Discussion Papers In Economics And Business

How Does Opt-in Work? A Field Experiment on Financial Incentives for Physical Activity

Hirofumi Kurokawa and Shusaku Sasaki

Discussion Paper 23-01

February 2023

Graduate School of Economics
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

How Does Opt-in Work? A Field Experiment on Financial Incentives for Physical Activity*

Hirofumi Kurokawa[a] Shusaku Sasaki[b]

February 1, 2023

Abstract

We examined the effectiveness of two different schemes for providing financial incentives to promote physical activity. We collaborated with a local government in Japan to conduct a field experiment with 498 residents randomly assigned to participate in one of three groups: an "opt-in" group (needed to apply to receive the incentive based on their daily steps), an "opt-out" group (received the incentive by default, but could request not to receive it), and a control group (no incentive). In the opt-in group, 31.1% of the participants applied to receive the incentive, while 100% of those in the opt-out group retained the default option and received it, indicating that their take-up rates depended heavily on the default settings. Our estimation results suggest that providing financial incentives in the opt-in scheme can be effective and efficient. The opt-in group, overall, showed a statistically significant increase of approximately 710 steps per day during the first half of the treatment period, and the number of steps was estimated to increase by 2,280 among the 31.1% of participants who opted-in to receive the incentive. The opt-out scheme did not show any significant increase in their number of daily steps.

JEL classification: C93, D90, I12

Keywords: field experiment, default, monetary incentive, health behavior, walking

---

* We would like to thank Takanori Ida, Ryota Nakamura, Rei Goto, Toshiaki Iizuka, and the seminar participants at the ABEF's 15th annual conference 2021, 2022 Virtual Asia Pacific ESA Meeting OSAKA, JEA's 2022 Spring Meeting, and the 17th Annual Conference of JHEA for their helpful comments. We thank the city of Kobe for collaborating in this study. We declare no conflict of interest. This study was approved by the Ethics Committee of the School of Economics and Management, University of Hyogo, and the Research Ethics Review Committee for Kobe City Health Services (No. 1382). This study has been pre-registered through the AEA RCT Registry (RCT ID AEARCTR-0006964). We gratefully acknowledge financial support from Urban Innovation Kobe (A20103). In preparing the manuscript, Sasaki was financially supported by Japan Science and Technology Agency (JPMJPR21R4).

[a] School of Economics and Management, University of Hyogo. E-mail: kurokawa@em.u-hyogo.ac.jp

[b] Center for Infectious Disease Education and Research (CiDER), Osaka University. E-mail: ssasaki.econ@cider.osaka-u.ac.jp

1. Introduction

Most traditional policy research using randomized controlled trials has measured the causal effects of mandatory policy assignment. However, implementing a policy intervention in a mandatory manner, as is done in academic research, is rare in the real world. This is because mandatory implementation requires a system that enables a policy to be applied to all individuals involved and monitors their adherence to it. In addition, a policy must be made mandatory by law, and implementation costs tend to be extremely high. In practice, policies are often applied to only those who choose to accept them, and the implementation costs will be relatively small. This can be accomplished in two ways. One is the "opt-out" method, in which a policy is applied to everyone by default, but individuals are allowed to opt-out upon request. The other is the "opt-in" method, in which a policy is not applied by default, but rather only upon request.

Under conditions with self-selection, overall policy impacts will vary depending on the heterogeneous effects across individuals and which individuals self-select to receive the policy. For example, if a policy is widely accepted by people for whom a large (or significant) positive policy effect appears, the overall policy impact will become larger than if the policy intervention was mandated, and thus the policy function more efficiently due to self-selection. Conversely, if those who are likely to experience small or negative effects choose to receive a policy intervention, the overall policy impact will become relatively small, and self-selection will prevent it from functioning efficiently.

Further, compared with opt-out schemes, opt-in schemes require people to incur procedural costs to receive policy interventions, and therefore generally have lower application rates (Madrian and Shea 2001; Johnson and Goldstein 2003); thus, who choose to receive the policy intervention could differ between the two schemes. To accurately understand the real-world implications of policy interventions, it is essential to ascertain the influence of self-selection on policy efficiency and how that influence differs across self-selection methods (e.g., opt-out and opt-in schemes). Recent field experimental studies have begun to measure policy intervention effects after accounting for self-selection, particularly in electricity markets (Wang et al. 2020; Fowlie et al. 2021; Ito et al. 2021; Ida et al. 2022).

The present study adds new evidence to this emerging literature stream. We collaborated with a local government in Japan to conduct a field experiment among the general public and measure the policy intervention effects of offering financial incentives on increasing physical activity, while taking self-selection into account. Specifically, we evaluated the impact of financial incentives on people's daily step counts under two schemes in which they could either opt-out of or opt-in for a program providing financial incentives linked to their daily step counts.

From a health policy perspective, daily physical activity is important for the general population because it helps them maintain good health (Reiner et al. 2013; Warburton and Bredin 2017; Warburton et al. 2006). Governments often employ programs that provide financial incentives to promote physical activity (Mitchell et al. 2020). The effectiveness of such financial incentive programs is supported by robust evidence, particularly for deterring addictive behaviors such as smoking (Notley et al. 2019). Over the short-term, financial incentives can encourage daily physical activity, such as walking; however, the linkage between financial incentives and habit formation or long-term health improvement remains under debate (Barte and

Wendel-Vos 2017; Mantzari et al. 2015), and thus experimental investigations continue. Further, as noted above, forcing the general public to engage in routine physical activity or accept policy interventions to encourage such activity is difficult in the real world. Therefore, when measuring the impact of a program that provides financial incentives for physical activity, considering whether the general population is willing to accept those incentives is essential.

Introducing a self-selection process could influence the observed effects of financial incentives on physical activity. If individuals who would significantly and continuously increase their daily steps counts by receiving the incentives mainly self-select to receive them, the program will have a large overall promotion effect. Conversely, if those whose daily steps would not either increase or decrease by receiving incentives mainly agree to receive them, the program's effectiveness and efficiency will be deteriorated. Investigations on how the opt-in self-selection process influences policy effectiveness and efficiency have begun also in medical- and health-related fields (Finkelstein and Notowidigdo 2019).

This study's novelty is that it experimentally evaluates the effects of two self-selection schemes, opt-out and opt-in, on physical activity. We conducted a field experiment using a mobile healthcare application during February and March 2021. Each participant was randomly assigned to one of three groups: an opt-out group, an opt-in group, and a control group. We set a two-week treatment period, in which the opt-out and opt-in groups had an opportunity to receive a financial incentive, while the control group did not. The opt-out group received the incentive by default, but could decline the offer in advance. The opt-in group was required to request the incentive in advance if they wished to receive it. No group received any financial incentives during either the pre-treatment period or the two-week post-treatment period.

First, we found a significant difference in the take-up rates for financial incentives between the opt-out (100%) and opt-in (31.1%) groups. This suggests that the default options strongly influenced whether participants received the incentives. Second, we found a statistically significant increase in daily step counts from the opt-in scheme, but not the opt-out. Those in the opt-in group, on average, showed an estimated increase of approximately 710 steps per day during the first half of the treatment period (intention-to-treat [ITT]). Further, the 31.1% of participants who opted-in showed an increase of approximately 2,280 steps (local average treatment effect [LATE]). Those in the opt-out group showed no significant increase in their number of daily steps. The results indicate that the opt-in scheme can be more efficient in offering financial incentives. Third, we found that the positive effect of the opt-in scheme on daily step counts was short-term, and was observed only during the first half of the treatment period. The effect did not reappear after the financial incentive ended. Further, we found no spillover effects, such as a treatment effect on body weight.

Our study contributes to two literature streams. First, we add evidence from the health field to the emerging literature on defaults, self-selection, and follow-on performance. Default settings are well-known to influence various economic choices (Jachimowicz et al. 2019), such as retirement savings plans (Madrian and Shea 2001; Choi et al. 2002; Thaler and Benartzi 2004), organ donation (Johnson and Goldstein 2003; Abadie and Gay 2006; Arshad et al. 2019), car insurance (Johnson et al. 1993), health insurance (Johnson et al. 2013), and energy saving (Wang et al. 2020; Fowlie et al. 2021; Ito et al. 2021; Ida et al. 2021). Recent studies have explored how default settings target people through self-selection and influence consequent performance. For example, in a comparison between two electricity pricing programs, time-of-use and critical-peak pricing (CPP), Fowlie et al. (2021) found that an opt-out scheme with CPP as the default resulted in higher acceptance of CPP than did an opt-in scheme, and reduced electricity consumption. However, the choices accepted in opt-out schemes do not always result in subsequent optimal economic behavior. In a comparison between flat-rate pricing and CPP, Wang et al. (2020) showed that an opt-in CPP scheme was more effective in saving electricity, compared with an opt-out scheme with CPP as the default.[1]

Second, our evidence adds new insight to the literature on how to incentivize health behaviors. The benefits of health behaviors appear in the future; however, the costs are incurred when the behavior is performed, which will likely lead to underinvestment. Providing incentives to lower the costs of current behaviors has been shown to be effective (Volpp et al. 2008; Charness and Gneezy 2009; Cawley and Price 2013; Acland and Levy 2015; Royer et al. 2015; Carrera et al. 2020). Some meta-analyses have shown that, at least in the short term, financial incentives are also effective for encouraging physical activity, including daily walks (Mitchell et al. 2020). Previous studies have focused on developing incentive structures, such as increasing price rates rather than keeping the incentive constant based on daily step counts (Bachireddy et al. 2019) and loss contracts, in which incentives are confiscated if the target number of steps is not reached (Adjerid et al. 2022). Other studies have used random and mandatory assignments of some incentive programs: however, the present study focused on how incentives are provided, rather than on incentive structure. We confirmed the positive effect on daily step counts of an opt-in scheme that allows people to self-select their participation in an incentive program, suggesting that introducing self-selection through an opt-in scheme when providing financial incentives can enhance treatment effects.

The remainder of this paper is organized as follows. Section 2 describes our experimental design and estimation strategy. Section 3 presents our main results, and Section 4 reports on the tests conducted to check the robustness of those results. Section 5 comprises the concluding statements.

2. Methods

2.1. Experimental Design

We collaborated with the city of Kobe, Japan, to conduct a field experiment on daily walking activity among city residents during the winter of 2020–2021. This experiment provided a treatment offering financial incentives through either an opt-out or opt-in scheme based on daily step counts. Kobe is a large city with a population of over 1.5 million and one of the 20 ordinance-designated cities in Japan. Several private companies, including foreign-affiliated firms, have headquarters in Kobe.

To improve people's health, Kobe provides a mobile healthcare management application that offers health-related information to residents and employees in the city. Users can link the app to their smartphone to automatically record their daily step counts. They can also record their body weight in the app for self-

[1] Some studies suggest that providing policy interventions in an opt-out scheme can have inefficient results for retirement savings (Beshears et al. 2022; Choukhamane 2019).

management.[2] At the beginning of the experiment, approximately 6,000 people had registered to use the app.

We targeted users of the app, and from December 1–14, 2020, we contacted existing users via app notifications, emails, and postcards to invite them to join the experiment. In addition, to include a wide range of individuals, we also recruited new users through Kobe's social media accounts and invited them to participate. In total, 660 individuals completed the participation process. Participants also completed a questionnaire regarding information such as socioeconomic status, health status, need for financial incentives, whether to set daily step count goals, and frequency of app use. Our analysis excluded users who were under 19 years of age, entered nonexistent IDs, or did not record their exact number of steps in the app (see Section 2.2 for details). Following this procedure, data from 498 individuals were ultimately used for our analysis. Our experiment was a randomized controlled trial conducted with individuals who agreed to participate in the experiment, which is common in health behavioral studies with the general public (Charness and Gneezy 2009; Acland and Levy 2015; Royer et al. 2015; Adjerid et al. 2021; Carrera et al. 2020; Giuntella et al. 2022). Therefore, the external validity of the sample is an important question, which we explore below in Section 2.2.

We stratified randomization based on step counts at the time of entry and need for financial incentives, and assigned the 498 participants to one of three groups as follows:

Control group (n = 174): Participants assigned to this group received 700 JPY (approximately 6.80 USD, based on the exchange rate at that time of 103 JPY to 1 USD) for participation and agreed to provide step count data.

Opt-out group (n = 160): Participants assigned to this group received 700 JPY for participation and agreed to provide step count data. Participants in this group automatically received a financial incentive based on their daily step counts. They could decline the incentive by submitting an application in advance; however, no participants opted out in the current experiment.

Opt-in group (n = 164): Participants assigned to this group received 700 JPY for participation and agreed to provide step count data. Participants in this group who wished to receive the incentive were required to apply in advance, and only those who completed the process were eligible. The opt-in rate in this group was 31.1% (51 out of 164 participants).

We provided financial incentives based on daily step counts from February 8–21, 2021, for all participants in the opt-out group and those who completed the application procedure in the opt-in group. Specifically, 5 JPY was offered for every 1,000 steps, with a daily cap of 70 JPY (for 14,000 steps), enabling participants to earn up to 980 JPY (approximately 9.5 USD) in rewards over the two-week treatment period.

---

[2] People can earn "health points" by using this app, independent of the financial incentives provided in this experiment. For example, on completing 10,000 steps per day, people can earn 5 points. The points can be exchanged for rewards (e.g., 150 points = one gym ticket). In this experiment, all participants, regardless of group, had the opportunity to earn these health points. We measured the effect of the financial incentive in this experiment by offering a random financial incentive, conditioning on the impact of health points.

Previous studies have provided incentives for walking activities themselves (e.g., daily step counts; Bachireddy et al. 2019) or reaching a goal (e.g., 10,000 steps per day; Adjerid et al. 2022). We employed the former method because this study's purpose was to examine the impact of different schemes of providing financial incentives for daily step counts rather than of goal-setting and commitment toward achieving a certain number of step counts. Among the studies employing the former method, for example, Adams et al. (2013) provided financial incentives according to increases from each individual's step count at baseline; however, we did not adopt a similar approach and instead provided the incentives according to their daily step counts during the treatment period for two reasons. First, the previous approach could lead participants to manipulate the performance by intentionally lowering their baseline step count to obtain a higher financial incentive. Second, when implementing financial incentive programs for the general public, providing different monetary amounts based on individual baseline step counts is costly, and implementing a uniform policy intervention is generally easier. Our interest is in whether combining such a simple policy intervention and self-selection can successfully target those for whom financial incentives are more likely to increase their step counts.

The overall period of this experiment was February 1–March 7, 2021, including the treatment period of February 8–21. On January 26, 2021, participants in all groups were informed via the app that the walking event would begin on February 1. On January 29, participants in the opt-in group received a message informing them that they need to submit an application via online until January 31 to receive the financial incentive. Participants in the opt-out group also received a message on the same day regarding the procedure to refuse to receive the incentive. Outside of the treatment period (i.e., February 1–7 and February 22–March 7), every Monday, participants in all groups received simple messages encouraging them to walk (Figure A1a). During the treatment period, participants in the control group or those in the opt-in group who did not wish to receive the incentive received similar simple messages. Participants in the opt-in and opt-out groups who wished to receive the incentive were sent a message regarding how they could obtain the incentive and how the incentive was calculated (Figure A1b).[3]

We defined the step counts during December 22, 2020–January 3, 2021, as the baseline step counts prior to the experiment, because the step counts from January 4–February 7, 2021, just prior to the treatment period, may have been affected by the declaration of a state of emergency during the COVID-19 pandemic and experimental preparations. Initially, we intended to start this experiment on January 4; however, owing to a rapid increase in the number of people infected with COVID-19, the government declared a state of emergency in Kobe, and residents were asked to refrain from going out unnecessarily and adhere to infection control measures. After discussions with Kobe officials, we decided to change the original schedule of the experiment, and communicated this change and its reason to the participants. Thereafter, we and Kobe officials considered that physical activity to maintain and improve people's health is essential even during

[3] In the post-experiment period (i.e., March 8–21), the financial incentive was also offered to the control group in an opt-out scheme to ensure fairness. During this period, participants in the control group were able to earn a monetary incentive of 5 JPY for every 1,000 steps each day. They could decline the incentive by completing the opt-out procedure between March 8–10, but the actual opt-out rate was still 0%.

the pandemic, decided to begin the experiment on February 1, and conveyed the updated schedule to participants on January 26.[4] The declaration of a state of emergency and schedule changes may have affected participants' walking habits in various ways, resulting in unusual trends. In addition, on January 29, different messages regarding the procedure for receiving financial incentives were sent to the opt-out and opt-in groups, as mentioned above. These communications may have had different effects on daily step counts across groups between February 1–7. For example, participants who wanted to receive higher financial incentives during the treatment period may have increased their steps prior to the treatment period to strengthen their walking habits and fitness level. We ruled out these effects to the extent possible by defining the baseline step counts as those obtained from December 22, 2020–January 3, 2021.

2.2. Data

Our analysis included data from 498 participants (174 in the control group, 160 in the opt-out group, and 164 in the opt-in group). We conducted a balance check using the step counts at the time of entry and answers to the entry survey and determined that the three groups are homogeneous. As Table 1 shows, we found no statistically significant differences among the three groups in relation to sociodemographic variables (e.g., age, gender, education, and income), health-related variables (e.g., height, weight, and personal goal), or the step counts at the time of entry.

Using the following procedure, we restricted our sample to 498 out of the 660 individuals from whom we obtained consent to participate. First, we excluded one underage user and two users who provided nonexistent IDs in the entry survey. Second, we excluded 74 users whose apps were not properly linked to their smartphones and whose steps were not recorded in the app. During the entry survey and the experiment, we attempted to reduce the number of excluded users by guiding them several times on how to link the app to their smartphones. Third, we excluded 85 users who, although their apps were linked to their smartphones, were more likely to record inaccurate step counts. For example, users who did not always carry a smartphone or wear a smartwatch, or who did not regularly open the app, were more likely to record extremely low or missing step counts. Although we also repeatedly reminded participants of the need to consistently carry a smartphone or wear a smartwatch during the experiment, we observed that older adult users in particular did not always do so. Therefore, we excluded older adult users and restricted our analysis sample to adults between the ages of 20 and 64 years. Adults within that age range are expected to work during weekdays, leading to differences in their walking habits between weekdays and weekends; therefore, we focused particularly on weekday data.

Even in the restricted analysis sample, daily step counts were sometimes missing for various reasons. Missing step counts were observed also in the literature on promoting physical activity where step counts were measured with a smartphone or smartwatch (Bachireddy et al. 2019; Aggarwal et al. 2020; Adjerid et al. 2022). In the current study, we basically imputed zero step to the missing data. This imputation approach may undermeasure observed step counts, compared to true step counts, and underestimate the treatment

---

[4] The state of emergency was lifted on February 28, 2021.

effect of financial incentives. Thus, if we found a significant promoting effect even under such a condition, it could be strong evidence. We also used other approaches to address missing data for robustness checks.

As noted above, our experiment was a randomized controlled trial conducted with individuals who agreed to participate; thus, the external validity of the sample was an important question. To investigate this question, we compared the gender and age structure between of all the app users and of the analyzed sample. Of the approximately 6,000 app users, 56.4% are female, while 56.6% of the analysis sample is female. The age structures of the app users and the analysis sample, respectively, were as follows: 7.5% and 8.0% were between 20 and 29 years old, 16.3% and 19.7% were between 30 and 39, 25.4% and 28.3% were between 40 and 49, 28.3% and 32.7% were between 50 and 59, and 16.0% and 11.2% were between 60 and 69. These comparisons suggest that the experimental sample has a similar gender and age structure to the app user population.

2.3. Participants Who Opted-in

No participants declined to receive the incentive in the opt-out group[5]; however, 51 participants in the opt-in group completed the process to receive the incentive. Thus, the opt-out and opt-in groups had take-up rates of 100% and 31.1%, respectively. This difference in take-up rates suggests that participants' choices were strongly influenced by the default options available. This study's opt-in rate (31.1%) could be considered reasonable, given those in previous studies and the characteristics of the incentive measures in this study. For example, in Ito et al.'s (2021) field experiment on electricity markets, the CPP opt-in rate ranged from 16% (without information provision) to 31% (with information provision). Comparatively, the opt-in rate in this study was somewhat higher. Unlike with increases or decreases in electricity prices, participants in this study received additional financial incentives based on their daily step counts, but were not fined and did not lose money. This is thought to have contributed to the high opt-in rate.

Table 2 shows the differences in characteristics between individuals who opted-in (n = 51) and those who did not (n = 113). No significant differences were observed in socioeconomic characteristics, such as age and gender, or health status, such as body weight and subjective health, between both groups. However, individuals who had higher step counts at the time of entry and those who used the app more frequently were more likely to opt-in. This may reflect that individuals who were health-conscious by nature were attracted to the financial incentive. Those who answered that they participated in this event for obtaining the incentive were also more likely to opt-in, suggesting that the incentive may have triggered the intention to walk among some participants.

2.4. Estimation Strategy

We use a difference-in-differences (DID) approach to estimate treatment effects on participants' daily step counts—our main outcome. We assume strong ignorability, in which the treatment assignment is

[5] The part of the registration site where participants could decline the incentive was accessed 14 times in total. Therefore, it was unlikely that participants were unable to opt out because they could not access the site.

independent of other observable and unobservable variables that affect daily walking, and this assumption is satisfied by our successful random assignment. We also assume the stable unit treatment values assumption (SUTVA), in which a participant's daily walking is affected by whether the individual participant chose the incentive, not by whether other participants chose the incentive. If a participant walked with other participants, this assumption might not be valid; however, we told participants that this was an individual-centric walking event, and the government had requested that people avoid social contact to prevent the spread of infection. Thus, it is expected that many participants walked alone, and we can assume that the SUTVA is validated.

First, we identify the effect of treatment assignment on daily step counts. We use data from the baseline and treatment periods and estimate two sets of DID regression equations with each set of the control group and one of the two treatment groups:

$$Steps_{it} = \alpha_0 + \alpha_1 Treatment\ Assignment_i \times Week1_t + \alpha_2 Treatment\ Assignment_i \times Week2_t + \alpha_3 Individual_i \times Day\_of\_week_{it} + \alpha_4 Day_t + \varepsilon_{it}, \quad (1)$$

where i and t denote individual and day, respectively. $Steps_{it}$ is the number of steps walked by individual i on day t. The $Treatment\ Assignment_i$ indicator takes the value 1 for the opt-out or opt-in group and 0 for the control group. The $Week1_t$ ($Week2_t$) indicator takes the value 1 in the first (second) week of the treatment period and 0 otherwise. The coefficients of interest are α_1 and α_2, which capture the average difference in daily steps across the treatment and control groups. $Individual_i \times Day\_of\_Week_{it}$ is an interaction term between individual i and the day of the week, which represents a fixed effect for individual i and the day of the week. Day_t is a fixed effect for each day. We use cluster standard errors at the individual level to account for serial correlation for an individual.

All participants in the opt-out group accepted the financial incentive. Thus, the opt-out treatment assignment effect estimated in Equation (1) is consistent with the average treatment effect (ATE). In contrast, only 31.1% of those in the opt-in group requested the incentive. The opt-in treatment assignment is not fully consistent with the receiving incentive status; thus, the estimated opt-in treatment assignment effect becomes the intention-to-treat (ITT).

Second, we use the data from the opt-in and control groups, estimate a DID-instrumental variable (IV) specification, and identify the local average treatment effect (LATE) of signing up for the opportunity of receiving the financial incentive on daily step counts, which is the treatment-on-treated (TOT). Here, we assume the exclusion restriction, where our treatment assignment does not directly affect daily step counts while indirectly affecting the counts through the effect of signing up for the opportunity of receiving the incentive, and our randomized control trial confirms this assumption. In addition, we assume monotonicity, where our treatment assignment weakly increases and never decreases the likelihood of receiving the incentive, and the actual take-up rate confirms this assumption. Specifically, we use the opt-in treatment assignment as the instrument and estimate the following DID-IV specification:

$$Receiving\ Incentive_{it} = \beta_0 + \beta_1 Optin\ Assignment_i \times Week1_t + \beta_2 Optin\ Assignment_i \times Week2_t + \beta_3 Individual_i \times Day\_of\_week_{it} + \beta_4 Day_t + \varepsilon_{it}, \quad (2)$$

$$Steps_{it} = \gamma_0 + \gamma_1 \widehat{Receiving\ Incentive_i} \times Week1_t + \gamma_2 \widehat{Receiving\ Incentive_i} \times Week2_t + \gamma_3 Individual_i \times Day\_of\_week_{it} + \gamma_4 Day_t + \varepsilon_{it.} \quad (3)$$

The $Receiving\ Incentive_{it}$ indicator takes the value 1 for participants who signed up to receive the incentive during the treatment period and 0 otherwise. The $Optin\ Assignment_i$ indicator takes the value 1 for the opt-in group and 0 for the control group. Equation (2) is estimated as the first step, and a two-step estimation is performed by fitting the estimated predicted value $\widehat{Receiving\ Incentive_i}$ to Equation (3). $Steps_{it}, Week1_t, Week2_t, Individual_i, Day\_of\_week_{it},$ and Day_t are defined as in Equation (1). The coefficients of interest are γ_1 and γ_2, which capture the average increase in daily step counts among individuals who received the incentive.

Third, we use the data from the opt-out and opt-in groups, estimate a similar DID-IV specification, and identify the LATE of receiving a financial incentive on daily step counts for passive joiners who received it under the opt-out condition but not under the opt-in condition, which is the treatment-on-untreated (TOU). Based on Fowlie et al. (2021), our experimental design can classify the participants as active joiners, passive joiners, and active leavers. Active joiners are those who retained the incentive offer in the opt-out group and applied to receive the incentive offer in the opt-in group. Passive joiners are those who did not actively decline the incentive offer in the opt-out group and did not actively accept the incentive offer in the opt-in group. Active leavers are those who opted out of the incentive offer in the opt-out group and did not apply to receive the incentive offer in the opt-in group, although we did not observe them in this experiment. Thus, participants in the opt-out group who accepted the incentive comprise active joiners and passive joiners, while participants in the opt-in group who applied to receive the incentive are all active joiners. Taking the difference between the opt-out and opt-in groups, we can estimate the treatment effects for passive joiners. We use the data from the opt-out and opt-in groups and estimate Equations (2) and (3), where all variables are defined as above, except the IV $Optin_i$. Here, the IV is $Optout_i$, which is an indicator variable equal to 1 for the opt-out group and 0 for the opt-in group.

3. Basic Analysis

Columns (1) and (2) in Table 3 show the estimation results for treatment assignments. Column (1) uses the sample comprising the control and opt-out groups. The coefficients on $Week_1(Week_2) \times Treatment\ Assignment$ capture the average difference in daily steps across the opt-out and control groups in the first (second) week of the treatment period. In the first week, the daily step count for the opt-out group was 541 steps higher than that for the control group, but this was not statistically significant. In the second week, the daily step count for the opt-out group was 287 higher than that for the control group, but this was also not statistically significant. Although all participants in the opt-out group accepted the incentive, we did not find significant evidence that offering incentives in the opt-out scheme increased daily step counts on average.

In contrast, we found a significant increase in daily step counts for the opt-in group. Column (2) shows that, in the first week of the treatment period, the daily step count for the opt-in group was 709 steps higher than that for the control group (p < .05). However, this increase and its statistical significance disappeared in the second week of the treatment period. The take-up rate of the incentive is 31.1%; thus, the estimated result is the ITT.

Column (3) shows the estimation results for receiving the financial incentive in the opt-in group (TOT), based on the sample of opt-in and control groups. We found that, in the first week of the treatment period, daily step count increased by 2,281 steps among those who actively applied for the incentive ($p < .05$). The step counts' increase in the second week is 744 and positive, but not statistically significant.

Column (4) shows the estimation results for the treatment effect for passive individuals who received the incentive under the opt-out condition but did not under the opt-in condition (TOU), using the sample of opt-out and opt-in groups. The coefficients on $Week_1 \times Receiving\ Incentive$ and $Week_2 \times Receiving\ Incentive$ were negative or close to zero, respectively, implying that offering the incentives to passive individuals had no impact on increasing their daily step counts. Thus, our experimental evidence suggests that offering incentives in an opt-in scheme can efficiently increase people's daily step counts better than an opt-out scheme, although this is a short-term effect.

To estimate the effects after ending the treatment, we used the same estimation strategy and data from the baseline and post-treatment periods. The estimation results in Table 4 provide no evidence that daily step counts decrease during the post-treatment period. In Columns (1) – (4), the coefficients are partly negative but not statistically significant. We further found no evidence that the financial incentive treatments influenced participants' body weight (see Appendix 2).

4. Further Analysis

4.1. Heterogeneity by App Use Frequency

App use frequency affects the take-up rate, which could then lead to a heterogenous treatment effect. As shown in Table 2, which reports the association between the take-up rate and individual characteristics, participants who frequently used the app are likely to join the incentive program. Those who less frequently used the app may have missed messages on the app because of their inattention and may not have applied for the incentive program in the opt-in group, while such the participants may have unknowingly accepted the program in the opt-out group. Those who used the app more frequently may have seen the messages but not applied for the incentive program in the opt-in group because they felt it would be too bothersome, while in the opt-out group, they could have joined the program because they did not feel it would be a burden.

We classified participants as "heavy users" (n = 314) and "light users" (n = 184) based on the entry survey, which asked participants how frequently they used the app. We defined "heavy users" as participants who answered, "I keep records every day," "I keep records almost every day (sometimes I forget to keep records a couple of times a week)," or "I keep records a couple of times a week." "Light users" were defined as participants who answered, "I do not keep records but I check the app every day," "I do not keep records but I check the app a couple of times a week," or "I do not keep records and rarely check the app."

Table 5 shows the estimation results for heterogeneous treatment effects, dividing the participants by app use frequency. Among heavy users in the opt-out group, we found a weak increase in daily step count by approximately 800 in the first week of the treatment period (Column (1)), with similar findings for heavy users in the opt-in group (Columns (2)). The IV estimation showed that active joiners among heavy users could further increase their step counts by approximately 2,000 (Columns (3)), while the treatment effect for passive joiners among heavy users was close to zero and not significant (Columns (4)). We also found that the financial incentive treatment showed no significant effect for light users in both the opt-out and opt-in groups (Columns (5)–(8)). Although the daily step counts for the opt-in treatment group and passive joiners were relatively large, the standard error was also large, likely because of the small sample size. In particular, heavy users showed a tendency in which the financial incentive contributed to an increase in daily step counts through people's opt-in application to receive the incentive.

4.2. Robustness Check for Missing Data

Here, we verify the robustness of our results using other approaches and re-addressing the problem of missing data for daily step counts. As explained in Section 2.2, some daily step counts were missing from the data. In the above analyses, we conservatively assumed that all missing daily step counts were considered to be zero. In this subsection, we first use the missing daily step counts as missing data and analyze unbalanced panel data. Second, we impute the average daily step counts at the baseline period for each individual and day of the week into the missing daily step count data during the treatment period.[6]

Table 6 shows the estimation results for the DID specification using data from the baseline and treatment periods. Columns (1)–(4) report the results of the unbalanced panel estimation, and Columns (5)–(8) report the results of the second approach for missing-data imputation. We found that participants in the opt-out group reported higher step counts per day, approximately 630 in Column (1) and 565 in Column (5), than those in the control group during the first week of the treatment period. Subsequently, we observed no statistically significant differences in daily step counts between the control and opt-out groups during the second week of the treatment period. We also found that participants in the opt-in group report higher step counts per day, approximately 718 in Column (2) and 643 in Column (6), than those in the control group, and daily step counts further increased by 2,023 (Column (3)) and 1,911 (Column (7)) among those who actively applied to receive the incentive. We observed no evidence that the treatment also increased daily step counts for passive individuals (Columns (4) and (8)).

In summary, we found that short-term increases in step counts may or may not be observed in the opt-out group, depending on which approach we employ for the missing data problem. However, the common results across the approaches are as follows: (1) short-term step counts increased in the opt-in group, (2) step counts further increased among those who completed the opt-in procedure, and (3) even if providing financial incentives to passive individuals who did not complete the opt-in procedure, it would not change their step counts. Thus, all our approaches confirmed the effects on increasing daily step counts in an opt-in scheme that allows people to self-select their participation in the incentive program, suggesting that introducing self-selection through an opt-in scheme when providing financial incentives can enhance treatment effects.

[6] When calculating the baseline average, missing step counts at the baseline period were not included.

5. Discussion and Conclusions

Most traditional policy research using randomized controlled trials has measured the causal impacts of mandatory policy assignment. However, in the real world, policies are often only applied to those who choose to accept them. Under self-selection conditions, overall policy impacts will vary depending on heterogeneous effects across individuals and which individuals self-select to receive the policy treatment.

This study aimed to maintain and improve the health status of the general public and measured the treatment effects of providing financial incentives on their walking behavior. Self-selection can significantly influence treatment effectiveness and efficiency in this context. While medical institutions and staff can guarantee the delivery of medical and preventive interventions to ill patients, those who remain healthy can decide by themselves whether they engage in physical activity as well as whether they receive treatments to promote it. Based on this, we collaborated with the local government of Kobe, Japan, to conduct a field experiment to identify the effects of providing financial incentives in two self-selection schemes: opt-in and opt-out.

Our estimation results suggest that the opt-in scheme works efficiently. Only 31.1% of participants in the opt-in group completed the application process to receive financial incentives, and this group overall increased their number of steps by an average of 710 step per day compared to the control group (ITT). In the opt-out group, none of the participants refused the financial incentives; however, no effect of increasing step counts was observed. Among the 31.1% of participants in the opt-in group who requested the incentive, we found that the financial incentives increased their step counts by 2,280 per day (TOT). We also found that even if the rest of those who did not apply had received the financial incentives, their step counts would not have changed (TOU). In this experiment, we leveraged participants' self-selection through an opt-in scheme, and were then able to target and deliver the financial incentive treatments primarily to those who experienced positive treatment effects.

Previous studies on electricity markets and retirement savings have reported that opt-in schemes may work efficiently or opt-out schemes may not (Wang et al. 2020; Beshears et al. 2022; Choukhamane 2019). Interest in self-selection has also grown in medical and health fields (Finkelstein and Notowidigdo 2019). To the best of our knowledge, this study is the first to add evidence to the medical and health literature by conducting a field experiment to identify the effects of providing financial incentives through two self-selection schemes, opt-in and opt-out, on physical activity.

The participants in our study self-selected to join the event before self-selecting to receive financial incentives, and thus had a higher level of interest in improving their physical activity. This dual self-selection could have contributed to targeting those likely to experience positive treatment effects and delivering financial incentives to them. This feature is shared with Wang et al. (2020), who compared CPP with flat-rate pricing and found that an opt-in CPP scheme was more effective in saving electricity, compared with an opt-out scheme with CPP as the default. Their finding that active CPP choice was more effective than passive CPP choice is consistent with our observation of increased step counts when incentives were offered in an opt-in scheme.

Our results also indicate room for improvement in the opt-in scheme used in this study. The opt-in scheme contributed to allocating the financial incentive treatment more efficiently to increase people's step counts. However, considering who opted in, we notice that it failed to consider "equity," which is a limitation of this study and an issue for future research. Specifically, we paid an average of 316 JPY per person in incentives in the opt-out group, compared to an average of 133 JPY per person in the opt-in group. The cost-effectiveness of paying incentives resulted in increases of 131 and 353 daily steps per 100 JPY for the opt-out and opt-in groups, respectively.[7] However, the opt-in scheme may not be desirable from an equity-related perspective. Table 2 shows that those who applied to receive the incentive in the opt-in scheme and effectively increased their step counts were already health-conscious individuals who regularly used the app prior to the experiment. Since the health-conscious people more walked and improved their health during this study, the opt-in scheme may have expanded health inequalities. In public policies, physical activity promotion is important also among those with low health consciousness to prevent their health deterioration. However, our results indicate that those who are passive to receiving financial incentives would have shown little change in their daily step counts even if they received the incentives. Thus, future research is needed to develop treatments that can promote physical activity among less health-conscious individuals. To improve the general population's health while considering health disparities, delivering different interventions for people with high and low health awareness will be necessary.

The opt-in scheme still has a limitation from an efficiency perspective, which is also an issue for future research. The positive effects of the opt-in scheme were short term and did not lead to body weight loss. Those who voluntarily completed the application to receive financial incentives significantly increased their step counts initially; however, this tendency did not persist. Future research will need to identify another treatment to promote continued physical activity, which can be added to financial incentives. However, we also need to consider the possibility that the COVID-19 pandemic may have influenced this short-term effect. Giuntella et al. (2021) provided financial incentives without any self-selection scheme for step counts during the COVID-19 pandemic and found the incentives increased step counts to pre-pandemic levels; however, after the incentives are removed, their step counts returned to the pandemic levels. Although how to promote their physical activity is an essential question during a situation including a pandemic when people tend to be inactive, we need to examine whether the positive effects of the opt-in scheme would be more sustainable after the COVID-19 pandemic has completely ended.

In this experiment, we provided financial incentives based on individuals' daily step counts during two treatment weeks. While the experimental features, including rules for providing incentives and length of treatment period, are similar to related studies conducted with the general public (Adjerid et al. 2021, Carrera et al. 2020, Giuntella et al. 2022), the features have varied across medical studies with unhealthy people (Mitchell et al. 2020). Follow-up studies should examine whether the efficiency of the opt-in scheme depends on different ways of providing financial incentives and different treatment periods. Hence, this study has great potential to extend and develop this emerging literature stream by providing the first experimental

[7] The opt-in scheme is more efficient as measured by the incremental cost-effectiveness ratio (opt-out: 0.76 JPY/day; opt-in: 0.28 JPY/day).

evidence on the interaction between financial incentive treatments and self-selection in a medical and health context, while considering effectiveness, efficiency, and equity.

References

Abadie, A., & Gay, S. (2006). The impact of presumed consent legislation on cadaveric organ donation: a cross-country study. *Journal of Health Economics*, *25*(4), 599–620.

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science*, *61*(1), 146–160.

Adams, M. A., Sallis, J. F., Norman, G. J., Hovell, M. F., Hekler, E. B., & Perata, E. (2013). An adaptive physical activity intervention for overweight adults: a randomized controlled trial. *PloS One*, *8*(12), e82901.

Adjerid, I., Loewenstein, G., Purta, R., & Striegel, A. (2022). Gain-loss incentives and physical activity: the role of choice and wearable health tools. *Management Science*, *68*(4), 2642–2667.

Aggarwal, S., Dizon-Ross, R., & Zucker, A. D. (2020). Incentivizing behavioral change: the role of time preferences. *National Bureau of Economic Research*, (No. w27079).

Arshad, A., Anderson, B., & Sharif, A. (2019). Comparison of organ donation and transplantation rates between opt-out and opt-in systems. *Kidney International*, *95*(6), 1453–1460.

Bachireddy, C., Joung, A., John, L. K., Gino, F., Tuckfield, B., Foschini, L., & Milkman, K. L. (2019). Effect of different financial incentive structures on promoting physical activity among adults: a randomized clinical trial. *JAMA Network Open*, *2*(8), e199863.

Barte, J. C., & Wendel-Vos, G. W. (2017). A systematic review of financial incentives for physical activity: the effects on physical activity and related outcomes. *Behavioral Medicine*, *43*(2), 79–90.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Skimmyhorn, W. L. (2022). Borrowing to save? The impact of automatic enrollment on debt. *The Journal of Finance*, *77*(1), 403–447.

Carrera, M., Royer, H., Stehr, M., & Sydnor, J. (2020). The structure of health incentives: evidence from a field experiment. *Management Science*, *66*(5), 1890–1908.

Cawley, J., & Price, J. A. (2013). A case study of a workplace wellness program that offers financial incentives for weight loss. *Journal of Health Economics*, *32*(5), 794–803.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, *77*(3), 909–931.

Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2002). Defined contribution pensions: plan rules, participant choices, and the path of least resistance. *Tax Policy and the Economy*, *16*, 67–113.

Choukhamane, T. (2019). Default options and retirement saving dynamics. Working Paper, MIT.

Fowlie, M., Wolfram, C., Baylis, P., Spurlock, C. A., Todd-Blick, A., & Cappers, P. (2021). Default effects and follow-on behaviour: evidence from an electricity pricing program. *The Review of Economic Studies*, *88*(6), 2886–2934.

Finkelstein, A., & Notowidigdo, M. J. (2019). Take-up and targeting: experimental evidence from SNAP. *The Quarterly Journal of Economics*, *134*(3), 1505–1556.

Giuntella, O., Hyde, K., Saccardo, S., & Sadoff, S. (2021). Lifestyle and mental health disruptions during COVID-19. *Proceedings of the National Academy of Sciences*, *118*(9).

Ida, T., Ishihara, T., Ito, K., Kido, D., Kitagawa, T., Sakaguchi, S., & Sasaki, S. (2022). Choosing who chooses: selection-driven targeting in energy rebate programs. *National Bureau of Economic*

Research. (No. w30469).

Ito, K., Ida, T., & Tanaka, M. (2021). Selection on welfare gains: experimental evidence from electricity plan choice. *National Bureau of Economic Research*. (No. w28413).

Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019). When and why defaults influence decisions: a meta-analysis of default effects. *Behavioural Public Policy*, *3*(2), 159–186.

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, *302*(5649), 1338–1339.

Johnson, E. J., Hassin, R., Baker, T., Bajger, A. T., & Treuer, G. (2013). Can consumers make affordable care affordable? The value of choice architecture. *PloS One*, *8*(12), e81521.

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, *7*(1), 35–51.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics*, *116*(4), 1149–1187.

Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J. P., & Marteau, T. M. (2015). Personal financial incentives for changing habitual health-related behaviors: a systematic review and meta-analysis. *Preventive Medicine, 75*, 75–85.

Mitchell, M. S., Orstad, S. L., Biswas, A., Oh, P. I., Jay, M., Pakosh, M. T., & Faulkner, G. (2020). Financial incentives for physical activity in adults: systematic review and meta-analysis. *British Journal of Sports Medicine*, *54*(21), 1259–1268.

Notley, C., Gentry, S., Livingstone-Banks, J., Bauld, L., Perera, R., & Hartmann-Boyce, J. (2019). Incentives for smoking cessation. *Cochrane Database of Systematic Reviews*, (7).

Reiner, M., Niermann, C., Jekauc, D., & Woll, A. (2013). Long-term health benefits of physical activity–a systematic review of longitudinal studies. *BMC Public Health*, *13*(1), 1–9.

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics*, *7*(3), 51–84.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*(S1), S164–S187.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: a randomized trial. *JAMA*, *300*(22), 2631–2637.

Wang, W., Ida, T., & Shimada, H. (2020). Default effect versus active decision: evidence from a field experiment in Los Alamos. *European Economic Review*, *128*, 103498.

Warburton, D. E., & Bredin, S. S. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Current Opinion in Cardiology*, *32*(5), 541–556.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *CMAJ*, 174(6), 801–809.

Table 1

BALANCE CHECK

	CONTROL	OPT-OUT	OPT-IN	P-VALUE FROM THE JOINT ORTHOGONALITY TEST OF TREATMENT ARMS
Daily step counts	5336.555	5171.741	5305.133	0.943
(at the time of entry)	(351.197)	(373.686)	(355.624)	
Age	46.483	45.756	46.890	0.640
	(0.858)	(0.888)	(0.795)	
Female	0.557	0.600	0.543	0.559
	(0.038)	(0.039)	(0.039)	
Married	0.701	0.731	0.732	0.773
	(0.035)	(0.035)	(0.035)	
Years of schooling	15.319	15.381	15.345	0.957
	(0.136)	(0.151)	(0.160)	
Household income	701.616	659.826	685.481	0.529
	(24.939)	(25.841)	(28.364)	
Missing income	0.167	0.119	0.195	0.168
	(0.028)	(0.026)	(0.031)	
Worker	0.833	0.869	0.878	0.458
	(0.028)	(0.027)	(0.026)	
Body weight	61.507	61.148	60.118	0.773
	(1.331)	(1.482)	(1.452)	
BMI	22.003	22.501	22.490	0.307
	(0.251)	(0.281)	(0.261)	
Subjective health	3.822	3.725	3.689	0.390
	(0.064)	(0.077)	(0.074)	
No daily step target	0.230	0.250	0.256	0.842
	(0.032)	(0.034)	(0.034)	
Daily step target	6908.046	6525.000	6780.488	0.670
	(301.841)	(303.277)	(317.108)	
App use frequency	2.937	3.063	2.976	0.832
	(0.149)	(0.149)	(0.150)	
Need for a financial incentive	2.201	2.225	2.128	0.583
	(0.067)	(0.067)	(0.070)	
Missing daily step counts	0.149	0.169	0.128	0.590
(at the time of entry)	(0.027)	(0.030)	(0.026)	
Number of participants	174	160	164	

NOTE. —BMI = body mass index. Standard errors clustered by individual are in parentheses. Daily step counts at the time of entry are the mean daily step counts from December 7–15, 2020 (missing step counts are treated as zero). Household income is defined as the average annual income after proxying the average annual income for those who did not respond to the question about household income.

Table 2

DEMOGRAPHIC CHARACTERISTICS OF PARTICIPANTS WHO CHOSE TO OPT-IN

	OPT-IN	NOT OPT-IN	P-VALUE FROM THE JOINT ORTHOGONALITY TEST OF TREATMENT ARMS
Daily step counts	6433.667	4795.795	0.033
(at the time of entry)	(703.982)	(399.780)	
Age	47.941	46.416	0.376
	(1.388)	(0.969)	
Female	0.549	0.540	0.914
	(0.070)	(0.047)	
Married	0.706	0.743	0.619
	(0.064)	(0.041)	
Years of schooling	15.255	15.385	0.708
	(0.305)	(0.188)	
Household income	745.275	658.494	0.157
	(63.961)	(29.225)	
Missing income	0.196	0.195	0.984
	(0.056)	(0.037)	
Worker	0.882	0.876	0.911
	(0.046)	(0.031)	
Body weight	61.353	61.188	0.933
	(1.649)	(1.090)	
BMI	22.580	22.449	0.816
	(0.426)	(0.328)	
Subjective health	3.863	3.611	0.116
	(0.125)	(0.091)	
No daily step target	0.314	0.230	0.259
	(0.066)	(0.040)	
Daily step target	6960.784	6699.115	0.704
	(644.714)	(358.411)	
App use frequency	4.843	3.655	0.000
	(0.219)	(0.183)	
Need for a financial incentive	2.451	1.982	0.002
	(0.102)	(0.088)	
Missing daily step counts	0.059	0.159	0.076
(at the time of entry)	(0.033)	(0.035)	
Observations	51	113	

NOTE. —Standard errors clustered by individual are in parentheses. Individual characteristics include age, female dummy, married dummy, schooling, income, worker dummy, body weight, and subjective health. We omitted participants who set a target of 0–4000 daily steps, and who do not provide input and check the app. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3

EFFECTS OF TREATMENT ASSIGNMENT AND INCENTIVE ON DAILY WALKING DURING THE TREATMENT PERIOD

	OLS		IV	
	Opt-out	Opt-in	Active	Passive
Take-up rate:	100.0%	31.1%	Joiners	Joiners
	(ATE)	(ITT)	(TOT)	(TOU)
	(1)	(2)	(3)	(4)
$Week_1 \times Treatment\ Assignment$	541.099	709.363**		
	(353.849)	(347.132)		
$Week_2 \times Treatment\ Assignment$	286.818	231.307		
	(380.717)	(374.698)		
$Week_1 \times Receiving\ Incentive$			2281.090**	-244.207
			(1115.715)	(542.240)
$Week_2 \times Receiving\ Incentive$			743.811	80.565
			(1196.032)	(579.896)
Constant	6612.186***	6654.480***		
	(88.293)	(87.161)		
Fixed effect	Yes	Yes	Yes	Yes
Number of participants	334	338	338	324
Number of observations	6346	6422	6422	6156

NOTE.—OLS = ordinary least squares; IV = instrumental variable; ATE = average treatment effect; ITT = intention-to-treat; TOT = treatment effects on the treated; TOU = treatment effects on the untreated. Standard errors clustered by individual are in parentheses. All models include fixed effects for day and individuals × day of the week. Column (1) uses the sample of the control and opt-out groups. Columns (2) and (3) use the sample of the control and opt-in groups. Column (4) uses the sample of the opt-out and opt-in groups. Columns (3) and (4) report the IV estimation results using opt-in and opt-out group dummies as instrument variables for receiving the incentive, respectively. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Column (1) uses the sample of the control and opt-out groups. The coefficients on $Week_1(Week_2) \times Treatment\ Assignment$ capture the average difference in daily steps across the opt-out and control groups in the first (second) week of the treatment period. In the first week of the treatment period, the daily step count for the opt-out group is 541 steps higher than that for the control group; however, this is not statistically significant. In the second week of the treatment period, the daily step count for the opt-out group is 287 higher than that for the control group; however, this is also not statistically significant. Although all participants in the opt-out group accepted the incentive, we did not find significant evidence that offering incentives in the opt-out scheme increases daily step counts on average.

In contrast, we found a significant increase in daily step counts from the opt-in treatment assignment. Columns (2) uses the control and opt-in groups and reports that, in the first week of the treatment period, the daily step count for the opt-in group is 709 steps higher than that for the control group ($p < .05$), although this increase and its statistical significance disappear in the second week of the treatment period. Since the take-up rate of the incentive is 31.1%, this estimated result is the ITT.

Column (3) shows the estimation results for receiving the financial incentive in the opt-in scheme (TOT), which uses the sample of opt-in and control groups. We found that, in the first week of the treatment period, daily step count increased by 2,281 steps among those who actively applied to receive the incentive ($p < .05$). The step counts' increase in the second week is 744 and positive, but not statistically significant.

Column (4) shows the estimation results for the treatment effect for passive individuals who received the incentive under the opt-out condition but not under the opt-in condition (TOU), which uses the sample of opt-out and opt-in groups. The coefficients on $Week_1 \times Receiving\ Incentive$ and $Week_2 \times Receiving\ Incentive$ are negative or close to zero respectively, implying that offering the incentives to passive individuals has no increasing impact on their daily step counts. In summary, our experimental evidence suggests that offering incentives in an opt-in scheme can efficiently increase people's daily step counts better than an opt-out scheme, although its positive effect is short-term.

Table 4

EFFECTS OF TREATMENT ASSIGNMENT AND INCENTIVE ON DAILY WALKING DURING THE POST-TREATMENT PERIOD

	OLS		IV	
	Opt-out	Opt-in	Active	Passive
Take-up rate:	100.0%	31.1%	Joiners	Joiners
	(ATE)	(ITT)	(TOT)	(TOU)
	(1)	(2)	(3)	(4)
$Post\ Week_1 \times Treatment\ Assignment$	184.203	-19.113		
	(359.059)	(358.129)		
$Post\ Week_2 \times Treatment\ Assignment$	-548.622	-367.538		
	(371.251)	(381.833)		
$Post\ Week_1 \times Receiving\ Incentive$			-61.462	295.078
			(1152.182)	(567.774)
$Post\ Week_2 \times Receiving\ Incentive$			-1181.887	-262.813
			(1242.285)	(541.151)
Constant	6725.666***	6759.484***		
	(81.442)	(84.476)		
Fixed effect	Yes	Yes	Yes	Yes
Number of participants	334	338	338	324
Number of observations	6346	6422	6422	6156

NOTE.—OLS = ordinary least squares; IV = instrumental variable; TOT = treatment effects on the treated; TOU = treatment effects on the untreated; =; ITT = intention-to-treat. Standard errors clustered by individual are in parentheses. All models include the fixed effects for day and individuals × day of the week. Column (1) uses the sample of the control and opt-out groups. Columns (2) and (3) use the sample of the control and opt-in groups. Column (4) uses the sample of the opt-out and opt-in groups. Columns (3) and (4) report the IV estimation results using opt-in and opt-out group dummies as instruments for receiving the incentive, respectively. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5

EFFECTS OF TREATMENT ASSIGNMENT AND INCENTIVES ON DAILY STEP COUNTS DURING THE TREATMENT PERIOD BY INDIVIDUAL APP USE FREQUENCY

	HEAVY USERS				LIGHT USERS			
	OLS		IV		OLS		IV	
Take-up rate:	Opt-out 100.0% (ATE)	Opt-in 39.3% (ITT)	Active Joiners (TOT)	Passive Joiners (TOU)	Opt-out 100.0% (ATE)	Opt-in 15.8% (ITT)	Active Joiners (TOT)	Passive Joiners (TOU)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Week_1 \times Treatment\ Assignment$	814.919*	782.453*			100.268	634.653		
	(442.640)	(443.965)			(590.584)	(554.106)		
$Week_2 \times Treatment\ Assignment$	433.951	384.721			45.384	-41.448		
	(494.736)	(476.554)			(598.855)	(616.833)		
$Week_1 \times Receiving\ Incentive$			1993.393*	53.443			4019.468	-634.582
			(1124.176)	(765.223)			(3550.989)	(744.498)
$Week_2 \times Receiving\ Incentive$			980.122	81.040			-262.501	103.112
			(1208.161)	(817.720)			(3926.251)	(806.289)
Constant	7259.673***	7314.691***			5554.430***	5505.681***		
	(113.307)	(115.428)			(141.860)	(132.076)		
Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of participants	207	214	214	207	127	124	124	117
Number of observations	3933	4066	4066	3933	2413	2356	2356	2223

NOTE.—OLS = ordinary least squares; IV = instrumental variable; TOT = treatment effects on the treated; TOU = treatment effects on the untreated; ITT = intention-to-treat. Standard errors clustered by individual are in parentheses. All models include the fixed effects for day and individuals × day of the week. We define *heavy users* as participants who answered "I keep records every day," "I keep records almost every day (sometimes I forget to keep records a couple of times a week)," and "I keep records a couple of times a week," and *light users* as participants who answered "I do not keep records but I check the app every day," "I do not keep records but I check the app a couple of times a week," and "I do not keep records and rarely check the app." Columns (1) and (5) use the sample of the control and opt-out groups. Columns (2), (3), (6), and (7) use the sample of the control and opt-in groups. Columns (4) and (8) use the sample of the opt-out and opt-in groups. Columns (3) and (7) report the IV estimation results using opt-in dummies as instruments for receiving the incentive.

Columns (4) and (8) report the IV estimation results using the opt-out dummy as the instrument for receiving the incentive. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6

EFFECTS OF TREATMENT ASSIGNMENT AND INCENTIVE ON DAILY WALKING WITH ADDRESSED MISSING VALUES DURING THE TREATMENT PERIOD

	UNIMPUTED DAILY WALKING STEPS				IMPUTED DAILY WALKING STEPS			
	OLS		IV		OLS		IV	
Take-up rate:	Opt-out 100.0% (ATE)	Opt-in 34.5% (ITT)	Active Joiners (TOT)	Passive Joiners (TOU)	Opt-out 100.0% (ATE)	Opt-in 31.1% (ITT)	Active Joiners (TOT)	Passive Joiners (TOU)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Week_1 \times Treatment\ Assignment$	630.451**	717.655**			565.338**	642.877**		
	(303.838)	(306.399)			(282.452)	(285.867)		
$Week_2 \times Treatment\ Assignment$	394.944	308.473			317.255	260.412		
	(342.422)	(337.210)			(314.059)	(312.557)		
$Week_1 \times Receiving\ Incentive$			2023.418**	-129.298			1911.237**	-115.933
			(840.905)	(500.268)			(824.176)	(449.671)
$Week_2 \times Receiving\ Incentive$			811.164	131.658			716.982	81.744
			(934.526)	(547.408)			(907.478)	(484.525)
Constant	7513.984***	7615.263***			7441.752***	7550.744***		
	(76.372)	(76.443)			(73.602)	(74.167)		
Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of participants	321	323	323	312	329	334	334	323
Number of observations	5525	5552	5552	5307	5832	5862	5862	5656

NOTE.—OLS = ordinary least squares; IV = instrumental variable; TOT = treatment effects on the treated; TOU = treatment effects on the untreated; ITT = intention-to-treat. Standard errors clustered by individual are in parentheses. All models include the fixed effects for day and individuals × day of the week. Columns (1) and (5) use the sample of the control and opt-out groups. Columns (2), (3), (6), and (7) use the sample of the control and opt-in groups. Columns (4) and (8) use the sample of the opt-out and opt-in groups. Columns (3) and (7) report the IV estimation results using the opt-in dummy as the instrument for receiving the incentive. Columns (4) and (8) report the IV estimation results using the opt-out dummy as the instrument for receiving the incentive. The dependent variables in Columns (1)–(4) are treated as missing when the daily step count is 0. The dependent variables in Columns (5)–(8) are imputed as the individual-level mean value for each day of the week prior to the treatment period for missing values. If the daily step counts were missing before the treatment period for which the mean is calculated, they are not included in the calculation of the mean. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix 1. Messages

Figure A1a. Messages during the experiment (control conditions).

Translations:

① We are pleased to bring you “Information from the City of Kobe.”

② Let’s beat COVID-19! Health Promotion Event 2021

③ Walk keeping distance between yourself and others!

④ Check out last week’s step counts. Walk this week, wearing a mask and keeping distance between yourself and others.

⑤ Check out last week’s step counts. Would you like to walk at a different time of day this week? Even on your usual path, you may see a different view at a different time.

Figure A1b. Messages during the experiment (treated conditions).

Translations:

① For two weeks, from February 8 to February 21, you will earn an additional 5 JPY gift certificate for every 1,000 step counts you take.

② If you make more effort than usual and walk 10,000 steps every day, you will earn an additional 700 JPY gift certificate!

③ From today through February 21, you can earn an additional 5 JPY Amazon gift certificate for your every 1,000 steps (the daily upper limit is 14,000 steps). Check out last week’s step counts. Walk this week, wearing a mask and keeping distance between yourself and others.

④ For the remaining week until February 21, you can earn an additional 5 JPY Amazon gift certificate for your every 1,000 steps (the daily upper limit is 14,000 steps). Check out last week’s step counts. Would you like to walk at a different time of day this week? Even on your usual path, you may see a different view at a different time.

Appendix 2. Effects on Body Weight

Here, we examine the treatment effects on body weight. An increase in daily steps could result in a loss of body weight. Our outcome measure is body weight recorded in the app. The app automatically recorded participants' daily step counts; however, they were required to enter their body weight themselves. To encourage participants to provide their body weight information, we paid an additional incentive of 100 JPY for each participant who entered their weight at least once during each of the following three periods: January 4–14, February 22–28 (first week of the post-treatment period), and March 1–7 (second week of the post-treatment period).

We estimate a DID specification using average body weight per week as the outcome variable. Participants did not enter their body weight every day; therefore, we first conduct the estimation by calculating the average using only the entered body weight. Second, if a body weight is not recorded, we assume no changes in body weight up to that date and impute the weight entered in the past for the analysis.

Tables A2a and A2b show the treatment effects on body weight during the treatment and post-treatment periods, respectively.[8]

We observe no evidence of decreasing body weight from any treatment assignment group, either during or after the treatment period (Columns (1)–(4) in Tables A2a and A2b). This finding does not change when we analyze the results by imputing the most recently entered weight for the missing weight (Columns (5)–(8) in Tables A2a and A2b). Although we find providing a financial incentive increased daily step counts in the opt-in group, this increase was insufficient to cause weight loss. This finding is consistent with the fact that the effect of the financial incentive on increasing step counts was short-term.

[8] Tables A2a and A2b show the results of the analysis using body mass index instead of body weight, which show no differences from the results using body weight.

Table A2a.

EFFECTS OF TREATMENT ASSIGNMENT AND INCENTIVE ON BODY WEIGHT DURING THE TREATMENT PERIOD

	UNIMPUTED BODY WEIGHT				IMPUTED BODY WEIGHT			
	OLS		IV		OLS		IV	
	Opt-out	Opt-in	Active	Passive	Opt-out	Opt-in	Active	Passive
Take-up rate:	100.0%	41.6%	Joiners	Joiners	100.0%	35.8%	Joiners	Joiners
	(ATE)	(ITT)	(TOT)	(TOU)	(ATE)	(ITT)	(TOT)	(TOU)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Week_1 \times Treatment\ Assignment$	0.105				0.093			
	(0.102)				(0.089)			
$Week_2 \times Treatment\ Assignment$	0.140				0.115			
	(0.118)				(0.095)			
$Week_1 \times Receiving\ Incentive$		-0.066	-0.154	0.105		-0.053	-0.132	0.093
		(0.089)	(0.204)	(0.102)		(0.080)	(0.184)	(0.089)
$Week_2 \times Receiving\ Incentive$		-0.006	-0.010	0.140		-0.009	-0.009	0.115
		(0.092)	(0.207)	(0.118)		(0.082)	(0.190)	(0.095)
Constant	60.664***	60.693***			60.361***	60.587***		
	(0.033)	(0.029)			(0.030)	(0.028)		
Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of participants	203	212	212	193	269	283	283	260
Number of observations	575	603	603	548	745	775	775	718

NOTE.—OLS = ordinary least squares; IV = instrumental variable; TOT = treatment effects on the treated; TOU = treatment effects on the untreated; ITT = intention-to-treat. Standard errors clustered by individual are in parentheses. All models include the fixed effects for day and individuals × day of the week. Columns (1) and (5) use the sample of the control and opt-out groups. Columns (2), (3), (6), and (7) use the sample of the control and opt-in groups. Columns (4) and (8) use the sample of the opt-out and opt-in groups. Columns (3) and (7) report the IV estimation results using the opt-in dummy as the instrument for receiving the incentive. Columns (4) and (8) report the IV estimation results using the opt-out dummy as the instrument for receiving the incentive. Columns (1)–(4) and (5)–(8) use the unimputed body weight and imputed body weight for missing data to the latest body weight, respectively. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A2b.

EFFECTS OF TREATMENT ASSIGNMENT AND INCENTIVE ON BODY WEIGHT DURING THE POST-TREATMENT PERIOD

	UNIMPUTED BODY WEIGHT				IMPUTED BODY WEIGHT			
	OLS		IV		OLS		IV	
	Opt-out	Opt-in	Active	Passive	Opt-out	Opt-in	Active	Passive
Take-up rate:	100.0%	41.6%	Joiners	Joiners	100.0%	35.8%	Joiners	Joiners
	(ATE)	(ITT)	(TOT)	(TOU)	(ATE)	(ITT)	(TOT)	(TOU)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Week_1 \times Treatment\ Assignment$	0.143				0.153			
	(0.119)				(0.106)			
$Week_2 \times Treatment\ Assignment$	0.119				0.114			
	(0.128)				(0.111)			
$Week_1 \times Receiving\ Incentive$		0.017	0.051	0.143		-0.027	-0.040	0.153
		(0.096)	(0.212)	(0.119)		(0.090)	(0.205)	(0.106)
$Week_2 \times Receiving\ Incentive$		-0.079	-0.183	0.119		-0.081	-0.204	0.114
		(0.101)	(0.221)	(0.128)		(0.096)	(0.221)	(0.111)
Constant	60.635***	60.904***			60.603***	60.868***		
	(0.039)	(0.033)			(0.036)	(0.033)		
Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of participants	223	234	234	217	289	304	304	283
Number of observations	621	651	651	602	785	817	817	764

NOTE.—OLS = ordinary least squares; IV = instrumental variable; TOT = treatment effects on the treated; TOU = treatment effects on the untreated; ITT = intention-to-treat. Standard errors clustered by individual are in parentheses. All models include the fixed effects for day and individuals × day of the week. Columns (1) and (5) use the sample of the control and opt-out groups. Columns (2), (3), (6), and (7) use the sample of the control and opt-in groups. Columns (4) and (8) use the sample of the opt-out and opt-in groups. Columns (3) and (7) report the IV estimation results using the opt-in dummy as the instrument for receiving the incentive. Columns (4) and (8) report the IV estimation results using the opt-out dummy as the instrument for receiving the incentive. Columns (1)–(4) and (5)–(8) use unimputed body weight and imputed body weight for missing data to the latest body weight, respectively. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.