE has been extended in duration in many countries in response to the economic downturn resulting from the spread of Covid-19. For example, since March 2020, the US Federal Reserve has implemented an unlimited QE policy.

However, from the perspective of the standard textbook model with complete markets without frictions and infinitely lived rational decision makers, this policy should have no effects on any macroeconomic variables (Eggertsson and Woodford, 2003). Nevertheless, a number of empirical studies provide evidence that QE has indeed affected the market price of financial assets. For instance, Corbet et al. (2019) find the Federal Open Market Committee announcement has significantly and substantially increased stock market volatility, McLaren et al. (2014) and Grosse-Rueschkamp et al. (2019) indicate that the QE announcements from central banks has reduced bond yields, implying there exists a QE announcement effect. In particular, D'Amico and King (2013) and McLaren et al. (2014) observe a "local supply" effect, i.e., the yield curve within a particular maturity sector has responded more to changes in the total value of outstanding bonds in that sector than to similar changes in the other sectors, indicating a segmentation or an imperfect substitution within the Treasury bond market.

Penalver et al. (2020) offer, based on an experimental analysis, a perspective based on the limited rationality of investors *a la* level-K (Nagel, 1995) or the cognitive hierarchy (Camerer et al., 2004) model; namely, a possibility of some market participants (maybe naively) believing that QE will raise the prices of bonds because QE is a commitment by the central bank, who is an unusual participant to the market, to buy a large amount of bonds. Such a possibility would indeed result in QE raising the prices of bonds because other, more sophisticated, participants optimally respond to the existence of such naive participants.[1]

The experiment of Penalver et al. (2020) is similar to that of Bostian and Holt (2009) and extends the framework of Smith et al. (1988).[2] In the Benchmark setting, participants trade a set of risk-free bonds using experimental currency (cash) for a known finite number of periods. Bonds and cash, when carried over to the next period, generate a dividend and an interest income, respectively. The bond matures at the end of the final trading period and is converted into cash. The interest on cash, the dividend, and the value of the bond at maturity are set so that the fundamental value of the bond is constant, and equal to the value of the bond at maturity in all trading

[1] Farhi and Werning (2019) analyze, theoretically, an implication of similar limited rationality on the effects of monetary policy in the New Keynesian setup. Namely, the authors show that level-k thinking together with incomplete credits markets, idiosyncratic risks that cannot be insured, and borrowing constraints that bind occasionally, can potentially explain the "forward guidance puzzle"; i.e., monetary policy being too effective.

[2] See, among others, Palan (2013) and Powell and Shestakova (2016), for surveys of the large literature on the experimental asset market pioneered by Smith et al. (1988).

periods. In the treatment with QE intervention, after a preannounced number of periods, the computer (acting as a QE operator) intervenes to buy a prespecified quantity of bonds from participants through a discriminatory auction. Penalver et al. (2020) find that such a market intervention not only raises the postintervention prices of bonds, but also, after participants experience it once, significantly raises *preintervention* market prices when the experiment is repeated.[3]

The effect of market intervention observed in the experiments by Penalver et al. (2020) and Haruvy et al. (2014) are observed in an environment where no trade is predicted under a risk neutral rational expectations equilibrium (REE, for short). As reported by Lei et al. (2001), in such an environment, asset prices tend to deviate from their fundamental value more because the main rationale for participants to trade the asset is to speculate. The same conclusion is also obtained by Crockett et al. (2019), who find the mispricing of assets is significantly decreased when the speculate motivation is removed by giving participants an incentive to trade, such as consumption smoothing motive. Hence, the effect of intervention observed by Penalver et al. (2020) and Haruvy et al. (2014) may be exaggerated compared with situations where participants have nonspeculative motives to trade the asset.

In this study, therefore, we reexamine the effect of preannounced mar-

[3]Haruvy et al. (2014) also show, in a similar but a different experimental design, that assets being (re-)purchased by computer from market participants raises their prices. Haruvy et al. (2014) did not investigate, however, how participants experiencing the effect of (repurchase) intervention once influences the preintervention price by repeating the experiment.

ket intervention considered by Penalver et al. (2020) in a new experimental framework introduced by Asparouhova et al. (2016) and Crockett et al. (2019), where participants trade assets to smooth their consumption across periods. This new experimental framework is designed based on the Lucas asset pricing model (Lucas Jr, 1978)[4], and has the following characteristics: (1) there are an indefinite number of periods, instead of a known finite number of periods; (2) participants receive incomes, which can be used to consume or to trade the assets, that fluctuate across periods; and (3) participants are paid based on their consumption at the end of periods, instead of the final value of their portfolio at the end of the last period. In this setting, the asset will be traded to smooth consumption in the REE.

We investigate, in this new framework, the effect of preannounced market intervention (which we call QE operations below), as well as the effect of the existence of an additional method of consumption smoothing; i.e., the possibility to directly save cash. Compared with Penalver et al. (2020) and Haruvy et al. (2014), such a framework reduces the effect of the speculate motivation which may exaggerate the impact of market intervention in the

[4]To our knowledge, only three experimental studies adopted this model for different purposes besides our research. Asparouhova et al. (2016) is the first research to adopt the experiment based on the Lucas model. This research aims to examine the model's features in the laboratory. Compared with Asparouhova et al. (2016), Crockett et al. (2019) aims to clarify the relationship between the existence of trading motivation and mispricing through modifying the participants' exchange function. Finally, the most recent research is Carbone et al. (2021), which extends the framework of Crockett et al. (2019) from a two-period cyclical world to a three-period cyclical world, aiming to test the robustness of the features of Lucas model and the results of Crockett et al. (2019). Appendix A shows a specific comparison of the experimental setting of these studies, including this research.

laboratory.

Our data show that the existence of QE operations significantly increase the magnitude of overpricing of the asset (which we call bonds below) relative to the REE level regardless of the existence of a saving possibility even before participants having experienced the effect of the QE operation. However, the participants' welfare, measured by the discounted sum of per period payoffs, are not significantly worsened by the larger mispricing of assets caused by the preannounced intervention.

The rest of the paper is organized as follows. Section 2 describes the model and experimental design. Section 3 outlines the theoretical predictions and hypotheses. Section 4 provides an analysis of the data. Finally, Section 5 offers some conclusions.

2 Experiments

2.1 Basic experimental design

Our experimental design in the Benchmark (B) treatment builds upon the previous experiments by Crockett et al. (2019) and Asparouhova et al. (2016). Appendix A compares experimental settings between ours and other studies based on the Lucas asset pricing model. We consider an indefinite horizon economy with a nonstorable consumption good and an indefinitely lived asset. In the experiment, the nonstorable consumption good and the indefinitely lived asset are represented by the experimental currency (francs) and the bond, respectively. Time is discrete. At the beginning of the economy ($t = 1$), each trader is endowed with some units of bond and some francs, which they can use to trade among themselves. Subsequently, at the beginning of period $t \geq 2$, each trader receives y_t^i francs of income as well as $D = 2$ francs per unit of bonds they hold as the dividend payment. While the aggregate income, $\sum_i y_t^i$, is constant across periods, individual income, y_t^i, fluctuates between odd- and even-numbered periods. Let y_o^i and y_e^i denote trader i's income in odd- and even-numbered periods, respectively.

Specifically, in the experiment, there are eight traders in an economy. These eight traders are equally divided into two types of traders, Type 1 and Type 2, who differ in their endowment and income streams. The endowments are 92 francs and one unit of bonds for Type 1 traders and 32 francs and four units of bonds for Type 2 traders, respectively. Incomes are set so that $y_o^1 = y_e^2 = 90$ francs and $y_e^1 = y_o^2 = 24$ francs, where superscripts denote trader's type.

At the end of each period ($t \geq 1$), each trader consumes the francs he/she has. Consumption, c, is converted into JPY (Japanese Yen) according to the following increasing and strictly concave function:

$$u_{JPY}(c) = 3573.50 - 64872.01 \times (c + 40.5)^{-0.7478}. \tag{1}$$

In the experiment, the "consumption" occurs as follows: each participant's end-of-period franc balance is forcibly converted into JPY through the exchange function shown in Eq. 1.

Once consumption takes place, the economy continues to the next period with probability $\pi = 5/6$ or ends with probability $1 - \pi = 1/6$. (Appendix B gives a visual description of the timeline of each period for all treatments.) Recall that while francs are nonstorable and cannot be carried over to the next period, the bond is storable and can be carried over to the next period. However, if the economy ends, all the bond holdings are lost without any compensation. In the experiment, participants, who are acting as traders, are paid based on *their final consumption before the economy ends* according to a prespecified conversion rate (Eq. 1). The risk neutral fundamental value of a unit of bonds in period t is $FV_t = D\pi/(1 - \pi) = 10$ francs for all t.

Under this experimental setting, traders can obtain a higher expected payoff by smoothing their consumption across periods. Note, however, to smooth the consumption, traders need to trade bonds. Thus, unlike many experiments employing the framework of Smith et al. (1988) including Haruvy et al. (2014) and Penalver et al. (2020), trade will take place in the current experimental setup under the REE. Crockett et al. (2019) demonstrated that such a consumption smoothing motive of trading reduces the price deviation from the fundamental value.

2.2 Trading mechanism

The experiment adopts an open-book continuous double-auction mechanism. The trading period lasts 120 seconds. As noted above, at the beginning of period t, before the market opens, trader i who hold K_t^i units of bonds receives $y_t^i + 2K_t^i$ units of cash (francs) as his/her income and the dividend payment.

Once trading begins, traders can submit a bid order (a buy order) and/or an ask order (a sell order) for a unit of bonds in continuous time. Traders can trade as many units of bonds as they wish during the 120 seconds within their budget constraint. No borrowing of cash or short-selling of bonds is permitted. Orders are sorted according to price and the time of the order submission. A transaction takes place when the best bid and the best ask cross, at the price determined by whichever is submitted earlier. Once a transaction takes place, cash and bond holdings are immediately updated and all the outstanding orders submitted by the two traders who just traded are automatically canceled.

2.3 Treatments

We use a two-by-two between-subjects design, in which we vary the existence of QE operations (with v.s. without QE operations), and the existence of an additional way of smoothing consumption (with v.s. without the possibility of saving francs). Table 1 summarizes these four treatments. And Appendix B shows the differences in the experiment process in each period among treatments.

Regarding the QE operations, before the experiment begins, participants are informed that, just before the trading in period 4 and after participants

Table 1: Four treatments

	Without saving	With saving
without QE	B	BS
with QE	QE	QES

receive their income and dividend, a maximum of six units of bonds will be bought by the computer through a uniform price auction. During this intervention, participants who want to sell their bonds to the computer, place an order by specifying the minimum price at which they wish to sell their bonds and the number of units they wish to sell.

Then, the computer buys up to six units of bonds from the lowest-priced orders from the participants in each group. The transaction price of the intervention is the highest price among those accepted orders in the group.

In the treatments with saving (BS and QES), francs become storable. Namely, just after the trading in each period, participants can determine how much of their remaining francs to save and to carry over to the next period. The saved francs generate an interest earning with a 20% interest rate, which is paid at the beginning of the next period. If the economy ends, however, just as the bond is lost, the saved francs are lost. Interest earnings are rounded down to the nearest integer. We set the interest rate at 20% to make it equal to the return of bonds at the rational expectations equilibrium.

2.4 Termination, timing, and payment

To secure sufficient data and to ensure that market intervention occurs in the QE and QES treatments, we adopted the method of block random termination (BRT, Fréchette and Yuksel, 2017). Similarly to the standard random termination, under BRT, the computer rolls a die at the end of each period to determine whether an economy (called a round) ends or not. However, participants must at least experience the round for a fixed number of periods (one block). Whether the round has ended or not is disclosed only at the end of each block. If the round has ended during the current block, participants are told in which period it has happened, and their corresponding payoffs at that period. Recall that participants' payoff (for the round) is computed based on their consumption at the final period of the round. Otherwise, they are told that the round has not ended yet, and they start a new block. In this experiment, at the end of each period, the round ends with a probability 1/6, and each block consists of six periods. Appendix C summarizes this procedure.

At the end of a round, if fewer than 35 minutes have passed since the beginning of the first round, a new round begins; otherwise, the experiment ends. Therefore, each session of the experiment consists of at least one round. Furthermore, each round consists of at least one block of six periods. At the start of each round, the endowments of all participants are reset. After the experiment ends, the computer randomly selects one round to calculate the participant's earning. Participants are paid in cash based on their earnings in the chosen round in addition to the 1000 JPY show-up fee.

To exclude the effect from variations in duration of rounds, and to control the length of the experiment, we use a computer program to randomly determine the number of periods (thus, the number of the block) in each round in advance.[5] Then, we used this predetermined number of periods across rounds for all the sessions in our experiment. Therefore, each round has the same number of periods (blocks) in all the sessions, and all the sessions consist of at least two rounds in our experiment. In particular, all of the rounds in the four treatments consist of one block of six periods, and each session of all the treatments except for B consists of two rounds, while each session of B consists of three rounds. We did not provide participants with this information.[6]

3 Hypothesis

Let us derive the REE bond price under the representative agent framework. At period 0, the agent faces the following maximization problem.

$$\max_{\{c_t,k_{t+1}\}_{t=0}^{\infty}}[\sum_{t=0}^{\infty}\pi^t v(c_t)] \tag{2}$$

[5]Although there is so far no evidence showing that variations in duration of rounds affect asset pricing in the macroeconomics experiment, some studies document that such variations indeed affect the experimental results. For instance, Bo and Fréchette (2011) and Engle-Warnick and Slonim (2006) report that such variations can affect the extent of cooperative behavior in repeated prisoner's dilemma game experiments and repeated trust game experiments.

[6]Such a design is also adopted by Duffy and Puzzello (2014), Duffy and Puzzello (2020), and Fréchette and Yuksel (2017).

subject to

$$y_t + (D + P_t)k_t = c_t + P_t k_{t+1}, \tag{3}$$

where c_t is the consumption of francs, P_t is the bond price, and k_t is the number of the bonds that the agent owns at the beginning of period t. $\pi \in (0, 1)$ is the continuation probability. $v(c_t) = (1 - \pi)u_{JPY}(c_t) + \pi 0 = (1 - \pi)u(c_t)$ is the expected monetary reward in period t by "consuming" c_t.

We define the Lagrangian as follows:

$$L = \sum_{t=0}^{\infty} \left(\pi^t v(c_t) + \lambda_t \left(y_t + (D + P_t)k_t - c_t - P_t k_{t+1} \right) \right). \tag{4}$$

The first order condition with respect to c_t is

$$\frac{\partial L}{\partial c_t} = \pi^t v'(c_t) - \lambda_t = 0 \text{ for all } t. \tag{5}$$

The first order condition with respect to k_{t+1} is

$$\frac{\partial L}{\partial k_{t+1}} = -\lambda_t P_t + \lambda_{t+1}(D + P_{t+1}) = 0 \text{ for all } t. \tag{6}$$

Combining Eq. 6 and Eq. 5 for c_t and c_{t+1} gives us

$$\pi^t v'(c_t) P_t = \pi^{t+1} v'(c_{t+1})(D + P_{t+1}).$$

Thus,

$$P_t^* = \frac{v'(c_{t+1})}{v'(c_t)} \pi (P_{t+1} + D), \tag{7}$$

where P_t^* is the equilibrium price of the bond at period t. By applying the law of iterated expectations, Eq. 7 can be rewritten as

$$P_t^* = \sum_{\tau=1}^{\infty} \pi^{\tau} \frac{v'(c_{t+\tau})}{v'(c_t)} D, \tag{8}$$

with $v'(c_t) = (1 - \pi)u'_{JPY}(c_t)$ for all t. As $u_{JPY}(c_t)$ is strictly concave, in treatment B, consumption satisfies $c_1 = c_2 = ... = c_t = ...$ Thus, $\frac{v'(c_{t+\tau})}{v'(c_t)} = 1$. Therefore,

$$P_t^{*,B} = \sum_{\tau=1}^{\infty} \pi^{\tau} D = \frac{\pi D}{1 - \pi} = FV \text{ for all } t. \tag{9}$$

We now consider the REE price of the QE treatment. In the experiment, the QE operation occurs at the beginning of period 4. Because up to six units of bonds are purchased during the QE operation, the dividend loss of the purchased bonds leads to a decrease in consumption after period 4. Thus,

$$c_4 > c_1 = c_2 = c_3 > c_5 = c_6 = ... \tag{10}$$

and

$$v'(c_{t|t\geq 5}) > v'(c_{t|t\leq 3}) > v'(c_4). \tag{11}$$

Given that consumption will be constant, in the equilibrium, after period 5, we have

$$P_{t|t\geq 5}^{*,QE} = FV. \tag{12}$$

The equilibrium prices in period 4 are

$$P_4^{*,QE} = \frac{v'(c_5)}{v'(c_4)}\pi(FV + D) = \frac{v'(c_5)}{v'(c_4)}FV, \tag{13}$$

which depends on c_4; i.e., the outcome of the QE operation. The equilibrium price in period 3 is

$$P_3^{*,QE} = \pi\frac{v'(c_4)}{v'(c_3)}(P_4^{*,QE} + D) = \pi\frac{v'(c_5)}{v'(c_3)}FV + \pi\frac{v'(c_4)}{v'(c_3)}D. \tag{14}$$

In period 4, if the agent responds to the QE operation, the expected payoff must be no less than when the agent can expect from not doing so. This means that the consumption stream must satisfy the following inequality

$$\sum_{\tau=1}^{\infty} \pi^{\tau}[v(c_3) - v(c_{**})] \leq v(c_4) - v(c_3), \tag{15}$$

where c_{**} is the optimal per period consumption from period 5 on.

The left part of the above inequality is the discounted future loss from permanently lower income (because of lost dividend income) when the agent responds to the QE operation, and the right part is the short-term gain in welfare from doing so. If the competition among agents drives the QE price down to the point of indifference, this implies

$$\frac{v(c_3) - \pi v(c_{**})}{1 - \pi} = v(c_4). \tag{16}$$

This identifies the equilibrium QE intervention price. Although we could not derive the relationship between $P_3^{*,QE}$ and FV under general conditions, we can at least numerically derive the specific value of $P_3^{*,QE}$ for the specific parameter used in our experiment. Assuming, consistent with the representative agent assumption, that all the agents have the same level of consumption in the REE, when the computer purchases six units of bonds in the QE operation, we $c_{t|t\leq 3} = 62$ and $c_{t|t\geq 5} = 60.5$. Moreover, substituting these values in Eq. 16, we have $c_4 = 70.1216$ under the representative agent assumption (so that each agent sells 6/8 units of bonds to the computer during the QE operation), with the equilibrium intervention price being

$$p_{inter}^{*,QE} = (70.1216 - 62)/\frac{6}{8} = 10.82. \tag{17}$$

With $u_{JPY}(c_t)$ and $P_5^{*,QE} = FV = 10$, we obtain

$$P_{t=3}^{*,QE} = 10.00959 > FV = P_{t=3}^{*,B}. \tag{18}$$

Given Eq. 7, this means

$$P_{t=3}^{*,QE} > P_{t=2}^{*,QE} > P_{t=1}^{*,QE} > 10; \tag{19}$$

in particular, $P_{t=2}^{*,QE} = 10.005$, $P_{t=1}^{*,QE} = 10.0042$. Thus, the announced QE operation slightly raises the prices in preintervention periods under the REE.

In the BS and QES treatments, because participants can carry over their francs into the next period by saving with an interest rate of 20%, a 10 francs saving can perfectly substitute for holding one unit of bonds when the market price is FV. Thus, there is no reason that participants buy or sell the bonds at a market price higher or lower than FV under the REE. Furthermore, in the QES treatment, for the same reason, the intervention price converges to FV when the QE operation is fully competitive.

While the REE bond prices are (slightly) different across the four treatments, we hypothesize, based on the REE,

Hypothesis 1 *The magnitudes of mispricing are the same across the four treatments.*

Alternatively, if, as proposed by Penalver et al. (2020), some participants naively anticipating profits to be made during the QE operation causes prices to deviate from the REE, the magnitude of mispricing would be larger in the treatments with QE than those without, especially, in the periods prior to the QE operation.

Furthermore, because of the competition during the QE operation, we expect, based on the REE,

Hypothesis 2 *The ex ante expected payoffs are the same across the four treatments.*

Alternatively, while QE operations may raise the consumption level in period 4 (and if saving is possible in the later periods) and thus make participants better off, speculative trades in periods prior to it may disturb consumption smoothing and thus may make participants worse off. The overall effect is, however, not clear and the ex ante expected payoff may differ across treatments.

4 Results and Discussions

The experiment was programmed using z-Tree (Fischbacher, 2007) and was run at the experimental laboratory of the Institute of Social and Economic Research at Osaka University from October 2020 to June 2021. All participants were students enrolled in the school and recruited by the ORSEE recruiting system (Greiner, 2015).

In total, 304 students participated in the experiment across 13 sessions. In all the sessions, the instruction movie was played to the participants. The participants had a printed handout at hand. The participants' understanding of the rules of the experiment, including how their payoffs are computed, was checked with a quiz. To ensure that participants understood the rules, the experiment started only after all the participants had answered all the questions correctly. There are nine groups in Benchmark, QE, and BS treatments and 11 Groups in the QES treatment.[7] The sessions lasted between

[7]In the QES treatment, however, the computer failed to purchase the bonds in three groups in one session because of a programming mistake. Subsequently, the error in the program was corrected, and we excluded the data of these three groups from the analyses. Thus, data from 35 groups (nine groups each in the Benchmark, QE, and BS treatments, and eight in the QES treatment) are used for subsequent analysis.

one and a half to two hours.[8] The average payoff was 2456 JPY ($\approx$ 22.99 USD, based on the exchange rate at the time experiments were conducted).

4.1 Price dynamics and mispricing

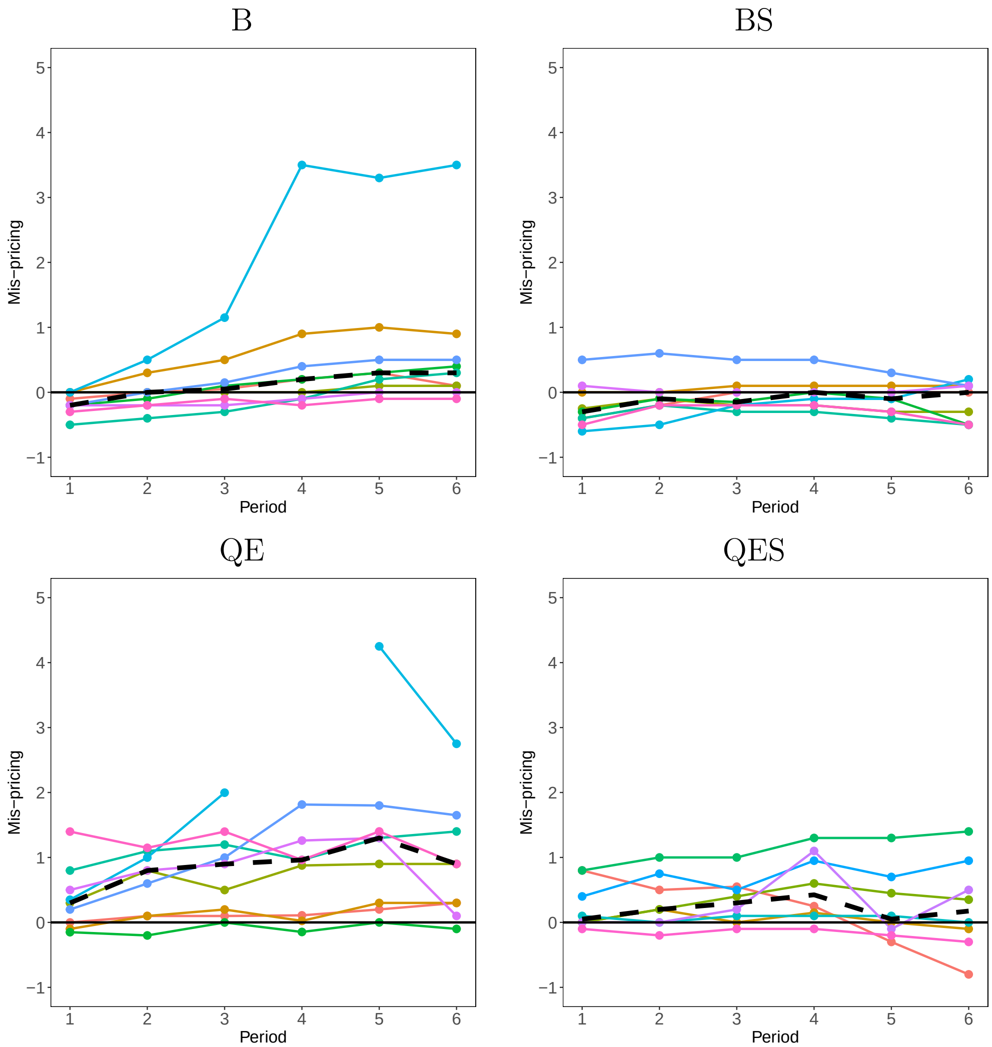
Did the intervention affect the bond transaction prices? Figures 1 and 2 show the dynamics of mispricing observed in each treatment in Rounds 1 and 2, respectively. The mispricing in period t of round r of group g is calculated by $MP_t^{g,r} = \frac{1}{N_t^{g,r}} \sum_n^{Ng,r} (P_t^{g,r,n} - p_t^{*,r})/p_t^{*,r}$, where $N_t^{g,r}$ is the number of transactions in the group for period t of round r, $P_t^{g,r,n}$ is the realized price of the n-th transaction in period t of round r for the group, and $p_t^{*,r}$ is the REE price in period t of round r for the treatment as derived in the previous section.

Each solid line shows the dynamics of the within period median mispricing of a group. The dashed lines correspond to the dynamics of the across-group median for the treatments. As mentioned in Section 2.4, because all sessions adopted the same sequence of predetermined dice numbers, the duration of each round is the same. Precisely, each round consists of one block (six periods).

Figure 1 reveals that, in Round 1, mispricing tends to be lower in B compared with QE, and in BS compared with QES. In particular, this tendency is observed not only after the QE operation that takes place in period 4

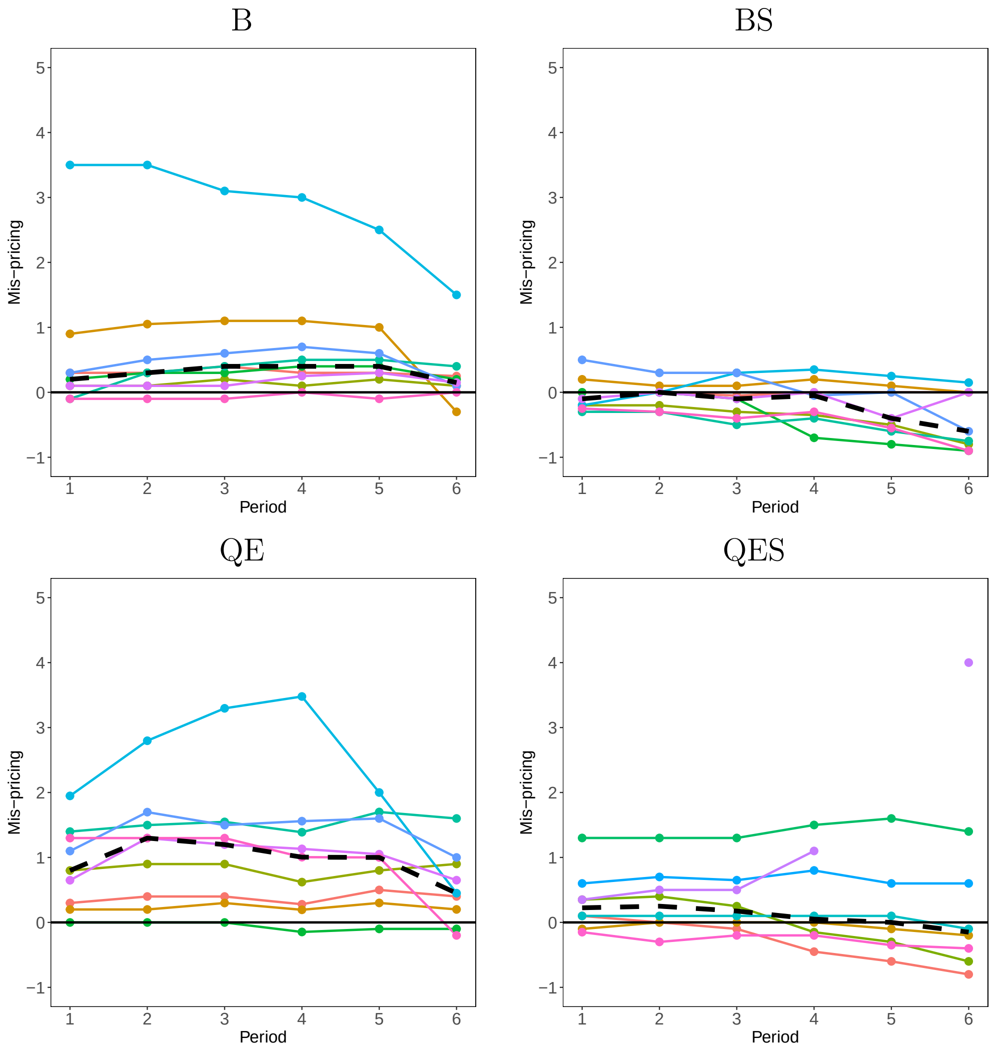
[8] We recruited participants for two hours, but our sessions all ended within two hours, so as to avoid any possible end game effects.

Figure 1: Dynamics of mispricing: Round 1



Note: Each line represents the dynamics of the normalized median prices of a group. The dashed lines represent the dynamics of the median normalized prices for the treatment in the round.

Figure 2: Dynamics of mispricing: Round 2



Note: Each line represents the dynamics of the normalized median prices of a group. The dashed lines represent the dynamics of the median normalized prices for the treatment in the round.

(periods 4–6), but also before it takes place (periods 1–3). Furthermore, the mispricing in QES tends to be smaller than that in QE. Between B and BS, a part from the latter half of the round, the mispricings are similar. A similar tendency is observed in Round 2, shown in Figure 2, except that mispricing in BS tends to be smaller than the one in B.

To formally compare the degree of mispricing, we compute the geometric deviation (GD) introduced by Powell (2016).[9] For group g in round r, $GD^{g,r}$ is defined as

$$GD^{g,r} = \exp\left(\frac{1}{N^{g,r}}\sum_{n=1}^{N^{g,r}}\frac{p_n^{g,r}}{p_n^*}\right) - 1, \tag{20}$$

where $N^{g,r}$ is the number of transactions that occurred in round r, $p_n^{g,r}$ is the realized price of the n-th transaction, and p_n^* are the REE prices in the period when the n-th transaction occurred. As GD takes the direction of price deviation into account, the bonds are overpriced (underpriced) relative to the REE prices if GD is significantly higher (lower) than 0.

We calculate the GDs for two subperiods separately: periods 1–3 denote the preintervention periods, and periods 4–6 denote the postintervention periods. Table 2 summarizes the median GDs in Round 1 for each treatment in the preintervention and postintervention periods (top panel), as well as the $p-values$ of the pairwise treatment comparisons based on the Wilcoxon rank-sum test (bottom panel). These p-values are corrected for multiple comparisons using the Bonferroni method.

Table 2 shows that, in Round 1, the preintervention GDs in the QE treat-

[9]See Appendix D for results based on other measures of mispricing.

Table 2: Median GDs in Round 1

Median GDs

	B	QE	BS	QES
Preintervention	-0.052	0.467	-0.208	0.101
Postintervention	0.166	0.871	-0.054	0.252
p-values* (H_0: Pre = Post)	**0.004**	**0.074**	0.910	0.547
Observations	9	9	9	8

p-values$^+$ for treatment comparisons

Preintervention periods				Postintervention periods			
	B	QE	QES		B	QE	QES
QE	**0.047**	-	-	QE	0.969	-	-
QES	0.164	1.000	-	QES	1.000	0.833	-
BS	1.000	**0.034**	**0.047**	BS	0.113	**0.011**	1.000

$*$: $p - values$ from the Wilcoxon signed-rank (WSR) test are reported.
$+$: $p{-}values$ from the pairwise Mann–Whitney U (MW) test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.

ment are significantly higher than those in the B treatment ($p = 0.047$), and similarly, the preintervention GDs in the QES treatment are significantly higher than in the BS treatment ($p = 0.047$). Thus, unlike the results reported in Penalver et al. (2020), the announced intervention raises the degree of mispricing even before the intervention actually takes place, regardless of the existence of the saving possibility.

Observation 1 *In Round 1, the announcement of the QE intervention results in significantly larger overpricing of the bond in the preintervention periods than the treatments without the intervention.*

Moreover, although the median postintervention periods GDs are higher in the QE treatments than in the B treatments (and also in the QES treatments than in the BS treatments), as reported in the bottom panel of Table 2, these differences are not statistically significant. Thus, unlike in Penalver et al. (2020), the intervention does not significantly affect the magnitude of mispricing in the postintervention periods compared with the cases without it.

Observation 2 *In Round 1, the intervention does not result in a statistically significant difference in the magnitude of mispricing in the postintervention periods compared with the cases without the intervention.*

A potential reason for the absence of the significant differences in the postintervention GDs between the B and QE treatments is the significant increase in the GDs in the postintervention periods compared with the preintervention periods observed in the B treatment. We have, however, no clear explanation for the upward trend in the mispricing observed in this treatment.

Let us now turn to the outcomes of Round 2. Penalver et al. (2020) report that once the effect of intervention is experienced in Round 1, the bond is significantly overpriced in the preintervention periods in the treatment with preannounced intervention in Round 2.

Table 3 shows the median preintervention and postintervention GDs in Round 2 for each treatment. While the median preintervention GD is higher in the QE than in the B treatments (higher in the QES than in the BS treatments) as shown in the top panel, there is no statistically significant

Table 3: Median GDs in Round 2

Median GDs

	B	QE	BS	QES
Preintervention	0.275	0.860	-0.056	0.198
Postintervention	0.264	0.764	-0.331	-0.059
p-values* (H_0: Pre = Post)	0.654	**0.027**	0.129	0.742
Observations	9	9	9	8

p-values+ for treatment comparisons

Preintervention periods				Postintervention periods			
	B	QE	QES		B	QE	QES
QE	0.969	-	-	QE	0.815	-	-
QES	1.000	0.278	-	QES	1.000	1.000	-
BS	**0.047**	**0.002**	0.355	BS	**0.011**	**0.005**	1.000

$*$: $p - values$ from the Wilcoxon signed-rank (WSR) test are reported.
$+$: $p{-}values$ from the pairwise Mann–Whitney U (MW) test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.

difference in preintervention or postintervention *GDs* between the QE and the B treatments (and in the QES treatments and BS treatments) in Round 2. Thus, unlike the results of Penalver et al. (2020), the effect of intervention does not persist when participants repeat the same experiment with the same group of participants under the same market conditions.

It should be noted that preintervention *GDs* are significantly higher in Round 2 than in Round 1 for the QE treatment ($p = 0.008$, WSR test) suggesting that having experienced the high preintervention and postintervention market prices in Round 1, participants in the QE treatment traded the bond at an even higher price before the intervention in Round 2 than in Round 1. However, as noted above, probably because of the upward price trend observed in the B treatment in Round 1, its preintervention *GDs* in Round 2 are also significantly higher than those in Round 1 ($p = 0.004$, WSR test).[10] In Penalver et al. (2020), and many other experiments using the Smith et al. (1988) paradigm, the magnitude of mispricing decreases as participants gain experience in the baseline treatment. We believe that the increase in the magnitude of mispricing in Round 2 under the B treatment is the reason for the absence of the persistence of the QE intervention when the saving possibility is absent. We also note that in Round 2, the *GDs* are significantly higher in the B treatment than in the BS treatments both for the preintervention ($p = 0.047$, using the MW test) and for the postintervention ($p = 0.011$, using the MW test). Thus, in the absence of the intervention, the saving possibility significantly reduces the magnitude of overpricing among experienced participants.[11]

4.2 Welfare and Consumption Smoothing

Note that in our experiment, the participants' expected payoffs do not directly depend on the transaction price in each period but on their consumption paths across several periods. Here, we first compare participants' welfare, measured by the discounted sum of the payoffs over six periods (for par-

[10] There is no statistically significant difference in the preintervention *GDs* between Round 1 and Round 2 for the BS and QES treatments ($p = 0.496$ in the BS treatment and $p = 0.641$ in the QES treatment, using the WSR test). Furthermore, there is no significant difference (at the 5% significance level) in the postintervention *GD*s between Rounds 1 and 2 for all the treatments (p-values are 1.000, 0.570, 0.098, 0.461, for the B, QE, BS, and QES treatments, respectively, using the WSR test.

[11] See Appendix E for analyses on trading volumes.

ticipant i, it is defined as $w^i = \sum_{t=1}^{6} \pi^{t-1} u_{JPY}(c_t^i))$, across four treatments.[12] We use the within-group mean of w^i, W^g, as an independent observation.

Table 4 shows the across-group median W^g in the four treatments in two rounds. In Round 1, the W^gs are similar between the B and QE treatments ($p = 1.000$), but the W^g in the BS treatment is significantly smaller than that in the QES treatment ($p = 0.047$). Thus, despite the significantly larger preintervention mispricing observed in the QE and QES treatments than in the B and BS treatments, the discounted payoffs are not significantly worsened. On the contrary, in the presence of a saving possibility, the QE intervention increased welfare.

Table 4 also shows that the W^gs are significantly higher in Round 2 than in Round 1 at the 5% significance level in all the treatments except for QES. In Round 2, there is no longer significant differences in the W^gs between the BS and QES treatments.[13]

As noted above, the QE intervention may lower the payoffs in the preintervention period because of increased mispricing, while improving the postintervention payoffs because of the increased consumption induced by the cash injection. However, as reported in Appendix G, the anticipated intervention does not result in a significant change in (nondiscounted) mean payoffs either

[12]Here, we do not regard participant payment as welfare since it can't well reflect participant performance throughout the experiment, such as the performance of consumption smoothing.

[13]In Appendix F, we separately analyze discounted payoffs for two types. The results show that significant treatment differences in the discounted payoffs are mainly because of the variation in the discounted payoff of Type 2 players.

Table 4: Median W^g

Median W^g	B	QE	BS	QES
Round 1	5723.84	5747.33	5507.73	5798.26
Round 2	5851.31	5892.67	5838.08	5955.09
p-values* (H_0: R1 = R2)	**0.008**	**0.004**	**0.004**	0.055
Observations	9	9	9	8

p-values$^+$ for treatment comparisons

Round 1				Round 2			
	B	QE	QES		B	QE	QES
QE	1.000	-	-	QE	1.000	-	-
QES	1.000	1.000	-	QES	0.560	1.000	-
BS	0.240	0.302	**0.047**	BS	1.000	1.000	0.560

$*$: $p - values$ from the Wilcoxon signed-rank (WSR) test are reported.
$+$: $p - values$ from the pairwise Mann–Whitney U (MW) test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.

in the preintervention or postintervention periods at the 5% significance level in both rounds.

Observation 3 *The intervention does not significantly affect welfare, measured by the discounted payoffs, in Round 2. It improves it in the presence of a saving possibility in Round 1.*

We also compute the Gini coefficient of w^i for each group to investigate whether the QE intervention has increased the within-group inequality in the discounted payoffs. Table 5 shows the median Gini coefficient in each treatment. While there is no statistically significant difference between any relevant pairs of treatments, we do observe significant reduction in the Gini coefficient in Round 2 compared with Round 1 for the BS treatment. This re-

Table 5: Median Gini coefficient based on w^i

Median Gini coefficient

	B	QE	BS	QES
Round 1	0.030	0.044	0.068	0.050
Round 2	0.030	0.041	0.032	0.037
p-values* (H_0: R1 = R2)	0.91	0.652	**0.012**	0.641
Observations	9	9	9	8

p-values$^+$ for treatment comparisons

Round 1				Round 2			
	B	QE	QES		B	QE	QES
QE	1.000	-	-	QE	0.970	-	-
QES	1.000	1.000	-	QES	1.000	1.000	-
BS	0.110	0.460	1.000	BS	1.000	0.970	1.000

$*$: $p-values$ from the Wilcoxon signed-rank (WSR) test are reported.
$+$: $p-values$ from the pairwise Mann–Whitney U (MW) test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.

duction in the Gini coefficient in the BS treatment is because of the improvement in the discounted payoff of Type 2 participants in Round 2 compared with Round 1.[14]

Observation 4 *The intervention does not significantly affect within-group inequality in terms of their discounted payoffs.*

4.3 Prices in the intervention

As reported in Section 4.1, the magnitudes of the overpricing in the preintervention periods are significantly larger in the presence of intervention than

[14] In fact, while the mean discounted payoffs are significantly higher in Round 2 than in Round 1 for Type 2 players at the 5% significance level in all the treatments except for QES, there is no significant increase for Type 1 participants. See Appendix F.

Table 6: Comparison of the mispricing during the intervention between QE and QES treatments and rounds.

	QE	QES	$p-values$
Round 1	1.032***	0.850***	0.726
Round 2	1.309***	0.300**	0.525
$p-values$	0.195	0.125	
Observations	9	8	

Note: The median mispricing of each treatment in each round is reported. The fourth column reports the p-values of the differences between treatments from the Wilcoxon rank-sum test. The third row reports the p-values of the differences between rounds within treatments from the Wilcoxon signed-rank test. **, and *** indicate a significant difference from 0 at the 5 and 1% significance levels using the Wilcoxon rank-sum test.

in the absence of it. This indicates that participants expect the bonds to be sold at high prices during the intervention, pushing the preintervention prices up. Are their expectations fulfilled? Here we investigate the following two questions: (1) do participants sell the bonds to the computer at a price higher than the competitive equilibrium intervention price? and (2) is the intervention price influenced by the possibility of saving?

Table 6 compares the magnitude of mispricing during the intervention between the QE and QES treatments for Rounds 1 and 2. The mispricing is computed as $(P^{g,r}_{inter} - p^*_{inter})/p^*_{inter}$, where $P^{g,r}_{inter}$ is the realized computer purchasing price for group g in round r, and p^*_{inter} is the competitive equilibrium price derived in Section 3. The fourth column reports the p-values of differences between treatments. The last row reports the p-values of differences between rounds within treatments. The "*"s indicate whether the mispricing is significantly different from zero.

As shown in Table 6, the mispricings during the intervention are significantly higher than 0 in both treatments in both rounds. This result is consistent with Penalver et al. (2020), who suggest that the price competition was not strong enough during the QE operation.[15]

Observation 5 *The intervention prices are significantly higher than the competitive equilibrium intervention prices, regardless of the existence of the saving possibility.*

Besides, Table 6 (the fourth column) shows that the magnitude of mispricing is not significantly different between the QE and QES treatments. Thus, the saving condition does not influence the intervention prices. Table 6 (the last row) also fails to show any statistically significant difference between Round 1 and Round 2 within both treatments, implying participants observing a high intervention in Round 1 do not promote the price competition during the QE operation of Round 2.

5 Conclusions

In this study, we conducted an experiment to examine the effect of market intervention in the presence of a consumption smoothing motive to trade. Existing experimental studies that investigate the impact of market inter-

[15]In Bertrand price competition experiments, for example, conducted by Dufwenberg and Gneezy (2000) and Baye and Morgan (2004), overpricing is also observed. Dufwenberg and Gneezy (2000) explain the reason for participants not competing aggressively enough *a la* the level-k model.

vention (Haruvy et al., 2014; Penalver et al., 2020) have employed variants of the Smith et al. (1988) paradigm where no trade is expected under the risk neutral rational expectations equilibrium, and a large mispricing has been observed. By employing the new experimental framework based on the Lucas asset pricing model (Lucas Jr, 1978) proposed by Asparouhova et al. (2016) and Crockett et al. (2019), we reexamine the effect of market intervention in the framework where assets are traded to smooth consumption. In this new framework, Crockett et al. (2019) report that the magnitude of mispricing tends to be smaller than those observed in the Smith et al. (1988) framework. To the best of our knowledge, this study is the first study that examines the effect of market intervention under the Lucas asset pricing model in a lab.

Furthermore, we investigated the effect of the market intervention under the two conditions: with and without the possibility to save. Under a no-saving condition, participants need to smooth consumption only by trading assets, and under the saving condition, they can smooth consumption not only by trading assets, but also by saving.

The results show a significant effect of market intervention on overpricing of the asset before the intervention actually occurs, regardless of the existence of the saving possibility. This result is consistent with that of Penalver et al. (2020). However, contrary to Penalver et al. (2020), which shows the effect becomes larger as participants repeat the same experiment, in our experiment, the effect of market intervention on the mispricing becomes statistically insignificant in the second round. Surprisingly to us, despite the significant effect on the overpricing of the asset, the intervention did not significantly worsen participants' payoffs. On the contrary, in the presence of a saving possibility, it improved it, although the effect was observed only in the first round.

However, there may exist some questions of external validity in this study. In our experiment, following Penalver et al. (2020), the computer purchased up to 30 percent of the bond from each market through the intervention. In the real world, however, an intervention of such a scale may cause hyperinflation, and thus not a realistic scenario to be considered. Clarifying the relationship between the scale of the intervention and its effects on the market outcome is a fruitful future research.

Another question that we consider interesting is how an intervention on one asset influences the pricing of other assets in a setting with multiple assets being traded simultaneously. Because experimental analyses with multiple assets markets are still scarce, and, apart from Asparouhova et al. (2016), they use the Smith et al. (1988) framework (Charness and Neugebauer, 2019; Duffy et al., 2021), we believe such an exercise will be very fruitful.

References

Asparouhova, E., P. Bossaerts, N. Roy, and W. Zame (2016): ""Lucas" in the laboratory," *Journal of Finance*, 71, 2727–2779.

Baye, M. R. and J. Morgan (2004): "Price Dispersion in the Lab and on the Internet: Theory and Evidence," *The RAND Journal of Economics*, 35, 449–466.

Bo, P. D. and G. Fréchette (2011): "The evolution of cooperation in infinitely repeated games: Experimental evidence," *American Economic Review*, 101, 411–429.

Bostian, A. A. and C. A. Holt (2009): "Price bubbles with discounting: A web-based classroom experiment," *Journal of Economic Education*, 40, 27–37.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004): "A cognitive hierarchy model of games," *Quarterly Journal of Economics*, 119, 861–898.

Carbone, E., J. Hey, and T. Neugebauer (2021): "An Experimental Comparison of Two Exchange Economies: Long-Lived Asset vs. Short-Lived Asset," *Management Science*, 67, 6496–6962.

Charness, G. and T. Neugebauer (2019): "A Test of the Modigliani-Miller Invariance Theorem and Arbitrage in Experimental Asset Markets," *The Journal of FINANCE*, 74, 493–529.

Committee on the Global Financial System (2019): "Unconventional monetary policy tools: a cross-country analysis," CGFS Papers 63.

Corbet, S., J. J. Dunne, and C. Larkin (2019): "Quantitative easing announcements and high-frequency stock market volatility: Evidence from the United States," *Research in International Business and Finance*, 48, 321–334.

Crockett, S., J. Duffy, and Y. Izhakian (2019): "An experimental test of the Lucas asset pricing model," *Reivew of Economic Studies*, 86, 627–667.

D'Amico, S. and T. B. King (2013): "Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply," *Journal of Financial Economics*, 108, 425–448.

Duffy, J. and D. Puzzello (2014): "Experimental Evidence on the Essentiality and Neutrality of Money in a Search Model," *Research in Experimental Economics*, 17, 259–311.

——— (2020): "The Friedman Rule: Experimental Evidence," Working paper.

Duffy, J., J. P. Rabanal, and O. A. Rud (2021): "The impact of ETFs in secondary asset markets" Expeirmenta evidence," *Journal of Economic Behavior and Organization*, 188, 674–696.

Dufwenberg, M. and U. Gneezy (2000): "Price competition and market concentration: an experimental study," *International Journal of Industrial Organization*, 18, 7–22.

Eggertsson, G. B. and M. Woodford (2003): "The Zero Bound on

Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 34, 139–235.

Engle-Warnick, J. and R. L. Slonim (2006): "Inferring repeated-game strategies from actions: Evidence from trust game experiments," *Economic Theory*, 28, 603–632.

Farhi, E. and I. Werning (2019): "Monetary Policy, Bounded Rationality, and Incomplete Markets," *American Economic Review*, 109, 3887–3928.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10, 171–178.

Fréchette, G. R. and S. Yuksel (2017): "Infinitely repeated games in the laboratory: four perspectives on discounting and random termination," *Experimental Economics*, 20, 279–308.

Greiner, B. (2015): "An Online Recruitment System for Economic Experiments," *Journal of the Economic Science Association*, 1, 114–125.

Grosse-Rueschkamp, B., S. Steffen, and D. Streitz (2019): "A capital structure channel of monetary policy," *Journal of Financial Economics*, 133, 357–378.

Haruvy, E., C. N. Noussair, and O. Powell (2014): "The impact of asset repurchases and issues in an experimental market," *Review of Finance*, 18, 681–713.

Lei, V., C. N. Noussair, and C. R. Plott (2001): "Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality," *Econometrica*, 69, 831–859.

Lucas Jr, R. E. (1978): "Asset prices in an exchange economy," *Econometrica: Journal of the Econometric Society*, 1429–1445.

McLaren, N., R. N. Banerjee, and D. Latto (2014): "Using changes in auction maturity sectors to help identify the impact of QE on gilt yields," *The Economic Journal*, 124, 453–479.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economics Review*, 85, 1313–1326.

Palan, S. (2013): "A Review of bubbles and crashes in experimental asset markets," *Journal of Economic Surveys*, 27, 570–588.

Penalver, A., N. Hanaki, E. Akiyama, Y. Funaki, and R. Ishikawa (2020): "A quantitative easing experiment," *Journal of Economic Dynamics and Control*, 119.

Powell, O. (2016): "Numeraire independence and the measurement of mispricing in experimental asset markets," *Journal of Behavioral and Experimental Finance*, 9, 56–62.

Powell, O. and N. Shestakova (2016): "Experimental asset markets: A survey of recent developments," *Journal of Behavioral and Experimental Finance*, 12, 14–22.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988): "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56, 1119–1151.

A Experimental setting comparison of the studies based on the Lucas model.

Table A.1: Experimental setting comparison (1)

Study	Purpose	The number of participant per market	Assets
Our study	To test the effect of QE policy on the financial asset price	8 participants per market	A nonstorable asset (experimental currency/cash/token) and a risk-free infinitely lived asset (bond)
Crockett et al. (2019)	To investigate the relationship between the existence of trading motivation and mispricing	12 participants per market	A nonstorable asset (experimental currency/cash/token) and a risk-free infinitely lived asset (asset)
Asparouhova et al. (2016)	To examine the features of the Lucas model in the laboratory.	12 to 30 participants per market	A nonstorable asset (experimental currency/cash/token), a risk-free infinitely lived asset (bond), and a risky asset (tree)
Carbone et al. (2021)	To test the robustness of the features of Lucas model to a three-period cyclical world	12 participants per market	A nonstorable asset (Experimental currency), a risk-free infinitely lived asset (asset), and a risk-free short-lived asset (credit)

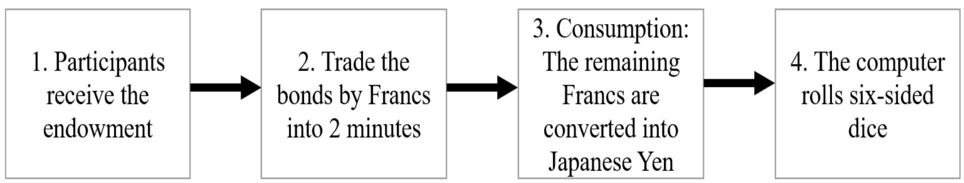
Table A.2: Experimental setting comparison (2)

Study	The features of the Exchange function	Termination	Payment
Our study	Concave	Random block termination. The market continues with a probability of 5/6	Final payment is the earning of the final period of a round randomly selected+1,000 Japanese Yen
Crockett et al. (2019)	Concave and linear	Random termination. The market continues with a probability of 5/6	Payoffs are earned from every period of every round
Asparouhova et al. (2016)	Linear	Random termination. The market continues with a probability of 5/6	Pay for 2 of the replications (round) randomly chosen after conclusion of the experiment
Carbone et al. (2021)	Concave and step	Random termination. The market continues with a probability of 5/6	Payoffs are earned from every period of every round

B The details on experiment timeline in each period.

Figure B.1: The timeline of Period 1

Benchmark and QE



BS and QES

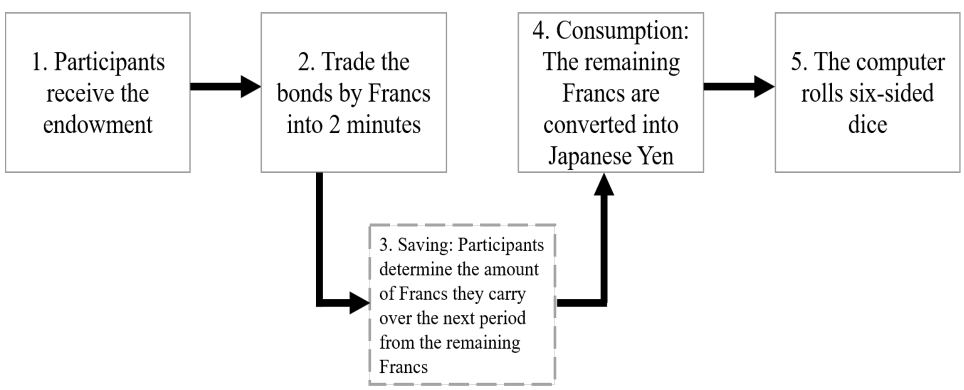
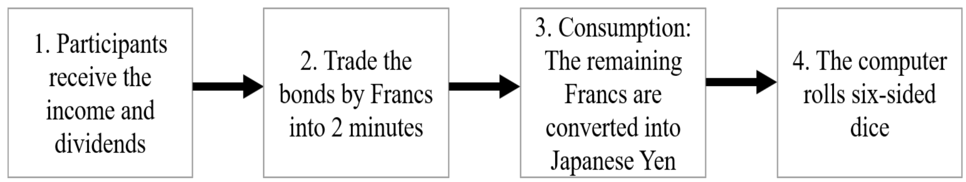


Figure B.2: The timeline of the period after period 1

Benchmark and QE



BS and QES

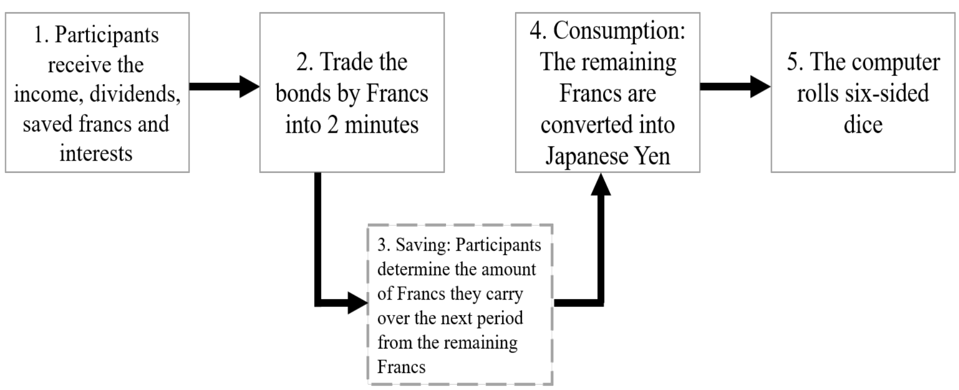
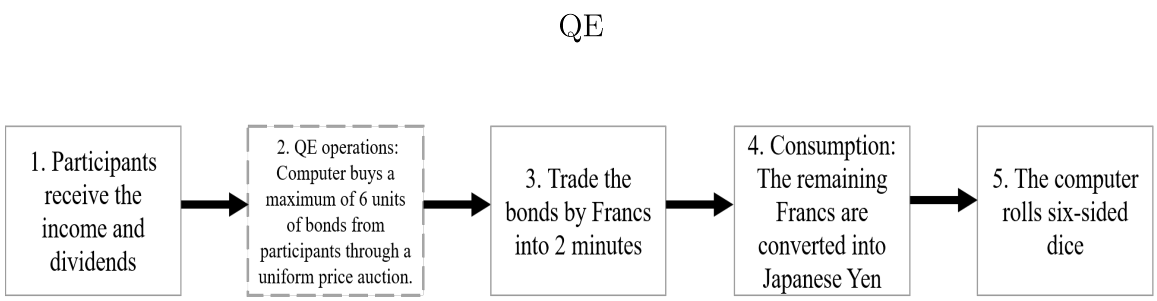
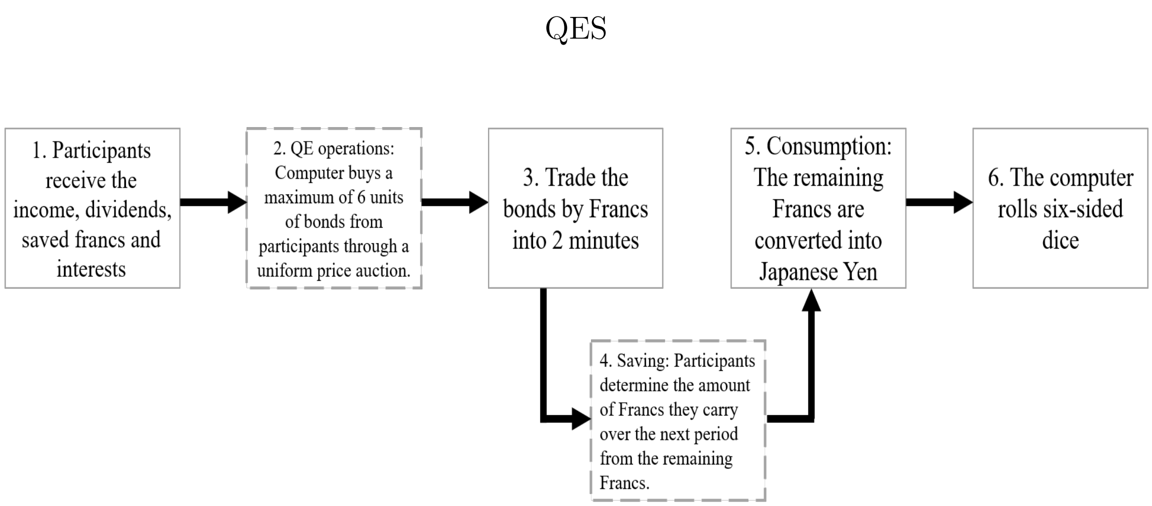


Figure B.3: The timeline of period 4 in QE and QES



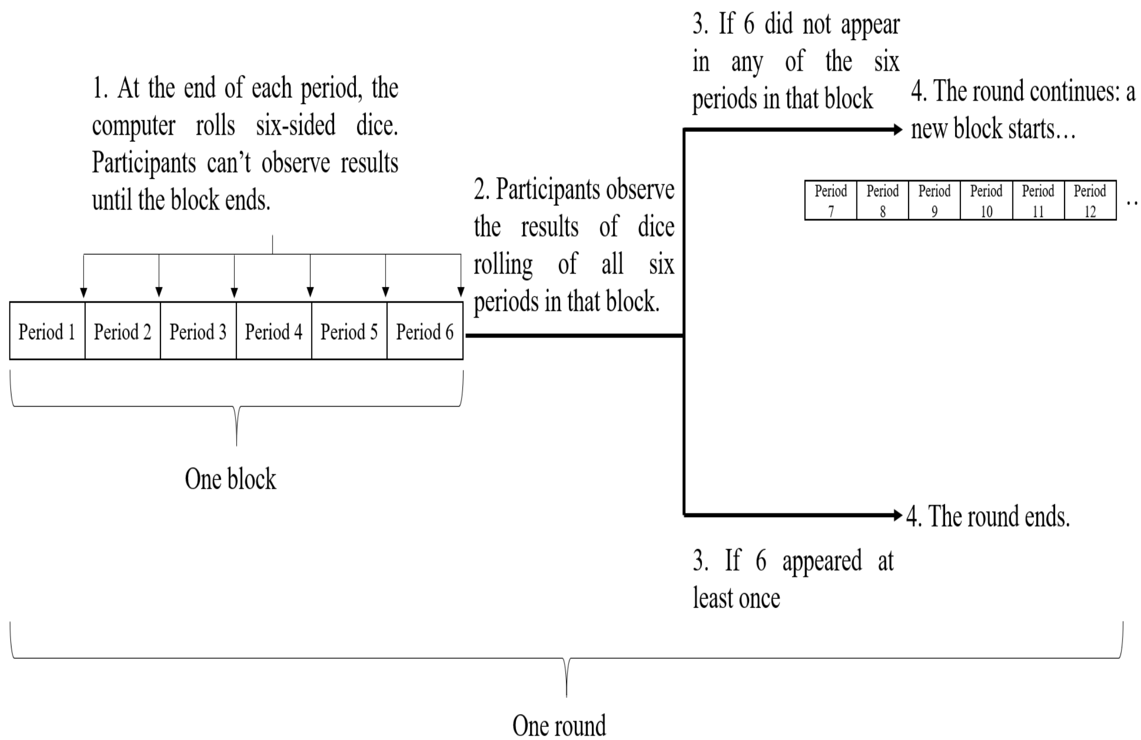


C The details of block random termination and payment.

As described in Sub-session 2.4, to secure sufficient data and to ensure that market intervention occurs in the QE and QES treatments, we adopted the method of block random termination (BRT, Fréchette and Yuksel, 2017). Figure C.1 reports the details of BRT in this experiment. Participants must at least experience the round for a block consisting of six periods. At the end of each period, the computer rolls six-sided dice, whereas participants can't observe the results until the block ends. At the end of each block, whether the round continues depends on the results of dice rolling of each period: the round continues to a new block if 6 did not appear in any of the six periods in that block; otherwise, the round ends. At the end of a round, if less than 35 minutes have passed since the beginning of the first round, a new round begins; otherwise, the experiment ends.

Then, a participant's payoff is determined as follows. Once the experiment ends, participants are paid in cash based on their earnings in the randomly chosen round in addition to the 1,000 JPY show-up fee. The earnings of each round of a participant equal the amount of Japanese Yen based on his/her consumption at the end of the "final period" of the round. Note that, the "final period" does not mean the final period of the final block but the earliest period in which the roll of a dice resulted in six. For ease of understanding, we give the following example in Figure C.2 to show how we

Figure C.1: The experiment process of each round and the implementation of the random block termination.



determine a participant's payment.

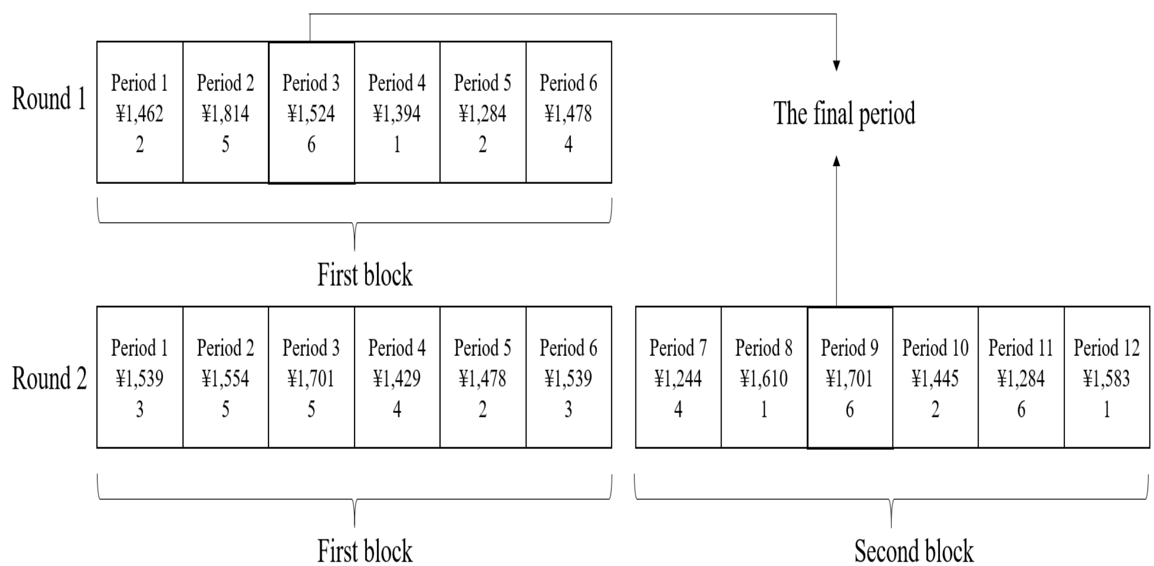
In Figure C.2, the numbers below the period numbers are the amount of Japanese Yen based on the consumption of a participant, and the numbers below the Japanese Yen are the results of dice rolling at the end of each period.

In round 1, since the dice rolling resulted in 6 for the first time in period 3, this round ends when the first block ended. The "final period" of this round was period 3, and the earning of the participant in this round was 1, 524 JPY.

In round 2, since no dice rolling resulted in six in any of the six periods in the first block, the round continued to the second block. In the second block, since the dice rolling resulted in six for the first time in period 9, this round ended when this block ended. Although the dice rolling resulted in six also in period 11, because the earliest period that dice rolling resulted in six was period 9, this period was the "final period" of round 2, and the participant earning of this round was 1,701 JPY.

Suppose more than 35 minutes have passed since the beginning of the first round when round 2 ended. Then session ends when round 2 ends. The computer will randomly select one round from round 1 and round 2 to calculate this participant's final payment. If round 1 is selected, the final payment of this participant is 1,524 + 1,000 = 2,524 JPY; otherwise, if round 2 is selected, the final payment of this participant is 1,701 + 1,000 = 2,701 JPY.

Figure C.2: The example of how we determine a participant's earning of each round.



D Other measures of market outcomes

Table D.1: Definitions of the measures of mispricing.

Relative absolute deviation (RAD)	$\frac{1}{N}\sum_{n=1}^{N}\lvert\frac{p_n-p_t^*}{p_n^*}\rvert$
Relative deviation (RD)	$\frac{1}{N}\sum_{n=1}^{N}\frac{p_n-p_t^*}{p_n^*}$
Geometric absolute deviation (GAD)	$exp(\frac{1}{N}\sum_{n=1}^{N}\lvert ln\frac{p_n}{p_n^*}\rvert)-1$
Geometric deviation (GD)	$exp(\frac{1}{N}\sum_{n=1}^{N}ln\frac{p_n}{p_n^*})-1$

This section presents the comparison results for mispricing among treatments based on the four measures proposed in the literature, including the geometric deviation (GD) reported in the main text. Table D.1 summarizes the definition of each measure. Table D.2 displays the median value of each measure of each treatment in each round and reports the results of comparisons between the treatments with and without intervention.

In Table D.2, consistent with the GD results, in the preintervention periods of the first round, we observe that there also exists a significant difference in relative deviation (RD) between the B and QE treatments ($p-value=0.047$) and between the BS and QES treatments ($p-value=0.047$) at the 5% significance level. However, we do not find any difference in relative absolute deviation (RAD) or geometric absolute deviation (GAD) in the preintervention periods. The reason is that RAD and GAD do not differentiate between overpricing and underpricing. While in the preintervention periods, bonds tend to be underpriced in the B and BS treatments, and they tend to be overpriced in the QE and QES treatments.

Table D.2: Median value (standard deviation) of each measure of bond mispricing.

	B	QE	BS	QES	B = QE[a]	BS = QES[a]
Round 1, Periods 1 - 3. Preintervention periods						
RAD	0.180 (0.154)	0.631 (0.440)	0.261 (0.177)	0.197 (0.316)	p = 0.375	p = 1.000
RD	-0.046 (0.259)	0.546 (0.497)	-0.200 (0.323)	0.117 (0.354)	**p = 0.033**	**p = 0.033**
GAD	0.221 (0.237)	0.596 (0.426)	0.388 (0.319)	0.201 (0.303)	p = 0.815	p = 1.000
GD	-0.524 (0.243)	0.467 (0.493)	-0.208 (0.318)	0.101 (0.328)	**p =0.047**	**p = 0.047**
Round 1, Periods 4 - 6. Postintervention periods						
RAD	0.233 (1.061)	0.920 (1.918)	0.164 (0.133)	0.479 (0.432)	p = 0.815	p = 0.447
RD	0.170 (1.096)	0.875 (1.934)	-0.050 (0.232)	0.285 (0.565)	p = 0.969	p = 0.684
GAD	0.237 (1.051)	0.871 (1.401)	0.232 (0.243)	0.591 (0.494)	p = 0.969	p = 0.684
GD	0.166 (1.091)	0.871 (1.433)	-0.054 (0.233)	0.252 (0.596)	p = 0.969	p = 1.000
Round 2, Periods 1 - 3. Preintervention periods						
RAD	0.277 (1.024)	0.908 (0.778)	0.200 (0.138)	0.251 (0.420)	p = 1.000	p = 1.000
RD	0.276 (1.041)	0.902 (0.778)	-0.053 (0.227)	0.200 (0.482)	p = 0.969	p = 0.355
GAD	0.276 (1.017)	0.860 (0.759)	0.248 (0.213)	0.272 (0.415)	p = 1.000	p = 1.000
GD	0.276 (1.036)	0.860 (0.759)	-0.055 (0.227)	0.198 (0.483)	p = 0.969	p = 0.355
Round 2, Periods 4 - 6. Postintervention periods						
RAD	0.333 (0.643)	0.786 (0.612)	0.392 (0.255)	0.507 (2.173)	p = 1.000	p = 1.000
RD	0.268 (0.648)	0.775 (0.645)	-0.240 (0.358)	-0.048 (2.34)	p = 0.969	p = 1.000
GAD	0.331 (0.602)	0.764 (0.544)	0.686 (1.260)	0.702 (1.291)	p = 1.000	p = 1.000
GD	0.264 (0.607)	0.764 (0.574)	-0.331(0.376)	-0.059 (1.314)	p = 0.815	p = 1.000

a: p-values are based on the pairwise Wilcoxon rank-sum test among the four treatments

E Analysis of trading volume

Besides the mispricing and consumption, we also consider whether the intervention influences the trading volume. Intuitively, in the QES treatment, if some participants expect to sell bonds to the computer at a high market price during the QE operation, they have an incentive to trade the bonds. That means the trading volume in the QES treatment may be higher than that in the BS treatment in the preintervention periods. Moreover, because participants have to smooth consumption by trading in the B and QE treatments, the trading volume in the B and QE treatments should be higher than in the BS treatment.

Table E.1 reports the median of Turnover in each treatment (top panel), as well as the results of comparisons among treatments (bottom panel). Turnover of group g, TO^g, is defined as

$$TO^g = \frac{1}{T}\sum_{t=1}^{T} \frac{q_t}{OS_t}, \tag{E.1}$$

where q_t is the trading volume at period t, and OS_t is the total number of the tradable bonds at period t.

Contrary to our intuition, from Table E.1, we observe that while Turnover is higher in the B treatment than in the QE treatment, and also in the B treatment than in the BS treatment, there is no statistically significant difference in Turnover among treatments either in the preintervation or the postintervention periods. This may be because the intervention may not only

Table E.1: Median of Turnover in each treatment.

	Preintervention periods				Postintervention periods			
	B	QE	BS	QES	B	QE	BS	QES
Round 1	0.50	0.40	0.40	0.38	0.42	0.40	0.38	0.35
Round 2	0.43	0.38	0.37	0.30	0.40	0.48	0.33	0.48

The comparison across treatments.

		Preintervention periods			Postintervention periods			
		B	QE	QES		Benchmark	QE	QES
Round 1	QE	1.000	-	-	QE	1.000	-	-
	QES	0.600	1.000	-	QES	1.000	1.000	-
	BS	0.250	1.000	1.000	BS	1.000	1.000	1.000
		B	QE	QES		Benchmark	QE	QES
Round 2	QE	1.000	-	-	QE	0.790	-	-
	QES	0.550	1.000	-	QES	0.540	1.000	-
	BS	1.000	1.000	1.000	BS	1.000	0.150	0.180

Note: The $p-values$ from the pairwise Wilcoxon rank-sum test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.

increase the participants' incentive to buy the bonds, but also increase the incentive to keep bonds, and neutralizes the variation of the trading volume.

F Discounted payoffs of each type

In this appendix, we compare discounted payoff, $w^i = \sum_{t=1}^{6} \pi^{t-1} u_{JPY}(c_t^i))$, for two types (1 and 2) separately across four treatments. Here, we use the within-group mean of w^i_{type} for each type ($\in \{1, 2\}$) , W^g_{type}, as an independent observation.

Table F.1 shows the across-group medians W^g_1 and W^g_2 in the four treatments for Rounds 1 and 2. It shows that W^g_1 and W^g_2 are significantly different in the BS treatment in both rounds. Furthermore, it shows that significant differences across treatments are observed only for Type 2 in both rounds. Finally, a statistically significant increase in W^g_2 in Round 2 compared with Round 1 is observed at the 5% significance level in all the treatments except for QES, but not for W^g_1.

The significantly lower discounted payoffs for Type 2 compared with Type 1 in the BS treatment is because of the difficulty Type 2 players had in trading the bond. When saving is possible, Type 1 players whose initial endowment consists of a larger amount and one unit of bonds can easily carry their surplus cash to the next period without purchasing the bond from Type 2 players who are initially endowed with four units of bonds and a smaller amount of cash. As a result, Type 2 players' consumption becomes lower in early periods, and thus the lower discounted payoffs.

Table F.1: Median W_1^g and W_2^g

The median W_1^g and W_2^g in Round 1

	B	QE	BS	QES
Type 1 (W_1^g)	5629.78	5727.28	5894.95	5783.58
Type 2 (W_2^g)	5833.97	5684.92	5235.36	5856.08
p-values* (H_0: $W_1^g = W_2^g$)	0.359	0.496	**0.039**	0.945

p-values$^+$ for treatment comparisons for Round 1

Type 1				Type 2			
	B	QE	QES		B	QE	QES
QE	1.000	-	-	QE	1.000	-	-
QES	1.000	1.000	-	QES	1.000	0.684	-
BS	1.000	1.000	1.000	BS	**0.034**	**0.024**	**0.033**

Median W_1^g and W_2^g in Round 2

	B	QE	BS	QES
Type 1(W_1^g)	5687.36	5800.10	5996.34	5930.37
Type 2 (W_2^g)	6036.74	5861.82	5679.81	6019.47
p-values* (H_0: $W_1^g = W_2^g$)	**0.039**	0.129	**0.008**	0.945

p-values$^+$ for treatment comparisons for Round 2

Type 1				Type 2			
	B	QE	QES		B	QE	QES
QE	1.000	-	-	QE	1.000	-	-
QES	0.091	0.556	-	QES	1.000	1.000	-
BS	0.189	0.969	1.000	BS	**0.001**	0.085	**0.047**

p-values* comparing Round 1 and Round 2

	B	QE	BS	QES
Type 1(R1=R2)	0.820	0.734	0.074	0.148
Type 2 (R1=R2)	**0.020**	**0.004**	**0.004**	0.055

$*$: $p-values$ from the Wilcoxon signed-rank (WSR) test are reported.
$+$: $p-values$ from the pairwise Mann–Whitney U (MW) test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.

G Nondiscounted payoffs in preintervention and postintervention periods

In this section, we investigate further the effect of intervention by separately considering the preintervention and postintervention periods. Namely, we compare nondiscounted payoffs in preintervention (periods 1 to 3), $\widetilde{w}^i_{pre} = \frac{1}{3}\sum_{t=1}^{3} u_{JPY}(c^i_t)$, and in postintervention (periods 4 to 6), $\widetilde{w}^i_{post} = \frac{1}{3}\sum_{t=4}^{6} u_{JPY}(c^i_t)$. We use within-group means as an independent observation. Table G.1 reports these measures for the four treatments in Rounds 1 and 2.

Table G.1 shows that in Round 1, $\widetilde{w}^i_{pre}$ are not significantly different (at the 5% significance level) between the B and QE treatments, as well as between the BS and QES treatments. The same is true for $\widetilde{w}^i_{post}$. Thus, the QE intervention does not significantly affect the nondiscounted payoffs. Table G.1 also shows that $\widetilde{w}^i_{pre}$ is significantly lower ($p = 0.034$), while $\widetilde{w}^i_{post}$ is significantly higher ($p = 0.003$), in the BS treatment than in the B treatment. Thus, the saving possibility, in the absence of QE intervention, results in significantly lower payoffs in periods 1 to 3, while it results in significantly higher payoffs in periods 4 to 6. In Round 2, contrary to Round 1, the saving possibility does not significantly lower the payoffs in periods 1 to 3. However, it results in significantly higher payoffs in the postintervention periods both in the presence (QE vs QES, $p = 0.022$) and in the absence (B vs BS, $p < 0.001$) of the intervention.

Table G.1: Median $\widetilde{w}^i_{pre}$ and $\widetilde{w}^i_{post}$

Median $\widetilde{w}^i_{pre}$ and $\widetilde{w}^i_{post}$ in Round 1

	B	QE	BS	QES
Preintervention ($\widetilde{w}^i_{pre}$)	4358.03	4202.81	3880.03	4168.39
Postintervention ($\widetilde{w}^i_{post}$)	4277.37	4326.12	4719.25	4869.29
p-values* (H_0: Pre = Post)	0.734	**0.004**	**0.004**	**0.008**

p-values+ for treatment comparisons in Round 1

Preintervention periods				Postintervention periods			
	B	QE	QES		B	QE	QES
QE	0.462	-	-	QE	0.969	-	-
QES	0.556	1.000	-	QES	**0.022**	0.216	-
BS	**0.034**	**0.011**	0.091	BS	**0.003**	0.113	1.000

Median $\widetilde{w}^i_{pre}$ and $\widetilde{w}^i_{post}$ in Round 2

	B	QE	BS	QES
Preintervention ($\widetilde{w}^i_{pre}$)	4413.53	4402.24	4278.49	4347.75
Postintervention ($\widetilde{w}^i_{post}$)	4371.66	4461.73	4733.41	4756.16
p-values* (H_0: Pre = Post)	0.652	0.359	**0.004**	**0.008**

p-values+ for treatment comparisons in Round 2

Preintervention periods				Postintervention periods			
	B	QE	QES		B	QE	QES
QE	1.000	-	-	QE	1.000	-	-
QES	1.000	1.000	-	QES	$<$**0.001**	**0.022**	-
BS	0.064	0.302	1.000	BS	$<$**0.001**	**0.024**	1.000

$*$: $p-values$ from the Wilcoxon signed-rank (WSR) test are reported.
$+$: $p-values$ from the pairwise Mann–Whitney U (MW) test are reported. Bonferroni method is used for correcting p-values for multiple comparisons.