Incentives to Persevere

Elif Incekara-Hafalir
University of Technology Sydney
elif.incekarahafalir@uts.edu.au

Grace HY Lee
Monash University (Malaysia)
grace.lee@monash.edu

Audrey KL Siah
Monash University (Malaysia)
audrey.siah@monash.edu

Erte Xiao*
Monash University
erte.xiao@monash.edu

Abstract
Achieving success often requires persistent effort. We study the effectiveness of two reward mechanisms, all-or-nothing and piece-rate, to incentivize full completion of repeated tasks over time. Our theoretical analysis shows that exogenously imposing the all-or-nothing mechanism can be ineffective due to the potential discouragement effect. In contrast, empowering individuals to choose between the two reward mechanisms can significantly improve the full completion rate. Data from a series of field experiments and follow up replications provide robust evidence that the all-or-nothing mechanism is effective only when it is presented as an option. Our results highlight the importance of choice in incentivizing persistent effort.

Keywords: perseverance; incentives; self-control; field experiments
JEL codes: C93, D91, D82, I20

Acknowledgements. We thank seminar participants at Advances with Field Experiments 2017, University of Chicago; Department of Economics, Monash University; Department of Economics, University of Technology Sydney; Department of Economics, University of Melbourne; Department of Economics, George Mason University; Behavioural Economics: Foundations and Applied Research workshop at the University of Sydney; Behavioural and Experimental Economics and Finance Workshop at the University of Sydney; Foundations of Utility and Risk 2018, Tsinghua Conference on Behavioral, Experimental and Theoretical Economics 2019, University of York; Economic Science Association World Meeting 2021, Humboldt-University Berlin; Department of Economics, RMIT for valuable comments. This research is supported by Monash University (Malaysia) Better Teaching Better Learning Education Grant (MUM-BTBL-2016- 010), Australia-Malaysia 2021-2022 Research Collaboration Development Scheme (Monash Faculty of Business and Economic), and University of Technology Sydney Business School.

1. Introduction

Perseverance plays an essential role in achieving success and accomplishing goals (Beattie et al. 2018, Duckworth 2016, Heckman and Kautz 2012; Heckman et al. 2006). In organizations, completing a project typically requires continuous effort for an extended period; further, the value of an unfinished project is often disproportionally lower than that of a completed project. For instance, developing a computer software package requires considerable hard work – it can involve writing codes and testing software for months or even years. Incomplete or malfunctioning software has little or no market value. In a sales setting, a salesperson needs to continually meet and work with potential clients to sell products and reach their sales quota. In education, to master a subject or a skill, participants need to engage in the whole program, and in some fields, they may be required to practice for several months or even years. Perseverance, however, is by its very definition hard. Studies show that non-cognitive skills and psychological factors such as self-control, grit, conscientiousness, and growth mindset are important for perseverance (Almlund et al. 2011, Claro et al. 2016, Tangney et al. 2004). While some people are able to rely on their self-discipline and have the self-control required to succeed, others may not possess such attributes. It is therefore important to explore whether external mechanisms can be designed to incentivize persistent effort.

We study both theoretically and experimentally the behavioral differences between three mechanisms: 1) piece-rate; 2) all-or-nothing; and 3) self-select. The first two are commonly used output-based schemes. One is a continuous scheme, a "piece-rate" mechanism that provides a reward for each task completed. The other is a discrete scheme, an "all-or-nothing" mechanism, that rewards effort if, and only if, all the required tasks are completed. The piece-rate can indirectly promote persistence by incentivizing each effort without imposing a cost on giving up. In contrast, the all-or-nothing directly promotes persistent efforts by increasing the cost of giving up, and in doing so, providing stronger incentives for overcoming potential self-control problems (Giné et al. 2010, Kaur et al. 2015, Royer et al. 2015). By comparing the all-or-nothing to the piece-rate, we can learn about the effects of raising the price of giving up on persistence.

Both the piece-rate and the all-or-nothing mechanisms have been discussed in the literature in relation to workplace compensation and incentives (Chung et al. 2014, Jain 2012, Kishore et al. 2013, Schöttner 2017). Examples of the piece-rate mechanisms include sales compensation or payment to manufacturing workers based on the quantity of their output. All-or-nothing mechanisms are usually observed in jobs that focus on a particular level of output. Contractors are often paid under such a discrete scheme. For example, many jobs listed on Task Rabbit require workers to complete a specific task before being paid a fixed payment amount (Lazear 2018). The advantages and the disadvantages of the two types are mostly examined in the context of the sales workforce compensation (Chung et al. 2014, Joseph and Kalwani 1998). For instance, quota-based bonuses are similar to the all-or-nothing mechanism where salespeople receive a lump-sum bonus once sales quotas have been met. Sales commissions, on the other hand, are often paid in piece-rates. While the quota bonus may motivate workers to achieve targets that workers otherwise may not attain, it can also be risky in that salespeople may be discouraged when they are unlikely to reach the target, or they may stop making an effort once the quotas have been met (Darmon 1997).

A novel contribution of this paper is that we introduce a self-select mechanism, where agents have the option to choose between the piece-rate and the all-or-nothing reward mechanisms. In the context of designing compensation packages for a sales workforce, for example, the manager could ask the employees to choose between sales commissions and quota-based bonuses. The self-select mechanism is motivated by the hypothesis that reward mechanisms (which is designed to help people overcome self-control problems) may be more effective when agents are able to choose what suits them best. In practice, given the risk involved in the all-or-nothing mechanism, the self-select mechanism might also be more acceptable to individuals and therefore more feasible to implement.

To obtain clear evidence on the behavioral differences of the three mechanisms, we start with the simplest form of a continuous scheme (i.e., piece-rate) and a discrete scheme (i.e., all-or-nothing). These schemes are comparable in all aspects except that the former scheme provides rewards for each completed task and the latter scheme provides rewards only when all tasks are completed. In particular, under the all-or-nothing mechanism, agents have to complete all tasks to receive the payment equal to the payment received by piece-rate agents when all tasks are completed. This design allows us to test whether individuals in the self-select mechanism will choose the all-or-nothing scheme because of its commitment value, which is the key for the self-select mechanism to achieve a higher full completion rate, e.g. share of people completing all tasks, than the piece-rate based on our theoretical framework. Moreover, using the simple all-or-nothing mechanism allows us to align our design with the predictions of full completion based on a simple theory.

In our theoretical model, a decision-maker has self-control problems and is fully aware of this (i.e., the decision-maker is sophisticated) in an environment with uncertainty associated with the costs of completing future tasks. First, compared with the piece-rate mechanism, the all-or-nothing mechanism can be a double-edged sword. On one hand, it helps the decision-maker to overcome their self-control problem by incentivizing full completion of all the tasks (i.e., an “encouragement effect”). On the other hand, under this mechanism, those who are not certain that the future tasks will be completed may quit at the beginning (i.e., a “discouragement effect”). Thus, the full completion rate in the all-or-nothing mechanism can be either higher or lower compared with the piece-rate mechanism. In contrast, the self-select mechanism allows participants to choose the mechanism that works best for them. As a result, the self-select mechanism avoids the potential backfire of the all-or-nothing mechanism, as those who were likely to quit under the all-or-nothing mechanism are able to simply choose the piece-rate mechanism instead. Meanwhile, those who perform better under the all-or-nothing mechanism may benefit from having the commitment option available. We show that the self-select mechanism performs better than the piece-rate mechanism as long as there are encouraged decision makers who self-select the all-or-nothing mechanism.

We conduct field experiments to provide empirical evidence on the effectiveness of the all-or-nothing and self-select mechanisms at promoting full completion compared with the piece-rate mechanism.[1] In particular, based on our findings from theoretical analysis, we investigate the following questions: Does the all-or-nothing mechanism promote or inhibit the full completion rate compared to the piece-rate mechanism? Do people choose the all-or-nothing reward mechanism when given the option? Is the self-select mechanism effective at promoting persistent effort compared to the piece-rate mechanism?

We investigate these questions in the setting of university coursework where: persistent effort is important for the learning outcome (e.g. completion of every assignment throughout the semester); and students may suffer from self-control problems. The tasks were weekly quizzes. We randomly assigned students to one of three treatments. In all treatments, students were told that the quizzes were worth 10 marks in total. Students in the baseline (*Piece-Rate*) treatment were told that the 10 marks would be spread equally across each quiz.[2] In the *All-Or-Nothing* treatment, students were awarded 10 marks only if they completed all quizzes, and they would receive no marks if they missed one quiz or more. In the *Self-Select* treatment, students were given a choice between the two marking mechanisms.

We find, compared with the baseline *Piece-Rate* treatment, the *All-Or-Nothing* treatment did not have a higher full completion rate. Interestingly, in the *Self-Select* treatment, more than half of the students chose the all-or-nothing mechanism in spite of the risk of receiving no marks. Further, the full completion rate is significantly higher in the *Self-Select* treatment than the *Piece-Rate* treatment. To test the robustness of the positive effect of the self-select mechanism, we conduct additional replication trials in different courses in the following years. Although the proportion of students who choose the all-or-nothing mechanism varies across the trials, the self-select mechanism always achieves a higher full completion rate compared to the piece-rate mechanism.

Our paper contributes to the literature investigating the effects of different incentives on helping people achieve their long-term goals, particularly goals requiring self-control (Aggarwal et al. 2020, Bryan et al. 2010, Carrera et al. 2020, Himmler et al. 2019, Kullgren et al. 2016, Toussaert 2018). We can classify the previously analyzed incentives as either piece-rate or all-or-nothing types. All-or-nothing types are those where the reward explicitly requires persistent effort over a period of time. With such incentives, the marginal return for an effort can be zero or significantly reduced due to either a subsequent or preceding failure. For example, Kaur et al. (2015) study a non-linear dominated contract at a workplace where, each day, workers can choose a target and receive only half of the piece-rate if they fail to reach the target. Likewise, under a deposit contract for smoking cessation (Giné et al. 2010) and for gym attendance (Royer et al. 2015), people lose their entire

[1] Registered RCT ID: AEARCTR-0004518

[2] To differentiate between the treatment and the mechanism, the treatment names are in italics but the reward mechanism names are not.

deposit if they give in to temptation even once during a set timeframe. A one-time failure means none of the previous days of abstinence or of gym attendance is rewarded. Another example is a lump-sum payment conditional on the full completion of a certain number of gym visits (Acland and Levy 2015, Charness and Gneezy 2006).

In contrast, another type of incentive is similar to the piece-rate mechanism where there is a positive reward for each success in taking the desired action; bundling an instantly gratifying experience (e.g. listening to a page-turning, low-brow audio novel) with actions requiring self-control (e.g. exercising) decreases the cost of exerting an effort at each instance (Milkman et al. 2014). Other examples include: buying a flat-rate gym membership which results in a decrease in the pay-per-visit mental cost of gym visits for *every gym visit* (DellaVigna and Malmendier 2006); or signing up for a savings scheme with deposit collection services which decreases the transaction cost of making future deposits into a savings account – which is a reward for each deposit (Ashraf et al. 2006). Like the *Piece-Rate* treatment in our paper, a common feature of these examples is that individuals are rewarded for each success, while a failure has little impact on either the size of reward for previous successes or on subsequent actions.

To the best of our knowledge, there is no study systematically comparing the effectiveness of the all-or-nothing mechanism and the piece-rate mechanism at promoting persistent efforts, nor one that investigates the effect of choosing between the two mechanisms. We fill this gap by providing theoretical analysis and empirical evidence for the comparison between the two mechanisms: both when the mechanisms are exogenously enforced and when participants are empowered to choose the reward mechanism.[3]

Our paper also contributes to the recent research in education examining the effectiveness of external interventions on perseverance after a failure, such as classroom education programs or peer pressure (Alan et al. 2019, Bettinger et al. 2018, Buechel et al. 2018). This line of research focuses on the resilience aspect of perseverance, i.e., the willingness to undertake challenging tasks and/or to continue working on tasks after an initial failure. However, even if the task itself is easy, individuals often fail to complete a task if it requires repeated engagement over a long period of time. Everyday people face all kinds of temptations distracting them from continuing what they have started, especially when they have little personal interest in a task or have self-control problems. The effort

[3] To date, the importance of choice has not received much attention in research on designing incentive mechanisms to overcome self-control problems. Two exceptions are Ariely and Wertenbroch (2002) and Bisin and Hyndman (2020), who study the effect of self-imposed deadlines on outcomes. Ariely and Wertenbroch (2002) find performance under self-imposed deadlines is lower than performance under exogenously imposed, evenly spaced deadlines due to the suboptimal timing of chosen deadlines. On the other hand, self-imposed deadlines achieve a better outcome than those under the longest delayed deadlines. Conversely, Bisin and Hyndman (2020) find that neither the exogenous nor the self-imposed deadlines improve the outcomes. Both Ariely and Wertenbroch (2002) and Bisin and Hyndman (2020) ask their participants to set their own deadlines for multiple tasks in the self-imposed deadlines condition. It is probably much harder to set optimal multiple deadlines as compared to making an optimal choice between two marking mechanisms that have the exact same criteria (getting at least 50% correct answers) in our study.

costs of tasks may also be occasionally high, which increases the chance of giving up. This important aspect of perseverance has received relatively less attention. Our study fills this gap.

The rest of the paper is organized as follows. Section 2 presents a theoretical analysis and comparison of the mechanisms. Section 3 lays out the experimental design. Section 4 reports the findings. Section 5 explores individual differences in the take-up decisions and discusses other possible mechanisms that may explain the results. Section 6 concludes and discusses the implications of the findings.

2. Theoretical analysis

We compare the full completion rate (i.e., the likelihood of completing all the tasks) under the exogenously assigned piece-rate mechanism and the all-or-nothing mechanism, as well as under the self-select mechanism. We construct a simple theoretical framework built on the quasi-hyperbolic discounting model (Laibson 1997, O'Donoghue and Rabin 1999b, Phelps and Pollak 1968).[4] A risk neutral decision maker (henceforth DM) discounts the future utility streams starting with today by $\{1, \beta\delta, \beta\delta^2, \ldots\}$ with δ and β capturing impatience and present bias respectively. There are three periods in the model. The timeline of events is summarized in Figure 1.

There are two tasks to be completed in Period 1.[5] The effort cost of each task may be low or high, $c \in \{c_L, c_H\}$ where probability of c_L is p and c_H is 1-p. We assume p>0.5 as our interest is in the case when the task per se is not difficult. The uncertainty of cost can be attributed to exogenous factors, such as getting sick or new tempting activities arising. These factors might play a key role in failing to persist on tasks. A commitment device is valuable to deal with self-control problems, however, the uncertainty of cost creates a trade-off between demand for commitment and flexibility (Amador et al. 2006, Augenblick et al. 2015, Carrera et al. 2020, Laibson 2015). We assume δ=1 for simplicity and focus on sophisticated DMs, e.g. a DM's belief on the value of β is correct, and therefore the DM values commitments (Basu 2011, Kaur et al. 2015, O'Donoghue and Rabin 1999a). We restrict our focus to cases where the per task reward is always higher than cost ($c_L < c_H < r = 1$). We also make an additional technical assumption, $(1 + p)/p < c_H/c_L$, to restrict the number of cases to analyze (see the online appendix for the solutions of the other cases, which are qualitatively the same).

[4] Other well-known intertemporal choice models include: the planner-doer model (Shefrin and Thaler 1988); the dual-self model (Fudenberg and Levine 2006); the temptation model (Gul and Pesendorfer 2001); and the internal commitment model (Benhabib and Bisin 2005). We use the quasi-hyperbolic discounting model because of its tractability and prevalence to analyze this type of self-control problem.

[5] More than three tasks will be less trackable. We group the two tasks in Period 1 to simplify the analysis. The grouping will not change the qualitative model predictions, as long as the choice of the mechanism takes place before task completion and the reward is given after the tasks.

Figure 1. Timeline of events

Period 0	Period 1	Period 2
DM is assigned to (or chooses) a mechanism (piece rate or all-or-nothing)	DM decides whether to complete each of two tasks. Each task cost is c $\in \{c_L, c_H\}$ where probability of c_L is p>0.5 and c_H is 1-p	DM receives rewards from task completion: • In piece rate: a reward r ($c_L < c_H < r = 1$) for each completed task. • In all-or-nothing: full reward of $2r$ only if both tasks are completed and zero otherwise.

2.1 Piece-Rate versus All-Or-Nothing

The following function determines whether a DM will complete a task under the piece-rate mechanism:

$$e_i^* = \underset{e_i \in \{0,1\}}{\operatorname{argmax}}\, e_i(-c + \beta) \tag{1}$$

where $c \in \{c_L, c_H\}$ is the effort cost and e_i is the binary effort decision for each task *i*. Solving the DM's problem under the piece-rate mechanism shows three types of behavior:

- T_1 type, $\beta \geq c_H$, always completes a task;
- T_2 type, $\beta < c_L$, never completes a task; and
- T_3 type, $c_L \leq \beta < c_H$, completes each task only if the cost is low.

In contrast to the piece-rate mechanism, the all-or-nothing mechanism requires the DM to use backward induction to solve the problem because the decision on the first task is affected by the expected action on the second task. The utility maximization problem for the second task is:

$$e_2^*(e_1, c_2) = \underset{e_2 \in \{0,1\}}{\operatorname{argmax}}\, e_2(-c_2 + 2\beta e_1) \tag{2}$$

where $c_2 \in \{c_L, c_H\}$ is the cost for the second task. We write the utility maximization for the first task as:

$$e_1^* = \underset{e_1 \in \{0,1\}}{\operatorname{argmax}}\, e_1[(-c_1 + pe_2^*(e_1, c_L)(2\beta - c_L) + (1-p)e_2^*(e_1, c_H)(2\beta - c_H)]. \tag{3}$$

The first task decision depends not only on the cost of the current task but also on the prospect of whether the second task will be performed. It is easy to see that the T_1 type will complete both tasks and the T_2 type will never complete a task under the all-or-nothing mechanism, as they would do under the piece-rate mechanism. Solving (2) and (3) together, the T_3 type can be further divided into four types that vary on whether the DM is encouraged or discouraged by the expected performance in the second task (see details in Appendix).

Depending on the parameter values, for some DMs, completing the first task will lead them to complete the second task regardless of cost, which in turn encourages them to complete the first task.

We call this an "encouragement effect". When the encouragement effect is large enough, the DM will complete the first task regardless of cost. We call this type the "encouraged DM". When the encouragement effect is not large enough to convince the DM to complete the first task regardless of cost, she completes it only when the cost is low. We call this type the "conditionally encouraged DM".

In contrast to the encouragement effect, after completing the first task some other types of DMs will go on to complete the second task only if the cost is low. As a result, when deciding whether or not to complete the first task, the DM will be discouraged by the possibility of the high cost of the second task. We call this a "discouragement effect". If the discouragement effect is strong enough, the DM will not complete the first task even if the cost is low. We call this type the "discouraged DM". If the discouragement is not as strong, the DM will take the risk and complete the first task when the cost is low, hoping that the cost for the next task will also be low. We call this type the "conditionally discouraged DM".

In sum, the all-or-nothing mechanism can improve the full completion rate for the encouraged types, but this mechanism can backfire for the discouraged types, which we state in the Lemma below. Parameter values such as effort costs (c_L, c_H), the probability of low cost (p), and the present bias (β), will determine the DM's type and thereby the difference in the full completion rate between the all-or-nothing and the piece-rate mechanisms. We summarize parameter values of each type and likelihood of full and zero completion under two mechanisms in Table 1 (the details are in Appendix).

Lemma 1: Encouraged types are more likely but the discouraged type is less likely to complete both tasks under all-or-nothing compared to those under piece-rate.
Proof. See Appendix.

It is important to highlight that the encouragement and discouragement effects occur because of the interdependence of the tasks under the all-or-nothing mechanism, not because of the present bias (β).[6] The present bias, however, creates a conflict between the different selves. Specifically, a DM's period-zero-self wants to make more patient decisions and chooses the all-or-nothing mechanism for this reason – stripping away the reward from partial completion to force the period-one-self to complete both tasks. This is similar to the idea of many commitment models that restrict the choice set to commit a DM to a better outcome (Giné et al. 2010, O'Donoghue and Rabin 2001). On the other hand, for some other types of DMs, the period-zero-self prefers the non-commitment

[6] Note that the encouragement and discouragement effects differ from the optimism prevalent in naïve people, e.g., incorrectly believing that they will complete a task as planned, and the pessimism characteristic of the sophisticated, e.g., correctly perceiving the future self-control problems (O'Donoghue and Rabin 1999a). For example, sophistication in O'Donoghue and Rabin (1999a) allows DMs to take an optimal plan of action today by taking the future self-control problems into account. In our case, sophistication will enable DMs to use the all-or-nothing mechanism as a commitment device only when this mechanism creates the encouragement effect.

option (the piece-rate mechanism) as the all-or-nothing mechanism discourages their future self from performing the tasks. We elaborate below how the DM chooses between two mechanisms.

Table 1: Summary of types and predictions

<table>
<tr><th rowspan="3">Types</th><th rowspan="3">Cutoffs for parameters</th><th colspan="6">Likelihood of completion</th><th rowspan="3">Choice between PR and AON</th></tr>
<tr><th colspan="3">Piece-rate (PR)</th><th colspan="3">All-or-nothing (AON)</th></tr>
<tr><th>Two tasks (Full)</th><th>One task (Partial)</th><th>No tasks (Zero)</th><th>Two tasks (Full)</th><th>One task (Partial)</th><th>No tasks (Zero)</th></tr>
<tr><td>Always</td><td>$c_H \leq \beta$</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>PR</td></tr>
<tr><td>Encouraged</td><td>$\frac{(2-p)c_H + pc_L}{2} \leq \beta < c_H$</td><td rowspan="4">$p^2$</td><td rowspan="4">$2p(1-p)$</td><td rowspan="4">$(1-p)^2$</td><td>1</td><td>0</td><td>0</td><td>AON</td></tr>
<tr><td>Cond. Encouraged</td><td>$\frac{c_H}{2} \leq \beta < \frac{(2-p)c_H + pc_L}{2}$</td><td>$p$</td><td>0</td><td>$(1-p)$</td><td>PR</td></tr>
<tr><td>Cond. Discouraged</td><td>$\frac{(1+p)c_L}{2p} \leq \beta < \frac{c_H}{2}$</td><td>$p^2$</td><td>$p(1-p)$</td><td>$(1-p)$</td><td>PR</td></tr>
<tr><td>Discouraged</td><td>$c_L \leq \beta < \frac{(1+p)c_L}{2p}$</td><td>0</td><td>0</td><td>1</td><td>PR</td></tr>
<tr><td>Never</td><td>$\beta < c_L$</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>PR</td></tr>
</table>

2.2 Piece-Rate versus Self-Select

We first analyze how the DM at period zero chooses between the piece-rate and the all-or-nothing mechanisms by comparing the expected lifetime utility of each. We show in both mechanisms that the T_1 type DMs will always complete all the tasks and the T_2 types will never commence any tasks. Thus, these two types will be indifferent to the two mechanisms. We assume that the DM chooses the piece-rate mechanism when indifferent because it is the status quo familiar to them (e.g. students usually receive credit for the work done for each assignment).

We demonstrate above that the T_3 type will perform each task if the cost is low under the piece-rate mechanism. However, they act differently under the all-or-nothing mechanism depending on whether they are encouraged or discouraged, which we elaborate below. We summarize their choice in the last column of Table 1.

- *Encouraged DM:* The period-one-self completes all tasks regardless of the cost. Anticipating this, the period-zero-self strictly prefers the all-or-nothing mechanism.

- *Conditionally encouraged DM*: Under the all-or-nothing mechanism, the period-one-self will complete the second task as long as the first task is complete, and they will complete the first task only if the cost is low. Thus, the likelihood of full completion under the all-or-nothing mechanism is higher than under the piece-rate mechanism. However, the period-one-self will not complete any task if the cost of the first task is high under the all-or-nothing mechanism. In contrast, under the piece-rate mechanism, zero completion occurs only when the cost of both tasks is high, making the partial completion (completing one task) more likely. Therefore, the all-or-nothing mechanism increases the likelihood of zero completion compared to the piece-rate mechanism ($1–p$ vs. $(1–p)^2$). Overall, the expected utility under the piece-rate mechanism is higher than for that under the all-or-nothing mechanism. Thus, the period-zero-self will choose the piece-rate mechanism.
- *Discouraged DM*: The DM will not engage in either of the tasks under the all-or-nothing mechanism. Thus, the period-zero-self will choose the piece-rate mechanism.
- *Conditionally discouraged DM*: We have demonstrated above that this type of DM achieves the same full completion rate under both mechanisms (p^2). Furthermore, under the all-or-nothing mechanism, the period-one-self will not complete the first task with high cost. Thus, the probability of zero completion is higher under the all-or-nothing mechanism than under the piece-rate mechanism ($1–p$ vs. $(1–p)^2$). Overall, the expected utility under the piece-rate mechanism is higher than under the all-or-nothing mechanism, causing the period-zero-self to choose the piece-rate mechanism.

In sum, our analysis shows that encouraged DMs strictly prefer the all-or-nothing mechanism and always achieve full completion, whereas the full completion rate for them would be only p^2 under the piece-rate mechanism. All other types choose the piece-rate mechanism. The self-select mechanism eliminates the risk of discouragement from an exogenously imposed all-or-nothing mechanism. Hence, the self-select mechanism can achieve a higher full completion rate than the piece-rate mechanism as long as there are encouraged DMs in the population. We write the predictions of our model in the proposition below.

Proposition 1: The *Self-Select* treatment improves the full completion rate when encouraged types are present. The *All-Or-Nothing* treatment, on the other hand, does not necessarily increase the full completion rate and may even backfire if the discouragement effect dominates the encouragement effect.

Proof. See Appendix.

2.3 Discussion of the assumptions and predictions of the model

We focus on the sorting effect as the channel through which the self-select mechanism may improve outcomes. One may consider other channels, such as the autonomy effect and signaling. We discuss these possible channels in Section 5.

We assume sophistication when exploring the sorting effect. If the DM is naïve, then they would not choose the all-or-nothing mechanism. Consequently, the self-select mechanism would not increase the full completion rate. If the DM was partially naïve, one who underestimates their self-control problem, they may choose the wrong mechanism, and under certain circumstances they may achieve suboptimal outcomes in the *Self-Select* treatment. For example, a partially naïve DM may incorrectly believe that they are the encouraged type although in reality they are (conditionally) discouraged or T_2 types. In that case, the DM would choose the all-or-nothing mechanism, but fail to complete all tasks. On the other hand, partial naïveté is not necessarily worse off in the *Self-Select* treatment. For example, if the DM incorrectly believes that they are the T_1 type although in reality they are the encouraged type or any other type we mentioned above, they will choose the piece-rate, and the self-select mechanism will neither help nor hurt the DM.

For simplicity, we assume that there is no intrinsic reward for the tasks in that the only reward for completing the tasks is the external incentive provided by the mechanism. The intrinsic value will change the cutoffs for the parameter values of each type. However, our prediction that the *Self-Select* treatment improves the full completion rate still holds.[7]

It is also worth pointing out that our theoretical analysis suggests that the choice of the all-or-nothing mechanism does not increase linearly with β (see the related work of Carrera et al. (2020) and John (2020) on non-monotonic demand for commitment). First, both very low and very high β DMs may choose not to select the all-or-nothing mechanism, but for different reasons. Very low β DMs may not choose the all-or-nothing mechanism because of its discouragement effect, whereas very high β DMs may not choose it because they do not need it. Second, according to our model, the choice of the all-or-nothing mechanism is not only determined by β but also by the belief in the future cost distribution for task completion (see Laibson (2015), Carrera et al. (2020), and Strack and Taubinsky (2021) for similar theoretical models). To elaborate, consider a DM with low β. When expecting an adverse cost distribution, they may be discouraged and, as a result, not choose the all-or-nothing mechanism. But when expecting a favorable cost distribution, they may be encouraged and choose the all-or-nothing mechanism. Likewise, a DM with high β may be encouraged and motivated to choose all-or-nothing under an adverse cost distribution. In contrast, under a favorable distribution, they may not need a commitment device, and consequently not choose the all-or-nothing mechanism. These implications of our theoretical analysis are in line with previous studies that show no empirical

[7] Formal analysis is available from the authors upon request.

relationship between the commitment take-up rate and individual characteristics, such as time preferences (Carrera et al. 2020, John 2020, Laibson 2015).

Nevertheless, the theoretical analysis above provides insight into how the share of the encouraged type may change with particular parameters while holding the other parameters constant. Note that, based on the proposition, the difference between the upper and the lower bound of β for the encouraged type is $\frac{p}{2}(c_H - c_L)$. An implication is that, as the difference between the high and low cost $(c_H - c_L)$ or the probability of low cost (p) decreases, the share of encouraged types choosing all-or-nothing will be smaller. Although this prediction is not the focus of this paper and cannot be tested in our current study, it would be interesting to design a future experiment to examine how the choice of the all-or-nothing mechanism varies with the changes to these parameters.

3. Experiment design

Our first experiment (Trial 1) involved 343 undergraduate students enrolled in the introductory microeconomics course in Malaysia between July and October 2016. This is a compulsory course for the Bachelor of Business and Commerce Degree. During the semester, the lecturer created pre-lecture content to facilitate learning at home. Each week, the lecturer gave students an online quiz that tested their understanding of the course content prior to the lecture. There were 9 quizzes that together provided 10% (10 marks) towards each student's possible final grade. Each online quiz involved 10 multiple-choice questions that were available for five days through the online learning platform. Students needed to score at least 50% to earn credit in each quiz. They were given up to five attempts. In each attempt, they had 20 minutes to complete the quiz and they saw their score at the end of each attempt.

We implemented randomization at the tutorial level.[8] We randomly assigned each tutorial session to one of the three treatments. In the *Piece-Rate* treatment, the 10 marks were shared equally across each quiz and students received credit for each quiz that they completed. In the *All-Or-Nothing* treatment, students received 10 marks only if they completed all of the online quizzes, and zero marks otherwise. In the *Self-Select* treatment, we gave students a choice between the piece-rate and the all-or-nothing options. Students in the *Self-Select* treatment made their decisions during the first tutorial session without discussing their decision with their peers. To ensure consistency of treatment implementation, we provided each tutor a script to read during the tutorial. The script explained how the quizzes would be marked (see Online Appendix B).

[8] The university has an online class allocation system to allow students to specify times for their tutorials. Students are required to nominate six preferences for the tutorial in order of priority. The system then carries out a sorting procedure that allocate all students into classes according to their preferences wherever possible, taking into account each student's overall timetable. Students were not allowed to switch to alternative tutorials after their assignments.

We carried out two additional trials (Trial 2 in 2017 and Trial 3 in 2018) to see whether the effect of the self-select mechanism could be replicated in different courses and in different years (see Table 2 for the summary). We did not conduct the *All-Or-Nothing* treatment in these two replication trials because the main contribution of this study is the self-select mechanism and also because, as we report below, the all-or-nothing mechanism does not have any significant impact on the full completion rate. The total number of students and the courses involved in each trial are listed in Table 2. The design and implementation followed the same protocol as the first trial. We report the power for our tests separately for each individual trial datum and pooled data in Table C.1 in the online appendix.

Table 2. Summary of the design of each trial

	Trial 1 (2016)	Trial 2 (2017)	Trial 3 (2018)
# of quizzes	9	10	10
Courses	Introductory microeconomics	Introductory macroeconomics Introductory accounting Public sector economics	Introductory microeconomics
# of tutorials	16	13	14
Treatments (# of students)	Piece-Rate (103) All-Or-Nothing (86) Self-Select (154)	Piece-Rate (103) Self-Select (131)	Piece-Rate (114) Self-Select (198)
Total # of students	343	234	312

Note: In each trial, quizzes were worth 10% of the final mark. Students received full marks in each quiz as long as they got 50% of the answers correct.

4. Results

To test our hypothesis that the self-select mechanism can achieve a higher full completion rate, we focus on the treatment effect on the proportion of students who completed all the tasks. We first present the results from Trial 1 and then examine the robustness of the positive effect of the self-select mechanism using data from the subsequent two replication trials. Next, we extend our analysis to other outcomes that might be of interest. These include the average completion rates (the ratio of the total number of completed quizzes to all assigned quizzes) and the average of the percentage of correct answers across all assigned quizzes. Lastly, we examine students' performance in other assignments and courses to explore whether the mechanism had any spillover effects. Examining these other measures of performance would be particularly important if the *Self-Select* treatment somehow harms the performance measured by the other indicators. Since the randomization was done at the tutorial level, we provide statistical tests for the treatment differences by conducting regression analysis with standard errors clustered at the tutorial level. We also check the robustness of the findings by controlling for gender and the nature of the courses (e.g. compulsory or not).

4.1 Trial 1

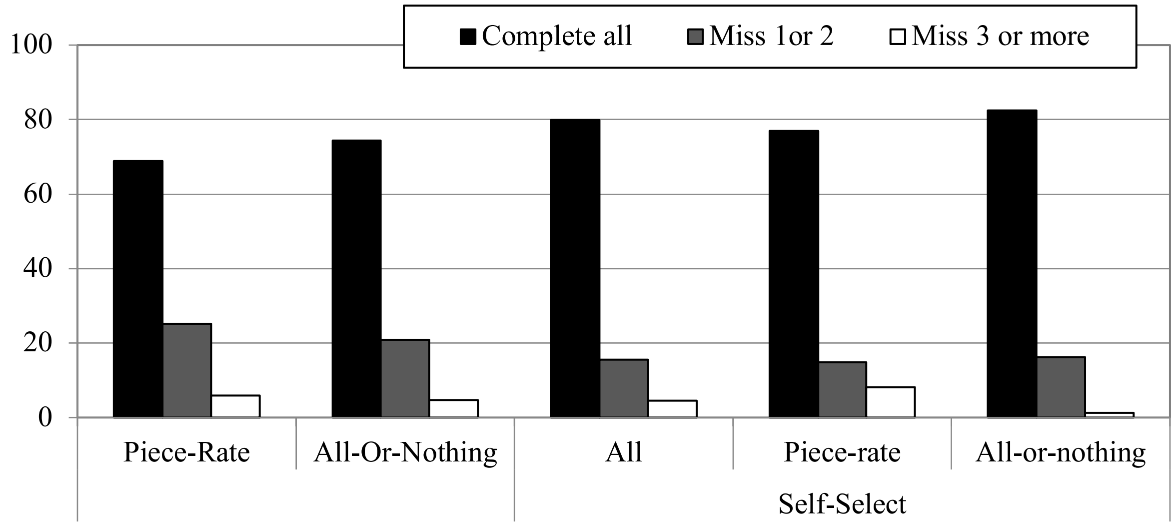
In the *Self-Select* treatment, 80 of 154 (52%) students selected the all-or-nothing marking mechanism. Figure 2 plots the distribution of the number of quizzes missed. Across all treatments, very few students missed three or more quizzes. Thus, we pool those students together in Figure 2. Compared with the baseline (*Piece-Rate*) treatment, the full completion rate is slightly higher in the *All-Or-Nothing* treatment and the difference is not significant (74.42% vs. 68.93%, p=0.299).[9] In contrast, the full completion rate in the *Self-Select* treatment is significantly higher than in the baseline (79.87% vs. 68.93%, p=0.048). The increase in the full completion rate is 16%. The effect size is comparable to the findings in previous studies which is around 3% and sometimes increases up to 28% depending on the incentive sizes and the targets to achieve (e.g. Angrist et al. 2014; Dellavigna and Pope, 2018).

Figure 2 shows that such an increase of the full completion rate can be mostly attributed to the 52% of students who self-selected the all-or-nothing mechanism. The full completion rate of this group of students is 82.5%, which is much higher than the completion rate of 68.93% in the *Piece-Rate* treatment (p=0.033). For those who selected the piece-rate mechanism, the full completion rate is higher than, but not significantly different from, that of the *Piece-Rate* treatment (77.03% vs. 68.93%, p=0.256). [10] These results are consistent with our theoretical framework. Those who self-selected the all-or-nothing mechanism should achieve a higher completion rate than those in the *Piece-Rate* treatment. On the other hand, the full completion rates between the self-selected piece-rate groups and those in the *Piece-Rate* treatment do not necessarily differ (see the online appendix).

[9] Throughout the paper, the reported p values comparing full completion rate across treatments are from a probit regression analysis where the dependent variable is whether the student completed all the tasks, and the independent variables are treatment dummies. We then use the Wald Chi-squared test to compare completion rate between treatments. Standard errors are clustered at the tutorial level.

[10] Some students registered for the course late and only got access to the online course materials after the first-week lecture and/or tutorial. In those cases, students were told that only the quizzes starting from Week 2 would affect their mark. Given this, our data analysis includes only quizzes starting from Week 2 and we count the quiz in Week 2 as the first task. As an alternative analysis, we dropped the latecomers and used Quiz #1 as the first quiz. The effect size of the *Self-Select* treatment is similar (the full completion rate is 78.57% in the *Self-Select* and 68.04% in baseline) although the difference is not significant (p=0.171). Similarly, the full completion rate of the students who self-selected the all-or-nothing mechanism is significantly higher than that in the *Piece-Rate* treatment (83.78% vs. 68.04%, p=0.015). Those who selected the piece-rate mechanism have a slightly higher full completion rate but not significantly different from the full completion rate of the *Piece-Rate treatment* (72.73% vs. 68.04%, p=0.731).

Figure 2. Distribution of quiz completion rate by treatment (Trial 1)



We conduct further probit regression analysis of the full completion rate to see whether the positive effect of the *Self-Select* treatment is robust when controlling for gender, and whether the course is compulsory for the student. Standard errors are clustered at the tutorial level. Table 3 reports regression results with marginal effects reported separately for each regression. As shown in Table 3, being assigned to the *Self-Select* rather than the *Piece-Rate* treatment has a marginal effect of increasing the probability of full completion by more than 10 percentage points in all specifications.

Table 3. Probit regression analysis of the full completion (Trial 1)

	Dependent variable: Full completion					
	(1)		(2)		(3)	
	Coef.	Marg. Eff.	Coef.	Marg. Eff.	Coef.	Marg. Eff.
All-Or-Nothing	0.162 (0.152)	0.051 (0.049)	0.220 (0.149)	0.068 (0.047)	0.263* (0.147)	0.081* (0.047)
Self-Select	0.343** (0.171)	0.108** (0.054)	0.331* (0.179)	0.102* (0.055)	0.352** (0.172)	0.108** (0.052)
Female			0.562*** (0.193)	0.174*** (0.060)	0.579*** (0.194)	0.178*** (0.060)
Compulsory					0.682*** (0.213)	0.210*** (0.062)
Constant	0.494*** (0.131)		0.235 (0.159)		–0.433** (0.207)	
Pseudo R^2	0.010		0.046		0.058	
# of obs.	343		343		343	

Notes:

a. Robust standard errors in parenthesis are clustered at the tutorial level.

b. * Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level.

Our results show that the self-select mechanism can significantly improve the full completion rate. Conversely, the all-or-nothing mechanism, if exogenously imposed (*All-Or-Nothing* treatment), does not improve the full completion rate. Observing more than half the students choosing the all-or-nothing mechanism suggests that the all-or-nothing mechanism is encouraging for a significant number of students who use this mechanism as a commitment device. As predicted in our model, those who choose the all-or-nothing mechanism drive up the full task completion rate. Meanwhile, no difference between the piece-rate and the all-or-nothing conditions suggests the existence of discouraged students, canceling out the positive effect on the encouraged ones. We next report data from the two additional replication trials to see whether the positive effect of the self-select mechanism can be replicated in different courses and in different years.

4.2 Robustness checks in replication trials

Overall, we observe similar effects of the self-select mechanism on the full completion rate. In Trial 2, 25.95% (34/131) of students chose all-or-nothing in the *Self-Select* treatment.[11] Compared with the baseline (i.e., the *Piece-Rate* treatment), the *Self-Select* treatment has a significantly higher full completion rate (49.62% vs. 20.39%, p=0.000). For those who selected the all-or-nothing mechanism, the full completion rate is significantly higher than those who were assigned to the piece-rate mechanism (64.71% vs. 20.39%, p=0.000). For those who selected the piece-rate mechanism, the full completion rate is also significantly higher than that of the *Piece-Rate* treatment (44.33% vs. 20.39%, p=0.004). In Trial 3, a relatively smaller proportion, 15.15% (30/198), of students chose the all-or-nothing mechanism in the *Self-Select* treatment. The *Self-Select* treatment still outperformed the baseline although the difference is not statistically significant probably due to the relatively low take-up rate of the all-or-nothing scheme (60.1% vs. 47.37%, p=0.251). Consistent with the findings in previous trials, those who chose the all-or-nothing mechanism achieve the highest full completion rate: 70% who self-selected the all-or-nothing mechanism compared to 47.37% who were assigned to the *Piece-Rate* treatment (p=0.048). The full completion rate for those who self-selected the piece rate is higher than the baseline but the difference is not significant (58.33% vs. 47.37%, p=0.329).

Lastly, we conduct regression analysis using the pooled data from the *Piece-Rate* and the *Self-Select* treatments in all three trials, adding controls such as gender, dummy for compulsory courses, and course dummies (see Table 4). Pooled regression results show that the *Self-Select* treatment increases the probability of full completion by at least 17 percentage points.

---

[11] We excluded 19 students who were in Trial 1 and who were also in classes in Trial 2 from the study. We checked whether the sessions that included those experienced students had different take-up rates in the all-or-nothing marking scheme (25 out of 93 students) than those without any experienced students (9 out of 38 students). We did not find any evidence suggesting this to be the case (26.88% vs. 23.68%, Z-test, p =0.705). No students in Trial 3 had participated in any of the previous two trials because they were all first-year students in their first semester.

Table 4. Probit regression analysis of the full completion (all three trials pooled)

	Dependent variable: Full completion							
	(1)		(2)		(3)		(4)	
	Coef.	Marg. Eff.	Coef.	Marg. Eff.	Coef.	Marg. Eff.	Coef.	Marg. Eff.
Self-Select	0.457**	0.180**	0.439**	0.172**	0.437**	0.172**	0.437***	0.172***
	(0.181)	(0.072)	(0.178)	(0.071)	(0.177)	(0.070)	(0.133)	(0.053)
Female			0.468***	0.184***	0.466***	0.183***	0.441***	0.173***
			(0.093)	(0.037)	(0.093)	(0.037)	(0.101)	(0.040)
Compulsory					0.384*	0.151*	0. 378*	0.148*
					(0.206)	(0.081)	(0.211)	(0.083)
Course[a]							Yes	Yes
Constant	-0.110		-0.311**		-0.671***		-0.129	
	(0.157)		(0.154)		(0.234)		(0.218)	
Pseudo R^2	0.023		0.047		0.051		0.118	
# of obs.	803		803		803		803	

Notes:

a. Controls for the five different courses in different years across these trials. As each trial involves different courses in different years, it means this regression also controls for different trials.
b. The data include all the data of the *Piece-Rate* and the *Self-Select* treatments in the three trials.
c. Robust standard errors in parenthesis are clustered at the tutorial level.
d. * Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level.

One interesting observation is that the percentage of students who chose the all-or-nothing mechanism in the *Self-Select* treatment is much higher in the first trial in 2016 than both of the replication trials. One possibility for this difference is that the three cohorts of students may somehow differ. This is unlikely because students are admitted using the identical entry requirements. The demographics (gender and international student status) of the students in each trial are also similar (see Table C.2 in Online Appendix).

We speculate the change in the take-up rate may be associated with the culture of communication between junior and senior students. For example, even before a new semester begins, the university's student associations use social media platforms to publish the contact details of student representatives for all discipline majors. Incoming students can seek course or program-related information from their seniors. Students in the replication trials might learn about the cost distribution of tasks from seniors in the first cohort.[12] Note that the first time weekly online quizzes were introduced to the department was in Trial 1. Prior to this, quizzes were never conducted to test the learning objectives on a weekly basis. Trial 1 was also the first time students were offered the option to choose between the piece-rate and all-or-nothing marking schemes. Thus, it is possible that students in the subsequent trials obtain relevant information when communicating with students from previous years.

[12] This speculation is consistent with a survey conducted at the department among first year students. About two thirds of students indicate that some of their peers communicate with students of previous years about the courses. More than 40% of students answered that their friends or classmates seek advice on how to approach assignments or exams from the previous cohorts.

As discussed in the theoretical analysis, the choice of the all-or-nothing is determined by cost distribution and present bias β. While the present bias β is more of a personal trait and unlikely to be affected by communication with other students, it is likely that students update their knowledge about the probability of completing each task after acquiring further information (e.g. learning how busy the semester gets towards the end). If, for example, the students in the subsequent semesters estimate the probability of the high cost to be higher (i.e., p is smaller) after talking with the previous cohorts, we would expect a relatively low take-up rate in Trials 2 and 3 than in Trial 1. In other words, the lower sign-up rate in the replication trials could be due to information spillover from previous cohorts.

4.3. Other measures of performance

So far we show that the *Self-Select* treatment can increase the full completion rate. However, one concern is that the self-select mechanism may have negative effects on other outcomes. In particular, we consider four additional measures. First, we check the average completion rate of the quizzes. Those who self-select the all-or-nothing mechanism may give up completely when they miss one task. As a result, the mechanism may achieve a higher full completion rate at the expense of a lower average completion rate.[13] Second, we compare the percentages of correct answers averaged across all the quizzes. This is because students receive full marks as long as they get 50% of answers correct. Thus, those who self-selected the all-or-nothing mechanism may focus too much on finishing the quizzes (getting 50% of answers correct) rather than on getting as many correct answers as possible. Third, we examine the overall mark of all assessments excluding the quizzes in the treated courses. The self-select mechanism intervention may lead students to allocate more than optimal time on the incentivized tasks by sacrificing the time that they would otherwise devote to other assignments of the same course. Fourth, we check the performance in other courses since it is possible that students reduce their time spent on other courses during the semester.

We report the descriptive data of the above four measures by treatments in Table 5. To provide a statistical test for the effect of the *Self-Select* treatment on each measure, we conduct regression analysis using the pooled data of the *Piece-Rate* and the *Self-Select* treatments and the same regression specification as reported in Regression (4) of Table 4. Figure 3 plots the regression coefficients of the *Self-Select* treatment with 95% confidence intervals (Jann 2014). The details of the regressions are provided in Table C.3 in the online appendix. As shown in Table 5 and Figure 3, the *Self-Select* treatment performs significantly better in all four measures compared to the *Piece-Rate* treatment.

[13] In the online appendix we show that the average completion rate is expected to be higher under the self-select mechanism compared with under the piece-rate mechanism in a simple two-task model, which easily follows from Proposition 1.

Table 5. Other measures of performance by treatments

	Average completion rate			Percentage of correct answers			Overall mark excluding the quizzes			Average mark in other courses		
	Trial 1 (2016)	Trial 2 (2017)	Trial 3 (2018)	Trial 1 (2016)	Trial 2 (2017)	Trial 3 (2018)	Trial 1 (2016)	Trial 2 (2017)	Trial 3 (2018)	Trial 1 (2016)	Trial 2 (2017)	Trial 3 (2018)
Baseline (Piece-Rate)	92.60 (1.54)	74.85 (2.48)	88.25 (1.47)	72.45 (1.54)	65.65 (2.38)	78.34 (1.85)	53.41 (1.61)	60.45 (1.91)	55.00 (1.59)	66.96 (1.30)	60.66 (1.51)	63.34 (1.17)
All-Or-Nothing	94.04 (1.61)	-	-	73.13 (1.64)	-	-	56.86 (1.65)	-	-	67.09 (1.26)	-	-
Self-Select	95.13 (0.98)	89.47 (1.34)	92.73 (0.86)	74.14 (1.19)	78.47 (1.47)	84.98 (1.08)	58.64 (1.34)	66.96 (1.17)	58.77 (1.08)	69.65 (1.05)	65.63 (1.12)	67.05 (0.71)
# of obs.	343	234	312	343	234	312	340	234	311	343	228	308

Notes:

a. The average completion rate is the ratio of the total number of completed quizzes to all assigned quizzes.

b. The percentage of correct answers is calculated across all assigned quizzes on a 0–100 scale.

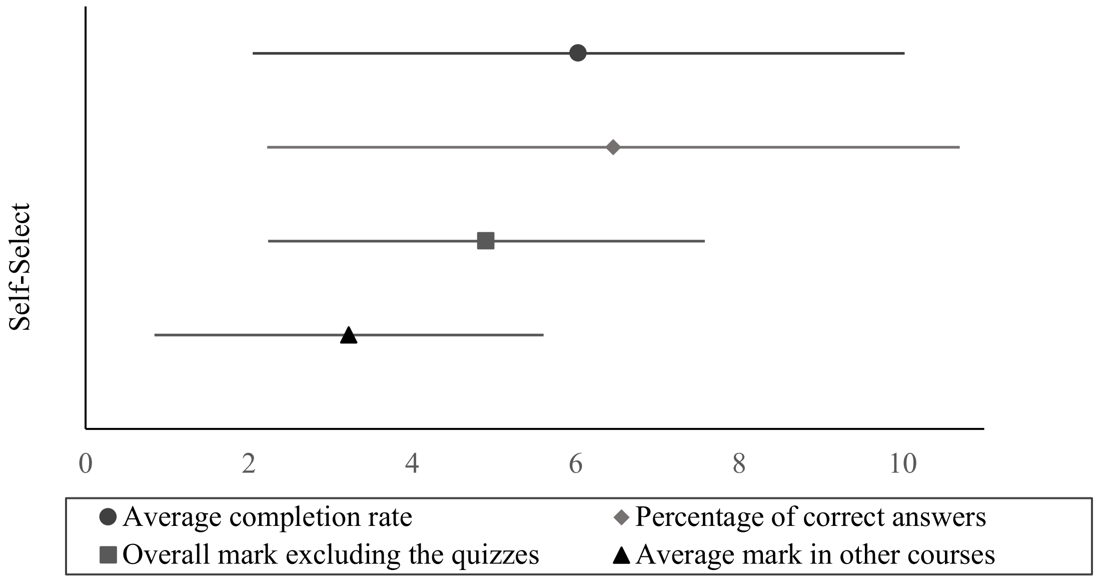
c. The overall mark excluding the quizzes is the average marks of all assessments except the quizzes in the treated subjects and is converted to a 0–100 scale.

d. The average mark in other courses includes all the untreated courses the students took during the intervention semester.

e. In Trial 1 and 3, there are 3 and 1 missing values in the "Overall mark excluding the quizzes" respectively. These students withdrew from the course after the end of the semester. In Trial 2 and 3, there are 6 and 4 missing values in the "Average mark in other courses" respectively. These students either withdrew from the course or did not enrol in any courses other than the treated course during the semester.

f. Standard errors are in parentheses.

Figure 3: Regression coefficients with 95% confidence intervals



Notes:

a. The markers show the percentage point increase in the *Self-Select* relative to the *Piece-Rate* treatment (expressed as percentage point differences).

b. Coefficients that are significant at the 5 percent level are those where the confidence interval does not cross the vertical line at zero.

It is interesting to observe that the self-select mechanism has a potential positive spillover effect on students' performance in other courses. Such a positive effect is not unique to our study. Kuang et al. (2019) and Mochon et al. (2017) find similar spillover effects in online knowledge exchange platforms and in the health context respectively. One explanation may be that all the other courses also used the same online teaching platform (in this case, Moodle). Once students sign into the platform, they are probably more likely to study the materials and complete any online tasks in other courses. By increasing the overall frequency of working on online tasks, the mechanism can help students to overcome the initial cost of commencing study. To explore this explanation, we conduct a regression analysis of the marks in other courses by controlling for the average completion rate of the quizzes. This analysis indicates how often students visit the teaching platform. As shown in Table 6, the results are consistent with this explanation. The *Self-Select* treatment is positively significant in Regression (1) but has no significant impact on the marks in other courses once we control for the average completion rate as shown in Regression (2). The significant positive coefficient of the average completion rate suggests that those who complete the online quizzes more frequently also tend to perform better in other courses. These results are robust when we add other control variables.

Table 6: Regression analysis of average mark in other courses (all three trials pooled)

	Dependent variable: Average mark in other courses				
	(1)	(2)	(3)	(4)	(5)
Self-Select	3.835**	1.426	1.331	1.346	1.252
	(1.446)	(1.036)	(1.026)	(1.043)	(0.855)
Average completion rate		0.358***	0.346***	0.347***	0.349***
		(0.028)	(0.030)	(0.029)	(0.024)
Female			2.666***	2.683***	2.690***
			(0.894)	(0.886)	(0.871)
Compulsory				-2.334	-2.551
				(2.138)	(2.277)
Course[a]					Yes
Constant	63.671***	32.907***	32.837***	34.894***	36.020***
	(1.209)	(2.430)	(2.531)	(3.048)	(2.839)
R-squared	0.022	0.228	0.238	0.240	0.270
# of obs.	793	793	793	793	793

Notes:

a. Controls for the five different courses in different years across these trials. As each trial involves different courses in different years, it means this regression also controls for different trials.
b. The data include all the data of the *Piece-Rate* and the *Self-Select* treatments in the three trials.
c. There are ten missing values in the "Average mark in other courses" because ten students did not take any other courses during the treated semester.
d. Robust standard errors in parentheses are clustered at the tutorial level.
e. * Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level.

Inspired by the positive spillover effect of the self-select mechanism on students' performance during that intervention semester, we explore whether it may have any impact for students' long-run academic performance. To do so, we examine students' academic performance after the intervention. For each student, we separately calculate the average grades Grade Point Average (GPA) students received in the courses they took in the first, second and third semesters after the intervention semester. The data is reported in Table 7. We find the positive spillover effect of the *Self-Select* treatment sustains in the first semester after the intervention (2.69 vs. 2.91, p=0.001)[14]. However, the effect diminishes in the subsequent semesters and is not statistically significant (2.86 vs. 2.93, p=0.323; 2.97 vs. 3.03, p=0.360). These results suggest that the *Self-Select* treatment implemented in one semester alone has a limited long-term effect.

[14] The reported p values are from an Ordinary Least Square (OLS) regression analysis where the dependent variable is GPA and the independent variables are the two treatment dummies. We then use an F-test to compare GPA between treatments. Standard errors are clustered at the tutorial level.

Table 7 Participants' GPA (all three trials pooled)

	1st semester after intervention		2nd semester after intervention		3rd semester after intervention	
	GPA	# of obs.	GPA	# of obs.	GPA	# of obs.
Baseline	2.69	281	2.86	243	2.97	201
	(0.04)		(0.05)		(0.05)	
Self-Select	2.91	438	2.93	380	3.03	306
	(0.03)		(0.04)		(0.04)	
# of obs.		719		623		507

Notes:

a. The (GPA) is the average result of all grades students achieved in a semester after the intervention. It is calculated on a four-point grade scale where 4.0 is the highest and 0.00 is the lowest achievement. We do not have the GPA record for some students after the treated semester as they either transferred to another campus/university, enrolled in a student exchange program, took a gap semester/year, withdrew, or graduated (especially those in Trial 2). The number of missing values for Trial 1, 2 and 3, respectively, are as follows:
   - 1st semester after treatment: 23, 43 and 18;
   - 2nd semester after treatment: 44, 74 and 62; and
   - 3rd semester after treatment: 94, 104 and 98.

b. The data are only available up to the 3rd semester after the intervention for Trial 3, after which an Academic Safety Net was put in place in response to Covid-19, which deemed that students who fail can choose to withdraw, and their fail grades will not be recorded on academic transcripts nor count towards the final GPA.

c. Standard errors are in parentheses.

Taken together, our data suggests that the self-select mechanism performs better than the baseline – not only in terms of the full completion rate, but also by other measures of performance during the intervention semester. One may raise the concern that the higher marks in other courses under the self-select mechanism suggest students assigned to the *Self-Select* treatment are somehow better than those assigned to the baseline. We think this is unlikely for several reasons. First, our randomization method in each trial should minimize such a possibility; second, we examine whether there is any difference across treatments in students' academic performance before the intervention. We calculate students' GPA prior to the intervention semester for Trial 2. Students in Trials 1 and 3 are mostly first year students in their first semester. Thus, there is no GPA record before the intervention semester. For these two trials, we examine the marks of the first test that took place after the completion of the first online quiz in Trial 1 and the second online quiz in Trial 3.[15] Presumably, if the positive spillover effect is due to any selection problem, we should also observe such a difference in the first test or GPA before the intervention. We report the data in Table C.4 in the

[15] Trial 2 included several different courses with the first test conducted well after half of the quizzes had been given. Thus, the treatment may affect the first quiz score in this trial.

online appendix. We find no difference between the *Self-Select* treatment and the baseline treatment in the early semester or before the intervention, which suggests no evidence for the selection problem.

5. Discussion

We show that the *Self-Select* treatment achieves a higher full completion rate compared to the baseline *Piece-Rate* treatment, and this positive effect is mainly driven by those who chose the all-or-nothing mechanism. It would be interesting to investigate the individual differences in the take-up decisions. One may hypothesize the choice to be correlated with present bias. However, as we explained in Section 2.3, the relationship between the present bias and the choice also depends on other parameters, such as the cost distribution, which cannot be measured easily in the current design.

Previous studies suggest gender differences in individual characteristics such as self-regulation, which may contribute to school achievement (Duckworth and Seligman 2006, Weis et al. 2013). Moreover, female students often respond to incentives in education differently than male students (Angrist and Lavy 2009, Campos-Mercade and Wengström 2020). We do not observe any gender difference in the choices of the all-or-nothing mechanism (Trial 1: 53.01% female vs. 50.70% male; Trial 2: 25.86% female vs. 26.03% male; Trial 3: 15.63% female vs. 14.71% male).

Although we do not have any theoretical predictions, we explore whether the choice correlates with students' academic ability. We first compare international and domestic students because the schoolwide average of GPA for the latter group is usually higher.[16] Using the pooled data of the three trials, we find no significant difference in the choice of the all-or-nothing mechanism between international students and domestic students (25.53% vs. 31.18%; Z-test, p=0.286). Next, we compare students' academic data before the intervention. As we explained in Section 4.3, we have students' GPA prior to the intervention semester for Trial 2. Students in Trial 1 and Trial 3 were mostly first year students in their first semester. Thus, we do not have their GPA record before the intervention. For these two trials, we examine the marks of the first assessment that took place after the completion of the first online quiz in Trial 1 and the second online quiz in Trial 3. We assume the *Self-Select* treatment would not have any significant effect on performance at this early stage. We find that the prior GPA in Trial 2 and early assessment marks in Trial 1 and 3 do not differ between students who select the all-or-nothing mechanism and the piece-rate mechanism (Trial 2: 2.92 vs. 2.83, t-test, p=0.630; Trial 1: 64.11 vs. 66.76, t-test, p=0.324; Trial 3: 63.67 vs. 67.64, t-test,

---

[16] For example, we compare the GPA of 56 international students and 139 domestic students from the Bachelor of Business and Commerce who joined the university during the same year as students in Trials 1 and 2 and who were not in the treated courses. We find the domestic students have higher GPA than the international students (2.76 vs. 2.45, t-test, p=0.008). If we extend the sample to include non-business-major students who joined during the same years, domestic students continue to perform significantly better than the international students (average GPA from 454 domestic students=2.98 vs. average GPA from 127 international students=2.65, t-test, p=0.000).

p=0.229). This result is interesting as it suggests that our self-select mechanism provides equal opportunities to those who do not perform well and those who are already high performers.

When evaluating the mechanisms, we focus on their impact on the full completion rate. In practice, other considerations may affect the desirability of the two mechanisms. For instance, in theory the all-or-nothing mechanism leads to both higher full completion and zero completion rate and thereby can produce more variance and inequality in outcomes. In this regard, the self-select mechanism not only helps to achieve a higher full completion rate but it also mitigates inequality. Future research exploring important parameters that could make one mechanism preferable over the other, both from the perspective of the principal and the agent, would be fruitful. For example, we assume risk neutrality in the theoretical analysis, however, the preferences between the all-or-nothing and the piece-rate mechanisms may well be affected by risk attitudes.

Our theoretical framework is based on the sorting effect. However, self-selecting may also work through autonomy (Bowles and Polanía-Reyes 2012, Deci et al. 1999). Self-determination theory identifies autonomy as one of the three important psychological needs that must be satisfied for self-motivation (Ryan and Deci 2006). For example, Zuckerman et al. (1978) show that participants enjoyed tasks more when they could choose what tasks to work on and how much time to allocate. Duckworth (2016) argues that people are more likely to persevere when they feel in control. Accordingly, the option provided in the self-select mechanism allows autonomy and thereby helps to motivate participants to complete tasks – irrespective of what incentive mechanism they choose. We report above, those who self-select the all-or-nothing mechanism achieve a significantly higher full completion rate than those in the baseline. When examining the difference between those who self-selected the piece-rate mechanism and those in the baseline who were assigned the piece-rate, we find the full completion rates are statistically higher in Trial 2 (2017) but not in Trial 1 (2016) and Trial 3 (2018). These results seem to suggest some weak evidence of the autonomy effect. However, there are two points worth noting.

First, the result of the higher completion rate of those who self-select the piece-rate mechanism is not inconsistent with our theoretical framework. According to our theory, the comparison of the full completion rate between the self-selected piece-rate groups and the assigned piece-rate groups can be in either direction (see the online appendix for the derivations of this result). It depends on the relative share of DMs who complete regardless of cost (T_1 type), and DMs who never complete (T_2 type) and the probability of low cost (p) (see Appendix for the details). Second, the autonomy effect alone does not offer any insights about why/how people choose the all-or-nothing mechanism and therefore does not provide a clear prediction on the performance difference between those who choose the all-or-nothing mechanism and those in the baseline condition. Nevertheless, the potential autonomy effect would be an additional reason for adopting the self-select mechanism instead of exogenously imposing either the piece-rate mechanism or the all-or-nothing mechanism.

Another potential mechanism is that some DMs choose the all-or-nothing mechanism to signal to third parties. For example, it is possible that some students chose the all-or-nothing mechanism simply to signal to the professors their commitment or their intention to work hard, or their willingness to take the challenge. However, it is not clear what the signaling channel would predict in terms of full completion rate without making additional assumptions about the relation between willingness to send a signal and the student type.

6. Conclusion

Perseverance is critical in achieving success. Daily temptation or extra duties arriving stochastically often distract people from completing tasks. Hence, achieving long-term goals that require persistent effort is difficult for many. External incentive mechanisms can be designed to motivate individuals to be more persistent in their effort. In this paper, we compare the effectiveness of two reward mechanisms on promoting persistent effort over time. One is the conventional piece-rate mechanism that provides rewards for the completion of each task and, thus, does not target perseverance directly. Another is an all-or-nothing mechanism that explicitly requires continual effort by rewarding only the completion of all tasks. Previously studied incentive mechanisms often take the form of one of the two. The all-or-nothing mechanism is useful for dealing with self-control problems and achieving higher targets, but it comes with the risk that people might be discouraged from trying at all, or might lead people to withdraw completely after a single failure. Indeed, we find the all-or-nothing mechanism, when exogenously imposed, is not effective.

A self-select mechanism where individuals can freely choose between the two types of rewards shows promise in promoting full completion. Theoretically, only people who can maximize the all-or-nothing mechanism would choose it, and those who would have been discouraged by the all-or-nothing mechanism can select the piece-rate mechanism instead. Indeed, data from a series of field experiments show that a significant number of individuals are willing to choose the risky all-or-nothing mechanism. Those who self-select the all-or-nothing mechanism achieve a significantly higher full completion rate than people assigned to the piece-rate mechanism, which also leads to the overall improved outcome under the *Self-Select* treatment.

The findings from our experiment offer useful policy insights in domains, such as organizational management, and education and health, where persistent effort over a long period of time is important. Persistent effort may be especially hard when tasks are boring and involve repeated tedious work. It is important to consider how to design external rewards to help achieve the long-term goal. Research on compensation in the workplace has focused on the pros and cons of various incentives such as continuous, as opposed to discrete, schemes. Our study shows that allowing agents to self-select the reward mechanism can improve performance. The self-select mechanism can also have practical advantages in that introducing a new mechanism as an option could presumably avoid potential resistance when individuals are used to the existing piece-rate reward mechanism.

In recent decades, much research has been conducted on helping children and young people do well academically. One key challenge is determining how to encourage students to persevere. Different monetary reward mechanisms have been designed and tested, but with mixed results (Burger et al. 2011, Fryer 2011, Leuven et al. 2010, Levitt et al. 2016). These mixed results indicate that the problem may not be the lack of monetary incentives. Students, especially those who grow up in adverse circumstances, are already likely to have sufficient material incentives to get a good education and change their life prospects. Yet, they may not know how to make decisions to achieve success. Our study shows that simple changes in incentive mechanisms may help. A mechanism, such as the all-or-nothing marking mechanism, can be quite helpful when students are provided with an *option*.

We take a step towards understanding the working of the all-or-nothing mechanism and how the option of such an incentive mechanism affects behavior. Many interesting and important questions remain for further investigation. For example, our current simple design does not allow us to know how those who chose the all-or-nothing mechanism would have performed with piece-rate and vice versa. Do people choose the commitment device optimally? In this study, we observe participants who chose the all-or-nothing mechanism and eventually failed to receive any reward. How would such a failure affect future decisions? Would they no longer choose the commitment, or would they learn from the negative experience and try harder next time? We find the *Self-Select* treatment wins mainly because of those who chose the all-or-nothing mechanism. Would it be helpful to design mechanisms that *nudge* more to "self-select" the all-or-nothing mechanism? We are conducting follow-up studies to address these interesting questions.

References:

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

Aggarwal S, Dizon-Ross R, Zucker AD (2020) Incentivizing behavioral change: The role of time preferences. *NBER Working Paper*.

Alan S, Boneva T, Ertac S (2019) Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit. *Quarterly Journal of Economics* 134(3):1121–1162.

Almlund M, Duckworth AL, Heckman J, Kautz T (2011) Personality Psychology and Economics. *Handbook of the Economics of Education*. 1–181.

Amador M, Werning I, Angeletos GM (2006) Commitment vs. flexibility. *Econometrica*.

Angrist J, Lavy V (2009) The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American Economic Review*.

Angrist J, Oreopoulos P, Williams T (2014) When Opportunity Knocks, Who Answers?: New Evidence on College Achievement Awards. *Journal of Human Resources* 49(3):572–610.

Ariely D, Wertenbroch K (2002) Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science* 13(3):219–224.

Ashraf N, Karlan D, Yin W (2006) Deposit Collectors. *Advances in Economic Analysis & Policy*.

Augenblick N, Niederle M, Sprenger C (2015) Working over Time: Dynamic Inconsistency in Real Effort Tasks. *The Quarterly Journal of Economics* 130(3):1067–1115.

Basu K (2011) Hyperbolic discounting and the sustainability of rotational savings arrangements. *American Economic Journal: Microeconomics* 3(4):143–171.

Beattie G, Laliberté JWP, Oreopoulos P (2018) Thrivers and divers: Using non-academic measures to predict college success and failure. *Economics of Education Review* 62:170–182.

Benhabib J, Bisin A (2005) Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption-saving decisions. *Games and Economic Behavior* 52(2):460–492.

Bettinger E, Ludvigsen S, Rege M, Solli IF, Yeager D (2018) Increasing perseverance in math: Evidence from a field experiment in Norway. *Journal of Economic Behavior and Organization* 146:1–15.

Bisin A, Hyndman K (2020) Present-bias, procrastination and deadlines in a field experiment. *Games and Economic Behavior*.

Bowles S, Polanía-Reyes S (2012) Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature* 50(2):368–425.

Bryan G, Karlan D, Nelson S (2010) Commitment devices. *Annual Review of Economics* 2(1):671–698.

Buechel B, Mechtenberg L, Petersen J (2018) If I can do it, so can you! Peer effects on perseverance. *Journal of Economic Behavior and Organization* 155:301–314.

Burger N, Charness G, Lynham J (2011) Field and online experiments on self-control. *Journal of Economic Behavior and Organization* 77(3):393–404.

Campos-Mercade P, Wengström E (2020) *Threshold incentives and academic performance*

Carrera M, Royer H, Stehr MF, Sydnor JR, Taubinsky D (2020) Who chooses commitment? Evidence and welfare implications. *NBER Working Paper*.

Charness G, Gneezy U (2006) Incentives to Exercise. *Econometrica* 77(3):909–931.

Chung DJ, Steenburgh T, Sudhir K (2014) Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans. *Marketing Science*.

Claro S, Paunesku D, Dweck CS (2016) Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences of the United States of America* 113(31):8664–8668.

Darmon YR (1997) Selecting appropriate sales quota plan structures and quota-setting procedures. *Journal of Personal Selling and Sales Management*.

Deci EL, Ryan RM, Koestner R (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*.

DellaVigna S, Malmendier U (2006) Paying not to go to the gym. *American Economic Review* 96(3):694–719.

DellaVigna S, Pope D (2018) What motivates effort? Evidence and expert forecasts. *Review of Economic Studies*.
Duckworth AL (2016) *The Power of Passion and Perseverance* (Scribner/Simon & Schuster, New York, NY, US).
Duckworth AL, Seligman MEPP (2006) Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores. *Journal of Educational Psychology* 98(1):198–208.
Fryer RG (2011) Financial incentives and student achievement: Evidence from randomized trials. *Quarterly Journal of Economics* 126(4):1755–1798.
Fudenberg D, Levine DK (2006) A dual-self model of impulse control. *American Economic Review*.
Giné X, Karlan D, Zinman J (2010) Put your money where your butt is: A commitment contract for smoking cessation. *American Economic Journal: Applied Economics* 2(4):213–235.
Gul F, Pesendorfer W (2001) Temptation and self-control. *Econometrica* 69(6):1403–1435.
Heckman JJ, Kautz T (2012) Hard evidence on soft skills. *Labour Economics* 19:451–464.
Heckman JJ, Stixrud J, Urzua S (2006) The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24(3):411–482.
Himmler O, Jäckle R, Weinschenk P (2019) Soft commitments, reminders, and academic performance. *American Economic Journal: Applied Economics* 11(2):114–142.
Jain S (2012) Self-control and incentives: An analysis of multiperiod quota plans. *Marketing Science*.
Jann B (2014) Plotting regression coefficients and other estimates. *Stata Journal* 14(4).
John A (2020) When commitment fails: Evidence from a field experiment. *Management Science*.
Joseph K, Kalwani MU (1998) The Role of Bonus Pay in Salesforce Compensation Plans. *Industrial Marketing Management*.
Kaur S, Kremer M, Mullainathan S (2015) Self-control at work. *Journal of Political Economy* 123(6):1227–1277.
Kishore S, Rao RS, Narasimhan O, John G (2013) Bonuses versus commissions: A field study. *Journal of Marketing Research*.
Kuang L, Huang N, Hong Y, Yan Z (2019) Spillover Effects of Financial Incentives on Non-Incentivized User Engagement: Evidence from an Online Knowledge Exchange Platform. *Journal of Management Information Systems* 36(1):289–320.
Kullgren JT, Troxel AB, Loewenstein G, Norton LA, Gatto D, Tao Y, Zhu J, et al. (2016) A Randomized Controlled Trial of Employer Matching of Employees' Monetary Contributions to Deposit Contracts to Promote Weight Loss. *American Journal of Health Promotion*.
Laibson D (1997) Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics* 112(2):443–478.
Laibson D (2015) Why don't present biased agents make commitments? *American Economic Review*. 267–272.
Lazear EP (2018) Compensation and Incentives in the Workplace. *Journal of Economic Perspectives*.
Leuven E, Oosterbeek H, van der Klaauw B (2010) The effect of financial rewards on students' achievement: Evidence from a randomized experiment. *Journal of the European Economic Association* 8(6):1243–1265.
Levitt SD, List JA, Neckermann S, Sadoff S (2016) The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy* 8(4):183–219.
Milkman KL, Minson JA, Volpp KGMM (2014) Holding the Hunger Hames Hostage at the Gym: An Evaluation of Temptation Bundling. *Management Science* 60(2):283–299.
Mochon D, Schwartz J, Maroba J, Patel D, Ariely D (2017) Gain without pain: The extended effects of a behavioral health intervention. *Management Science* 63(1):58–72.
O'Donoghue T, Rabin M (1999a) Doing it now or later. *American Economic Review* 89(1):103–124.
O'Donoghue T, Rabin M (1999b) Incentives for procrastinators. *Quarterly Journal of Economics* 114(3):769–816.
O'Donoghue T, Rabin M (2001) Choice and Procrastination. *Quarterly Journal of Economics* 116(1):121–160.
Phelps ES, Pollak RA (1968) On Second-Best National Saving and Game-Equilibrium Growth. *Review of Economic Studies* 35(2):185–199.

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*.

Ryan RM, Deci EL (2006) Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality* 74(6):1557–1585.

Schöttner A (2017) Optimal sales force compensation in dynamic settings: Commissions vs. bonuses. *Management Science*.

Shefrin HM, Thaler RH (1988) The Behavioral Life-Cycle Hypothesis. *Economic Inquiry* 26(4):609–643.

Strack P, Taubinsky D (2021) Dynamic Preference "Reversals" and Time Inconsistency. *NBER working paper*.

Tangney JP, Baumeister RF, Boone AL (2004) High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success. *Journal of Personality* 72(2):271–324.

Toussaert S (2018) Eliciting Temptation and Self-Control Through Menu Choices: A Lab Experiment. *Econometrica*.

Weis M, Heikamp T, Trommsdorff G (2013) Gender differences in school achievement: The role of self-regulation. *Frontiers in Psychology*.

Zuckerman M, Porac J, Lathin D, Deci EL (1978) On the Importance of Self-Determination for Intrinsically-Motivated Behavior. *Personality and Social Psychology Bulletin*.

Appendix

In this appendix, we provide proofs of Lemma 1 and Proposition 1. Before providing the proofs, some preliminary discussion is in order.

We model a decision-maker (DM) with a present bias by using quasi-hyperbolic discounting. A DM solves an intertemporal choice problem over three periods. Period zero, $t = 0$, is to choose (or to be assigned to) a mechanism. Period 1, $t = 1$, is to make task completion decisions, and period 2, $t = 2$, is to receive the reward. For simplicity, we assume that two tasks take place in period one.

The task cost is low, c_L with probability $p > 0.5$, and high, c_H otherwise. The reward is fixed and higher than the high cost, $c_H < r$, for each task completion. This is because we focus on analyzing task completion for a DM who is interested in completing the tasks but is unable to do so either because of self-control problems or because of exogenous shocks. We assume $r = 1$ without loss of generality. Under the all-or-nothing (A-or-N) mechanism, the reward is 2 if the agent completes both tasks and *zero* otherwise. The DM has a preference for immediate gratification, $0 < \beta < 1$, and they are sophisticated about their self-control problem, meaning that they have a correct belief about the value of β. For simplicity we assume there is no exponential discounting.

Proof of Lemma 1. Under piece-rate it is easy to see that there are three types of DMs:

T1 type:	$c_H \leq \beta,$	always completes;
T2 type:	$\beta < c_L,$	never completes; and
T3 type:	$c_L \leq \beta < c_H,$	completes if cost is low.

Our goal is to analzye whether these three types act differently under the all-or-nothing mechanism. The problem under all-or-nothing is more complex. We can solve the problem under three different conditions covering all cases: (1) $\frac{1+p}{p} < \frac{c_H}{c_L}$; (2) $2 < \frac{c_H}{c_L} \leq \frac{1+p}{p}$; and (3) $\frac{c_H}{c_L} \leq 2$. We provide the solution for the first case for conciseness. Solutions for the others are provided in the online appendix.

Note that $\frac{c_L}{2} < c_L < \frac{c_H}{2} < c_H$ which follows from the assumption that $p > 0.5$ under the first case. Below, we analyze four cases depending on where β stands with respect to c_H, $\frac{c_H}{2}$, c_L, and $\frac{c_L}{2}$.

A necessary condition for the DM to complete the second task is the completion of the first task. We will not restate this condition for each case to keep the writing concise.

1. If $c_H \leq \beta$, the DM completes both tasks irrespective of the cost. This is the type who *always* completes the tasks regardless of the mechanism, $T1$ type as above.
2. If $\frac{c_H}{2} \leq \beta < c_H$, the DM completes the second task regardless of the cost. We analyze the DM's decisions for the first task in two subcases as follows:

(a) If the DM faces a high cost, c_H, they complete the task if the expected cost is less than the benefit, e.g. $\frac{c_H+pc_L+(1-p)c_H}{2} \leq \beta < c_H$. The possibility of a low cost in the second task decreases the expected cost even when the first task cost is high, consequently making leading to the DM who would complete the first task only when the cost is low under the piece-rate to complete the first task with a high cost. We call them the $T3a$ *encouragement* type.

(b) If the DM faces a low cost, c_L, they complete the first task if $\frac{c_L+pc_L+(1-p)c_H}{2} < \beta$. We call this the $T3b$ *conditional encouragement* type. The cutoffs for the parameters for this type is $\frac{c_H}{2} \leq \beta < \frac{c_H+pc_L+(1-p)c_H}{2}$ since $\frac{c_L+pc_L+(1-p)c_H}{2} < \frac{c_H}{2}$. Such DMs complete the second task regardless of the cost, however the task completion is conditional on a low cost for the first task.

3. If $c_L \leq \beta < \frac{c_H}{2}$, the DM completes the second task only if the cost is low. Given this, we analyze the first task completion in two subcases as follows:

(a) If the DM faces a high cost, they exert an effort only if $\frac{c_H+pc_L}{2p} < \beta$. However, no DM satisfies this condition since $\frac{c_H}{2} < \frac{c_H+pc_L}{2p}$.

(b) If the DM faces a low cost, they exert an effort only if $\frac{c_L+pc_L}{2p} \leq \beta < \frac{c_H}{2}$. That is to say, they expect to complete the second task if the cost is low. We call them the $T3c$ *conditional discouragement* type. This type completes the first and the second tasks only if the cost is low.

(c) If $c_L \leq \beta < \frac{c_L+pc_L}{2p}$, they don't do the first task because of the risk of not doing the second task. We call this type the $T3d$ *discouragement* type.

4. If $\beta < c_L$, the DM acts exactly the same as the $T3d$ type creating the $T2$ *never* type.

We summarize six types and conditions under which they occur below:

$$
\begin{aligned}
T1\text{, always if} \quad & c_H \leq \beta \\
T3a\text{, encouragement if} \quad & \frac{c_H + pc_L + (1-p)c_H}{2} \leq \beta < c_H \\
T3b\text{, conditional encouragement if} \quad & \frac{c_H}{2} \leq \beta < \frac{c_H + pc_L + (1-p)c_H}{2} \\
T3c\text{, conditional discouragement if} \quad & \frac{c_L + pc_L}{2p} \leq \beta < \frac{c_H}{2} \\
T3d\text{, discouragement if} \quad & c_L \leq \beta < \frac{c_L + pc_L}{2p} \\
T2\text{, never if} \quad & \beta < c_L.
\end{aligned}
$$

As elaborated above, types $T1$ and $T2$ behave the same way under the piece-rate and all-or-nothing mechanisms. $T3$ types, who only complete the task when the cost is low under the piece-rate mechanism, are divided into four subtypes under the all-or-nothing mechanism. Types $T3a$ and $T3b$ (encouraged types) may complete a task even when the associated cost with that task is high, and types $T3c$ and $T3d$ (discouraged types) may not complete a task even when the associated cost with that task is low.

Hence, we establish that encouraged types are more likely to complete both tasks under all-or-nothing compared to those under piece-rate. The discouraged type, on the other hand, is less likely to complete both tasks under all-or-nothing compared to those under piece-rate.

■

Proof of Proposition 1.

We first start by calculating the expected utility at period zero under each mechanism for each DM type. In the table below, we provide the expected utilities of different types of DMs under piece-rate and all-or-nothing mechanisms, and also indicate which of the utilities is higher by "$>, <, =$" signs.

	$EU^{piece-rate}$		$EU^{all-or-nothing}$
$T1$	$\beta[2p^2(1-c_L)$ + $2(1-p)^2(1-c_H)$ + $2p(1-p)(2-c_L-c_H)]$	$=$	$\beta[2p^2(1-c_L)$ + $2(1-p)^2(1-c_H)$ + $2p(1-p)(2-c_L-c_H)]$
$T3a$	$\beta[2p^2(1-c_L)$ + $2p(1-p)(1-c_L)]$	$<$	$\beta[2p^2(1-c_L)$ + $2(1-p)^2(1-c_H)$ + $2p(1-p)(2-c_L-c_H)]$
$T3b$	$\beta[2p^2(1-c_L)$ + $2p(1-p)(1-c_L)]$	$>$	$\beta[2p^2(1-c_L)$ + $p(1-p)(2-c_L-c_H)]$
$T3c$	$\beta[2p^2(1-c_L)$ + $2p(1-p)(1-c_L)]$	$>$	$\beta[2p^2(1-c_L)]$
$T3d$	$\beta[2p^2(1-c_L)$ + $2p(1-p)(1-c_L)]$	$>$	0
$T2$	0	$=$	0

As it can easily be seen in the second line of the above table, the encouragement type, $T3a$, prefers all-or-nothing under self-select treatment since the expected utility is higher.

Next, we summarize the probability of full (f), zero (n), and partial (p) completion for each treatment and for each type of DM.

	Piece-rate			All-or-nothing			Self-select		
	f_i^P	n_i^P	p_i^P	f_i^A	n_i^A	p_i^A	f_i^S	n_i^S	p_i^S
$T1$	1	0	0	1	0	0	1	0	0
$T3a$	p^2	$(1-p)^2$	$2p(1-p)$	1	0	0	1	0	0
$T3b$	p^2	$(1-p)^2$	$2p(1-p)$	p	$(1-p)$	0	p^2	$(1-p)^2$	$2p(1-p)$
$T3c$	p^2	$(1-p)^2$	$2p(1-p)$	p^2	$(1-p)$	$p(1-p)$	p^2	$(1-p)^2$	$2p(1-p)$
$T3d$	p^2	$(1-p)^2$	$2p(1-p)$	0	1	0	p^2	$(1-p)^2$	$2p(1-p)$
$T2$	0	1	0	0	1	0	0	1	0

Naturally, the full completion rate among DMs depends on the share of different types in the population. Let s_T denote the share of DMs of type $T \in \{1, 2, 3a, 3b, 3c, 3d\}$. Next, we write the full task completion rate under each mechanism using the terminology s_T.

Let f^P, f^A, and f^S denote the full task completion rate in the population for piece-rate, all-or-nothing and self-select treatments, respectively.

$$f^P = s_1 + s_{3a}p^2 + s_{3b}p^2 + s_{3c}p^2 + s_{3d}p^2$$
$$f^A = s_1 + s_{3a} + s_{3b}p + s_{3c}p^2$$
$$f^S = s_1 + s_{3a} + s_{3b}p^2 + s_{3c}p^2 + s_{3d}p^2$$

Comparing self-select to piece-rate, we write the difference between them as:

$$f^S - f^P = s_{3a}(1 - p^2) \tag{1}$$

This expression is positive if the share of encouraged types in the population is positive, meaning the self-select mechanism will be effective if there are encouraged types.

Hence, we establish that the self-select mechanism improves upon the full completion rate over the piece-rate mechanism as long as there are encouraged type DMs in the population.

Lastly, comparing the all-or-nothing to piece-rate, we get:

$$f^A - f^P = s_{3a}(1 - p^2) + s_{3b}(p - p^2) + s_{3d}(-p^2) \tag{2}$$

This difference may be positive or negative depending on the values of s_{3a}, s_{3b}, and s_{3d}. It is easy to see that if s_{3d} is high enough (or s_{3a} and s_{3b} are low enough), in other words if the population has comparatively more discouraged types than encouarged types, $f^A - f^P$ may be negative and hence the all-or-nothing treatment may result in worse full completion rates. ■