Selling Self-Control*

Fuhai Hong[†] Wei Huang[‡]

March 2025

Abstract

This paper presents theoretical analysis on a monopolistic market selling commitment devices for solving self-control problems. The commitment contract increases the buyer's non-compliance cost with a personal goal. When the buyer's investment return is high, the optimal contracts achieve the first best, even with incomplete information. When it is low, asymmetric information leads to a second-best separating equilibrium in which the seller distorts the commitment contract for buyers with weak self-control and causes over-investment. Furthermore, we show that mandating sellers to use non-monetary penalty or to transfer penalty payments to third parties leads to more severe over-investment and reduces welfare.

Keywords: time inconsistency, self-control market, commitment device, penalty, screening, contract design

JEL Codes: D86; D91.

*We would like to thank Mark Armstrong, Tat How Teh, Songfa Zhong, participants in workshops at Shandong University and SUSTech, and conference participants in APIOC 2024 (Seoul), 2024 Econometric Society Australasia Meeting (Melbourne), and 2024 Lingnan International Conference on Competition Policy (Hong Kong) for helpful comments. Declaration: no conflict of interest.

[†]Department of Economics, Lingnan University, Hong Kong (fuhaihong@ln.edu.hk).

[‡]CUHK Business School, Chinese University of Hong Kong, Hong Kong (weihuang@cuhk.edu.hk).

1 Introduction

A self-control market has arisen in recent years, providing consumers with commitment devices to help them combat self-control deficiencies. The understanding of the seller's pricing strategies, contractual designs, and profit models in the market are still insufficient due to the nascent development of the market and its unique features distinct from traditional markets. Additionally, its regulatory policies are still in embryo. This paper makes a first attempt to analyze the market by examining the incentives and efficiency of the commitment contract design, especially when the buyer's self-control deficiency is not observable, and provides some policy implications.

Self-control is broadly recognized as a pivotal personal trait that significantly correlates with success across various life domains (Moffitt et al., 2011). However, it is well-documented in economic and psychological literature that individuals often exhibit imperfect self-control, which hampers their ability to resist temptations. The challenge of maintaining self-control is particularly acute in contemporary society, which is replete with constant and varied temptations, ranging from digital distractions and consumerist lures to more traditional vices such as unhealthy eating and substance abuse.

People are often aware of their self-control problems and proactively seek ways to manage their self-control. Commitment devices, designed to tackle self-control problems and encourage better choices, have been extensively explored both in scholarly research and in practice. These devices, including deadlines, punishment etc., allow individuals to voluntarily restrict their future choice set or increase the costs of certain potential future actions (Bai et al., 2021). Individuals are found to demand commitment devices in a variety of contexts, ranging from savings (Ashraf, Karlan, and Yin, 2006), smoking cessation (Giné, Karlan, and Zinman, 2010), alcohol consumption (Schilbach, 2019), fertilizer use (Duflo, Kremer, and Robinson, 2011), work effort (Kaur, Kremer, and Mullainathan, 2015), to exercise behavior (Royer, Stehr, and Sydnor, 2015), among others.

The demand of commitment devices also creates business opportunities. In recent years, a market supplying these commitment devices is emerging, particularly facilitated by the rapid development of mobile internet technology. This burgeoning sector includes platforms like stickK.com, beeminder.com, WeChat Reading, etc., each designed to help users manage their habits and achieve personal goals. stickK.com, founded by two economists Ian Ayres and Dean Karlan in 2007, utilizes commitment contracts to enhance self-control. Users set personal goals and face penalties for non-compliance, which can be a monetary loss or the embarrassment of having family or friends be notified of the failure. So far, they have successfully attracted $66 million on the line, provided over 624,000 commitment contracts, and helped one million workouts completed successfully and 69 million cigarettes not smoked. Beeminder.com is another goal-tracking tool that uses commitment contracts to help users achieve their goals. Users set a goal and stake their own money on following a specific path to reach it. If they deviate from this path, they pay a monetary penalty. Beeminder also provides some monthly-paid premium plans with richer functions. WeChat Reading, an extension of the popular messaging app, encourages reading habits by providing self-selected paid commitment contracts with a two-part tariff feature. Different contracts have different prices and provide different incentives of completing tasks. Besides the internet platforms and apps, wearable devices like Pavlok (see shop.pavlok.com for reference) use mild electric shocks as non-monetary punishment to help users break bad habits. These tools and platforms leverage technology to create markets that provide users with commitment devices to achieve their personal improvement goals, in finance, health, self-education, and others.

This paper aims to provide theoretical analysis on the burgeoning market of commitment devices by examining the incentives and efficiency of the commitment contract design. An individual with present-biased preference, is endowed with an investment project that has uncertain cost. He is aware of his under-investment problem arising from the present-biased preference, and so demands a commitment device that would increase his cost of non-investment (i.e. imposing a penalty if he does not invest). A seller, who may or may not observe the individual's degree of self-control deficiency, designs and sells a commitment contract to this individual, fine-tuning the penalty level and charging a price for this contract. The penalty of non-compliance (non-investment) is monetary and collected by the seller in our basic analysis. We find that, when the investment return is large enough, the seller designs first-best commitment contracts, and no buyer type receives an information rent, even if the buyer's self-control ability is his private information. However, when the investment return is not sufficient to justify investment under high cost, although a seller who observes the buyer's type still offers first-best commitment contracts and extracts all surplus, a seller who does not observe the buyer's type supplies only second-best commitment contracts. There is a canonical tradeoff between extraction and efficiency for the seller. The seller inflates the low-self-control buyer's penalty level to reduce (or even eliminate) the high-self-control buyer's information rent. The inflated penalty level for the low-self-control buyers, consequently, causes an over-investment problem on them. The degree of this distortion depends on the distribution of the two buyer types.

The screening problem of the self-commitment devices thus generates insights that are different from conventional screening models with traditional products. In the latter (see, e.g. Tirole (1988), Chapter 3), the seller offers a menu of quality-price pairs and the design of the menu always involves an extraction-efficiency tradeoff, where the seller distorts the low type's product quality to reduce the information rent obtained by high-type buyers. In our model with commitment devices, the analogous tradeoff (the seller distorting the low type's penalty level to reduce the high type's information rent) arises only when the investment return is not large. When the investment return is large enough, the seller designs first-best contracts and extracts all the surplus. The intuition can be understood as follows. In conventional screening models, both buyer types prefer higher quality; however, in our model, the penalty in the commitment contract should be "personalized" and features a "bliss point"; once the penalty reaches the "bliss point", the buyer has no incentive to pursue higher penalty levels. This is the case when the investment return is large enough, where the seller offers each type his "bliss point" penalty; the high type thus has no incentive to mimic the low type's contract which has a higher penalty and a higher price. However, when the investment return is low, to avoid over-investment, the seller provides the high-type buyer with a penalty level lower than his "bliss point"; the high-type buyer thus has an incentive to mimic the low type, causing the canonical extraction-efficiency tradeoff for the seller. Overall, our analysis suggests that the extraction-efficiency tradeoff is milder for the seller in the emerging market of commitment devices than in traditional markets.

In reality, the principal who designs commitment contracts can not only be a profit-maximizing firm, but can also be a non-profit organization, or even the buyer's family/friend, who does not pursue economic profits but aims to improve welfare. Our analysis next shows that when the profit-maximizing seller sets second-best contracts in facing the extraction-efficiency tradeoff, having a benevolent seller be the principal to design commitment contracts improves welfare. The benevolent contract designer offers first-best contracts that causes neither under-investment nor over-investment, even if she cannot observe the buyer's self-control ability. The intuition is that the benevolent principal is not motivated by "extraction" and so does not face the extraction-efficiency tradeoff. The above result thus highlights the limits and boundaries of relying on profit-driven business in selling commitment contracts.

Finally, the seller's collection of penalty payment upon customers' non-compliance may raise eyebrows from the public because of its "exploitative nature". We then extend our analysis to the case when the penalty is either non-monetary or is transferred to some third party other than the seller. The use of non-monetary penalty or the transfer of penalty payment to third parties does not break the efficiency result when the investment return is high enough, but gives rise to some different results when the investment return is low. When the return is not high enough to justify investment under high cost, the seller, who does not benefit from the penalty payment of non-compliance, chooses high penalty levels to fully commit the buyer to investment. This causes a broader and more severe over-investment problem compared to the case of monetary penalty collected by the seller, because here, both buyer types are subject to the full commitment to investment. Our analysis thus provides the policy implication that banning the "exploitative feature" of the contract (preventing sellers from collecting penalty payments upon non-compliance) can reduce welfare.

In a nutshell, our paper makes a first attempt to perform standard economic analysis for the growing market of commitment devices. The model employs the analytical paradigm of behavioral industrial organization (behavioral IO) and comprehensively investigates the monopolist seller's contractual strategies under complete and incomplete information. It also provides food for thought for the formulation of public policy relating to this emerging market, the regulation of which is still largely absent.

The literature in behavioral IO explores how firms design contracts to exploit or accommodate consumers' nonstandard preferences. For example, several studies investigate how consumers' self-control deficiency and their naivety affect the firm's contract design, with implications for social welfare and firm profitability (e.g. DellaVigna and

Malmendier, 2004, 2006; Eliaz and Spiegler, 2006; Gottlieb and Zhang, 2021; Gao and Guo, 2024). While this literature (more carefully reviewed in Section 2) focuses on traditional markets, our paper specifically analyzes the market supplying commitment contracts. There are distinct features in this emerging market that are different from the tradtional markets, which require special attention and separate analysis. First, traditional markets, the focus of the literature, supply real products or services that both time-inconsisent and time-consisent consumers demand. In such markets, traditional IO components such as production technology and cost necessarily affect the analysis. By contrast, in the market we analyze, the buyer has a personal goal, the attainment of which does not necessarily require the seller's engagement; the contract supplied by the seller is a pure commitment device with zero production cost, on which time-consistent consumers have no demand. Second, the market of commitment contracts stands out for its strong need for customization. Customers have different self-control challenges, as shown in platforms like stickK.com where they set their personal goals. Given the myriad personal goals set by heterogenous customers, the seller is thus unlikely to observe the customer's self-control deficiency on the specific issue, making analysis based on incomplete information of different self-control types imperative for the modelling of this market. Such analysis is the focus of our paper, and as discussed above, generates new insights that are different from the standard screening models on conventional markets.

The rest of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 presents the model setup and some preliminary results about the demand side of the market, which lays the foundation of our anlaysis. We investigate a monopolistic seller's contratual problem in Section 4, assuming that the seller collects a monetary penalty from the buyer upon the latter's non-compliance. Section 5 extends the analysis to non-monetary penalties or penalty transfer to third parties, and draws policy implications. Section 6 concludes.

2 Relation to the Literature

There are two strands of economics literature that study mechanisms that help solve individuals' self-control problems. One strand of the literature investigates the demand for and the effectiveness of external commitment devices (for reviews, see e.g. Bryan, Karlan, and Nelson, 2010; Kremer, Rao, and Schilbach, 2019). Commitment devices allow individuals to voluntarily restrict their future choice set or increase the costs of certain potential future actions (Bai et al., 2021). The commitment contract considered in our study, which imposes a penalty on non-investment behavior, belongs to the latter category. Individuals who are aware of their self-control problems demand commitment devices in a variety of contexts, ranging from savings, smoking cessation, alcohol consumption, fertilizer use, work effort, to exercise behavior, among others. Commitment devices are shown to be effective in managing self-control problems in such contexts as health, saving, etc. (Wertenbroch, 1998; Thaler and Benartzi, 2004; Ashraf, Karlan, and Yin, 2006). While this (mostly empirical) literature focuses on individuals' demand for commitment devices and studies their effectiveness, our paper theoretically investigates the contracting problem of designing commitment devices in a market context from an industrial organization perspective.

Another strand of the literature focuses on motivated beliefs as intrinsic commitment devices in resolving self-control problems. These studies show how people mitigate their self-control problems by (selectively) processing information to generate favorable beliefs, such as avoiding information strategically (Carrillo and Mariotti, 2000), creating memory biases (Bénabou and Tirole, 2002; Chew, Huang, and Zhao, 2020), and self-signaling (Bénabou and Tirole, 2004, 2011; Hong, Huang, and Zhao, 2019). The commitment contract considered in our model is an extrinsic device.

The literature of behavioral industrial organization investigates how profit-maximizing firms respond to consumers who exhibit nonstandard preferences or behavioral biases, including reference dependence (Zhou, 2011), prominence (Armstrong, Vickers, and Zhou, 2009; Armstrong and Zhou, 2011), overconfidence (Grubb, 2009), confusion (Chioveanu and Zhou, 2013), etc. For reviews, see, e.g. Kőszegi (2014), Grubb (2015), and Heidhues and Kőszegi (2018). Within the literature, there is a line of studies on the contracting with time inconsistent consumers, including but not restricted to DellaVigna and Malmendier (2004), Eliaz and Spiegler (2006), Esteban, Miyagawa, and Shum (2007), Gottlieb (2008), Li, Yan, and Xiao (2014), Galperti (2015), Gottlieb and Zhang (2021), Heidhues, Kőszegi, and Murooka (2024) in economics, Jain (2012) and Gao and Guo (2024) in marketing, and Li and Jiang (2022) in operation research. Most studies analyze one-shot models, except Gottlieb and Zhang (2021) who investigate dynamic contracting in a long-term relationship.

In DellaVigna and Malmendier (2004), firms sell products to time inconsistent consumers, with a two-part tariff contract. The seller knows the consumers' degree of time inconsistency while consumers are partially naive about it. The sellers craft contracts that strategically exploit the consumer's naivety. Empirically, DellaVigna and Malmendier (2004, 2006) analyze data from gym memberships and show that contracts are structured to attract consumers with low self-control who over-estimate their future attendance, leading to profitability from underutilized memberships, which supports their theory.

Relatedly, Heidhues and Kőszegi (2010) analyze contracting problems with partially naive time inconsistent consumers in a competitive credit market, demonstrating how credit markets exploit consumer naivety about their self-control, resulting in contracts that lead to suboptimal borrowing behaviors and increased profits for lenders through fees and interest on unanticipated debt. Heidhues, Kőszegi, and Murooka (2024) explore how consumers' procrastination affects competition in service markets where switching providers is possible but costly. They develop a model showing that consumers often fail to switch to better offers due to present bias and naivety, leading to reduced competition and higher prices. While in these models, the principal knows better than the agent about the agent's preference, in our model, we assume that consumers are sophisticated about their degree of time inconsistency, which is the consumer's private information that the seller cannot observe.

Li, Yan, and Xiao (2014) extend DellaVigna and Malmendier (2004)'s model to a screening model in which the consumer's benefit is his private information. Relatedly, in the screening models of Eliaz and Spiegler (2006) and Gao and Guo (2024), agent types differ in their degree of sophistication. Put differently, the principal knows the agent's degree of inconsistency, but does not know whether he is aware of it. Gao and Guo (2024) reveal that offering a variety of contracts can help screen the consumers with different beliefs, leading to two-sided deviations from marginal cost pricing. They suggest that higher degrees of time inconsistency might reduce firm profits but increase social welfare, while reducing consumer naivety may have adverse effects on social welfare.

In our model, the agents' types being screened is their degree of time inconsistency. In this sense, we are closer to Esteban, Miyagawa, and Shum (2007) who consider monopolistic nonlinear pricing with consumers with self-control preferences *a la* Gul and Pesendorfer (2001). In their model, the consumer's temptation utility, which causes time inconsistency, is her private information. As in Li, Yan, and Xiao (2014), Esteban, Miyagawa, and Shum (2007), and Galperti (2015) (discussed below), in our model, agents are sophisticated about their own time inconsistency.

Unlike the above models in which sellers sell traditional products using various pricing strategies, the seller in our model does not sell any real product; instead, the seller sells a commitment contract. In such a market, the buyers are sophisticated in the sense that they are aware of their self-control problems – otherwise they won't demand commitment contracts. The major heterogeneity in the demand side of the market, arguably, is the consumers' degree of time inconsistency, which is not observable to the seller. Moreover, in the traditional product market, the tradeoff between commitment and flexibility is less likely a concern (Heidhues and Kőszegi, 2018, Footnote 35). The tradeoff is crucial to our model.

Galperti (2015) studies a mechanism design problem, with saving plans as a major application, on optimally providing flexible commitment devices to time inconsistent agents, in the presence of state uncertainty with (un)observable time inconsistency. He focuses on two types of consumers, time-consistent and time-inconsistent consumers. The device, on top of providing commitment value, also provides an indispensable facility to the consumers, so that the time-consistent consumers also demand it.[1] In this sense, the commitment device considered in his paper is closer to the products/services sold in traditional markets analyzed in the above mentioned papers. By contrast, the commitment contract in our model only provides a commitment function (as in the example of stickK) and the consumer's pursuit of goals does not necessitate the seller's engagement;[2] hence, a time-consistent consumer has no demand on the commitment contracts considered in our model and can achieve his goal by himself. We solve for commitment contracts that screen two types of consumers who are both time inconsistent but are different in their

[1]In the example of saving, the principal provides financial service (a device) to the agent; the agent cannot save without the device. Moreover, the contractible action (saving) is costly to the principal.

[2]And it does not cause any cost to the seller.

degree of time inconsistency. Moreover, in Galperti (2015), the "high" (time-consistent) type values any device more than the "low" (time-inconsistent) type (Proposition 4.1), which is not the case in our model. With the differences in settings between Galperti (2015) and our model, the two models produce qualitatively different results. For one prominent example, in Galperti (2015), the "high" (time-consistent) type's contract violates the usual "no distortion at the top" property, while in our model, this property is satisfied.

Relatedly, Amador, Werning, and Angeletos (2006) and Ambrus and Egorov (2013) study the optimal tradeoff between commitment and flexibility in a consumption-saving model for agents with self-control problems. Their optimal-commitment problem is to find the best subset of the individual's budget set from which she will be allowed to choose her saving level. While their commitment revolves around the restriction of choice sets, our commitment device increases the cost of a certain future choice.

3 Preliminaries: Demand for Commitment Contracts

3.1 Model Setup

We consider a model with three dates, $t = 0, 1, 2$. A risk-neutral individual ("he") is endowed with an investment project with cost c. The cost c is a random variable that is realized at $t = 1$. At $t = 0$, the individual only knows that c is distributed over a normalized interval, $[0, 1]$, with cumulative distribution function $F(c)$ and density function $f(c)$. At $t = 1$, after observing the realized c, the individual decides whether to invest. If he invests, the individual will receive a return $v > 0$ at $t = 2$, which can be greater or less than 1. When v exceeds 1, the project always deserves investment.

The individual has time-inconsistent preferences. We adopt the quasi-hyperbolic discounting model. Specifically, the individual at date t discounts future payoff at date $t+n$ with discount factor $\beta\delta^n$ for $n = 1, 2$, where $\delta = 1$ is the normal discount factor that discounts payoffs between any two adjacent periods, and on top of that $0 < \beta < 1$ further discounts future payoffs to the present. Put differently, β indicates the individual's self-control ability. We assume that the individual is aware of his β as in Esteban, Miyagawa, and Shum (2007), Li, Yan, and Xiao (2014), and Galperti (2015), though it may or may not be observable to others. In markets selling commitment devices, the customers are usually those who are aware of their own self-control problems. We also assume that $\beta v < 1$, and so the imperfect self-control leads to an under-investment problem; that is, when $\beta v < c < v$, the individual does not invest at $t = 1$ despite his initial intention at $t = 0$ to do so.

The individual (henceforth referred to as the "buyer") decides at $t = 0$, the contracting stage, whether to purchase a commitment contract provided by a seller ("she"). The contract consists of two components: p and b, where $p > 0$ is a penalty if the buyer does not invest, and $-\infty < b < +\infty$ is the lump-sum price of the contract. If the buyer decides to purchase the contract, the contract comes into effect at $t = 1$, when the buyer pays b to the seller. In the case that the buyer procures the contract but fails to invest at $t = 1$, he not only forfeits gains from the investment project but also incurs penalty $p > 0$ at $t = 2$. Conversely, if the buyer declines the contract at $t = 0$, he is neither obligated to pay b at $t = 1$ nor subjected to any penalty at $t = 2$. Figure 1 summarizes the timeline. The two-part tariff feature of the contract (if p is monetary and collected by the seller upon the buyer's non-compliance) and the timeline of the model are similar to DellaVigna and Malmendier (2004).

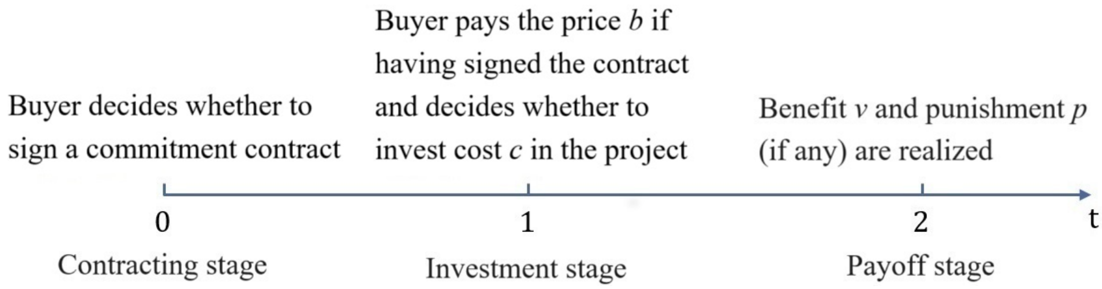

Figure 1: Timeline

The model assumes that the buyer's investment decision is contractible (observable and verifiable). This assumption applies to many real-world situations with self-control problems but not all. Goals such as losing weight, how often to go to the gym, doing certain types of exercise (facilitated by the development of mobile apps), and stopping playing certain video games, are observable and verifiable. For a student, spending how many days in the library might be contractible, but how much real effort to be made is not.

3.2 The Buyer

We make a tie-breaking assumption that the buyer does not invest if he is indifferent. Having not bought the contract, at $t = 1$, the buyer will invest if and only if

$$\beta v - c > 0.$$

Having bought the contract, the buyer will invest at $t = 1$ if and only if

$$\beta(v + p) - c > 0.$$

At $t = 0$, the buyer's expected payoff of not buying the contract is

$$\beta \int_0^{\beta v} (v - c) dF(c).$$

Not considering the price b, the expected payoff from buying the contract is

$$\beta \int_0^{\min\{\beta(v+p),1\}} (v - c) dF(c) - \beta \int_{\min\{\beta(v+p),1\}}^1 p dF(c).$$

Their difference, from the buyer's viewpoint at $t = 0$, is the net value of signing the commitment contract βV_C (not considering the price b), where

$$V_C(\beta, p) \equiv \int_{\beta v}^{\min\{\beta(v+p),1\}} (v-c)dF(c) - p\int_{\min\{\beta(v+p),1\}}^{1} dF(c). \tag{1}$$

If $\beta = 0$, i.e., when the agent's self-control ability is extremely weak, the first term in the right side of (1) equals zero, and so $V_C < 0$. If $\beta = 1$, i.e., when the agent has perfect self-control ability, $V_C \leq 0$, as the commitment contract causes over-investment for $c > v$. However, the value of the commitment device could be positive. Given our assumption $\beta v < 1$, $V_C = \int_{\beta v}^{1}(v-c)dF(c) > 0$ for $\frac{1}{v+p} < \beta < \frac{1}{v} \leq 1$.

For tractability, in the following analysis we assume that c is uniformly distributed. We have

$$\begin{aligned} V_C &= \int_{\beta v}^{\min\{\beta(v+p),1\}} (v-c)dc - p\int_{\min\{\beta(v+p),1\}}^{1} dc \\ &= \begin{cases} \int_{\beta v}^{1}(v-c)dc \text{ if } p \geq \frac{1}{\beta} - v \\ \int_{\beta v}^{\beta(v+p)}(v-c)dc - p\int_{\beta(v+p)}^{1} dc \text{ otherwise} \end{cases}. \end{aligned} \tag{2}$$

The following lemma summarizes the relationship between V_C and β.

Lemma 1 *$V_C(\beta, p)$ at any given $p > 0$ is increasing in β for $\beta \in \left(0, \frac{1}{v+p}\right]$, and decreasing in β for $\beta \in (\frac{1}{v+p}, \frac{1}{v})$.*

The proofs of lemmas and propositions are relegated to Appendix A.

When $\beta \in \left(0, \frac{1}{v+p}\right]$, the buyer with low self-control ability is likely to pay the penalty, and so the value of the commitment increases with the individual's self-control ability. For $\beta \in (\frac{1}{v+p}, \frac{1}{v})$, the penalty ensures that the individual always invests and so he does not pay the penalty. In this case, the lower the self-control ability, the larger the value of commitment is. The following lemma examines the relationship between V_C and p.

Lemma 2 *If $v \geq \frac{1}{(2-\beta)\beta}$, V_C is positive, increasing in p for $p \in [0, \frac{1}{\beta} - v]$, and keeps constant for $p \in [\frac{1}{\beta} - v, +\infty)$; if $v < \frac{1}{(2-\beta)\beta}$, V_C is negative and decreasing in p for $p \in [0, \frac{1}{(2-\beta)\beta} - v]$, increasing in p for $p \in \left(\frac{1}{(2-\beta)\beta} - v, \frac{1}{\beta} - v\right]$, and keeps constant (either positive or negative) for $p \in [\frac{1}{\beta} - v, +\infty)$.*

Lemma 2 can be illustrated in Figure 2. When the return v is high enough ($v \geq \frac{1}{(2-\beta)\beta} > 1$; Figure 2a), the commitment device is valuable in resolving the self-control problem. The higher the penalty p is, the more benefit there is from resisting temptation. When the penalty is sufficiently high, the individual fully commits to investment, and so the value of the commitment device is constant for larger punishment.

For lower return v ($v < \frac{1}{(2-\beta)\beta}$; Figure 2b), the commitment device is less helpful. The value of the commitment device with small penalty p is negative: it is not strong

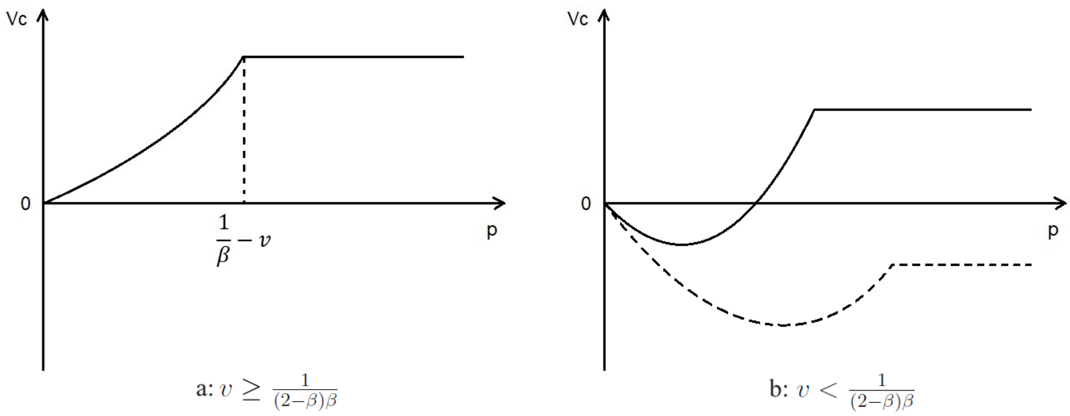

Figure 2: Value of the commitment contract

enough to change investment behavior and the buyer has to pay the penalty. Initially, the buyer bears more cost with higher penalty p. However, the commitment value starts to increase for a large enough p and keeps constant eventually (when the penalty is so high that the individual fully commits to investment). Depending on the value of v, the full commitment may have positive (the solid line in Figure 2b) or negative (the dashed line) values. The latter case is less interesting because the commitment device is never useful for the buyer (which is the case when v is very small). We will focus on the former case, i.e., $\int_{\beta v}^{1} (v-c)\,dc > 0$. Moreover,

$$\begin{aligned} &\int_{\beta v}^{1} (v-c)\,dc = (\beta v - 1)\left[\left(\frac{1}{2}\beta - 1\right)v + \frac{1}{2}\right] > 0 \\ \Leftrightarrow\, &(\beta v - 1)\left[\left(1 - \frac{1}{2}\beta\right)v - \frac{1}{2}\right] < 0 \\ \Leftrightarrow\, &\frac{1}{2-\beta} < v < \frac{1}{\beta}, \end{aligned} \tag{3}$$

where the last line is derived from the assumption $\beta v < 1$. Thus altogether, we assume:

Assumption 1 $\frac{1}{2-\beta} < v < \frac{1}{\beta}$.

Given Figure 2a or 2b (solid line, Assumption 1), V_C is maximized by any $p \geq \frac{1}{\beta} - v$, i.e. a high enough penalty that fully commits the buyer to investment. Penalty $p = \frac{1}{\beta} - v$ thus constitutes a "bliss point" for the buyer on or above which the commitment value is maximized.

4 Monopolistic Seller's Optimal Contracts

In line with much of the literature discussed in Section 2, we consider a monopolistic seller's design of optimal contracts. In reality, there are only a few firms selling self-control contracts all over the world, targeting somewhat different populations, and so they each have some market power. Moreover, if an individual is engaged with a gym club, a coach, or a social media platform in a long-term relationship, the latter would naturally have some monopolistic power in selling self-control devices to the individual.

In this section, we assume that, the penalty p is monetary and, in the case of non-compliance, is collected by the seller. The contract is thus akin to a two-part tariff. Beeminder's business model has this feature. Section 5 extends the analysis to non-monetary penalties or penalty transfer to third parties and discusses policy implications.

This section will discuss the cases of high and low investment returns separately, as they yield qualitatively different results. With high investment return ($v \geq 1$), investment is always worthwhile.

4.1 High Return ($v \geq 1$)

4.1.1 Benchmark: Complete Information

With complete information of buyers' β, the problem of the monopolistic seller is[3]

$$\begin{aligned} &\max_{p,b} b + p \int_{\min\{\beta(v+p),1\}}^{1} dc \\ &s.t.\ V_C(p) - b \geq 0 \end{aligned} \tag{4}$$

where the constraint is the buyer's individual rationality (IR) condition.

In the solution, $b = V_C(p)$. The problem becomes

$$\max_{p} \int_{\beta v}^{\min\{\beta(v+p),1\}} (v-c)dc, \tag{5}$$

where the objective function is in line with the "social welfare", i.e. the sum of the payoffs of the buyer and of the seller.

Given $v \geq 1$ (i.e. investment is always worthwhile), the seller will set

$$\begin{aligned} &\min\{\beta(v+p),1\} = 1 \\ &\Rightarrow \beta(v+p) \geq 1 \end{aligned}$$

implying

$$\begin{aligned} &p^c \geq \frac{1}{\beta} - v, \\ &b^c = V_C(p^c) = \int_{\beta v}^{1} (v-c)dc = v - \frac{1}{2} - \beta v^2 + \frac{\beta^2 v^2}{2}, \end{aligned} \tag{6}$$

where superscript c denotes complete information.

[3]We slightly abuse notations by using both $V_C(\beta,p)$ and $V_C(p)$. We use $V_C(p)$ for the analysis when β is given.

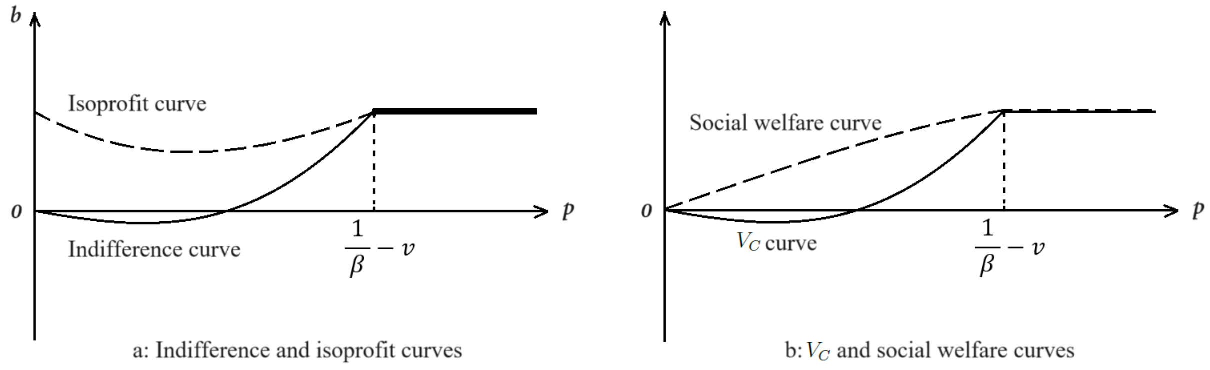

Figure 3: Profit maximizing contract with $v \geq 1$ (for the case of $1 \leq v < \frac{1}{(2-\beta)\beta}$)

Figure 3a illustrates the profit-maximizing contracts when $v \geq 1$ (for the case of $1 \leq v < \frac{1}{(2-\beta)\beta}$), with p at the horizontal axis and b at the vertical axis.[4] With V_C, one can draw the buyer's indifference curves for contract profiles (p, b):

$$b = V_C(\beta, p) - \overline{u}, \tag{7}$$

where $\overline{u}$ indicates the buyer's payoff level. The lower the indifference curve is, the better off the buyer is. The seller's isoprofit curves can be written by

$$b = \begin{cases} \overline{b} - p \int_{\beta(v+p)}^{1} dc \text{ when } p \leq \frac{1}{\beta} - v \\ \overline{b} \text{ otherwise} \end{cases},$$

where $\overline{b}$ denotes the seller's profit. A higher isoprofit curve indicates a higher profit. In Figure 3a, the solid curve, which passes the origin, is the buyer's zero-payoff indifference curve ($b = V_C(\beta, p)$). Thus the area on and below the indifference curve satisfies the IR constraint in (4). The dashed curve in Figure 3a is the highest isoprofit curve that the seller can achieve, given the buyer's IR constraint. The bold section of the buyer's indifference curve indicates the optimal contracts as specified in (6).

Figure 3b presents another illustration by drawing the V_C curve (solid line) and the social welfare curve (dashed line), with p at the horizontal axis (again, for the case of $1 \leq v < \frac{1}{(2-\beta)\beta}$). Both the V_C and social welfare curves are maximized by $p \geq \frac{1}{\beta} - v$; put differently, the social welfare is aligned with the commitment value in that the buyer's "bliss point" also maximizes the social welfare. The seller, who internalizes the social welfare given the buyer's binding IR constraint, chooses a penalty level that is on or above the "bliss point".

Here, since investment is always worthwhile ($v \geq 1$), the seller sets a high enough penalty to fully commit the buyer to investment. The buyer never pays a penalty to the seller; however, all the surplus from the commitment contract is extracted by the seller

[4] The case of $v \geq \frac{1}{(2-\beta)\beta}$ is similar except that the indifference curve does not cross the horizontal axis.

who sets a high enough price for the contract, b. This commitment contract achieves the first best.

The result we obtain here is not new, and is similar to a previous finding by DellaVigna and Malmendier (2004) that when the buyer's self-control ability is common knowledge, the optimal two-part tariff implements the first-best outcome, which perfectly solves the self-control problem. We next investigate whether asymmetric information about the buyer's self-control ability would affect this result.

4.1.2 Incomplete Information

In this subsection we assume that β is the buyer's private information. We will show that when the return from investment v is sufficiently high (when $v \geq 1$), incomplete information will not distort the equilibrium outcome. For simplicity, assume that there are two types in the population of buyers: β_H (called type-H) with probability $\alpha \in (0,1)$ and β_L (called type-L) with probability $1-\alpha$, where $0 < \beta_L < \beta_H < \min\{1, \frac{1}{v}\}$.

Definition 1 (Separating Equilibrium) *Commitment contracts offered by a seller* (b_L, p_L) *and* (b_H, p_H) *form a separating equilibrium if and only if*

(i) (b_L, p_L) *and* (b_H, p_H) *are different;*

(ii) buyers of type-L and -H take (b_L, p_L) *and* (b_H, p_H) *respectively;*

(iii) this set of contracts is optimal for the seller.

Figure 4a draws each buyer type's zero-payoff indifference curve, I_L^0 and I_H^0, respectively (for the case of $1 \leq v < \frac{1}{(2-\beta)\beta}$). Put differently, I_i^0 graphs $b = V_C(\beta_i, p)$ for $i \in \{H, L\}$. Lemmas 1 and 2 altogether imply that $V_C(\beta_H, p) < V_C(\beta_L, p)$ for $p > \frac{1}{\beta_L} - v > \frac{1}{\beta_H} - v$ (when p is large enough for both types' full commitment to investment). This is shown on Figure 4a, where I_L^0's horizontal section (when $p > \frac{1}{\beta_L} - v$) is above I_H^0's (when $p > \frac{1}{\beta_H} - v$). However, I_H^0 reaches its horizontal section earlier than I_L^0 ($\frac{1}{\beta_H} - v < \frac{1}{\beta_L} - v$). These results imply that type-L's zero-payoff indifference curve crosses the horizontal section of type-H's zero-payoff indifference curve from below at $p = \overline{p}$, where $\overline{p} \equiv \sup\{p | V_C(\beta_H, p) = V_C(\beta_L, p)\}$, i.e. the positive intersection of the two indifference curves, and[5]

$$\frac{1}{\beta_H} - v < \overline{p} < \frac{1}{\beta_L} - v. \tag{8}$$

Put differently, $\overline{p}$ is the positive solution of the following equation:[6]

$$\int_{\beta_L v}^{\beta_L(v+\overline{p})} (v-c)dc - \overline{p} \int_{\beta_L(v+\overline{p})}^{1} dc = \int_{\beta_H v}^{1} (v-c)dc.$$

[5]Note that this property does not rely on $v \geq 1$. It also applies to $v < 1$.

[6]The solution: $\overline{p} = \frac{1-2\beta_L v+\beta_L^2 v+\sqrt{1-(2-\beta_L)\beta_L[1+(2-\beta_H-\beta_L)(\beta_H-\beta_L)v^2]}}{(2-\beta_L)\beta_L}$.

We have

$$V_C(\beta_L, p) > V_C(\beta_H, p) \text{ if and only if } p > \overline{p}. \tag{9}$$

The following proposition identifies the separating equilibrium.

Proposition 1 *With $v \geq 1$, the seller chooses contract profile $\{(b_L^*, p_L^*), (b_H^*, p_H^*)\}$ that supports a separating equilibrium, where*

$$p_L^* \geq \frac{1}{\beta_L} - v, \ \frac{1}{\beta_H} - v \leq p_H^* \leq \overline{p} \tag{10}$$

$$b_i^* = V_C(p_i^*, \beta_i) = \int_{\beta_i v}^{1} (v - c)dc = v - \frac{1}{2} - \beta_i v^2 + \frac{\beta_i^2 v^2}{2} \text{ for } i \in \{H, L\}$$

for type-L and type-H respectively. The first best, where each type of buyers fully commits to investment, is achieved despite the information asymmetry.

Figure 4a illustrates the separating equilibrium using the two buyer types' indifference curves. Any contract profile with type-H's contract located on the bold part of the indifference curve I_H^0 and with type-L's contract located on the bold part of the indifference curve I_L^0 supports a separating equilibrium. For each type, the penalty level is high enough to ensure that the buyer type fully commits to investment (so no penalty is paid on the equilibrium path), and the price of the contract, b_i for $i \in \{H, L\}$, extracts all the surplus from each buyer type.

The only difference in terms of contract design between complete information and incomplete information is the following. With observable types, any high enough penalty for each type is optimal (see (6)). However, the unobservability of buyer types imposes an upper bound for type-H's contract: $p_H^* \leq \overline{p}$ as in (10); otherwise, type-L buyers would have incentive to mimic type-H (see Figure 4a).

Interestingly, asymmetric information in buyers' self-control ability does not cause any inefficiency in this case (when $v \geq 1$). Figure 3 shows that, given any buyer type, the penalty level chosen by the seller who internalizes the social welfare is in line with the buyer's "bliss point". On one hand, a buyer with self-control ability β_H does not have incentive to pay more money to pursue a punishment higher than p_H^*, which is already high enough to fully prevent slacking. On the other hand, the buyer with β_L does not prefer a contract with lower punishment $p_H^* \leq \overline{p} < p_L^*$, because the gain from the lower price b_H^* cannot compensate the loss from slacking. Therefore, no buyer has incentive to deviate from the commitment contract designed for him.[7]

[7]Anticipating a bit, the no-inefficiency result under incomplete information does not hold when the investment return is low ($v < 1$). There, the buyer's "bliss point" penalty is different from the seller's, who maximizes the social welfare, causing a buyer type's incentive to mimic another type's first-best contract.

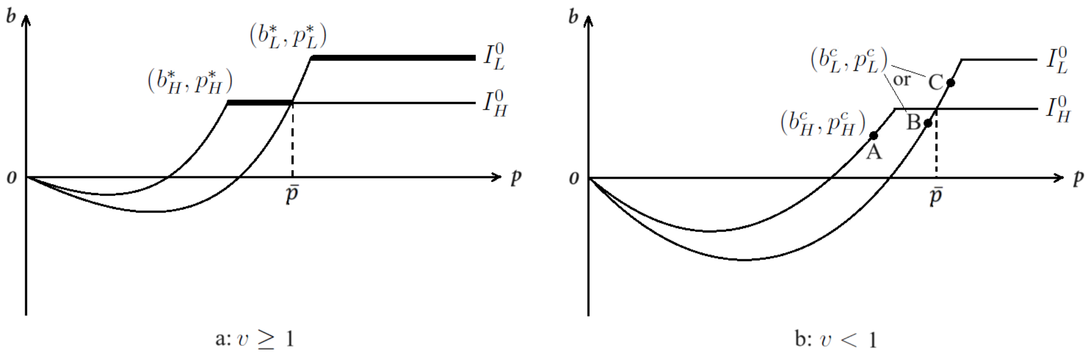

Figure 4: First-best separating equilibrium: Existence for $v \geq 1$ and (in)existence for $v < 1$

4.2 Low Return ($v < 1$)

4.2.1 Benchmark: Complete Information

Now we examine the situation when the investment return is low ($v < 1$). With complete information (problem (5)), given $v < 1$, the seller will set

$$\min\{\beta(v+p), 1\} = v$$
$$\Rightarrow \beta(v+p) = v,$$

implying

$$p^c = \left(\frac{1}{\beta} - 1\right) v, \tag{11}$$
$$b^c = V_C(p^c) = \int_{\beta v}^{v} (v-c)\, dc - \left(\frac{1}{\beta} - 1\right) v \int_v^1 dc = \left(\frac{1}{\beta} - 1\right) v \left(v + \frac{\beta}{2} v - \frac{\beta^2 v}{2} - 1\right).$$

Figure 5a illustrates the profit-maximizing contract for $v < 1$ as the tangent point of the buyer's indifference curve (solid curve) and the seller's isoprofit curve (dashed curve).[8]

When $v < 1$, investment is not always worthwhile. From the social point of view, the buyer should invest if and only if $c \leq v$. The seller, who internalizes the social welfare (the IR condition in (4) is binding), optimally sets the penalty level in (11) so as to induce the efficient level of investment (without under- or over-investment). The result again substantiates DellaVigna and Malmendier (2004)'s result that under complete information, the optimal two-part tariff implements the first-best outcome, which perfectly solves the self-control problem.

Figure 5b draws the V_C curve (solid line) and the social welfare curve (dashed line) with p at the horizontal axis. While the buyer has a "bliss point" at $p = \frac{1}{\beta} - v$, on or above which V_C is maximized, the social welfare, however, is maximized at $p = \left(\frac{1}{\beta} - 1\right) v$

[8]Technically, Assumption 1 ($\int_{\beta v}^{1} (v-c)\, dc > 0$, given (3)) implies that the tangent point in Figure 5a lies in the increasing section of the buyer's indifference curve.

(where the tangency in Figure 5a lies). The penalty p, being a pure transfer from the buyer to the seller, is the buyer's private cost, but is not part of the social cost.

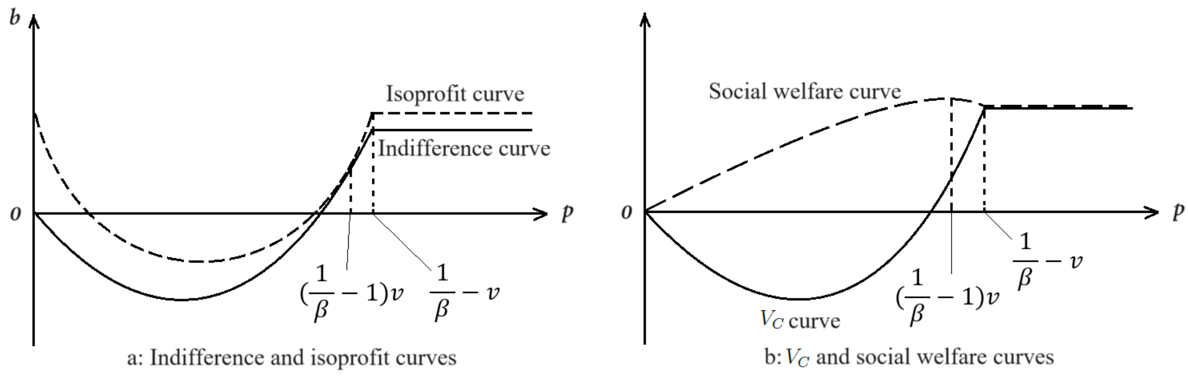

Figure 5: Profit maximizing contract with $v < 1$

4.2.2 Incomplete Information: Social Non-optimality

The previous analysis shows that when the investment return is high, the first best is achieved no matter whether or not information is complete. However, with small v, the penalty level that maximizes social welfare (chosen by the seller with complete information) is lower than the buyer's "bliss point". Consequently, under incomplete information, the type-H buyer may have an incentive to mimic type-L to get a higher penalty level. The following proposition shows that the first-best contracts in (11) may or may not support a separating equilibrium under incomplete information.

Proposition 2 *With $v < 1$, the first-best contract profile $\{(b_L^c, p_L^c), (b_H^c, p_H^c)\}$, where for $i \in \{H, L\}$,*

$$p_i^c = \left(\frac{1}{\beta_i} - 1\right) v, \; b_i^c = V_C(\beta_i, p_i^c), \tag{12}$$

supports a separating equilibrium if and only if $p_L^c \geq \overline{p}$, or equivalently, if and only if $v \geq \frac{\beta_L}{\beta_H}$ and

$$(1 - v)\left(\frac{2v}{\beta_L} - v - 1\right) \leq v^2 (\beta_H - \beta_L)(2 - \beta_H - \beta_L). \tag{13}$$

Proposition 2 points out that $p_L^c \geq \overline{p}$ is the necessary and sufficient condition that the first-best contract profile supports a separating equilibrium. The intuition is that when $p_L^c \geq \overline{p}$, both (b_H^c, p_H^c) and (b_L^c, p_L^c) are located above the other type's zero-payoff indifference curve as indicated in points A and C of Figure 4b respectively, and therefore no one has incentive to mimic the other type. Instead, if $p_L^c < \overline{p}$ so that type-L's first-best contract (b_L^c, p_L^c) is as shown by point B of Figure 4b, then type-H has an incentive to mimic type-L.

In Figure 6, we fix β_L and let the horizontal line represent $\beta_H \in (\beta_L, 1)$ and the vertical line represent $v > 0$. Assumption 1 restricts the parametric space to the area

ABC. In the blank area I, where $1 \leq v < \frac{1}{\beta_H} < \frac{1}{\beta_L}$, there is a first-best separating equilibrium (Proposition 1). The area above curve DF represents the area satisfying $v \geq \frac{\beta_L}{\beta_H}$ and condition (13) simultaneously.[9] Thus, in the blank area II, there is also a first-best separating equilibrium (Proposition 2).

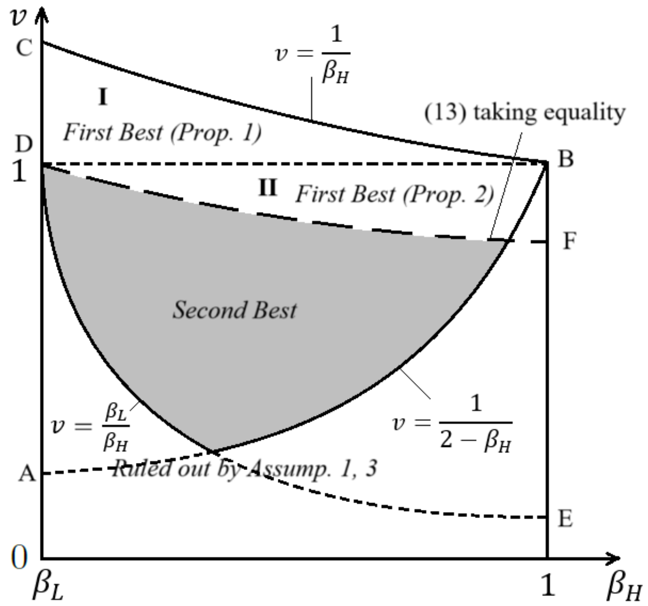

Figure 6: (In)Existence of first-best separating equilibrium with $v < 1$

4.2.3 Monopolistic Screening: Second-best Contracts

In this subsection, we investigate the second-best screening contracts by assuming the first-best contracts in (12) cannot be sustained in a separating equilibrium, i.e., the sufficient and necessary condition for the first-best separating equilibrium in Proposition 2 does not hold:

Assumption 2 $p_L^c = (\frac{1}{\beta_L} - 1)v < \overline{p}$.

Note that in our model, the two buyer types' indifference curves do not satisfy the commonly-used single-crossing assumption globally.[10] We make the following additional assumptions to rule out some (locally) double-crossing cases, making the problem tractable.

Assumption 3 $v \geq \frac{\beta_L}{\beta_H}$.

[9] There is also a parametric space satisfying (13), but it lies below curve DE, which violates $\frac{\beta_L}{\beta_H} \leq v$.

[10] For example, in Figure 4, the two types' indifference curves cross each other at $p = 0$ (the origin) and at $p = \overline{p}$. In the literature, Araujo and Moreira (2010) and Schottmüller (2015) study screening models without single crossing with continuous types. Some other papers study perfect-competition insurance markets with discrete types and without single crossing (e.g., Smart, 2000; Wambach, 2000; Netzer and Scheuer, 2010).

This assumption states that the two types are sufficiently different so that the type ratio is small enough: $\frac{\beta_L}{\beta_H} \leq v < 1$. By Proposition 2, Assumption 2 implies that either condition (13) is violated, or $v < \frac{\beta_L}{\beta_H}$, or both. Assumption 3 rules out the case of $v < \frac{\beta_L}{\beta_H}$ and so implies that (13) is violated. The assumption is also equivalent to the assumption that type-L's first-best penalty p_L^c is (weakly) higher than type-H's "bliss point" penalty level:

$$\frac{\beta_L}{\beta_H} \leq v \Leftrightarrow \left(\frac{1}{\beta_L} - 1\right) v \geq \frac{1}{\beta_H} - v.$$

Altogether, Assumptions 1-3 restrict our attention to the gray area in Figure 6, where $v \geq \frac{\beta_L}{\beta_H}$ but (13) is violated and so the first-best contract profile cannot be sustained as a separating equilibrium. We will look for the second best.

Assumption 4 $V_C(\beta_H, p_H^c) + V_C(\beta_L, p_L^c) > V_C(\beta_L, p_H^c) + V_C(\beta_H, p_L^c)$.

This assumption says the sum of the commitment value when two buyer types choose their own first-best penalty is higher than that when they choose the other type's. It is also equivalent to say that if a type-H buyers is indifferent between a contract (b_H, p_H^c) and type-L' first-best contract (b_L^c, p_L^c), where p_H^c and (b_L^c, p_L^c) are determined by (12), then type-L buyers' utility of taking contract (b_H, p_H^c) is negative: I.e., with

$$V_C(\beta_H, p_H^c) - b_H \equiv V_C(\beta_H, p_L^c) - b_L^c \equiv V_C(\beta_H, p_L^c) - V_C(\beta_L, p_L^c), \tag{14}$$

we have

$$\begin{aligned} & V_C(\beta_L, p_H^c) - b_H \\ =& V_C(\beta_L, p_H^c) - (V_C(\beta_H, p_H^c) - V_C(\beta_H, p_L^c) + V_C(\beta_L, p_L^c)) \\ =& V_C(\beta_L, p_H^c) + V_C(\beta_H, p_L^c) - (V_C(\beta_H, p_H^c) + V_C(\beta_L, p_L^c)) \\ <& 0. \end{aligned} \tag{15}$$

Appendix B illustrates the double-crossing cases that are ruled out by Assumptions 3 and 4 respectively.

We will first characterize the optimal separating contracts, and then show that the optimal separating contract profile is the optimal contract profile for the seller and so supports a separating equlibrium, because *a*) it dominates any pooling contract (one single contract that attracts both buyer types), and *b*) it dominates any contract profile that only attracts one buyer type.

The monopolistic seller's separating contracting problem is

$$\begin{aligned}
&\max_{p_H,b_H,p_L,b_L} \alpha\pi_H + (1-\alpha)\,\pi_L \\
&\quad s.t.\ \ V_C(\beta_H, p_H) - b_H \geq 0 \text{ (IR-H)} \\
&\qquad\quad V_C(\beta_L, p_L) - b_L \geq 0 \text{ (IR-L)} \\
&\qquad\quad V_C(\beta_H, p_H) - b_H \geq V_C(\beta_H, p_L) - b_L \text{ (IC-H)} \\
&\qquad\quad V_C(\beta_L, p_L) - b_L \geq V_C(\beta_L, p_H) - b_H \text{ (IC-L)}
\end{aligned} \tag{16}$$

where for $i \in \{H, L\}$, π_i is the seller's profit from a type-i buyer. The first two constraints are the individual rationality (IR) constraints, while the latter two are the incentive compatibility (IC) constraints. The following lemma paves the way for solving the contracting problem.

Lemma 3 *In the solution of problem (16), IR-L and IC-H are binding,* $p_H = p_H^c = \left(\frac{1}{\beta_H} - 1\right) v$*, and* $p_L \in [p_L^c, \overline{p}]$*.*

Since type-L has no incentive to mimic type-H's first-best contract, the optimal contracts entails $p_H = p_H^c$, the first-best penalty level for type-H under complete information. This is in line with the "no distortion at the top" property in the screening literature.

Proposition 3 *With* $v < 1$ *and under Assumptions 1-4, the optimal separating contract profile (i.e. the solution of Problem (16)) is the following. There exists* $\overline{\alpha} \in (0, \frac{\beta_L}{2})$ *such that*

(a) when $\alpha \geq \overline{\alpha}$*, the seller offers*

$$\begin{aligned}
p_L &= \overline{p},\ b_L = V_C(\beta_L, \overline{p}), \\
p_H &= p_H^c = (\frac{1}{\beta_L} - 1)v,\ b_H = V_C(\beta_H, p_H^c),
\end{aligned} \tag{17}$$

which are taken by type L and type H respectively;

(b) when $\alpha < \overline{\alpha}$*, the seller offers*

$$\begin{aligned}
p_L &= \widehat{p} \equiv \frac{\alpha - \beta_L v - \alpha\beta_L v + \beta_L^2 v}{(2\alpha - \beta_L)\beta_L} \in (p_L^c, \overline{p})\,,\ b_L = V_C(\beta_L, \widehat{p})\,, \\
p_H &= p_H^c = (\frac{1}{\beta_H} - 1)v,\ b_H = \widehat{b}_H \equiv V_C(\beta_H, p_H^c) - [V_C(\beta_H, \widehat{p}) - V_C(\beta_L, \widehat{p})]\,,
\end{aligned}$$

which are taken by type L and type H respectively. $\widehat{p}$ *increases in* α*. When* $\alpha \to 0$*,* $p_L \to p_L^c$*, and* $b_L \to b_L^c$*.*

Figure 7 illustrates the optimal separating contracts for $\alpha \geq \overline{\alpha}$ (left figure) and $\alpha < \overline{\alpha}$ (right figure). To prevent type-H buyers from micmicking type-L buyers, the seller can either extract less surplus from the type-H buyers, or distort type-L's penalty level. Thus, her tradeoff between extracting more surplus from type-H buyers and giving type-L buyers

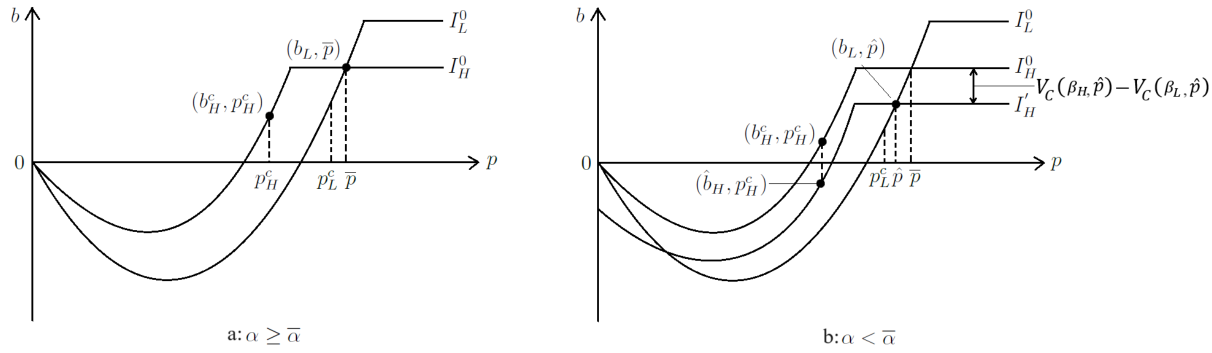

Figure 7: Optimal Separating Contracts ($v < 1$)

a more efficient penalty level determines the optimal separating contracts. When α is large enough, the proportion of type-L buyers $(1 - \alpha)$ is small, and so the efficiency loss from distorting type-L's contract is low. Thus, the extraction motives trumps the efficiency motive. Type-L's penalty level is distorted to such an extent that the seller can extract all the surplus from type-H: $p_L = \overline{p}$, a corner solution, and type-H buyers, as type-L buyers, receive no information rent (i.e. IR-H is also binding, with $p_H = p_H^c$ and $b_H = b_H^c$; Figure 7a).

When α is small, the proportion of type-L buyers becomes large, and so the efficieny loss from distorting type-L's contract is more substantial. In this case, the seller balances between the extraction motive and the efficiency motive: She still distorts p_L but to a smaller extent: $p_L = \widehat{p} \in (p_L^c, \overline{p})$, as in Figure 7b, an interior solution. Meanwhile, the seller reduces the type-H contract's price to $\widehat{b}_H$ in Figure 7b (by $[V_C(\beta_H, \widehat{p}) - V_C(\beta_L, \widehat{p})]$ relative to b_H^c) such that IC-H is satisfied (binding). The reduction in b_H leaves an information rent to type-H buyers. Moreover, when $\alpha \to 0$, because there are few type-H buyers, the efficiency motive trumps; i.e., the efficiency loss from type-L's contract is more of a concern than extracting surplus from type-H buyers. In this case, the contract for type-L buyers is no longer distorted and it converges to their first-best contract (b_L^c, p_L^c), as shown in Figure 8.

Figure 9 shows the evolution of p_L and p_H when α changes, in the optimal separating contracting profile. Type-H's penalty level p_H is fixed at p_H^c. Type-L's penalty level p_L starts with p_L^c near $\alpha = 0$, (continuously) increases as α increases (in which case type-H buyers receive an information rent), and reaches $\overline{p}$ for $\alpha \geq \overline{\alpha}$ (in which case type-H buyers receive no information rent). In all cases, each buyer type indeed pays penalty to the seller on the equilibrium path for high c, due to $p_i < \frac{1}{\beta_i} - v$ for all $i \in \{H, L\}$.

It remains to show that the optimal separating contract profile in Proposition 3 dominates pooling contracts (single contracts that attract both buyer types) and the contracts

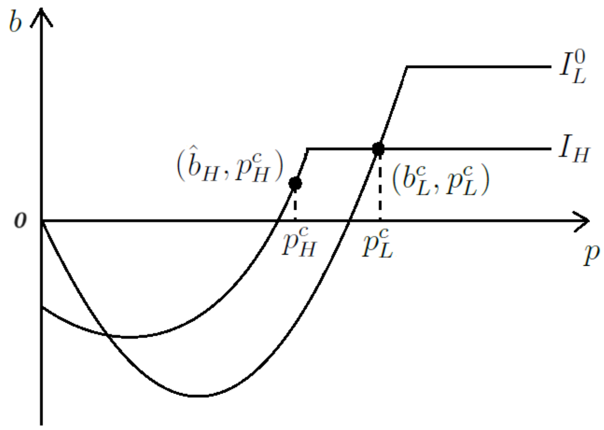

Figure 8: Optimal Separating Contracts when $\alpha \to 0$ (for $v < 1$)

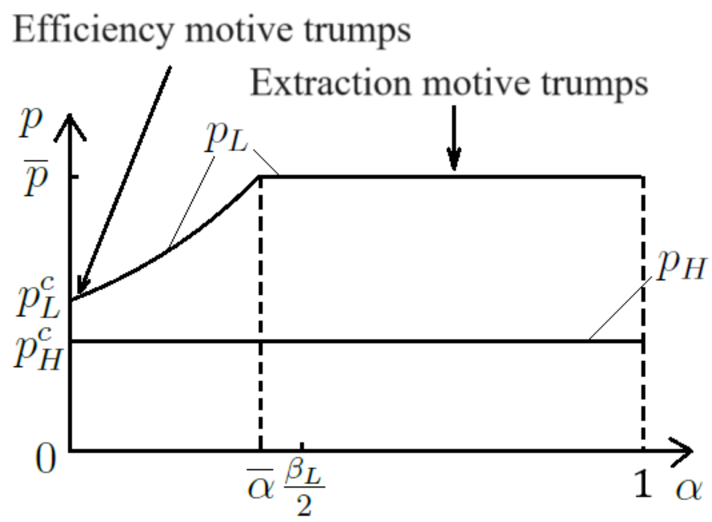

Figure 9: Relationship between p_L, p_H and α (for $v < 1$)

that only attract one buyer type. The optimal pooling contract solves

$$\begin{aligned} &\max_{p,b} \alpha\pi_H + (1-\alpha)\,\pi_L \qquad (18)\\ &s.t.\ V_C(\beta_H, p) - b \geq 0 \text{ (IR-H)}\\ &\quad\ V_C(\beta_L, p) - b \geq 0 \text{ (IR-L)} \end{aligned}$$

We establish the following two lemmas.

Lemma 4 *From the seller's viewpoint, the optimal pooling contract (i.e., the solution of problem (18)) is dominated by the optimal separating contract profile in Proposition 3.*

Lemma 5 *From the seller's viewpoint, the optimal separating contract profile dominates the case where the seller sells contracts to one buyer type only.*

The above two lemmas demonstrate that the optimal separating contract profile in Proposition 3 is indeed a separating equilibrium as defined in Definition 1, leading to the following proposition.

Proposition 4 *With $v < 1$, unobservable β and under Assumptions 1-4, the separating contract profile in Proposition 3 supports a separating equilibrium. The social welfare is lower than the complete-information case, and converges to the first-best case when $\alpha \to 0$.*

Proposition 4 demonstrates the existence of a second-best separating equilibrium for the parametric space of the gray area in Figure 6. The welfare comparison in the second half of the proposition is straightforward. In the separating equilibrium, the only inefficiency arises from type-L buyers' over-investment. Type-L buyers invest when $v < c < \beta_L (v + p_L)$, because the penalty level for type L is overly high: $p_L > p_L^c$. When $\alpha \to 0$, $p_L \to p_L^c$ (the first best), and so the welfare loss converges to 0.

To sum up, while under complete information, the optimal two-part tariff implements the first-best outcome and perfectly solves the self-control problem (Sections 4.1.1 and 4.2.1; also DellaVigna and Malmendier (2004)), the buyer's private information about his self-control ability may or may not cause inefficiency. The seller's monopolistic screening leads to a separating equilibrium in which the seller always attracts both types of buyers. With high return ($v > 1$), investment is always worthwhile, and so flexbility is not a concern; the only problem is commitment. In this case, the separating equilibrium achieves the first best, fully committing each buyer type to investment, and neither type has incentive to take the other's first-best contract. However, with lower investment return, investment is not always worthwhile, and so there is a tradeoff between flexibility and commitment in contract design. In this case, the separating equilibrium only achieves a second best, where type-L buyers' contract is distorted with an overly high penalty level so as to reduce (or even eliminate) type-H's information rent. The high penalty level for the type-L buyers, consequently, causes over-investment from these buyers.

4.2.4 Comparison with a Benevolent Seller

In reality, the principal who designs commitment contracts can not only be a profit-maximizing firm, but can also be a non-profit organization, or even the buyer's family/friend, who does not pursue economic profits but aims to improve welfare. In this subsection, we show that even if a profit-maximizing seller's contracts are the second best under incomplete information, a benevolent seller, who maximizes the social welfare, can still design contracts that achieve the first-best outcome without observing the buyer's type.

The prices b_H and b_L are pure transfers and do not enter the benevolent seller's consideration. The benevolent seller's problem is given by

$$\max_{p_H,b_H,p_L,b_L} \alpha \int_{\beta_H v}^{\min\{\beta_H(v+p_H),1\}} (v-c)dc + (1-\alpha) \int_{\beta_L v}^{\min\{\beta_L(v+p_L),1\}} (v-c)dc \tag{19}$$

$$\begin{aligned} s.t.\ \ & V_C\left(\beta_H, p_H\right) - b_H \geq 0 \text{ (IR-H)} \\ & V_C\left(\beta_L, p_L\right) - b_L \geq 0 \text{ (IR-L)} \\ & V_C\left(\beta_H, p_H\right) - b_H \geq V_C\left(\beta_H, p_L\right) - b_L \text{ (IC-H)} \\ & V_C\left(\beta_L, p_L\right) - b_L \geq V_C\left(\beta_L, p_H\right) - b_H \text{ (IC-L)} \end{aligned}$$

The seller may also have a resource constraint that the "total profit" from the two buyer types should be nonnegative:

$$\alpha\pi_H + (1-\alpha)\,\pi_L \geq 0, \tag{20}$$

meaning the seller herself does not put resources into the system.

Appendix C analyzes the benevolent seller's problem, whose decision is summarized in the following proposition.

Proposition 5 *With $v < 1$ and unobservable β, under Assumptions 1-4, the benevolent seller, with or without a resource constraint in (20), offers $p_i^s = p_i^c = (\frac{1}{\beta_i} - 1)v$ for $i \in \{H, L\}$, and proper b_L^s and b_H^s to achieve the first best, where superscript s denotes the benevolent seller (social planner).*

Propositions 4 and 5 altogether produce the following policy-relevant implication.

Corollary 1 *With $v < 1$ and unobservable β, under Assumptions 1-4, letting a benevolent seller design commitment contracts achieves higher welfare than letting a profit-maximizing seller design commitment contracts.*

With $v < 1$ and unobservable β, a profit-maximizing seller is faced with the extraction-efficiency tradeoff and designs second-best contracts. In contrast, a benevolent seller does not care about "extraction" and so is not faced with the extraction-efficiency tradeoff. Thus, she can use her degree of freedom in designing prices b_H and b_L to achieve "price discrimination" with the efficient penalty level for each buyer type.

5 Mandating Non-monetary Penalties or Penalty Transfer?

In our analysis above, the penalty p is monetary and is collected by the seller in the case of non-compliance. The commitment contract is thus akin to a two-part tariff.[11] In reality, the commitment contract may differ from this feature in two aspects.

[11]When $v < 1$, each buyer type indeed pays penalty to the seller on the equilibrium path of the separating equilibrium.

First, the penalty p can be monetary or non-monetary. stickK's customers may either lose their deposits if not meeting the goal (a monetary penalty), or have stickK send emails to their friends telling them that the customer has failed the goal (the embarrassment being a non-monetary penalty). Other examples of non-monetary penalty include shop.pavlok.com's mild electric shocks.

Second, in the case of monetary penalty, the seller may or may not collect the penalty payment herself. stickK allows customers to specify the beneficiary (a third party other than stickK) for their penalty payment, while Beeminder itself collects the penalty payment, in the case of customers' non-compliance of the goal.

The seller's collection of penalty payments may sound like taking advantage of customers' failure of achieving personal goals and, in reality, it leads to raised eyebrows because of this "exploitative nature".[12] This raises a question of whether the government should ban (or somehow discourage) the seller from collecting penalty payments from customers' failures of achieving personal goals. If the government does so, the penalty upon non-compliance should be either non-monetary (for example, the embarrassment of having friends or family informed of the failure) or is monetary but transferred to certain third parties. In eitehr case, the penalty is not part of the seller's profit. This section shows that this policy, which we call a no-profit-from-penalty (NPfP) policy, will reduce welfare.

The policy does not affect the analysis on the buyer, but affects the seller's profit. For our analysis, we maintain Assumption 1. With this assumption, we do not restrict our analysis to either $v \geq 1$ or $v < 1$.

With the NPfP policy, the seller's profit only depends on the contract's price b. Under the complete-information benchmark, the seller's problem with a buyer of type β is

$$\begin{aligned} &\max_{b,p} b \\ &s.t.\ V_C(\beta, p) - b \geq 0 \end{aligned} \tag{21}$$

which is equivalent to $\max\limits_{b,p} V_C(\beta, p)$ given the binding IR constraint. Thus, given Assumption 1, the seller offers $p^{NPfP} \geq \frac{1}{\beta} - v$, and charges $b^{NPfP} = V_C\left(\beta, p^{NPfP}\right)$, to completely extract the surplus from the buyer, where superscript $NPfP$ indicates the government's NPfP policy. The buyer fully commits to investment, even if $v < 1$.

Without the NPfP policy and under complete information, as noted in Section 4.1.1, the seller's profit is aligned with the social welfare, which is simply the change of the project value due to the commitment (Eq. (5)), so the seller sets $p \geq \frac{1}{\beta} - v$ to fully commit the buyer to investment for $v \geq 1$ (when investment is always worthwhile) but

[12]Representative opinions in online media/forums on Beeminder's taking penalty payment include: "No charity? Just two people trying to take your money?", "[S]eller intentionally design[s] contract in a way that makes you fail your commitment goal to earn your penalty money!", "[H]aving the company take the deposit creates perverse incentives, e.g. the company is basically betting on the user breaking their commitment" (sources include https://www.runnersworld.com/runners-stories/a20829762/need-more-motivation-try-behavioral-economics/, and https://news.ycombinator.com/item?id=36152956).

$p = \left(\frac{1}{\beta} - 1\right) v$ to avoid over-investment for $v < 1$. With the NPfP policy, however, the seller's profit is aligned with V_C, which, under Assumption 1, is maximized by $p \geq \frac{1}{\beta} - v$ (full commitment to investment; Figure 2). The policy prevents the seller from benefiting from the buyer's penalty payment, which induces her to set a high penalty to fully commit him to investment no matter whether $v \geq 1$ or $v < 1$; consequently, the buyer nevers incurs the penalty.

Incomplete information does not qualitatively change the result here. The seller offers contracts in (10) and achieves the same outcome as under complete information (full commitment to investment). Under these contracts, the type-H buyer has no incentive to mimic type-L, because a higher p is not beneficial (p_H^* in (10) has helped him fully commit to investment) and meanwhile by pretending to be type-L he has to pay more: $b_L^* = \int_{\beta_L v}^1 (v-c)dc > b_H^* = \int_{\beta_H v}^1 (v-c)dc$. Type-L has no incentive to mimic type-H either.[13] With the NPfP policy, incomplete information does not change the outcome nor the welfare.

We are ready to compare, with unobservable β, the social welfare with and without the NPfP policy. With the policy, given that the buyers fully commit to investment, the social welfare is given by

$$W^{NPfP} = \alpha \int_{\beta_H v}^1 (v-c)dc + (1-\alpha) \int_{\beta_L v}^1 (v-c)dc.$$

There is no penalty occurring on the equilibrium path, while the high penalty leads to over-investment for both types when $v < 1$.

Without the NPfP policy, the social welfare of the separating equilibrium, as stated in Proposition 1 or 4 depending on v, is

$$W = \begin{cases} \alpha \int_{\beta_H v}^1 (v-c)dc + (1-\alpha) \int_{\beta_L v}^1 (v-c)dc \text{ if } v \geq 1 \\ W_{v<1} \text{ if } v < 1 \end{cases},$$

where $W_{v<1}$ is defined by (28) in the proof of Proposition 4.

If $v \geq 1$ (investment is always worthwhile), the seller offers commitment contracts that realize the first-best solution, with or without the NPfP policy (Proposition 1).

If $v < 1$ (investment is not always worthwhile), neither scenario achieves the first best, due to the over-investment problem. With the NPfP policy, both buyer types fully commit to investment. Without the policy, the penalty level for type-H buyers is efficient and so type-H buyers do not over-invest; however, there is over-investment for type-L buyers: The threshold below which type-L buyers invest is $c = \beta_L (v + \hat{p})$ or $c = \beta_L (v + \hat{p})$, depending on the value of α (Proposition 3). Using $W_{v<1}$ in (28) and observing that

$$v < \beta_L (v + \hat{p}) < \beta_L (v + \overline{p}) < 1,$$

we get the following proposition.

[13] As shown in the proof of Proposition 1 (3rd para.), which does not rely on $v > 1$.

Proposition 6 *Under incomplete information, when $v \geq 1$, $W^{NPfP} = W$, which achieves the first best. When $v < 1$, $W^{NPfP} < W$. The no-profit-from-penalty (NPfP) policy reduces welfare.*

Without the NPfP policy, over-investment only occurs to type-L buyers and is mild, arising from the seller's extraction-efficiency tradeoff. With the NPfP policy, however, because each buyer type is induced to fully commit to investment, the over-investment problem is broader (occurring to both types) and more severe. Indeed, the "exploitative feature" of the contract leads to a higher social welfare. Proposition 6 thus provides policy implications for the emerging market of self-control contracts: Mandating non-monetary penalties or transfer of penalty payments to third partities is welfare-decreasing.

6 Concluding Remarks

A self-control market has arisen in recent years with the escalating challenges of self-control problems in modern society. This emerging market has some notable features distinct from traditional markets. First, unlike traditional markets that sell regular products or services which time-consistent consumers also demand, the emerging self-control market provides purely self-commitment devices only to help time-inconsistent customers achieve their goals, while the attainment of these goals is not directly tied to the seller's engagement. Second, unlike in traditional markets where products or services are more standardized, the self-control market has high demand for personalization and customization. Individuals have varying, usually unobservable, self-control needs, and a myriad of personal goals. Platforms like stickK.com and beeminder.com allow users to set their own goals and provide commitment contracts with varying commitment powers to help users meet those customized goals.

This paper is the first to formally analyze the emerging market for self-commitment devices. It investigates a monopolist seller's contractual problem with a buyer who demands self-control in an investment project. The seller may or may not observe the buyer's self-control deficiency. Our model identifies the optimal commitment contracts that increase the buyer's cost of non-investment. We find that the first-best outcome can be achieved when the investment return is high, even if the buyer's self-control ability is his private information. When the investment return is low, however, a profit-maximizing seller's second-best contracts induce a separating equilibrium, where some buyers over-invest. The over-investment problem is milder when the profit-maximizing seller collects monetary penalty upon the buyer's non-compliance, compared to the case where the penalty is non-monetary or is monetary but transferred to some third parties. In contrast, a non-profit-driven principal, even with incomplete information, designs first-best contracts.

The behavioral industrial organization literature focuses on traditional markets, where consumers with self-control problems consume conventional products or services. Meanwhile, a broader literature has well-documented the demand for and the effectiveness of commitment devices for individuals. The emergence of the self-control markets represents a synergy between market mechanisms and individuals' self-commitment demands. While the existent theoretical literature has yet to formally analyze this emerging market, our paper represents a first endeavor that explores this promising market, offering new insights on the understanding of the market's potentials and boundaries. It also generates significant implications for the formulation of public policy relating to the emerging market, the regulation of which is still in embryo.

Our model assumes that consumers in the self-control market are perfectly aware of their self-control deficiencies. In reality, individuals might lack such perfect awareness. Introducing partial naivety into the analysis complicates it, but may yield novel insights, meriting further investigation in the future. Another potential direction worth exploring in the future is to extend the analysis of this paper to alternative market structures, such as competitive settings.

A Proofs of Propositions and Lemmas

In the proofs, we use $u_i(b,p) \equiv V_C(\beta_i, p) - b$ for $i \in \{H, L\}$ to denote type-i buyer's payoff from contract (b,p), for exposition purpose.

A.1 Proof of Lemma 1

When $\beta \in \left(0, \frac{1}{v+p}\right]$, $\beta(v+p) \leq 1$. We have

$$V_C = \int_{\beta v}^{\beta(v+p)} (v-c)dc - p\int_{\beta(v+p)}^{1} dc = \left(1 - \frac{\beta}{2}\right)\beta\left[(v+p)^2 - v^2\right] - p. \tag{22}$$

Hence, $\frac{\partial V_C(\beta,p)}{\partial \beta} > 0$ since $\left(1 - \frac{\beta}{2}\right)\beta$ increases in β for $\beta \in (0,1)$.

When $\beta \in (\frac{1}{v+p}, \frac{1}{v})$, then $\beta v < 1 < \beta(v+p)$. We have

$$V_C(\beta, p) = \int_{\beta v}^{1} (v-c)dc = v - \frac{1}{2} - \beta v^2 + \frac{1}{2}\beta^2 v^2.$$

Hence, $\frac{\partial V_C(\beta,p)}{\partial \beta} = (\beta - 1)v^2 < 0$.

A.2 Proof of Lemma 2

By (2) and (22), we have

$$V_C = \begin{cases} \int_{\beta v}^{1} (v-c)dc \text{ if } p \geq \frac{1}{\beta} - v \\ \left(1 - \frac{\beta}{2}\right)\beta\left[(v+p)^2 - v^2\right] - p \text{ if } 0 < p < \frac{1}{\beta} - v \end{cases}. \tag{23}$$

For $0 < p < \frac{1}{\beta} - v$,

$$\frac{\partial V_C}{\partial p} = (2-\beta)\beta(p+v) - 1, \tag{24}$$

where $0 \leq (2-\beta)\beta \leq 1$ given $0 \leq \beta \leq 1$.

If $v \geq \frac{1}{(2-\beta)\beta} \geq 1$, then $\frac{\partial V_C}{\partial p} > 0$ for all $0 < p < \frac{1}{\beta} - v$. If $v < \frac{1}{(2-\beta)\beta}$, then $\frac{\partial V_C}{\partial p} < 0$ for small enough p, and $\frac{\partial V_C}{\partial p} \to 1 - \beta > 0$ when p converges to $\frac{1}{\beta} - v$ from below. Also note that $V_C(0) = 0$ and V_C is continuous for $p \geq 0$. Thus, if $v \geq \frac{1}{(2-\beta)\beta}$, V_C starts from 0, increases in p for all $0 \leq p \leq \frac{1}{\beta} - v$, and then remains constant at $\int_{\beta v}^{1}(v-c)dc$ for $p \geq \frac{1}{\beta} - v$. If $v < \frac{1}{(2-\beta)\beta}$, then V_C starts from 0, decreases first and then increases in p, for $0 \leq p \leq \frac{1}{\beta} - v$, and then remains constant at $\int_{\beta v}^{1}(v-c)dc$ for $p \geq \frac{1}{\beta} - v$.

A.3 Proof of Proposition 1

The penalty levels p_L^* and p_H^* are so high that each type of buyers fully commits to investment. The contract generates the same outcome as the first-best solutions in the complete information benchmark and extracts all the surplus from the buyers, and so is optimal for the seller given both buyer types take the contract designed for him. It remains to show that neither buyer type has incentive to deviate.

Type-H's incentive to deviate to (b_L^*, p_L^*) is

$$\begin{aligned}
&u_H(b_L^*, p_L^*) - u_H(b_H^*, p_H^*) \\
&= [V_C(\beta_H, p_L^*) - b_L^*] - [V_C(\beta_H, p_H^*) - b_H^*] \\
&= [V_C(\beta_H, p_L^*) - V_C(\beta_H, p_H^*)] - [b_L^* - b_H^*] \\
&= [V_C(\beta_H, p_L^*) - V_C(\beta_H, p_H^*)] - \left[\left(\frac{\beta_L}{2} - 1\right)\beta_L - \left(\frac{\beta_H}{2} - 1\right)\beta_H\right] v^2 < 0,
\end{aligned}$$

where the inequality holds because on the last line the first term $V_C(\beta_H, p_L^*) - V_C(\beta_H, p_H^*) = 0$ (by Lemma 2, $V_C(\beta_H, p)$ is constant for all $p \geq p_H^* = \frac{1}{\beta_H} - v$) and the second term in the brackets is positive (note that $\left(\frac{\beta}{2} - 1\right)\beta$ is decreasing in $\beta \in (0,1)$).

Type-L's incentive to deviate to (b_H^*, p_H^*) is

$$u_L(b_H^*, p_H^*) - u_L(b_L^*, p_L^*) = V_C(\beta_L, p_H^*) - b_H^* = V_C(\beta_L, p_H^*) - V_C(\beta_H, p_H^*) \leq 0,$$

where the first equality uses $u_L(b_L^*, p_L^*) = 0$, the second uses $b_H^* = V_C(\beta_H, p_H^*)$, and the inequality is due to (9) given $p_H^* \leq \overline{p}$.

Hence, no type has incentive to deviate and so (b_L^*, p_L^*) and (b_H^*, p_H^*) support a separating equilibrium.

A.4 Proof of Proposition 2

Firstly it can be shown that type-L never has incentive to deviate to (b_H^c, p_H^c). If type-L accepts (b_L^c, p_L^c), the payoff is 0. Showing that type-L has no incentive to deviate is equivalent to showing that his payoff from (b_H^c, p_H^c) is negative:[14]

$$\int_{\beta_L v}^{\frac{\beta_L}{\beta_H} v} (v - c)\, dc - p_H^c \int_{\frac{\beta_L}{\beta_H} v}^{1} dc - b_H^c < 0,$$

[14] With $p_H^c = \left(\frac{1}{\beta_H} - 1\right) v$, type-L invests if and only $c \leq \beta_L (v + p_H) \Leftrightarrow c \leq \frac{\beta_L}{\beta_H} v$.

We have

$$
\begin{aligned}
&\int_{\beta_L v}^{\frac{\beta_L}{\beta_H} v} (v-c)\, dc - p_H^c \int_{\frac{\beta_L}{\beta_H} v}^{1} dc - b_H^c \\
&= \int_{\beta_L v}^{\frac{\beta_L}{\beta_H} v} (v-c)\, dc - \left(\frac{1}{\beta_H} - 1\right) v \int_{\frac{\beta_L}{\beta_H} v}^{1} dc - \int_{\beta_H v}^{v} (v-c)\, dc + \left(\frac{1}{\beta_H} - 1\right) v \int_{v}^{1} dc \\
&= \int_{\beta_L v}^{\beta_H v} (v-c)\, dc - \int_{\frac{\beta_L}{\beta_H} v}^{v} \left(\frac{v}{\beta_H} - c\right) dc \\
&= \left(\frac{1-\beta_H^2}{\beta_H^2}\right) v^2 \left[\beta_L \left(1 - \frac{\beta_L}{2}\right) - \beta_H \left(1 - \frac{\beta_H}{2}\right)\right] < 0,
\end{aligned}
$$

where the inequality is due to $\beta\left(1-\frac{\beta}{2}\right)$ being increasing in β.

Next we characterize the conditions under which the type-H has incentive to deviate to (b_L^c, p_L^c). Type-H's incentive to deviate to (b_L^c, p_L^c) is

$$
u_H\left(b_L^c, p_L^c\right) - u_H\left(b_H^c, p_H^c\right) = u_H\left(b_L^c, p_L^c\right) = V_C(\beta_H, p_L^c) - b_L^c = V_C(\beta_H, p_L^c) - V_C(\beta_L, p_L^c) \le 0
$$

if and only if $p_L^c \ge \overline{p}$ due to the definition of $\overline{p}$ and (9).

Explicitly, type-H's incentive to deviate to (b_L^c, p_L^c) can be written as

$$
\begin{aligned}
u_H\left(b_L^c, p_L^c\right) &= \int_{\beta_H v}^{\min\{\frac{\beta_H}{\beta_L} v, 1\}} (v-c)\, dc - p_L^c \int_{\min\{\frac{\beta_H}{\beta_L} v, 1\}}^{1} dc - b_L^c \\
&= \int_{\beta_H v}^{\min\{\frac{\beta_H}{\beta_L} v, 1\}} (v-c)\, dc - \left(\frac{1}{\beta_L} - 1\right) v \int_{\min\{\frac{\beta_H}{\beta_L} v, 1\}}^{1} dc - \int_{\beta_L v}^{v} (v-c)\, dc + \left(\frac{1}{\beta_L} - 1\right) v \int_{v}^{1} dc.
\end{aligned}
$$

Suppose $v < \frac{\beta_L}{\beta_H}$, implying $\frac{\beta_H}{\beta_L} v < 1$. Then

$$
\begin{aligned}
u_H\left(b_L^c, p_L^c\right) &= \int_{\beta_H v}^{\frac{\beta_H}{\beta_L} v} (v-c)\, dc - \left(\frac{1}{\beta_L} - 1\right) v \int_{\frac{\beta_H}{\beta_L} v}^{1} dc - \int_{\beta_L v}^{v} (v-c)\, dc + \left(\frac{1}{\beta_L} - 1\right) v \int_{v}^{1} dc \\
&= \frac{(\beta_H - \beta_L)(1-\beta_L)(1+\beta_L)(2-\beta_H-\beta_L)\, v^2}{2\beta_L^2} > 0,
\end{aligned}
$$

implying type-H has incentive to mimic type-L.

Suppose $\frac{\beta_L}{\beta_H} \le v < 1$, implying $\frac{\beta_H}{\beta_L} v \ge 1$. Then

$$
\begin{aligned}
u_H\left(b_L^c, p_L^c\right) &= \int_{\beta_H v}^{1} (v-c)\, dc - \int_{\beta_L v}^{v} (v-c)\, dc + \left(\frac{1}{\beta_L} - 1\right) v \int_{v}^{1} dc \\
&= \frac{1}{2}(1-v)\left(\frac{2v}{\beta_L} - 1 - v\right) - \frac{1}{2} v^2 (\beta_H - \beta_L)(2 - \beta_H - \beta_L).
\end{aligned}
$$

Therefore, (13) and $\frac{\beta_L}{\beta_H} \le v$ are necessary and sufficent for $u_H\left(b_L^c, p_L^c\right) \le 0$.

A.5 Proof of Lemma 3

We organize the proof in several steps.

Step 1. In the solution of (16), IR-L is binding.

First, in the solution of (16), at least one type's IR constraint is binding. To see this, note that if no type's IR constraint is binding, the seller can always increase both b_L and b_H by some small amounts to earn more profit while keeping the four constraints still satisfied.

Next, suppose IR-L is not binding (i.e. $u_L(p_L, b_L) > 0$). Then, IR-H must be binding. The solution of (16) must have $p_H = p_H^c$ and $b_H = b_H^c$. To see this, note a) that (p_H^c, b_H^c) maximizes the profit from type-H; and b) that it satisfies IR-H, IC-L[15] and the satisfaction of the other two constraints is not affected by the choice of (p_H, b_H) given that IR-H is binding. However, the contract profile with $p_H = p_H^c$, $b_H = b_H^c$, and $u_L(p_L, b_L) > 0$ is not optimal, because by slightly increasing b_L, the seller can earn a higher profit from type-L while keeping the 4 constraints still satisified (since $u_L(p_L, b_L) > 0 > u_L(p_H^c, b_H^c)$).

Step 2. In the solution of (16), $p_L \leq \overline{p}$.

Suppose $p_L > \overline{p}$. First, moving type-L's contract from $p_L > \overline{p}$ to $p_L = \overline{p}$ along type-L's zero-payoff indifference curve (given by the binding IR-L) does not affect the satisfaction of the constraints: type-H buyers have no incentive to deviate to any $p_L \geq \overline{p}$ on type-L's zero-payoff indifference curve and so IC-H is satisfied. It does not affect IC-L given the binding IR-L. Of course, it does not affect IR-H.

Second, the change increases the seller's profit from type-L:

$$\begin{aligned}
\pi_L &= p_L \int_{\min\{\beta_L(v+p_L),1\}}^{1} dc + b_L \\
&= p_L \int_{\min\{\beta_L(v+p_L),1\}}^{1} dc + V_C(\beta_L, p_L) \\
&= p_L \int_{\min\{\beta_L(v+p_L),1\}}^{1} dc + \int_{\beta_L v}^{\min\{\beta_L(v+p_L),1\}} (v-c)dc - p_L \int_{\min\{\beta_L(v+p),1\}}^{1} dc \\
&= \int_{\beta_L v}^{\min\{\beta_L(v+p_L),1\}} (v-c)dc.
\end{aligned}$$

We have

$$\frac{d\pi_L}{dp_L} = \begin{cases} \beta_L[v - \beta_L(v+p)] > 0 \text{ when } p < p_L^c = (\frac{1}{\beta_L} - 1)v \\ 0 \text{ when } p = p_L^c = (\frac{1}{\beta_L} - 1)v \\ \beta_L[v - \beta_L(v+p)] < 0 \text{ when } (\frac{1}{\beta_L} - 1)v = p_L^c < p \leq \frac{1}{\beta_L} - v \\ 0 \text{ when } p > \frac{1}{\beta_L} - v \end{cases}.$$

Since $p_L^c < \overline{p} < \frac{1}{\beta_L} - v$ (by (8) and Assumption 2), the change from $p_L > \overline{p}$ to $p_L = \overline{p}$ increases π_L.

Step 3. In the solution of (16), IC-H is binding.

Suppose not. First we show that IR-H is not binding if IC-H is not binding. If IR-H is binding and IC-H is not binding, given that IR-L is also binding (Step 1), the only possibility is that $p_L > \overline{p}$, a contradiction with Step 2. So IR-H is not binding. Then the seller can increase profit by slightly increasing b_H which still satifies all the 4 contraints: IR-H and IC-H were not binding, and so are still satisfied with the slighltly larger b_H; IC-L is still satisfied because type-H's contract becomes less attractive, and the IR-L is not affected.

Step 4. In the solution of (16), if $p_L = p_L^c = (\frac{1}{\beta_L} - 1)v$, then $p_H = p_H^c = (\frac{1}{\beta_H} - 1)v$.

Suppose the seller offers the type-L buyers $p_L = p_L^c$. Then by the binding IR-L, $b_L = b_L^c = V_C(\beta_L, b_L^c)$. By the binding IC-H, the seller offers type-H p_H and b_H such that

$$\begin{aligned}
& V_C(\beta_H, p_H) - b_H = V_C(\beta_H, p_L^c) - b_L^c = V_C(\beta_H, p_L^c) - V_C(\beta_L, b_L^c) > 0 \\
\Leftrightarrow\ & b_H = V_C(\beta_H, p_H) - V_C(\beta_H, p_L^c) + V_C(\beta_L, p_L^c),
\end{aligned}$$

[15] $u_L(p_L, b_L) > 0 > u_L(p_H^c, b_H^c)$, where the second inequality is proved in the proof of Proposition 2.

where the inequality is from the assumption that type-H has incentive to mimic type-L if the two types are given the first-best contracts under complete information (Assumption 2). Due to the inequality, the IC-H gurantees that IR-H is satisfied. The seller's problem is thus

$$\max_{p_H} p_H \int_{\min\{\beta_H(v+p_H),1\}}^{1} dc + V_C\left(\beta_H, p_H\right) - V_C\left(\beta_H, p_L^c\right) + V_C(\beta_L, p_L^c)\ .$$

Because $-V_C\left(\beta_H, p_L^c\right) + V_C(\beta_L, p_L^c)$ is independent of p_H, the problem is equivalent to

$$\begin{aligned}
&\max_{p_H} p_H \int_{\min\{\beta_H(v+p_H),1\}}^{1} dc + V_C\left(\beta_H, p_H\right) \\
\Leftrightarrow &\max_{p_H} p_H \int_{\min\{\beta_H(v+p_H),1\}}^{1} dc + \int_{\beta_H v}^{\min\{\beta_H(v+p_H),1\}} (v-c)\, dc - p_H \int_{\min\{\beta_H(v+p_H),1\}}^{1} dc \\
\Leftrightarrow &\max_{p_H} \int_{\beta_H v}^{\min\{\beta_H(v+p_H),1\}} (v-c)\, dc,
\end{aligned}$$

the solution of which is $p_H = p_H^c = (\frac{1}{\beta_H} - 1)v$, implying

$$b_H = V_C\left(\beta_H, p_H^c\right) - V_C\left(\beta_H, p_L^c\right) + V_C(\beta_L, p_L^c) \equiv \widetilde{b}_H.$$

Finally, it remains to show that type-L buyers has no incentive to mimic. Given the binding IR-L, it is equivalent to show that type-L's payoff from taking $\left(p_H^c, \widetilde{b_H}\right)$ is negative:

$$V_C\left(\beta_L, p_H^c\right) - \widetilde{b}_H = V_C\left(\beta_L, p_H^c\right) - V_C\left(\beta_H, p_H^c\right) + V_C\left(\beta_H, p_L^c\right) - V_C(\beta_L, p_L^c) < 0,$$

which is satisfied given Assumption 4.

Step 5. In the solution of (16), $p_L \geq p_L^c = \left(\frac{1}{\beta_L} - 1\right) v$.

It is sufficient to show that choosing any $p_L < p_L^c$ (while holding the binding IR-L) gives the seller a lower profit than $p_L = p_L^c$. First, the profit from type-L is lower because p_L deviates from the optimal p_L^c. Second, the profit from type-H is lower with $p_L < p_L^c$. With $p_L = p_L^c$, the seller's profit from type-H is maximized by choosing $p_H = p_H^c$ (Step 4). Compared to $p_L = p_L^c$, with $p_L < p_L^c$, type-H's indifference curve moves to a lower one (by the binding IC-H), and the profit-maximizing contract on this indifference curve with $p_H = p_H^c$ may or may not satisfy the IC-L constraint. In the former case, the profit from type-H is lower because of a lower b_H; in the latter case, the seller has to move to a suboptimal $p_H \neq p_H^c$, which further reduces the profit from type-H. In either case, the profit from type-H is lower. And so, overall, the profit is lower with $p_L < p_L^c$ than with $p_L = p_L^c$.

Step 6. In the solution of (16), $p_H = p_H^c = \left(\frac{1}{\beta_H} - 1\right) v$.

Given type-L's contract with $p_L \in [p_L^c, \overline{p}]$, we can use the binding IC-H to identify type-H's indifference curve on which type-H's contract is located. Given this indifference curve, the seller maximizes her profit from type-H by choosing $p_H = \left(\frac{1}{\beta_H} - 1\right) v$, temporarily ignoring the constraints. This contract for type-H has the same p_H as – but (weakly) higher b_H than – the contract $\left(p_H^c, \widetilde{b}_H\right)$ in the proof of Step 4 given $p_L \in [p_L^c, \overline{p}]$. The proof of Step 4 shows that type-L with binding IR-L has no incentive to deviate to $\left(p_H^c, \widetilde{b}_H\right)$, and thus has no incentive to deviate to such a contract. So the constraint IC-L is satisfied. The other constraints are also satisfied.

A.6 Proof of Proposition 3

Lemma 3 shows binding IR-L and IC-H, $p_H = p_H^c$, and $p_L \in [p_L^c, \overline{p}]$ in the solution of the contracting problem. Given $p_L \in [p_L^c, \overline{p}]$, b_L is given by the binding IR-L: $b_L = V_C(\beta_L, p_L)$; b_H is then determined by the binding IC-H and given by (25) below. The binding IC-H determines a type-H buyer's payoff

$$V_C(\beta_H, p_H) - b_H = V_C(\beta_H, p_L) - V_C(\beta_L, p_L) \geq 0 \text{ for } p_L \in [p_L^c, \overline{p}],$$

showing that IR-H is satisfied. By the proof of Step 6 in Lemma 3's proof, IC-L is also satisfied. The optimal contracts are characterized by

$$\begin{aligned} & p_L \in [p_L^c, \overline{p}], \, b_L = V_C(\beta_L, p_L); \\ & p_H = p_H^c = (\frac{1}{\beta_L} - 1)v, \, b_H = V_C(\beta_H, p_H^c) - [V_C(\beta_H, p_L) - V_C(\beta_L, p_L)]. \end{aligned} \tag{25}$$

Given $p_L \leq \overline{p} < \frac{1}{\beta_L} - v$, the seller's profit as a function of p_L:

$$\begin{aligned} \pi(p_L) &= \alpha \left[p_H^c \int_v^1 dc + b_H \right] + (1-\alpha) \left[p_L \int_{\beta_L(v+p_L)}^1 dc + V_C(\beta_L, p_L) \right] \\ &= \alpha \left[p_H^c \int_v^1 dc + V_C(\beta_H, p_H^c) - V_C(\beta_H, p_L) + V_C(\beta_L, p_L) \right] + (1-\alpha) \left[p_L \int_{\beta_L(v+p_L)}^1 dc + V_C(\beta_L, p_L) \right] \\ &= \alpha \left[\int_{\beta_H v}^v (v-c)\, dc - V_C(\beta_H, p_L) \right] + (1-\alpha) p_L \int_{\beta_L(v+p_L)}^1 dc + V_C(\beta_L, p_L). \end{aligned}$$

We have

$$\begin{aligned} \frac{\partial \pi(p_L)}{\partial p_L} &= -\alpha \frac{\partial V_C(\beta_H, p_L)}{\partial p_L} + (1-\alpha)(1 - \beta_L v - 2\beta_L p_L) + \frac{\partial V_C(\beta_L, p_L)}{\partial p_L} \\ &= 0 + (1-\alpha)(1 - \beta_L v - 2\beta_L p_L) + [(2-\beta_L)\beta_L(p_L + v) - 1] \\ &= \left(2\alpha\beta_L - \beta_L^2\right) p_L - \left(\alpha - \beta_L v - \alpha\beta_L v + \beta_L^2 v\right), \end{aligned}$$

where in the second line $\frac{\partial V_C(\beta_H, p_L)}{\partial p_L} = 0$ is due to $p_L \geq p_L^c \geq \frac{1}{\beta_H} - v$ (Assumption 3), and $\frac{\partial V_C(\beta_L, p_L)}{\partial p_L} = [(2-\beta_L)\beta_L(p_L + v) - 1]$ is by (24). The interior solution, if any,

$$\widehat{p} \equiv \frac{\alpha - \beta_L v - \alpha\beta_L v + \beta_L^2 v}{(2\alpha - \beta_L)\beta_L}.$$

The second-order derivative:

$$\frac{\partial^2 \pi(p_L)}{\partial p_L^2} = 2\alpha\beta_L - \beta_L^2.$$

When

$$2\alpha\beta_L - \beta_L^2 = 0 \Leftrightarrow \alpha = \frac{\beta_L}{2},$$

we have

$$\frac{\partial \pi(p_L)}{\partial p_L} = -\left(\alpha - \beta_L v - \alpha\beta_L v + \beta_L^2 v\right) = -\frac{\beta_L}{2}[1 - (2-\beta_L)v] > 0$$

where the inequality is by (3). Then the profit-maximizing $p_L = \overline{p}$.

When

$$2\alpha\beta_L - \beta_L^2 > 0 \Leftrightarrow \alpha > \frac{\beta_L}{2},$$

$\pi(p_L)$ is strictly convex, and so the profit-maximizing p_L is a corner solution. Since

$$\widehat{p} - p_L^c = \widehat{p} - (\frac{1}{\beta_L} - 1)v = \frac{\alpha[1 - (2 - \beta_L)v]}{\beta_L(2\alpha - \beta_L)} < 0$$

given $\alpha > \frac{\beta_L}{2}$ and (3), the profit-maximizing $p_L = \overline{p}$ given $p_L \in [p_L^c, \overline{p}]$.

When

$$2\alpha\beta_L - \beta_L^2 < 0 \Leftrightarrow \alpha < \frac{\beta_L}{2},$$

$\pi(p_L)$ is strictly concave. We have

$$\widehat{p} - p_L^c = \widehat{p} - (\frac{1}{\beta_L} - 1)v = \frac{\alpha[1 - (2 - \beta_L)v]}{\beta_L(2\alpha - \beta_L)} > 0$$

where the denominator is negative given $\beta_L > 2\alpha$, and the numerator is also negative due to (3). Next we compare the values of $\widehat{p}$ and $\overline{p}$. First, since

$$\frac{\partial \widehat{p}}{\partial \alpha} = \frac{(2 - \beta_L)v - 1}{(\beta_L - 2\alpha)^2} > 0$$

given (3), $\widehat{p}$ is increasing in α for $\alpha \in (0, \frac{\beta_L}{2})$. We have

$$\lim_{\alpha \to 0} \widehat{p} = \frac{-\beta_L v + \beta_L^2 v}{-\beta_L^2} = \left(\frac{1}{\beta_L} - 1\right) v = p_L^c < \overline{p}, \tag{26}$$

and

$$\lim_{\alpha \to \frac{\beta_L}{2}} \widehat{p} = \infty.$$

Therefore, there exists $\overline{\alpha} \in (0, \frac{\beta_L}{2})$ such that $\widehat{p}(\overline{\alpha}) \equiv \overline{p}$. When $\alpha \in \left[\overline{\alpha}, \frac{\beta_L}{2}\right)$, $\widehat{p} \geq \overline{p}$, the profit-maximizing $p_L = \overline{p}$; when $\alpha \in (0, \overline{\alpha})$, $\widehat{p} < \overline{p}$, the profit-maximizing $p_L = \widehat{p}$.

To summarize, when $\alpha \geq \overline{\alpha}$, the profit-maximizing $p_L = \overline{p}$; when $\alpha < \overline{\alpha}$, the profit-maximizing $p_L = \widehat{p}$. In either case, inserting the profit-maximizing p_L into (25) gives the optimal separating contracts in the proposition, and the two contracts (b_L, p_L) and (b_H, p_H) are taken by type L and type H respectively.

When $\alpha \to 0$, (26) shows $p_L \to p_L^c$, and so $b_L \to b_L^c$.

A.7 Proof of Lemma 4

The only difference between problem (18) and problem (16) is that problem (18) imposes additional constraints of $p_H = p_L$ and $b_H = b_L$, and so the profit from problem (18) is weakly lower than the profit from problem (16). Moreover, given the uniqueness of the solution of problem (16) as in Proposition 3, in which $p_H \neq p_L$ and $b_H \neq b_L$, the profit from the solution of problem (16) is strictly higher.

A.8 Proof of Lemma 5

Given α, the profit from the optimal separating contract profile is at least as large as the profit from the contract profile in (17), denoted by $\pi(\overline{p}|\alpha)$ (where both IR-L and IR-H are binding):

$$\pi(\overline{p}) = \alpha \underbrace{\int_{\beta_H v}^{v} (v - c)dc}_{>0} + (1 - \alpha) \underbrace{\int_{\beta_L v}^{\beta_L(v + \overline{p})} (v - c)dc}_{>0}, \tag{27}$$

where $\int_{\beta_L v}^{\beta_L(v+\overline{p})}(v-c)dc > \int_{\beta_L v}^{1}(v-c)dc > 0$ due to $v < \beta_L(v+\overline{p}) < 1$. It suffices to show that for any α, $\pi(\overline{p})$ is larger than the profit when the seller only sells contracts to one-type buyers.

Suppose the seller offers one single contract to attract type-H but not type-L. Then the best she could offer is type-H's first-best complete-information contract (b_H^c, p_H^c), giving her profit $\alpha\int_{\beta_H v}^{v}(v-c)dc$, which is lower than (27).

Suppose the seller, instead, offers one single contract to attract type-L but not type-H. Then the best she could do is to offer a contract that lies on type-L's zero-payoff indifference curve, with p_L higher than (so that type-H wouldn't take it) but as close to $\overline{p}$ as possible (to maximize the profit from type-L; see the proof of Step 2 in Lemma 3's proof). The profit converges to, but is no higher than, $\int_{\beta_L v}^{\beta_L(v+\overline{p})}(v-c)dc$ and so is lower than (27).

A.9 Proof of Proposition 4

Under complete information, the social welfare is given by

$$W_{v<1}^c = \alpha\int_{\beta_H v}^{v}(v-c)dc + (1-\alpha)\int_{\beta_L v}^{v}(v-c)dc,$$

the first best, where the subscript $v<1$ indicates the case of $v<1$.

The social welfare of the separating equilibrium is

$$W_{v<1} = \begin{cases} W(\overline{p}) \equiv \alpha\int_{\beta_H v}^{v}(v-c)dc + (1-\alpha)\int_{\beta_L v}^{\beta_L(v+\overline{p})}(v-c)dc \text{ when } \alpha \geq \overline{\alpha} \\ W(\widehat{p}) \equiv \alpha\int_{\beta_H v}^{v}(v-c)dc + (1-\alpha)\int_{\beta_L v}^{\beta_L(v+\hat{p})}(v-c)dc \text{ when } \alpha < \overline{\alpha} \end{cases}. \tag{28}$$

Since $\beta_L(v+\overline{p}) > \beta_L(v+\hat{p}) > v$, $W(\overline{p}) < W(\widehat{p}) < W_{v<1}^c$, and thus $W_{v<1} < W_{v<1}^c$.

Moreover, since $\lim_{\alpha\to 0}\widehat{p} = p_L^c$, $\lim_{\alpha\to 0}\beta_L(v+\hat{p}) = v$, and so $\lim_{\alpha\to 0}W_{v<1} = W_{v<1}^c$.

B Ruling-out of Double-crossing Cases

Assumptions 3 and 4 rule out some (locally) double-crossing cases. Figure 10a shows the double crossing (at points A and B) when Assumption 3 is violated. There, type-L's first-best penalty level is smaller than type-H's "bliss point" penalty level: $p_L^c = \left(\frac{1}{\beta_L} - 1\right)v < \frac{1}{\beta_H} - v$. Figure 10b shows the double-crossing case when Assumption 4 is violated. Contract (b_H, p_H^c) (point C) and contract (b_L^c, p_L^c) (point A) are on the same indifference curve of type H (Equation (14)); however, type L likes point C (weakly) better than the zero-payoff contract at point A (i.e., (15) is violated). There are double crossings on the two types' indifference curves: points A and B.

C Benevolent Seller with Incomplete Information

To solve problem (19), the seller gives the two buyer types their own first-best punishment level, p_H^c and p_L^c, respectively, to achieve the first best. For type-L, the seller sets

$$b_L^s \leq V_C(\beta_L, p_L^c) \tag{29}$$

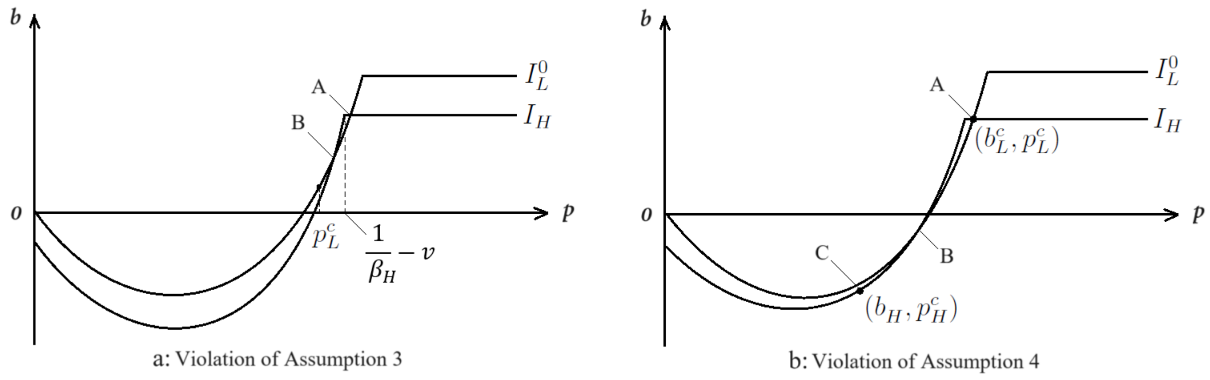

Figure 10: Violation of Assumption 3 or 4

to satisfy the IR-L, where superscript s denotes the benevolent seller (a social planner). For type-H, the seller sets

$$V_C(\beta_L, p_H^c) - V_C(\beta_L, p_L^c) + b_L^s \leq b_H^s \leq V_C(\beta_H, p_H^c) - V_C(\beta_H, p_L^c) + b_L^s \tag{30}$$

to safisfy IC-H and IC-L. By Assumption 4, the range of b_H^s in (30) is non-empty. The IR-H is also satisfied because the IC-H and the IR-L are satisfied:

$$V_C(\beta_H, p_H^c) - b_H^s \geq V_C(\beta_H, p_L^c) - b_L^s > V_C(\beta_L, p_L^c) - b_L^s \geq 0,$$

where the second inequality uses $V_C(\beta_H, p_L^c) > V_C(\beta_L, p_L^c)$.[16]

Thus, the contracts (b_H^s, p_H^c) and (b_L^s, p_L^c) characterized above solve problem (19) and achieve the first best.

To satisfy the resource constraint (20), the seller can, among other feasible configurations, let $b_L^s = V_C(\beta_L, p_L^c)$, and

$$b_H^s = V_C(\beta_H, p_H^c) - V_C(\beta_H, p_L^c) + b_L^s = V_C(\beta_H, p_H^c) - V_C(\beta_H, p_L^c) + V_C(\beta_L, p_L^c).$$

Such a configuration satisfies (29) and (30) and so satisfies the constraints in problem (19) and still achieves the first best (with the first-best penalty levels for both types). It remains to show that it satisfies (20). Under the contracts (p_H^c, b_H^s) and (p_L^c, b_L^s), $\pi_L > 0$ because IR-L is binding and p_L^c is the first best. Thus it remains to show that $\pi_H > 0$. We have

$$\begin{aligned}\pi_H &= b_H^s + p_H^c \int_v^1 dc \\ &= V_C(\beta_H, p_H^c) - V_C(\beta_H, p_L^c) + V_C(\beta_L, p_L^c) + p_H^c \int_v^1 dc \\ &= \int_{\beta_H v}^{v} (v - c)\, dc - \int_{\beta_H v}^{1} (v - c)\, dc + V_C(\beta_L, p_L^c) > 0,\end{aligned}$$

[16] Assumptions 2 and 3 imply $\frac{1}{\beta_H} - v \leq p_L^c < \overline{p}$, where the horizontal section of type-H's indifference curves start at $p = \frac{1}{\beta_H} - v$, so $V_C(\beta_H, p_L^c) = V_C(\beta_H, \overline{p})$. By definition, $V_C(\beta_H, \overline{p}) \equiv V_C(\beta_L, \overline{p})$. Furthermore, Assumption 1 (see also Footnote 8) and $p_L^c < \overline{p}$ imply $V_C(\beta_L, \overline{p}) > V_C(\beta_L, p_L^c)$. Put together,

$$V_C(\beta_H, p_L^c) = V_C(\beta_H, \overline{p}) = V_C(\beta_L, \overline{p}) > V_C(\beta_L, p_L^c).$$

where the third equality uses $V_C(\beta_H, p^c_H) = \int_{\beta_H v}^{v} (v-c)\,dc - p^c_H \int_v^1 dc$ and $V_C(\beta_H, p^c_L) = \int_{\beta_H v}^{1} (v-c)\,dc$ (by Assumption 3) and the inequality uses $\int_{\beta_H v}^{v} (v-c)\,dc > \int_{\beta_H v}^{1} (v-c)\,dc$ due to $v < 1$.

References

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. "Commitment vs. flexibility." *Econometrica* 74 (2):365–396.

Ambrus, Attila and Georgy Egorov. 2013. "Comment on commitment and flexibility." *Econometrica* 81 (5):2113–2124.

Araujo, Aloisio and Humberto Moreira. 2010. "Adverse selection problems without the Spence–Mirrlees condition." *Journal of Economic Theory* 145 (3):1113–1141.

Armstrong, Mark, John Vickers, and Jidong Zhou. 2009. "Prominence and consumer search." *The RAND Journal of Economics* 40 (2):209–233.

Armstrong, Mark and Jidong Zhou. 2011. "Paying for prominence." *The Economic Journal* 121 (556):F368–F395.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines." *The Quarterly Journal of Economics* 121 (2):635–672.

Bai, Liang, Benjamin Handel, Edward Miguel, and Gautam Rao. 2021. "Self-control and demand for preventive health: Evidence from hypertension in India." *Review of Economics and Statistics* 103 (5):835–856.

Bénabou, Roland and Jean Tirole. 2002. "Self-confidence and personal motivation." *The Quarterly Journal of Economics* 117 (3):871–915.

———. 2004. "Willpower and personal rules." *Journal of Political Economy* 112 (4):848–886.

———. 2011. "Identity, morals, and taboos: Beliefs as assets." *The Quarterly Journal of Economics* 126 (2):805–855.

Bryan, Gharad, Dean Karlan, and Scott Nelson. 2010. "Commitment devices." *Annual Review of Economics* 2 (1):671–698.

Carrillo, Juan D and Thomas Mariotti. 2000. "Strategic ignorance as a self-disciplining device." *The Review of Economic Studies* 67 (3):529–544.

Chew, Soo Hong, Wei Huang, and Xiaojian Zhao. 2020. "Motivated false memory." *Journal of Political Economy* 128 (10):3913–3939.

Chioveanu, Ioana and Jidong Zhou. 2013. "Price competition with consumer confusion." *Management Science* 59 (11):2450–2469.

DellaVigna, Stefano and Ulrike Malmendier. 2004. "Contract design and self-control: Theory and evidence." *The Quarterly Journal of Economics* 119 (2):353–402.

———. 2006. "Paying not to go to the gym." *American Economic Review* 96 (3):694–719.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya." *American Economic Review* 101 (6):2350–2390.

Eliaz, Kfir and Ran Spiegler. 2006. "Contracting with diversely naive agents." *The Review of Economic Studies* 73 (3):689–714.

Esteban, Susanna, Eiichi Miyagawa, and Matthew Shum. 2007. "Nonlinear pricing with self-control preferences." *Journal of Economic Theory* 135 (1):306–338.

Galperti, Simone. 2015. "Commitment, flexibility, and optimal screening of time inconsistency." *Econometrica* 83 (4):1425–1465.

Gao, Buqu and Liang Guo. 2024. "Optimal contracts for time-inconsistent consumers with heterogeneous beliefs." *Management Science* .

Giné, Xavier, Dean Karlan, and Jonathan Zinman. 2010. "Put your money where your butt is: a commitment contract for smoking cessation." *American Economic Journal: Applied Economics* 2 (4):213–235.

Gottlieb, Daniel. 2008. "Competition over time-inconsistent consumers." *Journal of Public Economic Theory* 10 (4):673–684.

Gottlieb, Daniel and Xingtan Zhang. 2021. "Long-term contracting with time-inconsistent agents." *Econometrica* 89 (2):793–824.

Grubb, Michael D. 2009. "Selling to overconfident consumers." *American Economic Review* 99 (5):1770–1807.

———. 2015. "Behavioral consumers in industrial organization: An overview." *Review of Industrial Organization* 47:247–258.

Gul, Faruk and Wolfgang Pesendorfer. 2001. "Temptation and self-control." *Econometrica* 69 (6):1403–1435.

Heidhues, Paul and Botond Kőszegi. 2010. "Exploiting naivete about self-control in the credit market." *American Economic Review* 100 (5):2279–2303.

———. 2018. "Behavioral industrial organization." *Handbook of Behavioral Economics: Applications and Foundations 1* 1:517–612.

Heidhues, Paul, Botond Kőszegi, and Takeshi Murooka. 2024. "Procrastination Markets." :Working Paper.

Hong, Fuhai, Wei Huang, and Xiaojian Zhao. 2019. "Sunk cost as a self-management device." *Management Science* 65 (5):2216–2230.

Jain, Sanjay. 2012. "Self-control and incentives: An analysis of multiperiod quota plans." *Marketing Science* 31 (5):855–869.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2015. "Self-control at work." *Journal of Political Economy* 123 (6):1227–1277.

Kőszegi, Botond. 2014. "Behavioral contract theory." *Journal of Economic Literature* 52 (4):1075–1118.

Kremer, Michael, Gautam Rao, and Frank Schilbach. 2019. "Behavioral development economics." In *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 2. Elsevier, 345–458.

Li, Li and Li Jiang. 2022. "How should firms adapt pricing strategies when consumers are time-inconsistent?" *Production and Operations Management* 31 (9):3457–3473.

Li, Sanxi, Jianye Yan, and Binqing Xiao. 2014. "Contract design and self-control with asymmetric information." *Economic Inquiry* 52 (2):618–624.

Moffitt, Terrie E, Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J Hancox, HonaLee Harrington, Renate Houts, Richie Poulton, Brent W Roberts, Stephen Ross et al. 2011. "A gradient of childhood self-control predicts health, wealth, and public safety." *Proceedings of the National Academy of Sciences* 108 (7):2693–2698.

Netzer, Nick and Florian Scheuer. 2010. "Competitive screening in insurance markets with endogenous wealth heterogeneity." *Economic Theory* 44:187–211.

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

Schilbach, Frank. 2019. "Alcohol and self-control: A field experiment in India." *American Economic Review* 109 (4):1290–1322.

Schottmüller, Christoph. 2015. "Adverse selection without single crossing: Monotone solutions." *Journal of Economic Theory* 158:127–164.

Smart, Michael. 2000. "Competitive insurance markets with two unobservables." *International Economic Review* 41 (1):153–170.

Thaler, Richard H and Shlomo Benartzi. 2004. "Save more tomorrow: Using behavioral economics to increase employee saving." *Journal of Political Economy* 112 (S1):S164–S187.

Tirole, Jean. 1988. *The Theory of Industrial Organization.* The MIT Press.

Wambach, Achim. 2000. "Introducing heterogeneity in the Rothschild-Stiglitz model." *Journal of Risk and Insurance* 67 (4):579–591.

Wertenbroch, Klaus. 1998. "Consumption self-control by rationing purchase quantities of virtue and vice." *Marketing Science* 17 (4):317–337.

Zhou, Jidong. 2011. "Reference dependence and market competition." *Journal of Economics & Management Strategy* 20 (4):1073–1097.