UNSW Economics Working Paper 2018-15

# Seemingly Exploitative Contracts *

Pei Cheng Yu †

October 27, 2018

**Abstract**

This paper studies sequential price discrimination of sophisticated present-biased consumers in the credit market. The optimal contract utilizes present bias to improve screening by inducing certain consumers to over-consume and over-accumulate debt without the presence of naiveté. This shows that the optimal contract can have seemingly exploitative features that cause certain consumers to experience ex-post welfare losses even when they are sophisticated. This has important policy implications. If the intention of firms is to screen and not exploit consumers, then financial regulations aimed at protecting consumers by eliminating seemingly exploitative features could introduce additional distortions. I also analyze the optimal contract for naïve consumers. The main difference between contracts for sophisticated and naïve consumers is the lack of a commitment mechanism in exploitative contracts, while the presence of teaser rates, late fees or overdraft fees does not necessarily make contracts exploitative. (*JEL* Code: D18, D82, D86, G28)

**Keywords:** Credit contract, Financial regulations, Non-linear pricing, Present bias, Sequential screening

*I thank David Rahman for many helpful comments and his encouragement and support. I am also grateful for the support from the participants of the Theory Workshop at the University of Minnesota and seminar participants at the University of Technology Sydney and University of Bristol. I also benefited greatly from conversations with Ken Binmore, Eugene Jeong, Anton Kolotilin, Hongyi Li, Keler Marku, Radek Paluszynski, Aldo Rustichini, Emil Temnyalov, Siyang Xiong, Benjamin Young and Jan Werner. I also wish to acknowledge the support of the University of Minnesota Doctoral Dissertation Fellowship.

†University of New South Wales (e-mail: *pei-cheng.yu@unsw.edu.au*)

# 1 Introduction

Several recent studies in behavioral economics have focused on exploitative contracts that take advantage of non-sophisticated consumers with features, such as teaser rates, back-loaded late fees or overdraft fees.[1] These features are considered exploitative, because consumers would have rejected these contracts if they were sophisticated. However, it is unclear whether similar features exist in optimally designed contracts for sophisticated consumers. If certain aspects of a contract are meant to take advantage of consumers' incorrect beliefs, then there may be sound reason to argue for regulatory oversight. On the other hand, if similar aspects can emanate from the screening of sophisticated consumers, then the case for government intervention is weaker.

This article departs from the literature by examining the optimality of seemingly exploitative features, such as teaser rates, in contracts for sophisticated agents. I examine the monopolist screening problem where sophisticated present-biased consumers sequentially learn their demand. In particular, I focus on the design of unsecured credit contracts.[2] Credit contracts are signed when consumers only have an informative signal of their demand. After demand is realized, consumers choose the level of consumption based on the contract they selected, while payment is due in the future. Present-biased consumers over-consume, because they under-weigh the cost of payment.

The main result of this article shows that credit contracts can have seemingly exploitative features even when agents are sophisticated in a sequential screening setting. A firm sequentially price discriminates present-biased consumers more efficiently by introducing such features. The optimal contract induces consumers with low expected demand but high realized demand to over-consume and over-accumulate debt. This deters consumers with high expected demand from mimicking consumers with low expected demand, since they are more likely to end up with higher demand. It is important to point out that sophisticated consumers sign the contract with knowledge that it may lead to losses in the future. This is in contrast to exploitative contracts, where sophisticated consumers would have rejected the contract.

This paper considers the sequential screening environment, because it is particularly suited for the study of credit contracts. There are two main reasons for this. First, borrowers do not usually know how often they need to access credit in the near future, but they

[1]Several papers have demonstrated this following Gabaix and Laibson (2006) using consumer unawareness. Others, such as Eliaz and Spiegler (2006) and Heidhues and Koszegi (2010), have highlighted how these features exploit misspecified beliefs.

[2]Unsecured credit contracts include credit cards, payday loans and bank overdrafts, which are often used by liquidity-constrained consumers.

typically know the possible frequency from past spending habits. Second, if sophisticated present-biased consumers know their exact demand when selecting a contract, then prefect price discrimination can be implemented with off-equilibrium-path threats (Esteban and Miyagawa, 2005). This stark result shows that despite the presence of asymmetric information, without sequential price discrimination, the screening problem is de facto complete information. As a result, a rigorous study of how creditors sequentially screen present-biased debtors is vital for the understanding of credit provision.

The sequential mechanism highlights how the firm utilizes the consumers' present bias to elicit their private signal in the contracting stage. Consider an environment with two signals: high and low. A high signal implies higher demand is more likely in the future. If consumers tend to misrepresent their signals downwards, the firm price discriminates by distorting the consumption of low-signal consumers *downwards* for *small* realized demand, which is standard. However, the consumers are also aware of their temptation to over-consume in the future. The firm can further decrease the information rents of high-signal consumers by distorting the consumption of low-signal consumers *upwards* for *large* realized demand. This can be interpreted as providing commitment to high-signal consumers, because if they had selected the contract for low signals, then they will likely end-up over-consuming and over-accumulating debt.

As an extension, I also address a technical issue in dynamic contracting with present-biased consumers and a continuum of signals. By taking advantage of consumers' present bias, the optimal consumption may not be monotonically increasing with respect to the signal. As a result, local incentive compatibility does not usually imply global incentive compatibility at the contracting stage, and I provide a condition such that it does.

The optimal credit contract can be implemented naturally with teaser rates, which are often considered exploitative. I show that the marginal price for high-signal consumers is lower for small loans but much higher for bigger loans. On the other hand, the marginal price for low-signal consumers exhibits the opposite phenomenon. For high-signal consumers, the lower marginal price for small loans acts as a teaser rate and the higher marginal price for big loans is akin to a commitment device. I also demonstrate the implementation of late fees and overdraft fees using a similar reasoning.

Since features that are exploitative for naïve agents could actually decrease the cost of screening sophisticated agents, this paper provides novel insights on policy and highlights a new cost to financial regulations. Policymakers have been advocating for regulations to protect consumers from potentially harmful financial contracts. For example, the Credit Card Accountability, Responsibility and Disclosure act of 2009 (Credit CARD act of 2009) was passed with bipartisan support in US Congress. The Credit CARD act placed restrictions on teaser rates, late fees and overdraft limits. To assess these regulations, I characterize the optimal regulations that increase consumer welfare. Then, I demonstrate that some of the current regulations are targeting contractual features that facilitate the provision of credit.

Finally, I show that the main difference between contracts for sophisticated and naïve consumers is the lack of commitment devices in exploitative contracts. I examine the optimal contract for naïvely present-biased consumers and demonstrate how creditors can exploit the consumers' misprediction of their future behavior. I find that exploitative contracts induce over-consumption for *all* agents and over-charges for the provision of credit, because commitment mechanisms are absent. Another difference is that exploitative contracts do not screen the signals. In other words, the firm offers the same contract for all naïve consumers regardless of expected demand, so the mechanism is essentially static.

This paper is related to the behavioral contracting literature. The recent research on contracting with non-sophisticated time-inconsistent consumers has yielded several new insights. Spiegler (2011) and Koszegi (2014) provide overviews of the exploitative contracting literature. In contrast, this paper focuses on sophisticated time-inconsistent consumers. Some papers have examined optimal contracts with sophisticated time-inconsistent consumers. Esteban and Miyagawa (2005) examines an optimal non-linear pricing model of sophisticated consumers with self-control and temptation preferences, á la Gul and Pesendorfer (2001). They find that despite the consumers having unobservable tastes, a firm is able to perfectly price discriminate when consumers' preference reversal is in a certain direction by using off-path threats. The use of off-path threats has been explored in other papers (Galperti, 2015; Bond and Sigurdsson, 2017; Yu, 2018). In particular, Galperti (2015) also examines a sequential screening environment where a provider of commitment devices initially screens time consistency before screening taste shocks. Galperti (2015) shows how off-path threats can be used when the range of possible marginal valuation of consumption varies among different levels of time inconsistency. This paper analyzes a sequential screening environment where agents share the same degree of present bias, so the range of marginal valuation of consumption is the same for all agents. Thus, off-path threats are not helpful in this paper.

This paper is also related to the sequential screening literature (Courty and Li, 2000; Eso and Szentes, 2007a,b; Krahmer and Strausz, 2015). The setup of the model closely follows Courty and Li (2000). The main departure of this paper from the literature is that consumers in my model are time inconsistent, which provides several new insights.

The paper closest in message with the current paper is Grubb (2015). In one of the main findings of Grubb (2015), 'bill-shocks' seem exploitative, but it can be used by firms to engage in price discrimination and regulating it could decrease welfare in certain environments. The current paper differs from Grubb (2015) in two distinct dimensions. First, Grubb (2015) focuses on inattentive consumers who lose track of their usage history, while I consider consumers who are time inconsistent. Second, price discrimination in Grubb (2015) is static, while this paper looks at price discrimination over time.

The paper is organized as follows. Section 2 introduces the model. Section 3 analyzes optimal sequential screening for sophisticated time-inconsistent consumers. Section 4 applies the results to explain characteristics often observed in credit contracts. Section 5 discusses some policy insights. Section 6 explores some extensions, including an examination of actual exploitative contracts with naïve consumers. Section 7 summarizes the paper. All proofs can be found in Appendix A.

# 2 Model

In this section, I introduce the sequential screening model in the spirit of Courty and Li (2000) with quasi-hyperbolic consumers (Laibson, 1997). I also discuss two benchmark environments where perfect price discrimination is implemented: the case without private information and the case with static screening.

I consider a three-period model: $t = 0, 1, 2$. The monopoly firm provides an unsecured credit contract for liquidity-constrained consumers to consume in $t = 1$ and pay in $t = 2$. The outside option is normalized to 0 if the contract is not signed.

Consumers learn about their demand over time. They receive a private signal $\sigma \in \Sigma$ at $t = 0$, with probability distribution $\gamma(\sigma)$ and cumulative distribution $\Gamma(\sigma)$. After receiving the signal, consumers choose whether to sign a contract with the firm. This decision is made before they privately learn their demand, which is represented by $\theta \in \Theta = [\underline{\theta}, \overline{\theta}] \subset \mathrm{R}_+$. The signal $\sigma$ is informative of the distribution of $\theta$, which is drawn from the differentiable conditional distribution with full support $\pi(\theta|\sigma)$ at $t = 1$, with cumulative function $\Pi(\theta|\sigma)$. Also, $\Pi(\theta|\sigma)$ strictly dominates $\Pi(\theta|\sigma')$ in terms of first-order stochastic dominance. In essence, $\Pi(\theta|\sigma) < \Pi(\theta|\sigma')$ for any $\theta \in (\underline{\theta}, \overline{\theta})$ if $\sigma > \sigma'$.

Consumers are present biased. At $t = 0$, after receiving $\sigma$, consumers have the following *ex-ante* utility from consuming $q$ units of the good bought at price $p$ :

$$U(q, p; \sigma) = \mathrm{E}[\theta u(q) - p|\sigma],$$

where $u(\cdot)$ is twice-differentiable, strictly increasing and strictly concave with $u(0) = 0$ and $\lim_{q\to\infty} u(q) \to \infty$. The consumer consumes $q$ at $t = 1$ but pays $p$ at $t = 2$. The ex-ante utility has a discount factor of 1. However, consumption decisions are made at $t = 1$ with a discount factor of $\beta < 1$. The *ex-post* utility at $t = 1$ after realizing $\theta$ is

$$V(q, p; \theta) = \theta u(q) - \beta p.$$

When $\beta \neq 1$, the consumer is dynamically inconsistent and does not consume the amount originally intended. In particular, with $\beta < 1$, the consumer is present biased and tends to over-consume, because the disutility from payment is discounted.

The firm produces $q \geq 0$ units of the good at constant marginal cost $c > 0$. The firm has full commitment and offers a contract for each $\sigma$ : $C_\sigma = (q_\sigma(\theta), p_\sigma(\theta))$, with $q_\sigma : \Theta \mapsto \mathbb{R}_+$ and $p_\sigma : \Theta \mapsto \mathbb{R}$.

Since information is revealed sequentially to the consumer, the firm needs to ensure that the contract is dynamically incentive compatible. The *ex-post incentive compatibility constraint* ensures that consumers report $\theta$ truthfully at $t = 1 : \forall \theta, \theta' \in \Theta$ and $\sigma \in \Sigma$,

$$\theta u(q_\sigma(\theta)) - \beta p_\sigma(\theta) \geq \theta u\left(q_\sigma(\theta')\right) - \beta p_\sigma(\theta') \tag{1}$$

The *ex-ante incentive compatibility constraint* ensures that consumers report $\sigma$ truthfully at $t = 0$: $\forall \sigma, \sigma' \in \Sigma$

$$\int_\Theta [\theta u(q_\sigma(\theta)) - p_\sigma(\theta)] \, \pi(\theta|\sigma) \, d\theta \geq \int_\Theta [\theta u(q_{\sigma'}(\theta)) - p_{\sigma'}(\theta)] \, \pi(\theta|\sigma) \, d\theta. \tag{2}$$

The contract is incentive compatible if both (1) and (2) are satisfied. The participation constraints are standard: $\forall \sigma \in \Sigma$,

$$\int_\Theta [\theta u(q_\sigma(\theta)) - p_\sigma(\theta)] \, \pi(\theta|\sigma) \, d\theta \geq 0. \tag{3}$$

The firm chooses a menu of contracts $C = \{C_\sigma\}_{\sigma \in \Sigma}$ to maximize expected profit $\mathrm{E}(p_\sigma(\theta) - c q_\sigma(\theta))$ subject to the participation and incentive compatibility constraints.

To summarize, the timing of the model is shown in Figure 1. The firm posts a menu $C$. Consumers choose the contract $C_\sigma \in C$ after receiving the signal $\sigma$. After realizing $\theta$, the consumers settle on $(q_\sigma(\theta), p_\sigma(\theta))$ to finalize on the quantity and payment.

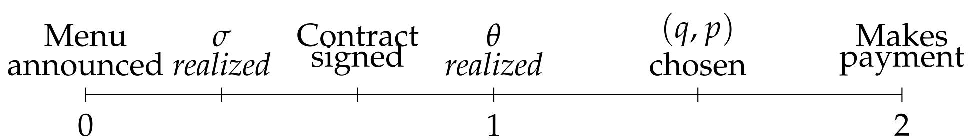

Figure 1: Timing of Events

## 2.1 Benchmark: No Private Information

In an environment without private information, the firm is able to observe the consumer's signal and demand, so perfect price discrimination can be implemented. This is implementable regardless of the presence of time inconsistency.

**Proposition 1** *Perfect price discrimination is implemented by* $(q^*_\sigma, p^*_\sigma)$ *such that for all* $\sigma \in \Sigma$, $U(q^*_\sigma, p^*_\sigma; \sigma) = 0$, *and* $\theta u'(q^*_\sigma(\theta)) = c$.

Hence, the first best profit of the firm is equivalent to the case without time inconsistency. In essence, since the consumer is time-inconsistent and aware of it, the firm provides a commitment contract for the consumer and is unable to exploit them. It is also worth noting that the optimal allocation, $q^*_\sigma$, is independent of $\sigma$ and the firm extracts rents from the consumers by varying price schedules, $p^*_\sigma$, with $\sigma$.

## 2.2 Benchmark: Static Screening of Present-Biased Consumers

The first best is also implementable if the consumer learns $\theta \in \Theta$ at $t = 0$ privately and receives no further information.[3] To see why, recall that a time-inconsistent consumer wants to prevent the future-self from betraying the present-self. As shown by Esteban and Miyagawa (2005), the firm takes advantage of the time inconsistency by introducing extraneous options in the contract designed to punish the time-inconsistent consumers only when they misreport in $t = 0$. Formally, a contract for $\theta$ is

$$C_\theta = \left\{ \left(q^R_\theta, p^R_\theta\right), \left(q^T_{\theta|\hat{\theta}}, p^T_{\theta|\hat{\theta}}\right)_{\hat{\theta}>\theta} \right\}_{\theta\in\Theta}.$$

Let $\left(q^R_\theta, p^R_\theta\right)$ denote the option the firm intends a $\theta$ consumer to choose, which is referred to as the *real allocation*. The extra options in $C_\theta$ are the *threat allocations*. The threat allocation $\left(q^T_{\theta|\hat{\theta}}, p^T_{\theta|\hat{\theta}}\right)$ is an off-path allocation meant to deter $\hat{\theta}$ consumers from misreporting as $\theta < \hat{\theta}$.[4] Following the notation of Yu (2018), let $C^*_\theta$ denote the set of options a truthful consumer at $t = 0$ chooses at $t = 1 : C^*_\theta = \arg\max_{(q'_\theta, p'_\theta)\in C_\theta} V(q'_\theta, p'_\theta; \theta)$. While if a $\theta$ consumer misreports to be $\theta' < \theta$, the consumer chooses an option from the set $C^*_{\theta'|\theta}$ in $t = 1$, in essence, $C^*_{\theta'|\theta} = \arg\max_{(q'_{\theta'}, p'_{\theta'})\in C_{\theta'}} V(q'_{\theta'}, p'_{\theta'}; \theta)$.

The contract is designed so that $\left(q^R_\theta, p^R_\theta\right) \in C^*_\theta$, but $\left(q^T_{\theta'|\theta}, p^T_{\theta'|\theta}\right) \in C^*_{\theta'|\theta}$. In other words, truthful $\theta$ consumers select $\left(q^R_\theta, p^R_\theta\right)$, but consumers who misreport as $\theta'$ end-up choosing

[3] See Esteban and Miyagawa (2005), Galperti (2015), Bond and Sigurdsson (2017) and Yu (2018)

[4] In the standard nonlinear pricing model, the relevant deviation is for high type consumers to misreport as low type consumers. Hence, the threat allocations are focused on deterring downward misreports.

the threat allocation $\left(q^T_{\theta'|\theta}, p^T_{\theta'|\theta}\right)$. The incentive compatibility constraints are for any $\theta \in \Theta$ and $\theta' < \theta$,

$$\theta u\left(q^R_\theta\right) - p^R_\theta \geq \theta u\left(q^T_{\theta'|\theta}\right) - p^T_{\theta'|\theta}.$$

The following theorem shows that the informational constraints are non-binding, so the firm achieves perfect price discrimination.

**Proposition 2** *With static information revelation, the firm achieves perfect price discrimination when $\beta < 1$, in essence, $\left(q^R_\theta, p^R_\theta\right) = \left(q^*_\theta, p^*_\theta\right)$, for all $\theta \in \Theta$.*

When $\beta < 1$, the consumer is afraid of over-consuming at $t = 1$, so there is a demand for commitment to prevent this. The firm takes advantage of the preference reversal by introducing an allocation that indulges the consumer's present bias along with a higher delayed payment: $q^T_{\theta'|\theta} > q^*_\theta$ and $p^T_{\theta'|\theta} > p^*_\theta$. The threat allocation is designed so that it is chosen if the consumer misreports downwards, which yields a lower utility from the consumers' ex-ante perspective. The mechanism is illustrated in Figure 2.

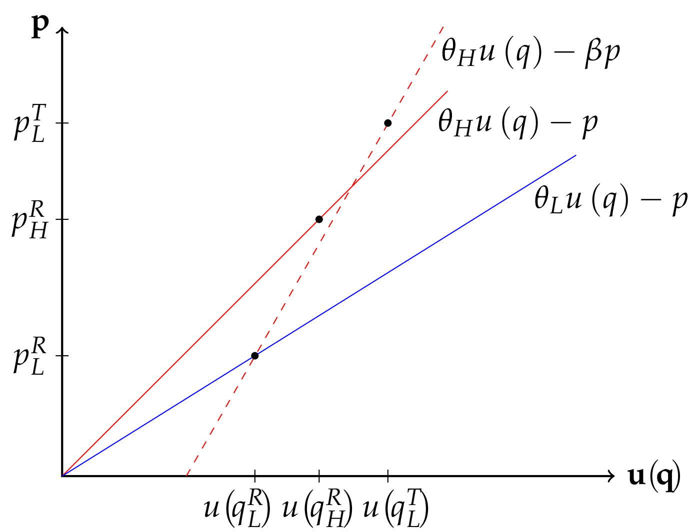


Figure 2: Static Screening with Time-Inconsistent Consumers

Figure 2 depicts the case with $\Theta = \{\theta_L, \theta_H\}$ and $\theta_H > \theta_L$. The firm only needs to deter the $\theta_H$ consumers from pretending to be $\theta_L$ consumers. As a result, the threat allocation needs to be part of $C_L$, in essence, $\left(q^T_L, p^T_L\right) \in C_L$. The threat allocation keeps the $\theta_H$ consumers honest. If $\theta_H$ consumers misreport as $\theta_L$, then $\left(q^T_L, p^T_L\right)$ is chosen over $\left(q^R_L, p^R_L\right)$, which gives a lower utility than truth-telling. More importantly, the threat allocation is also credible, since the $\theta_L$ consumers never choose $\left(q^T_L, p^T_L\right)$ over $\left(q^R_L, p^R_L\right)$.

# 3 Sequential Screening of Present-Biased Consumers

In this section, I characterize the optimal contract in a sequential screening model with present-biased consumers. I do so for the two signals case, where $\Sigma = \{L, H\}$ with $\Pi(\theta|H) < \Pi(\theta|L)$ for all $\theta \in (\underline{\theta}, \overline{\theta})$ and $\Pr(\sigma = H) = \gamma$.[5]

If information arrives sequentially, the threat allocations introduced in the static mechanism are no longer helpful in implementing perfect price discrimination. This is because in a dynamic screening environment, consumers can potentially coordinate a misreport in $t = 0$ with another in $t = 1$. To see this, consumers' information rents can be decomposed into two independent components: the ex-ante and ex-post private information. $\theta$ contains additional information not contained in $\sigma$. The signal $\sigma$ is referred to as the *ex-ante information* and the additional information as *ex-post information*. Following Eso and Szentes (2007a,b), define $\mu = \Pi(\theta|\sigma)$. Notice $\mu$ is distributed uniformly on the unit interval, $\mu \sim U[0,1]$, and is orthogonal to $\sigma$, so $\mu$ is the ex-post information. The consumer would know $\theta$ from learning $\sigma$ and $\mu$, so $\theta = \theta(\sigma, \mu) = \Pi^{-1}(\mu|\sigma)$. From the assumptions on the distribution, $\frac{\partial \theta}{\partial \sigma}, \frac{\partial \theta}{\partial \mu} > 0$. Note for any signals $\sigma, \sigma' \in \Sigma$ with $\sigma' > \sigma$ and $\mu \in [0,1]$ there is a unique $\mu^* \in [0,1]$ such that $\theta(\sigma', \mu) = \theta(\sigma, \mu^*)$. In essence, if the firm designs a threat allocation to deter a consumer with signal $\sigma'$ from mimicking a consumer with signal $\sigma$, it is unable to deter some truthful consumers at $t = 0$ from choosing the threat allocation. This is because even though a consumer with signal $\sigma'$ may be deterred by a threat allocation for type $\theta$ in $C_\sigma$, a truthful consumer with signal $\sigma$ and type $\theta = \theta(\sigma, \mu^*)$ would choose the threat allocation over the real allocation. Figure 3 illustrates the undetectability of ex-ante deviations.

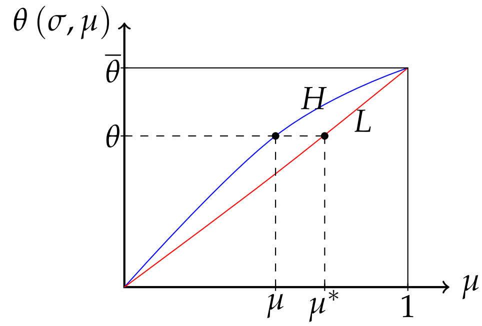


Figure 3: Full Support of Types

As a result, if threats are implemented, it is not possible for it to be off the equilibrium path. However, I show that the firm can still exploit consumers' time inconsistency by inducing consumers with lower signals to over-consume when $\theta$ is large.

[5] I examine the case with a continuum of signals in Section 6.2.

To simplify notation, let $V(\sigma,\theta) = V(q_\sigma(\theta), p_\sigma(\theta); \theta)$ and $U(\sigma'|\sigma) = U(q_{\sigma'}, p_{\sigma'}; \sigma)$, with $U(\sigma) = U(\sigma|\sigma)$. The following lemma characterizes the set of allocations that satisfy ex-post incentive compatibility (1).

**Lemma 1** *For any $\sigma \in \Sigma$, $C_\sigma$ is ex-post incentive compatible if and only if: (i.) $q_\sigma(\theta)$ is non-decreasing in $\theta$, and (ii.) $V(\sigma,\theta)$ is absolutely continuous in $\theta$, and $\frac{\partial V}{\partial \theta} = u(q_\sigma(\theta))$ when differentiable.*

The proof of Lemma 1 is standard, so it is omitted. From Lemma 1, revenue equivalence shows that the incentive compatible $p_\sigma$ is scaled upwards by $\frac{1}{\beta}$ due to present bias:

$$p_\sigma(\theta) = p_\sigma(\underline{\theta}) + \frac{1}{\beta}\left[\theta u(q_\sigma(\theta)) - \underline{\theta} u(q_\sigma(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} u(q_\sigma(t))\, dt\right].$$

The firm charges higher payments since consumers underestimate the cost of repayment at $t = 1$. The next lemma characterizes the ex-ante utility of consumers given a contract that is ex-post incentive compatible.

**Lemma 2** *If $C_\sigma$ is ex-post incentive compatible, then for any $\sigma, \sigma' \in \Sigma$,*

$$U(\sigma'|\sigma) = \frac{1}{\beta} V(\sigma', \underline{\theta}) + \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} [1 - \Phi(\theta|\sigma)]\, u(q_{\sigma'}(\theta))\, d\theta, \tag{4}$$

*where $\Phi(\theta|\sigma) = \Pi(\theta|\sigma) + \theta(1-\beta)\pi(\theta|\sigma)$.*

The expression $\Phi$ captures the disagreement between the consumer at $t = 1$ and $t = 0$ (Amador et al., 2006). If $\beta = 1$, then $\Phi(\theta|\sigma) = \Pi(\theta|\sigma)$, which is the usual expression of expected utility in sequential screening problems (Courty and Li, 2000). Even though there is a stochastic ordering for $\Pi(\theta|\sigma)$, no such ordering exists for $\Phi(\theta|\sigma)$. In other words, when $\beta < 1$, $\Phi(\theta|\sigma)$ is not monotonically ordered with respect to $\sigma$. This implies that $q_\sigma(\theta)$ is not usually monotonically increasing with $\sigma$. In fact, I will show that the novelty of contracting with present-biased consumers is that for some $\theta$, $q_\sigma(\theta)$ is monotonically decreasing in $\sigma$.[6] Furthermore, if $q_\sigma(\theta)$ is non-monotonic with respect to $\sigma$, then local ex-ante incentive constraints are no longer sufficient and the relevant ex-ante incentive constraints may not be the downward constraints.

**Lemma 3** *If $C_\sigma$ is ex-post incentive compatible, then for sufficiently large $\beta$ and for any $\sigma > \sigma'$, $U(\sigma') \geq 0$ and $U(\sigma) \geq U(\sigma'|\sigma)$ imply $U(\sigma) \geq 0$. On the other hand, for sufficiently small $\beta$, $U(\sigma) \geq 0$ and $U(\sigma') \geq U(\sigma|\sigma')$ imply $U(\sigma') \geq 0$.*

[6] It should be pointed out that incentive compatibility does not imply $q_\sigma(\theta)$ is monotonic with respect to $\sigma$ even when $\beta = 1$.

Lemma 3 shows how the incentives to misreport may differ when the degree of present bias varies. This is entirely due to the fact that $\Phi(\theta|\sigma)$ is not monotonic in $\sigma$. The rest of the paper will focus on the case with sufficiently large $\beta$, so that the usual downward incentive compatibility constraints bind and the individual rationality constraint for consumers with the lowest signal binds.

The analysis will focus on the two signal case $\Sigma = \{L, H\}$, so local and global incentive compatibility are equivalent. The two signal case provides us with the key intuition and most of the main results. I examine the continuous signal case and provide a condition such that local incentive compatibility is sufficient in Section 6.2.

Let $\Psi(\theta) = \frac{\Phi(\theta|L)-\Phi(\theta|H)}{\pi(\theta|L)}$, which is a crucial term that determines the structure of the optimal contract. The analysis proceeds with the following assumption.

**Assumption 1** *(i.) $\Psi(\theta)$ is non-increasing, and (ii.) $\underline{\theta} > \frac{\gamma}{\beta(1-\gamma)}\Psi(\underline{\theta})$.*

Assumption 1 (i.) is the time-inconsistent version of the regularity assumptions imposed to ensure monotonicity of $q_\sigma(\theta)$ with respect to $\theta$. Assumption 1 (ii.) guarantees that $L$-consumers consume a positive quantity in $t = 1$, which simplifies the analysis and is not essential to the core message of the following theorem.

**Theorem 1** *For sufficiently large $\beta$, the optimal contract satisfies*

$$\theta u'(q_H(\theta)) = c \text{ and } \left[\theta - \frac{\lambda}{\beta(1-\gamma)}\Psi(\theta)\right] u'(q_L(\theta)) = c,$$

*where $0 < \lambda \leq \gamma$, and $p_\sigma(\theta) = p_\sigma(\underline{\theta}) + \frac{1}{\beta}\left[\theta u(q_\sigma(\theta)) - \underline{\theta} u(q_\sigma(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} u(q_\sigma(t))\,dt\right]$ for any $\sigma \in \{L, H\}$. Furthermore, $U(L) = 0$ and $U(H) = U(L|H)$ and if $\lambda = \gamma$, then $U(H) \geq 0$, and if $\lambda < \gamma$, then $U(H) = 0$.*

By Theorem 1, $H$-consumers are assigned the socially efficient quantity, which is a standard result, but the distortions in the optimal allocation of $L$-consumers is novel. If consumers were time-consistent, the $L$-consumers experience a downward distortion for $\theta \in (\underline{\theta}, \overline{\theta})$ and no distortions at the top and bottom. (Note that when $\beta = 1$, then $\left[\theta - \frac{\gamma}{1-\gamma}\left[\frac{\Pi(\theta|L)-\Pi(\theta|H)}{\pi(\theta|L)}\right]\right] u'(q_L(\theta)) = c$. ) However, with present-biased consumers, the firm relaxes informational constraints by introducing upward distortions for $L$-consumers when $\theta$ is such that $\Phi(\theta|L) < \Phi(\theta|H)$. The following assumption guarantees the existence of $\theta^c \in (\underline{\theta}, \overline{\theta})$ such that $\Phi(\theta^c|L) = \Phi(\theta^c|H)$.

**Assumption 2** $\pi(\underline{\theta}|L) > \pi(\underline{\theta}|H)$ *and* $\pi(\overline{\theta}|L) < \pi(\overline{\theta}|H)$.

**Corollary 1** *If Assumption 2 holds, then there exists* $\theta^c \in (\underline{\theta}, \overline{\theta})$ *such that* $\Phi(\theta^c|L) = \Phi(\theta^c|H)$. *If* $\theta > \theta^c$*, then* $\theta u'(q_L(\theta)) < c$ *and if* $\theta < \theta^c$*, then* $\theta u'(q_L(\theta)) > c$.

Corollary 1 shows that in addition to the standard non-linear pricing result of compressing $q_L(\theta)$, the firm relaxes the informational constraints of the $H$-consumers by taking advantage of their present bias and inducing them to over-consume when they misreport $\sigma$ and $\theta$ is large. This works because the punishment from over-consumption is harsher in expectation for $H$-consumers due to first-order stochastic dominance.

More specifically, the distortion can be decomposed into two parts:

$$\Psi(\theta) = \underbrace{\left(\frac{\Pi(\theta|L) - \Pi(\theta|H)}{\pi(\theta|L)}\right)}_{\text{informativeness measure}} + \underbrace{\theta(1-\beta)\left[1 - \frac{\pi(\theta|H)}{\pi(\theta|L)}\right]}_{\text{detectability measure}}. \tag{5}$$

The first component is the informativeness of $\sigma$ on $\theta$. This component is present regardless of $\beta$, and it is strictly positive.[7] As $\sigma$ becomes more informative, the firm must provide the consumers with more information rents to prevent misreports, so downward distortion increases with informativeness.

The second component is unique to the environment with present-biased consumers.[8] It provides the firm with an additional instrument in screening consumers. A key concern in sequential screening is double deviation. The ex-ante incentive compatibility constraint has to take into account the possibility of consumers misreporting $\sigma$ and proceed to corroborate it with a misreport on the ex-post information $\mu = \Pi(\theta|\sigma)$. This concern is weakened when consumers are time-inconsistent. Double deviations become more difficult to coordinate as the disagreement between present and future-selves increases. When the present-self loses control over the future-self, the value of $\theta$ as a signal of $\sigma$ increases. As a result, the second component can be seen as resulting from a signal extraction problem. With time-inconsistent consumers, the firm can deter ex-ante deviation with ex-post inference. When $\theta$ is equally likely to be drawn from $H$ and $L$, then detectability measure recommends no additional distortions beyond the informativeness measure. When $\theta$ is more likely to be from $H$, then the detectability measure is negative and if it dominates the informativeness measure, the firm introduces upward distortion to decrease the incentive to misreport $\sigma$. Assumption 1 implies that over-consumption is less costly to induce when $\theta$ is large, because from the detectabiliy measure, $L$-consumers are less likely than $H$-consumers to have high $\theta$. Also, the firm can increase prices to cover the higher

[7] See Baron and Besanko (1984) and Courty and Li (2000).

[8] The associated likelihood ratio is commonly found in principal-agent models where the principal revises beliefs regarding the effort of the agent based on observed output (Holmstrom (1979)).

production cost with upward distortion when the consumers place a smaller weight on payment. This 'on-path threat' is analogous to the off-path threat in the static screening environment. When $\theta$ is likely from $L$, then downward distortion is optimal.

The nature of the distortions in this paper differs from Galperti (2015). In Galperti (2015), a principal sequentially screens agents' time inconsistency in $t = 0$ and then their taste shock $\theta$ in $t = 1$. Consider a fully flexible contract that curbs the present-biased agents' desire to over-consume, which must reward them for their restraint. However, time-consistent agents can obtain rents by pretending to be present-biased because they can exert self-control for free and receive the reward. As a result, the optimal contract restricts flexibility of consumption for present-biased agents by distorting consumption downwards for large $\theta$ and upwards for small $\theta$, which is different from the current paper.

By introducing over-consumption, it is possible for full rent extraction: $U(H) = U(L|H) = 0$. In this case, the firm can attenuate the distortion by lowering $\lambda$ to the point that both the ex-ante incentive compatibility and individual rationality constraints for the $H$-consumers bind. The following example demonstrates the mechanism of full rent extraction.

**Example 1** *Suppose $u(q) = \sqrt{q}$, $\gamma = \frac{1}{2}$, $\beta > \frac{1}{2}$ and $\theta \in \Theta = (8, 9)$ is drawn from a beta distribution: $\Pi(\theta|H) = (\theta - 8)^3$ and $\Pi(\theta|L) = (\theta - 8)^2$. Notice that $\Pi(\theta|H)$ first order stochastically dominates $\Pi(\theta|L)$ and Assumption 1 is satisfied when $\beta \in \left(\frac{1}{2}, \frac{21}{22}\right]$. For full rent extraction, it has to be the case that $U(L) = U(H) = U(L|H) = 0$, so the firm sets*

$$\lambda^{FRE} = \frac{\beta \int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi(\theta|L) - \Phi(\theta|H)\right] \theta d\theta}{2 \int_{\underline{\theta}}^{\overline{\theta}} \frac{\left[\Phi(\theta|L) - \Phi(\theta|H)\right]^2}{\pi(\theta|L)} d\theta}.$$

*When $\beta > \frac{1}{2}$, then $\int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi(\theta|L) - \Phi(\theta|H)\right] \theta d\theta > 0$ which implies $\lambda^{FRE} > 0$. The firm chooses $\lambda = \min\left\{\gamma, \lambda^{FRE}\right\}$. If $\lambda^{FRE} \leq \gamma$, then full rent extraction is possible. If $\lambda^{FRE} > \gamma$, then the firm must provide H-consumers with positive information rent. From this example, for full rent extraction to be possible ($\lambda^{FRE} \leq \gamma$), $\beta$ needs to be small ($\beta < 0.79$).*

*Figure 4 and Figure 5 illustrates this example for the cases when full rent extraction is possible and when it isn't respectively. Figure 4 illustrates the case where the firm chooses $\lambda^{FRE} < \gamma$ to attenuate the distortions above and below $\theta^c$. Here, when $\lambda = \gamma$, the distortions are too severe and $U(L|H) < 0$. The firm lowers $\lambda$ below $\gamma$ till $U(L|H) = 0$, which brings it closer to the socially efficient quantity. Figure 5 shows the case where full rent extraction is too costly and the firm chooses $\lambda = \gamma$ and H-consumers receive positive information rent.*

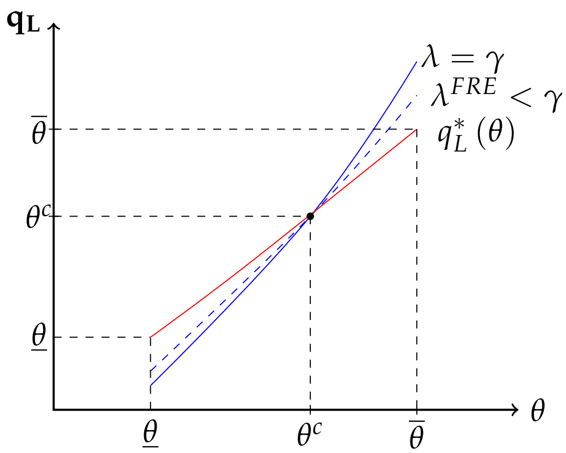


Figure 4: Firm chooses $\lambda^{FRE} < \gamma$

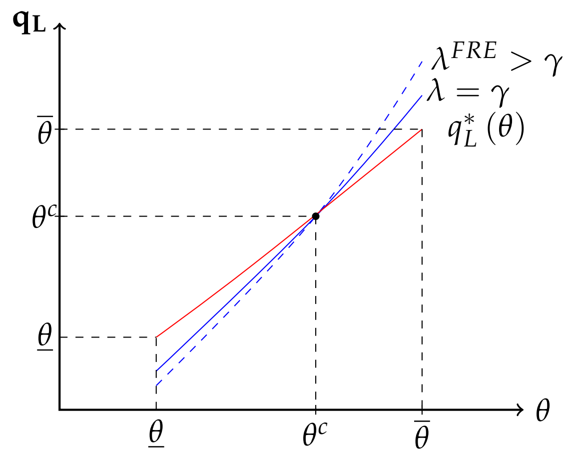


Figure 5: Firm chooses $\lambda = \gamma$

Full rent extraction is too costly when $\lambda > \gamma$, because it magnifies the distortion (over-consumption and under-consumption). In this case, the firm is better off with leaving positive information rents for the $H$-consumers ($U(H) = U(L|H) > 0$), so $\lambda \leq \gamma$. It is important to remember that full rent extraction here does not imply that the firm is able to achieve first best profits, because production for the $L$-consumers is still distorted.

# 4 Implementing Seemingly Exploitative Contracts

In this section, I demonstrate how the optimal contract can be implemented in realistic financial arrangements that seem exploitative but are not. This implies that consumers make decisions that appear to be sub-optimal, but are actually in their best interest. Though the applications are rather stylized, it helps identify seemingly exploitative features in dynamic contracts.

## 4.1 Teaser Rates

An immediate application of the general model is the interest rates of credit contracts. Consider a liquidity-constrained consumer who takes out a loan to make purchases in $t = 1$. The creditor loans $q$ to the consumer and stipulates a repayment scheme $p : \mathrm{R}_+ \to \mathrm{R}$, where $p(q)$ is the amount paid to the creditor for loan size $q$.[9] At $t = 0$, the borrower chooses from the following menu of repayment schemes: $\{p_L(q), p_H(q)\}$. By Theorem 1, recall the optimal repayment scheme induces the following behavior from sophisticated

[9]The analysis makes no distinction between the amount of loans with the amount of consumption.

borrowers: $q_H(\theta) \geq q_L(\theta)$, for $\theta \leq \theta^c$ and $q_H(\theta) < q_L(\theta)$, for $\theta > \theta^c$. Let $q^c = q_H(\theta^c) = q_L(\theta^c)$, this gives us the following corollary.

**Corollary 2** *The optimal repayment scheme satisfies (i.) $p'_H(q) \leq p'_L(q)$ for $q \leq q^c$ and (ii.) $p'_H(q) > p'_L(q)$ for $q > q^c$.*

Corollary 2 shows that for small loans ($q \leq q^c$), the repayment scheme encourages $H$-consumers more than $L$-consumers to take out additional loans. While for larger loans ($q > q^c$), it discourages $H$-consumers from taking out more loans. In a sense, the repayment scheme $p_H(q)$ acts as a 'teaser rate' when $q \leq q^c$, while it acts as a commitment device when $q > q^c$.

The change in $p_H(q)$ with respect to the loan size seems exploitative, because it is consistent with the theory that creditors take advantage of borrowers who underestimate the chances of them borrowing at the higher interest rate (Ausubel, 1999). However, it is not exploitative. The $H$-consumers would have likely over-borrowed and accumulate too much debt if they had chosen the contract for $L$-consumers with lower interest rates for high consumption levels. This contrasts with the empirical finding that suggests consumers are making sub-optimal financial decisions when they accumulate debt on high-interest credit cards.[10]

## 4.2 Late Fees

The theory introduced in the paper can also be extended to shed light on the design of penalty fees for late payments to screen heterogeneous consumer liquidity. To see how, let $p_t$ denote the amount paid at $t$ and consider consumers who sign a contract before $t = 0$ and take out a loan of $q$ at $t = 0$. The ex-ante utility, before $t = 0$, is $U(q, p_1, p_2; \sigma) = q - \mathrm{E}\left(\frac{\theta}{2}p_1^2 + p_2|\sigma\right)$. While at $t = 1$, the ex-post utility is $V(p_1, p_2; \theta) = -\frac{\theta}{2}p_1^2 - \beta p_2$. Here, $\theta$ captures the tightness of the consumers' liquidity in $t = 1$, and the convex repayment cost at $t = 1$ represents the difficulty in paying on-time. Present-biased consumers delay payment for any $\theta$, and $L$-consumers have an incentive to misreport as $H$-consumers. The creditor offers a contract, $\{p_{H,2}(p_1), p_{L,2}(p_1)\}$, where late fees $p_2(p_1)$ depend on the amount paid on-time $p_1$. Let $\Psi(\theta) = \frac{\Phi(\theta|L) - \Phi(\theta|H)}{\pi(\theta|H)}$ and assume that $-\frac{\beta\gamma}{1-\gamma} \leq \Psi'(\theta) \leq 0$. Also, denote $p^c = \frac{1}{\theta^c}$. The full derivation of the model is found in Appendix B.

**Proposition 3** *The optimal late fees satisfies (i.) $0 \geq p'_{\sigma,2}(p_1)$ for any $\sigma$ and $p_1 \geq 0$, (ii.) $p'_{H,2}(p_1) \leq p'_{L,2}(p_1)$ for $p_1 \leq p^c$ and (iii.) $p'_{H,2}(p_1) > p'_{L,2}(p_1)$ for $p_1 > p^c$.*

[10] For example, see Ausubel (1999), Shui and Ausubel (2005), Ponce et al. (2017).

Proposition 3 shows how late fees can act as a commitment device for consumers who expect to be more liquid, $L$-consumers, to pay on time. The optimal late fee encourages $L$-consumers more than $H$-consumers to pay on time when it is easy to do so (low $\theta$), while it encourages $H$-consumers more than $L$-consumers to pay on time when it is difficult to do so (high $\theta$). Since $L$-consumers expect to be more liquid, they are more likely to face incentives that encourage timely payment, which includes a punishment for delayed payment. This punishment can be implemented as a late fee, which encourages $L$-consumers to pay on schedule. $L$-consumers want to weaken the temptation of delaying payment, because they likely have the means to pay on time. Therefore, late fees serves as a valuable commitment device for liquid consumers.

### 4.3 Overdraft Fees

Overdraft limits of bank accounts or credit cards can work as a commitment device. Consider an environment where consumers have heterogeneous demand for commitment. To screen consumers who have high demand for commitment from those who have lower demand, the bank creates two types of accounts. Accounts with overdraft protection allow withdrawal even after the depletion of funds, but consumers incur an overdraft fee. On the other hand, owners of accounts *without* overdraft protection are *not* allowed to withdraw.

Consider a simple model where at $t = 0$, the value of having commitment is $a > 0$ when the future state in $t = 1$ is $\theta_H$ and zero when the state is $\theta_L$. However, consumers are time-inconsistent and at $t = 1$, the demand for commitment is zero regardless of the state. Hence, the ex-post incentive compatibility constraints are trivial because the bank is unable to screen $\theta$ at $t = 1$. At $t = 0$, consumers privately receive an informative signal $\sigma \in \{L, H\}$ of the state with $\Pr(H) = \gamma$. The conditional distribution is $\Pr(\theta_H|H) = \Pr(\theta_L|L) = \alpha > \frac{1}{2}$.

Suppose the accounts are designed such that $H$-consumers choose the account without overdraft protection and $L$-consumers choose the account with overdraft protection. Without overdraft protection, consumers are fully committed. While with overdraft protection, consumers only obtain a fraction $\chi \in (0, 1)$ of the value of commitment and incur an overdraft fee $o$ in state $\theta_H$.[11] The ex-ante incentive compatibility constraints are

$$\alpha a - f_{without} \geq \alpha(\chi a - o) - f_{with} \text{ and } (1-\alpha)(\chi a - o) - f_{with} \geq (1-\alpha) a - f_{without},$$

[11] Consumers have partial commitment, because the transaction still goes through after the depletion of funds, but they face overdraft fees which raises the marginal cost of withdrawals and helps deter overdraft.

where $f$ is the maintenance fee. The participation constraints are $\alpha a - f_{without} \geq 0$ and $(1-\alpha)(\chi a - o) - f_{with} \geq 0$. The bank maximizes the fee payment, so the optimal fees are $(f_{without}, f_{with}, o) = (\alpha a, 0, \chi a)$ and makes a profit of $a[\gamma\alpha + (1-\gamma)(1-\alpha)\chi]$.

If the population of consumers with demand for commitment is small, $\gamma < \frac{(1-\alpha)(1-\chi)}{\alpha-(1-\alpha)\chi}$, then it is not optimal for the bank to provide any consumers with overdraft protection. This is because if all accounts are without overdraft protection, then the $L$-consumers' participation constraint binds: $(1-\alpha)a = f_{without}$, which implies the bank's profit is $(1-\alpha)a$. This dominates the profit when the bank provides all consumers with overdraft protection, where the profit is $[\gamma\alpha + (1-\gamma)(1-\alpha)]\chi a$. Therefore, if there is a large enough population of $H$-consumers, $\gamma \geq \frac{(1-\alpha)(1-\chi)}{\alpha-(1-\alpha)\chi}$, then the bank screens its consumers by providing two accounts.

The analysis above implies that when there is sufficient demand for commitment (large $\gamma$), then the account with overdraft protection has a low maintenance fee ($f_{with} = 0$) but a large overdraft fee ($o = \chi a$). In essence, the bank screens consumers by exposing the ones with low-expected demand for commitment to a high risk of depleting their funds. Notice that the bank is able to achieve full surplus extraction by doing so. As a result, overdraft fees help screen commitment preferences. This is different from Heidhues and Koszegi (2017), where overdraft fees are evidence of exploitative contracting.

# 5 Regulatory Policies

Though financial contracts are not exploitative when consumers are sophisticated, the government or planner may still want to introduce regulatory policies that increase consumer welfare above the laissez faire level. In this section, I derive the optimal regulatory policy for credit contracts. The optimal policy raises consumer welfare by decreasing the distortion in the repayment schedule for $L$-consumers, so both the degree of under-consumption for when $\theta < \theta^c$ and over-consumption for when $\theta > \theta^c$ are smaller. This could be implemented using subsidies on repayments for $L$-consumers when the loan is small and taxes when the loan is large. Through the lens of the optimal policy, I discuss current and proposed regulations that place limits on financial contracts.

Let $\kappa \in (0,1)$ denote the weight a planner puts on firm profits $\Lambda$ and the remainder of the weight $1-\kappa$ on consumer welfare $\underline{U}$. The planner also puts $\eta \in (0,1)$ weight on $H$-consumers and the remainder $1-\eta$ on $L$-consumers. The planners is paternalistic, so it evaluates the consumers' welfare using the ex-ante utility, in essence, $\underline{U} = \eta U(H) + (1-\eta) U(L)$. In essence, it treats present bias as an error and attempts to correct it, because $\beta \neq 1$ reflects a self-control problem that consumers try to avoid in other periods. Thus, the welfare criterion is

$$\kappa\Lambda + (1-\kappa)\left[\eta U(H) + (1-\eta) U(L)\right].$$

The planner chooses policies that maximize welfare subject to the firm maximizing its profit. Clearly, if $\kappa = 1$, the economy is at the laissez faire benchmark, so the planner does not introduce any regulatory policies. When $\kappa < 1$, the planner may choose to enact policies that increase consumer welfare at the expense of firm profits. I assume that $\kappa > \frac{1}{2}$ and $\eta \leq \gamma$, or else the firm would either exit or only serve consumers of a certain signal. Let $\lambda^{LF}$ denote the laissez faire ($\kappa = 1$) level of distortion on $q_L(\theta)$ from Theorem 1.

**Theorem 2** *For sufficiently large $\beta$, the optimal allocation satisfies*

$$\theta u'(q_H(\theta)) = c \text{ and } \left[\theta - \frac{\lambda}{\beta(1-\gamma)}\Psi(\theta)\right] u'(q_L(\theta)) = c,$$

*where $0 < \lambda \leq \gamma - \left(\frac{1-\kappa}{\kappa}\right)\eta$ and the optimal payment satisfies $p_\sigma(\theta) = p_\sigma(\underline{\theta}) + \frac{1}{\beta}\left[\theta u(q_\sigma(\theta)) - \underline{\theta} u(q_\sigma(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} u(q_\sigma(t))\,dt\right]$ for any $\sigma \in \{L, H\}$. Also, $U(L) = 0$ and $U(H) = U(L|H) \geq 0$ with equality when $\lambda < \gamma - \left(\frac{1-\kappa}{\kappa}\right)\eta$.*

Theorem 2 shows that if anyone benefits from the regulation, it is the $H$-consumers. First, notice that the range of $\lambda$ is smaller than that under laissez faire, and this range is decreasing in $\eta$ and increasing in $\kappa$. This means that if the firm gives away rents to the $H$-consumers under laissez faire ($\lambda^{LF} = \gamma$), then the planner would introduce regulations that decreases the distortion on $q_L$, which increases the $H$-consumers' rent. This is because when the planner attempts to increase consumer welfare above the laissez faire benchmark, it is most efficient to do so through $H$-consumers while keeping the individual rationality constraint for $L$-consumers binding. On the other hand, if the firm implements full rent extraction under laissez faire ($\lambda^{LF} < \gamma$), then the planner would introduce regulations only if the welfare weight on $H$-consumers $(1-\kappa)\eta$ is sufficiently large. This is because the efficiency loss from increasing consumer welfare is too costly from the planner's viewpoint when it places insufficient weight on consumers. Surprisingly, $L$-consumers' individual rationality constraint remains binding even when the planner places more weight on them than $H$-consumers. This is because the planner decentralizes through the private market, and it is too costly for the firm to separate $H$-consumers from $L$-consumers when $U(L) > 0$. As a result, to increase the utility of $L$-consumers, the planner needs to either prohibit the price discrimination of ex-ante information or provide publicly funded credit and crowd out the private market.

Finally, when the utility of $H$-consumers goes up, the planner can decrease the distortion of the credit contract for the $L$-consumers, but it is not optimal to eliminate it. Thus, the credit contract for $L$-consumers is less distorted than the laissez faire benchmark when the planner introduces regulations. However, the planner still allows for upward distortion, so $L$-consumers would still over-borrow. This could have important implications for regulations that attempt to curb ex-post losses.

## 5.1 Discussion on Regulatory Policies

Here, I discuss recent financial regulations in light of the analysis above. In Section 4, I argued that certain features of credit contracts may seem exploitative, but could actually act as commitment devices and help firms lower the cost of screening. Also, an immediate implication of Theorem 2 is that these seemingly exploitative features are attenuated but still present when regulators try to raise consumer welfare.

It is important to remember that the aim of the paper is to consider the potential costs of paternalistic financial regulations. It does not discuss whether the cost regulations outweighs the benefit, which depends on the relative size of the sophisticated and the naïve population.[12] The goal of this paper is to point out that welfare evaluations that neglect the screening benefits of seemingly exploitative features might over-state the success of paternalistic financial regulations.

***Teaser rates:*** The Credit CARD act requires any teaser rates to remain in place for at least 6 months. There are proposals to further increase the duration of teaser rates (Bar-Gill and Bubb, 2012). By lengthening the duration of teasers, policymakers hope to increase the firm's cost of providing credit at these lower rates. This policy is potentially welfare improving for naively present-biased borrowers, consumers suffering from misspecified beliefs (incorrectly believing they will cancel after the teaser rates expire) or inattention (losing track of how long they have used the card).

For sophisticated agents, depending on the welfare weights $\kappa$ and $\eta$, such a regulatory intervention could lower welfare. Notice that a lengthening of the duration of teaser rates is similar to a regulatory increase of $q^c$ from Corollary 2. With the increased duration, consumers get to borrow more under the lower marginal interest rate. However, an increase in $q^c$ raises the firm's cost of screening ex-ante demand. This is because when the firm introduces the possibility of over-consumption to decrease the cost of screening, it does so through the detectability measure in (5). When $q^c$ increases, the firm's ability

[12] An example of quantitative analysis on the effects of financial regulations is Agarwal et al. (2015).

to use the detectability measure decreases and the ex-ante incentive compatibility constraints tighten. The firm has to increase information rent, which introduces additional inefficiency in the provision of credit. An extreme would be to consider the case when all teaser rates are banned, then $q^c = q\left(\overline{\theta}\right)$ and a creditor cannot induce over-consumption to lower screening cost, and by Theorem 2 this is not optimal.

In essence, instead of interpreting teaser rates and back-loaded fees as exploitative, the higher interest rates post-introductory offer can be viewed as a commitment device. By increasing the duration of the introductory offer, the contract becomes less appealing due to its loss in value as a commitment device.

***Late fees:*** One of the main purposes of the Credit CARD act was to regulate certain fees, and late fees was one of them. It significantly restricted the magnitude of the penalties and requires it to be proportional to the violation. Regulators placed an upper bound on these fees: $25 for the first violation and $35 for each future violation within the next six months. Indeed, the average late fees prior to the Credit CARD act was $33.08, and the average late fees immediately after its passage dropped to $23.13 (Consumer Financial Protection Bureau, 2013). This policy is potentially welfare improving for borrowers who are unaware of their present bias or are inattentive to the payment due dates.

An upper bound on late fees could introduce distortions when borrowers are sophisticated, and is not optimal by Theorem 2. Proposition 3 showed how late fees can be used to screen the expected tightness of consumers' liquidity. It provides more liquid borrowers a commitment device to pay on time. More liquid borrowers are willing to pay creditors more for this commitment device, because they expect to have the means to pay at $t = 1$ but are afraid that their future-selves would delay payment. Thus, the late fees relax the ex-ante incentive compatibility, and a restriction on them implies that creditors would have to raise the information rent to $L$-borrowers, which increases distortion.

***Overdraft fees:*** Overdraft or over-the-limit fees have been a perennial target of policy-makers. After the Credit CARD act, credit card companies cannot charge over-the-limit fees unless consumers consent to being taken over the limit. The Federal Reserve also prohibits banks from charging overdraft fees unless consumers opt-in to the service. More importantly, the Federal Reserve also does not allow financial institutions to discriminate against consumers who do not opt-in to overdraft protection, so all consumers are presented with the same account terms and features.

These measures limit the ability of financial institutions to take advantage of naïve borrowers, while still allowing sophisticated borrowers who may want to use overdraft protection for commitment purposes to opt-in. As a result, these regulations seem to satisfy the requirements for libertarian paternalism (Thaler and Sunstein, 2003).

However, if creditors are not allowed to discriminate between borrowers who opt-in from those who do not, then they cannot properly screen sophisticated agents with high-expected valuation for commitment from those with low-expected valuation. As was shown in Section 4.3, when the population of consumers with demand for commitment is large, restricting the ability of firms to discriminate based on borrowers' decision to opt-in is suboptimal. This is because the cost of screening these the borrowers could manifest elsewhere and firms may uniformly increase the cost for all consumers.

# 6 Extensions

## 6.1 Actual Exploitative Contracts

So far I have focused on sophisticated consumers. To identify which contracts are actually exploitative, in this section, I analyze a model where the consumers are fully naïve and the firm attempts to screen consumer demand. The optimal contract in this setting takes advantage of the consumers' naiveté, so it is exploitative. I also discuss the implementation of the exploitative contract and contrast it with the optimal contract for sophisticated consumers. Here, I consider a continuum of signals: $\Sigma = [\underline{\sigma}, \overline{\sigma}] \in \mathrm{R}_{+}$, with distribution $\gamma(\sigma)$ and cumulative distribution $\Gamma(\sigma)$.

At $t = 0$, naïve consumers believe the discount factor at $t = 1$ to be 1 when it is actually $\beta < 1$. In essence, at $t = 0$, consumers perceive ex-post utility to be $W(\sigma, \theta) = \theta u(q_\sigma(\theta)) - p_\sigma(\theta)$. For naïve consumers, the firm announces $\{C_\sigma^n\}_{\sigma \in \Sigma}$ where

$$C_\sigma^n = \left\{ \left(q_\sigma^R(\theta), p_\sigma^R(\theta)\right)_{\theta \in \Theta^R}, \left(q_\sigma^I(\theta), p_\sigma^I(\theta)\right)_{\theta \in \Theta^I} \right\}.$$

Following Eliaz and Spiegler (2006), the firm exploits the incorrect beliefs of naïve consumers with off-path promises via the imaginary allocations: $\left(q^I(\theta), p^I(\theta)\right)$. The naïve consumers believe they will choose the imaginary allocations, but unbeknownst to them they choose the real allocations $\left(q^R(\theta), p^R(\theta)\right)$ instead. This exploitation also alters the message space in a direct mechanism, so $\Theta^R, \Theta^I \neq \Theta$, where $\Theta^R$ and $\Theta^I$ denote the message spaces for the real and imaginary allocations respectively. To distinguish the utility of consuming the imaginary from the real allocation, denote the utility of consuming the imaginary allocation with a superscript $I$ and the real allocation with a superscript $R$. For example, $V^R(\sigma, \theta) = \theta u\left(q_\sigma^R(\theta)\right) - \beta p_\sigma^R(\theta)$.

The optimal contract has to satisfy the fooling constraints: $\forall \sigma \in \Sigma$,

$$\max_{\theta'} \left\{ \theta u \left( q^I_\sigma \left( \theta' \right) \right) - p^I_\sigma \left( \theta' \right) \right\} \geq \max_{\theta'} \left\{ \theta u \left( q^R_\sigma \left( \theta' \right) \right) - p^R_\sigma \left( \theta' \right) \right\}. \tag{6}$$

The fooling constraints (6) ensures that at $t = 0$, the naïve consumers believe they will prefer the imaginary allocation over the real allocation at $t = 1$. Note that though the real allocations are ex-post incentive compatible, it is not incentive compatible from the viewpoint of the naïve consumer.

The firm designs the contract such that consumers choose the real allocations at $t = 1$: $\forall \sigma \in \Sigma$ and $\forall \theta \in \Theta$,

$$\theta u \left( q^R_\sigma \left( \theta \right) \right) - \beta p^R_\sigma \left( \theta \right) \geq \max_{\theta'} \left\{ \theta u \left( q^I_\sigma \left( \theta' \right) \right) - \beta p^I_\sigma \left( \theta' \right) \right\}. \tag{7}$$

I refer to (7) as the executability constraint. It ensures the consumers choose the real allocations over the imaginary allocations at $t = 1$.

I focus on imaginary allocations that are *perceived ex-post incentive compatible:* $\forall \sigma \in \Sigma$ and $\forall \theta, \theta' \in \Theta$,

$$\theta u \left( q^I_\sigma \left( \theta \right) \right) - p^I_\sigma \left( \theta \right) \geq \theta u \left( q^I_\sigma \left( \theta' \right) \right) - p^I_\sigma \left( \theta' \right).$$

In essence, at $t = 0$, consumers believe they will report truthfully in $t = 1$. This restriction on imaginary allocations is imposed without loss of generality. Since the imaginary allocations are perceived ex-post incentive compatible, Lemma 1 applies. As a result, $q^I_\sigma \left( \theta \right)$ is non-decreasing in $\theta$ and revenue equivalence implies

$$p^I_\sigma \left( \theta \right) = p^I_\sigma \left( \underline{\theta} \right) + \theta u \left( q^I_\sigma \left( \theta \right) \right) - \underline{\theta} u \left( q^I_\sigma \left( \underline{\theta} \right) \right) - \int_{\underline{\theta}}^{\theta} u \left( q^I_\sigma \left( t \right) \right) dt. \tag{8}$$

Also, by Lemma 2, the expected utility of the naïvely present-biased consumer at $t = 0$ is

$$U^I \left( \sigma' | \sigma \right) = W^I \left( \sigma', \underline{\theta} \right) + \int_{\underline{\theta}}^{\overline{\theta}} \left[ 1 - \Pi \left( \theta | \sigma \right) \right] u \left( q^I_{\sigma'} \left( \theta \right) \right) d\theta. \tag{9}$$

The next lemma characterizes ex-ante incentive compatibility for a continuum of signals.

**Lemma 4** *If $C^n_\sigma$ is incentive compatible, then (i.)* $\frac{\partial U^I(\sigma)}{\partial \sigma} = -\int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial \Pi(\theta|\sigma)}{\partial \sigma} u \left( q^I_\sigma \left( \theta \right) \right) d\theta$, *and (ii.)* $-\int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial \Pi(\theta|\sigma)}{\partial \sigma} \frac{\partial q^I_\sigma(\theta)}{\partial \sigma} u' \left( q^I_\sigma \left( \theta \right) \right) d\theta \geq 0$.

Since agents are fully naïve, Lemma 4 is the same as Lemma 3.2 in Courty and Li (2000), and $\frac{\partial U(\sigma)}{\partial \sigma} > 0$. The next lemma simplifies the fooling and executability constraints.

**Lemma 5** *If the real allocations are ex-post incentive compatible and the imaginary allocations are perceived ex-post incentive compatible, then the fooling constraints are,* $\forall \sigma \in \Sigma$,

$$W^I(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\theta} u\left(q_\sigma^I(\theta)\right) d\theta \geq \frac{1}{\beta}\left[V^R(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\beta\theta} u\left(q_\sigma^R(t)\right) dt\right]. \tag{10}$$

*and the executability constraints are,* $\forall \sigma \in \Sigma$ *and* $\forall \theta \in \Theta$,

$$V^R(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\theta} u\left(q_\sigma^R(t)\right) dt \geq \beta\left[W^I(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\frac{\theta}{\beta}} u\left(q_\sigma^I(t)\right) dt\right]. \tag{11}$$

Lemma 5 characterizes the fooling and executability constraints when the allocations are ex-post incentive compatible. The executability constraints bind at the optimum. The following theorem characterizes the optimal exploitative contract.

**Theorem 3** *For all* $\sigma \in \Sigma$, $\frac{\theta}{\beta}u'\left(q_\sigma^R(\theta)\right) = c$ *for any* $\theta \in \Theta^R = \left[\beta\underline{\theta}, \overline{\theta}\right]$, *and* $q_\sigma^I(\theta) = 0$ *for all* $\theta \in \Theta^I = \left[\underline{\theta}, \frac{\overline{\theta}}{\beta}\right]$. *For all* $\sigma \in \Sigma$, *the payment scheme is* $p_\sigma^R(\theta) = \frac{\theta}{\beta}u\left(q_\sigma^R(\theta)\right)$ *for all* $\theta \in \Theta^R$ *and* $p_\sigma^I(\theta) = 0$ *for all* $\theta \in \Theta^I$.

First, Theorem 3 shows that exploitative credit contracts do not discriminate $\sigma$ at $t = 0$. In other words, the realization of $\sigma$ at the contracting stage is irrelevant. This is because to exploit the naïve consumers, the firm needs to manipulate the consumers into believing that they would consume less than they actually would. As a result, the best the firm can do is to manipulate all consumers into believing that they would not borrow in the future: $q_\sigma^I(\theta) = 0$ for any anticipated and realized demand, regardless of $\sigma$.

Second, for any $\theta$, consumers over-consume and over-pay relative to the optimum. As a result, a sophisticated consumer would never choose the exploitative contract. To see this, notice that at the contracting stage, the exploitative credit contract gives the sophisticated consumer an expected utility of $U^R(\sigma) = -\frac{1-\beta}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} \theta u\left(q_\sigma^R(\theta)\right)\pi(\theta|\sigma)\, d\theta < 0$. In other words, even though the consumers do not plan to execute the credit contract ($q_\sigma^I = 0$), the contract incentivizes their future-selves to over-borrow and then make a large payment at a later date.

### 6.1.1 Implementing Exploitative Contracts

This section serves as a compliment to Section 4 by discussing the implementation of actual exploitative contracts. The main difference between contracts for sophisticated and naïve consumers is the lack of a commitment device in exploitative contracts.

***Teaser rates:*** Exploitative contracts induce naïve borrowers to over-borrow. Too see this, let $p_n(q)$ denote the repayment scheme for naïve borrowers. By Theorem 3, not taking out a loan is what the borrowers imagine doing at $t = 1$, so naïve borrowers do not plan to use it, which implies $p_n(0) = 0$. However, they do end-up taking out loans and the repayment scheme always encourages them to take out more loans. Theorem 1 and Theorem 3 shows that naïve borrowers borrow more than sophisticated $H$-consumers: $p'_H(q) > p'_n(q)$ for all $q$.

This implies the key exploitative feature is the absence of a commitment mechanism that limits borrowing and not the presence of a teaser rates or introductory offers. In fact, $p_n(q)$ seems to dominate other offers since repayment is less sensitive to loan size, but this induces present-biased debtors to over-borrow and only non-sophisticated agents do not foresee this and are exploited.

***Late fees:*** The late fees for naïve borrowers induces them to delay payment, so they end-up paying high late fees. Let $p_{n,2}(p_1)$ denote the on-path late fees when borrowers repaid $p_1$ on time. Comparing the exploitative contract in Proposition 5 in Appendix B with the contract for sophisticated borrowers in Proposition 3, the optimal $p_{n,2}(p_1)$ provides less incentives for naïve borrowers to pay on-time than $p_{L,2}(p_1)$. More formally, $p_{n,2}(p_1)$ is chosen such that $p'_{L,2}(p_1) < p'_{n,2}(p_1)$ for all $p_1$. As a result, naïve borrowers pay less at $t = 1$ and incur higher late fees than sophisticated borrowers. Also, the optimal contract has a front-loaded imaginary repayment schedule and a back-loaded real or on-path repayment schedule. This is similar to Heidhues and Koszegi (2010), who show how contracts exploit non-sophistication by providing cheap access to loans for timely repayment, but delays in repayment are penalized with large fees.

***Overdraft fees:*** If naïve consumers do not believe that they need commitment in the future, then banks only provide overdraft protection accounts and charge high overdraft fees with low maintenance fees. More specifically, the bank does not discriminate and provides an account that has low maintenance fee $f_{with} = 0$ and large overdraft fee $o$.

There are two main differences between the exploitative account and the accounts for sophisticated consumers. First, the exploitative account has less variety in overdraft policies. The accounts for sophisticated consumers differ in the amount of commitment it provides, while exploitative accounts lack devices for commitment. Second, the exploitative account has higher overdraft fees. Though the exploitative account is similar to the account for sophisticated $L$-consumers, the bank charges higher overdraft fee to take advantage of consumers who do not internalize the overdraft cost.

## 6.2 Continuum of Signals

It is known that when screening time-inconsistent individuals, local incentive compatibility does not imply global incentive compatibility (Halac and Yared, 2014; Galperti, 2015). This makes the analysis of sequential screening with present-biased consumers difficult when the set of ex-ante signals is large.

To see why local incentive compatibility constraints are not sufficient, consider the case with $\Sigma = \{\sigma_1, \sigma_2, \sigma_3\}$ and $\sigma_3 > \sigma_2 > \sigma_1$. At the optimum, the participation constraint for $\sigma_1$ binds: $U(\sigma_1) = 0$. Suppose local incentive compatibility constraints bind, in essence, $U(\sigma_2) = U(\sigma_1|\sigma_2)$ and $U(\sigma_3) = U(\sigma_2|\sigma_3)$, then

$$U(\sigma_2|\sigma_3) = \frac{1}{\beta}\left[\int_{\underline{\theta}}^{\overline{\theta}} [\Phi(\theta|\sigma_1) - \Phi(\theta|\sigma_2)]\, u(q_{\sigma_1}(\theta))\, d\theta + \int_{\underline{\theta}}^{\overline{\theta}} [\Phi(\theta|\sigma_2) - \Phi(\theta|\sigma_3)]\, u(q_{\sigma_2}(\theta))\, d\theta\right],$$

and $U(\sigma_1|\sigma_3) = \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} [\Phi(\theta|\sigma_1) - \Phi(\theta|\sigma_3)]\, u(q_{\sigma_1}(\theta))\, d\theta$. Notice that

$$U(\sigma_2|\sigma_3) - U(\sigma_1|\sigma_3) = \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} [\Phi(\theta|\sigma_2) - \Phi(\theta|\sigma_3)]\, [u(q_{\sigma_2}(\theta)) - u(q_{\sigma_1}(\theta))]\, d\theta.$$

As a result, when $\beta = 1$ and $q_\sigma(\theta)$ is non-decreasing in $\sigma$, then local incentive compatibility is sufficient. However, the main message of this paper is that with sophisticated present-biased consumers, the firms would distort the optimal quantity such that $q_\sigma(\theta)$ is decreasing in $\sigma$ for a certain range of demand realizations. As a result, non-local deviations could be more tempting than local ones. I consider a continuum of signals below: $\Sigma = [\underline{\sigma}, \overline{\sigma}] \in \mathrm{R}_+$, with distribution $\gamma(\sigma)$ and cumulative distribution $\Gamma(\sigma)$.

**Assumption 3** *Let* $\Psi_\sigma(\theta) = -\frac{\partial\Phi(\theta|\sigma)/\partial\sigma}{\pi(\theta|\sigma)}$, *(i.)* $\Psi_\sigma(\theta)$ *is non-increasing in* $\theta$, *(ii.)* $\underline{\theta} > \frac{1-\Gamma(\underline{\sigma})}{\beta\gamma(\underline{\sigma})}\Psi_{\underline{\sigma}}(\underline{\theta})$, *and (iii.)* $\epsilon_\psi(\sigma,\theta) = \frac{\partial\Psi_\sigma(\theta)}{\partial\theta}\frac{\theta}{\Psi_\sigma(\theta)}$ *is finite.*

Assumption 3 is analogous to Assumption 1. As was already mentioned, due to the lack of single crossing with respect to $\sigma$, ex-ante incentive compatibility does not imply that $q_\sigma(\theta)$ is non-decreasing in $\sigma$. It also requires the elasticity of $\Psi_\sigma(\theta)$ with respect to $\theta$ to be finite. The following assumption ensures a monotonicity-like rule for the optimal quantity with respect to the signal.

**Assumption 4** *For* $\theta \leq \theta^c$, $\frac{1-\Gamma(\sigma)}{\gamma(\sigma)}\Psi_\sigma(\theta)$ *is strictly decreasing in* $\sigma$. *For* $\theta \geq \theta^c$, $\frac{1-\Gamma(\sigma)}{\gamma(\sigma)}\Psi_\sigma(\theta)$ *is strictly increasing in* $\sigma$.

Assumption 4 is the time-inconsistent counterpart to a similar assumption in Courty and Li (2000). In their paper, a sufficient condition for incentive compatibility is if $q_\sigma(\theta)$ was strictly increasing in $\sigma$ and $\theta$ and the time-consistent counterpart of Assumption 4 guarantees it. However, as was mentioned, with present-biased consumers, imposing conditions to ensure monotonicity of $q_\sigma(\theta)$ with respect to $\sigma$ misses the main features of an optimal contract with time-inconsistent consumers. Assumption 4 maintains these features while providing tractability. I show that with Assumption 4, the optimal quantity is monotonically increasing in $\sigma$ when $\theta < \theta^c$ and monotonically decreasing in $\sigma$ when $\theta > \theta^c$. However, Assumption 4 by itself is not enough to ensure ex-ante incentive compatibility, because $\theta^c$ could be different for various signals. The following assumption on the distribution combined with Assumption 4 are sufficient for the contract to be ex-ante incentive compatible. I refer to it as the *strong single crossing condition*.

**Assumption 5 (Strong Single Crossing Condition)** *There exists* $\theta^c \in (\underline{\theta}, \overline{\theta})$ *such that* $\Phi(\theta^c|\sigma) = \Phi(\theta^c|\sigma')$ *for any* $\sigma, \sigma' \in \Sigma$.

**Theorem 4** *For sufficiently large* $\beta$, $U'(\sigma) > 0$ *and the optimal contract has* $U(\underline{\sigma}) = 0$ *and*

$$\left[\theta - \frac{1-\Gamma(\sigma)}{\beta\gamma(\sigma)}\Psi_\sigma(\theta)\right] u'(q_\sigma(\theta)) = c.$$

*with* $p_\sigma(\theta) = p_\sigma(\underline{\theta}) + \frac{1}{\beta}\left[\theta u(q_\sigma(\theta)) - \underline{\theta} u(q_\sigma(\underline{\theta})) - \int_{\underline{\theta}}^{\theta} u(q_\sigma(t))\,dt\right]$ *.olut*

The optimal sequential mechanism with a continuum of signals is similar to the two-signal case. Theorem 4 focuses on the case for large $\beta$, because when $\beta$ is large, $U(\sigma)$ is increasing in $\sigma$ and only the participation constraint of $U(\underline{\sigma})$ binds.

# 7 Conclusion

A firm's optimal pricing strategy is sensitive to the consumer's preferences and sophistication. This paper explores the optimal pricing of a firm facing time-inconsistent consumers with dynamic private information. It showed how a firm can achieve perfect price discrimination in a static environment, and design contracts containing provisions that induce over-consumption in a sequential screening environment. The optimal contract contains features that explain seemingly exploitative clauses in a credit contract.

This paper suggests that policies restricting credit contracts may exacerbate adverse selection and introduce new distortions. It also suggests that policy evaluation that ignores the improved screening of sophisticated agents with seemingly exploitative features would over-state the benefits of paternalistic regulations. Though this paper highlights a new cost to imposing paternalistic financial regulations, there is still more work to be done before its welfare impact can be concretely evaluated.

This article also highlights theoretical properties of sequential screening with time-inconsistent preferences, which is useful for future work. For example, the design of optimal policies or insurance contracts for present-biased agents.

# References


**Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel**, "Regulating Consumer Financial Products: Evidence from Credit Cards," *Quarterly Journal of Economics*, 2015, *130* (1), 111–164.

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

**Ausubel, Lawrence**, "Adverse Selection in the Credit Card Market," *Unpublished Manuscript*, 1999.

**Bar-Gill, Oren and Ryan Bubb**, "Credit Card Pricing: The CARD Act and Beyond," *Cornell Law Review*, 2012, *97*, 967–1017.

**Baron, David and David Besanko**, "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1984, *1*, 267–302.

**Bond, Philip and Gustav Sigurdsson**, "Commitment Contracts," *Review of Economic Studies*, 2017, *85* (1), 194–222.

**Consumer Financial Protection Bureau**, "CARD Act Report: A Review of the Impact of the CARD Act on the Consumer Credit Card Market," 2013.

**Courty, Pascal and Hao Li**, "Sequential Screening," *Review of Economic Studies*, 2000, *67*, 697–717.

**Eliaz, Kfir and Ran Spiegler**, "Contracting with Diversely Naive Agents," *Review of Economic Studies*, 2006, *73*, 689–714.

**Eso, Peter and Balasz Szentes**, "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies*, 2007, *74*, 705–731.

_ **and** _ , "The Price of Advise," *RAND Journal of Economics*, 2007, *38*, 863–880.

**Esteban, Susanna and Eiichi Miyagawa**, "Optimal Menu of Menus with Self-Control Preferences," *NAJ Economics, Peer Reviews of Economics Publications*, 2005, *12.*

**Gabaix, Xaver and David Laibson**, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505540.

**Galperti, Simone**, "Commitment, Flexibility, and Optimal Screening of Time Inconsistency," *Econometrica*, 2015, *83* (4), 1425–1465.

**Grubb, Michael**, "Consumer Inattention and Bill-Shock Regulation," *Review of Economic Studies*, 2015, *82*, 219–257.

**Gul, Faruk and Wolfgang Pesendorfer**, "Temptation and Self-Control," *Econometrica*, 2001, *69* (6).

**Halac, Marina and Pierre Yared**, "Fiscal Rules and Discretion Under Persistent Shocks," *Econometrica*, 2014, *82* (5), 1557–1614.

**Heidhues, Paul and Botond Koszegi**, "Exploiting Naivete about Self-Control in the Credit Market," *American Economic Review*, 2010, *100* (5), 2279–2303.

_ **and** _ , "Naiveté Based Discrimination," *Quarterly Journal of Economics*, 2017, *132* (2), 1019–1054.

**Holmstrom, Bengt**, "Moral Hazard and Observability," *Bell Journal of Economics*, 1979, *10* (1), 74–91.

**Koszegi, Botond**, "Behavioral Contract Theory," *Journal of Economic Literature*, 2014, *52* (4), 1075–1118.

**Krahmer, Daniel and Roland Strausz**, "Ex Post Information Rents in Sequential Screening," *Games and Economic Behavior*, 2015, *90*, 257–273.

**Laibson, David**, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

**Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa**, "Borrowing on the Wrong Credit Card? Evidence from Mexico," *American Economic Review*, 2017, *107* (4), 1335–1361.

**Shui, Haiyan and Lawrence Ausubel**, "Time Inconsistency in the Credit Market," *Working Paper*, 2005.

**Spiegler, Rani**, "Bounded Rationality and Industrial Organization," *Oxford University Press*, 2011.

**Thaler, Richard and Cass Sunstein**, "Libertarian Paternalism," *American Economic Review, Paper and Proceedings*, 2003, *93* (2), 175 – 179.

**Yu, Pei Cheng**, "Optimal Retirement Policies with Present-Biased Agents," *Working Paper*, 2018.

# A Proofs

**Proof of Proposition 2:** It has to be the case that $\left(q_{\theta}^{R}, p_{\theta}^{R}\right) \in C_{\theta}^{*}$. Set $\left(q_{\theta}^{R}, p_{\theta}^{R}\right) = \left(q_{\theta}^{*}, p_{\theta}^{*}\right)$ for all $\theta \in \Theta$, so (3) automatically holds. For a threat mechanism, it remains to be shown that $\left(q_{\theta'|\theta}^{T}, p_{\theta'|\theta}^{T}\right)$ can be constructed so that $\left(q_{\theta'|\theta}^{T}, p_{\theta'|\theta}^{T}\right) \in C_{\theta'|\theta}^{*}$ and incentive compatibility is satisfied while $\left(q_{\theta}^{R}, p_{\theta}^{R}\right) \in C_{\theta}^{*}$. Since $\left(q_{\theta}^{R}, p_{\theta}^{R}\right) = \left(q_{\theta}^{*}, p_{\theta}^{*}\right)$ for all $\theta \in \Theta$, the threat option only needs to deter downward misreports.

Consider a type $\tilde{\theta} \in \Theta$. Construct the threat allocations such that $\forall \theta > \tilde{\theta}$,

$$U\left(q_{\tilde{\theta}|\theta}^{T}, p_{\tilde{\theta}|\theta}^{T}; \theta\right) = U\left(q_{\theta}^{*}, p_{\theta}^{*}; \theta\right) \tag{12}$$

and for all $\theta, \theta' \geq \tilde{\theta}$

$$\theta u\left(q_{\tilde{\theta}|\theta}^{T}\right) - \beta p_{\tilde{\theta}|\theta}^{T} \geq \theta u\left(q_{\tilde{\theta}|\theta'}^{T}\right) - \beta p_{\tilde{\theta}|\theta'}^{T}. \tag{13}$$

Constraint (13) guarantees incentive compatibility off-equilibrium path. In essence, if $\theta$ consumers misreported as $\tilde{\theta}$ in $t = 0$, they choose the threat allocation $\left(q_{\tilde{\theta}|\theta}^{T}, p_{\tilde{\theta}|\theta}^{T}\right)$, which is designed for $\theta$ consumers in $t = 1$. Constraint (12) is the incentive compatibility constraint, so consumers weakly prefer reporting truthfully in $t = 0$.

Let $q_{\tilde{\theta}|\tilde{\theta}}^{T} = q_{\tilde{\theta}}^{*}$ and $p_{\tilde{\theta}|\tilde{\theta}}^{T} = p_{\tilde{\theta}}^{*}$ and the standard mechanism design techniques yields the off-path revenue equivalence: $\forall \theta > \tilde{\theta}$,

$$p_{\tilde{\theta}|\theta}^{T} = \frac{1}{\beta}\left[\theta u\left(q_{\tilde{\theta}|\theta}^{T}\right) - \int_{\tilde{\theta}}^{\theta} u\left(q_{\tilde{\theta}|\hat{\theta}}^{T}\right) d\hat{\theta}\right] - \frac{1}{\beta}\theta u\left(q_{\tilde{\theta}}^{*}\right) + p_{\tilde{\theta}}^{*}.$$

Along with constraint (12) and the fact that the range of $u$ spans the whole positive reals, the threat allocations in $C_{\tilde{\theta}}$ are pinned down.

If $q_{\tilde{\theta}|\theta}^{T}$ is increasing in $\theta$, then the above construction satisfies (13). To check that $q_{\tilde{\theta}|\theta}^{T}$ is increasing in $\theta$, notice by Proposition 1 and (12), the off-path revenue equivalence can be rewritten as

$$\theta\left[u\left(q_{\tilde{\theta}|\theta}^{T}\right) - u\left(q_{\tilde{\theta}}^{*}\right)\right] = \frac{1}{1-\beta}\int_{\tilde{\theta}}^{\theta} u\left(q_{\tilde{\theta}|\hat{\theta}}^{T}\right) d\hat{\theta}.$$

Hence, $q_{\tilde{\theta}|\theta}^{T}$ is increasing in $\theta$ and (13) is satisfied.

The off-path threats are constructed so that all agents with $\theta \geq \tilde{\theta}$ prefer to report as $\theta$ rather than misreport as $\tilde{\theta}$. The process above can be repeated for all types $\hat{\theta} \geq \underline{\theta}$ so that all higher types, $\theta \geq \hat{\theta}$, are deterred to misreport as $\hat{\theta}$, which completes the proof. ■

**Proof of Lemma 2:** Lemma 1 implies $V(\sigma, \theta)$ is integrable in $\theta$. The ex-ante utility is

$$\begin{aligned} U(\sigma'|\sigma) &= \int_{\underline{\theta}}^{\overline{\theta}} \left[ \frac{1}{\beta} \theta u\left(q_{\sigma'}(\theta)\right) - p_{\sigma'}(\theta) - \left(\frac{1-\beta}{\beta}\right) \theta u\left(q_{\sigma'}(\theta)\right) \right] \pi(\theta|\sigma)\, d\theta \\ &= \frac{1}{\beta} V\left(\sigma', \underline{\theta}\right) + \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\theta} u\left(q_{\sigma'}(t)\right) \pi(\theta|\sigma)\, dt d\theta - \left(\frac{1-\beta}{\beta}\right) \int_{\underline{\theta}}^{\overline{\theta}} \theta u\left(q_{\sigma'}(\theta)\right) \pi(\theta|\sigma)\, d\theta \\ &= \frac{1}{\beta} V\left(\sigma', \underline{\theta}\right) + \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} \left[1 - \Pi(\theta|\sigma) - \theta(1-\beta)\pi(\theta|\sigma)\right] u\left(q_{\sigma'}(\theta)\right) d\theta. \end{aligned}$$

The second equality comes from Lemma 1. This completes the proof. ∎

**Proof of Lemma 3:** By Lemma 2, the ex-ante incentive compatibility constraints for $\sigma$ and $\sigma'$ with $\sigma > \sigma'$ are

$$U(\sigma) \geq U\left(\sigma'|\sigma\right) = U\left(\sigma'\right) + \int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi\left(\theta|\sigma'\right) - \Phi(\theta|\sigma)\right] u\left(q_{\sigma'}(\theta)\right) d\theta,$$

$$U\left(\sigma'\right) \geq U\left(\sigma|\sigma'\right) = U(\sigma) + \int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi(\theta|\sigma) - \Phi\left(\theta|\sigma'\right)\right] u\left(q_{\sigma}(\theta)\right) d\theta.$$

Integration by parts yields

$$\begin{aligned} &\int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi\left(\theta|\sigma'\right) - \Phi(\theta|\sigma)\right] u\left(q_{\sigma'}(\theta)\right) d\theta \\ &\qquad = \int_{\underline{\theta}}^{\overline{\theta}} \left[\Pi\left(\theta|\sigma'\right) - \Pi(\theta|\sigma)\right] \left[\beta u\left(q_{\sigma'}(\theta)\right) - (1-\beta)\theta u'\left(q_{\sigma'}(\theta)\right) \frac{\partial q_{\sigma'}(\theta)}{\partial \theta}\right] d\theta. \end{aligned}$$

By Lemma 1, $q_{\sigma'}(\theta)$ is non-decreasing. Therefore, if $\beta$ is large enough, then $\int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi\left(\theta|\sigma'\right) - \Phi(\theta|\sigma)\right] u\left(q_{\sigma'}(\theta)\right) d\theta \geq 0$, which implies that $U(\sigma) \geq 0$ when $U\left(\sigma'\right) \geq 0$ and $U(\sigma) \geq U\left(\sigma'|\sigma\right)$. By a similar process, if $\beta$ is small enough, then $\int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi(\theta|\sigma) - \Phi\left(\theta|\sigma'\right)\right] u\left(q_{\sigma}(\theta)\right) d\theta \geq 0$, which implies that $U\left(\sigma'\right) \geq 0$ when $U(\sigma) \geq 0$ and $U\left(\sigma'\right) \geq U\left(\sigma|\sigma'\right)$. ∎

**Proof of Theorem 1:** When $\beta$ is large, Lemma 3 implies that the individual rationality constraint for the $L$-consumers is binding and at least one of the following constraints is binding: the individual rationality constraint or the incentive compatibility constraint of the $H$-consumers. Conjecture the incentive compatibility constraint for the $L$-consumers is non-binding. Let $\lambda$ and $\xi$ denote the Lagrange multipliers on the incentive compatibility and the participation constraints for the $H$-consumers respectively. By Lemma 1, substituting in the expression for $p_\sigma(\theta)$, the Lagrangian is

$$
\begin{aligned}
&\frac{\gamma}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}\left\{\left[\theta-\frac{1-\Pi(\theta|H)}{\pi(\theta|H)}\right]u(q_H(\theta))-\beta c q_H(\theta)-V(H,\underline{\theta})\right\}\pi(\theta|H)\,d\theta\\
&+\frac{1-\gamma}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}\left\{\left[\theta-\frac{1-\Pi(\theta|L)}{\pi(\theta|L)}\right]u(q_L(\theta))-\beta c q_L(\theta)-V(L,\underline{\theta})\right\}\pi(\theta|L)\,d\theta\\
&+\frac{\lambda}{\beta}\left[V(H,\underline{\theta})-V(L,\underline{\theta})+\int_{\underline{\theta}}^{\overline{\theta}}[1-\Phi(\theta|H)]\left[u(q_H(\theta))-u(q_L(\theta))\right]d\theta\right]\\
&+\frac{\xi}{\beta}\left[V(H,\underline{\theta})+\int_{\underline{\theta}}^{\overline{\theta}}[1-\Phi(\theta|H)]\,u(q_H(\theta))\,d\theta\right],
\end{aligned}
$$

with $V(L,\underline{\theta})=-\int_{\underline{\theta}}^{\overline{\theta}}[1-\Phi(\theta|L)]\,u(q_L(\theta))\,d\theta$. This is a convex problem, because it can be reformulated such that the firm chooses $u_\sigma(\theta)=u(q_\sigma(\theta))$ and $q_\sigma(\theta)=u^{-1}(u_\sigma(\theta))$. Hence, the problem is strictly concave with a linear participation constraint for the $H$-consumers. The first order conditions are satisfied if

$$
\theta u'(q_H(\theta))=c \text{ and } \left[\theta-\frac{\lambda}{\beta(1-\gamma)}\left(\frac{\Phi(\theta|L)-\Phi(\theta|H)}{\pi(\theta|L)}\right)\right]u'(q_L(\theta))=c.
$$

Clearly, Assumption 1 guarantees $q_L(\theta)$ is non-decreasing in $\theta$.

Also, $\lambda>0$, because if $\lambda=0$ then the firm can increase $p_H(\underline{\theta})$. This implies $U(H)=U(L|H)$. From the first order condition on $V(H,\underline{\theta})$, $\lambda+\xi=\gamma$. Since $\xi\geq 0$, it must be that $\lambda\in(0,\gamma]$. By complementary slackness, if $\lambda=\gamma$ ($\xi=0$) then $U(H)\geq 0$. Also, if $\lambda<\gamma$ ($\xi>0$), then $U(H)=0$.

Finally, if $L$-consumers mimic $H$-consumers, the utility is

$$
U(H|L)=\frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}[\Phi(\theta|L)-\Phi(\theta|H)]\left[u(q_L(\theta))-u(q_H(\theta))\right]d\theta.
$$

By Assumption 1, if there is $\theta^c$ such that for any $\theta\geq\theta^c$, $\Phi(\theta|L)\leq\Phi(\theta|H)$, then the optimal contract has $q_L(\theta)\geq q_H(\theta)$ for any $\theta\geq\theta^c$. Similarly, for any $\theta\leq\theta^c$, $\Phi(\theta|L)\geq\Phi(\theta|H)$ and the optimal contract has $q_L(\theta)\leq q_H(\theta)$. This implies that $U(H|L)\leq 0$, so the incentive compatibility constraint for the $L$-consumers is satisfied. ■

**Proof of Corollary 1:** Let $g(\theta)=\Phi(\theta|L)-\Phi(\theta|H)$. Since $g$ is a continuous function and Assumption 2 is satisfied, the intermediate value theorem implies that there exists $\theta^c\in(\underline{\theta},\overline{\theta})$ such that $g(\theta)=0$. By Assumption 1, for any $\theta>\theta^c$, $\Phi(\theta|H)>\Phi(\theta|L)$, and if $\theta<\theta^c$, $\Phi(\theta|H)<\Phi(\theta|L)$. The result follows from Theorem 1. ■

**Proof of Corollary 2:** The optimal payment schedule satisfies consumers' first order condition: $\theta u'(q) = \beta p'(q)$. By Corollary 1, Theorem 3 and Lemma 1, the optimal payment schedules $p(q)$ have the following property: $p'_H(q) \le p'_L(q)$ for $q \le q^c$, $p'_H(q) > p'_L(q)$ for $q > q^c$. ■

**Proof of Theorem 2:** For large $\beta$, by Lemma 2, Lemma 3 and the fact that $\kappa > \frac{1}{2}$ and $\eta \le \gamma$, it is optimal to decrease $V(L, \underline{\theta})$ till $U(L) = 0$. Similarly, it is optimal to decrease $V(H, \underline{\theta})$ till $U(H) = U(L|H)$, so $U(H) = \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} [\Phi(\theta|L) - \Phi(\theta|H)] u(q_L(\theta)) d\theta$. The optimization problem can be rewritten as

$$\begin{aligned}&\frac{\kappa\gamma}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} \{\beta[\theta u(q_H(\theta)) - cq_H(\theta)]\pi(\theta|H) - [\Phi(\theta|L) - \Phi(\theta|H)] u(q_L(\theta))\} d\theta \\ &+ \frac{\kappa(1-\gamma)}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} \beta[\theta u(q_L(\theta)) - cq_L(\theta)]\pi(\theta|L) d\theta \\ &+ \frac{(1-\kappa)\eta + \xi}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} [\Phi(\theta|L) - \Phi(\theta|H)] u(q_L(\theta)) d\theta,\end{aligned}$$

where $\xi$ is the multiplier on the individual rationality constraint for $H$-consumers. The first order conditions are satisfied if

$$\theta u'(q_H(\theta)) = c \text{ and } \left[\theta - \frac{\lambda}{\beta(1-\gamma)}\left(\frac{\Phi(\theta|L) - \Phi(\theta|H)}{\pi(\theta|L)}\right)\right] u'(q_L(\theta)) = c,$$

where $\lambda = \gamma - \left(\frac{1-\kappa}{\kappa}\right)\eta - \frac{\xi}{\kappa}$ and $\xi \in [0, \kappa\gamma - (1-\gamma)\eta)$. ■

**Proof of Lemma 4:** Let sophistication $\hat{\beta} = \beta_n \in [\beta, 1]$. Lemma 2 and the ex-ante incentive compatibility constraint implies

$$U(\sigma') - U(\sigma) \ge \frac{1}{\beta_n}\int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi^n(\theta|\sigma) - \Phi^n(\theta|\sigma')\right] u(q_\sigma(\theta)) d\theta,$$

where $\Phi^n(\theta|\sigma) = \Pi(\theta|\sigma) + \theta(1-\beta_n)\pi(\theta|\sigma)$. Exchanging $\sigma'$ and $\sigma$ yields

$$U(\sigma') - U(\sigma) \le \frac{1}{\beta_n}\int_{\underline{\theta}}^{\overline{\theta}} \left[\Phi^n(\theta|\sigma) - \Phi^n(\theta|\sigma')\right] u(q_{\sigma'}(\theta)) d\theta.$$

By combining both inequalities and dividing them by $\sigma' - \sigma$ and let $\sigma' \to \sigma$, we get $\frac{\partial U(\sigma)}{\partial \sigma} = -\frac{1}{\beta_n}\int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial \Phi^n(\theta|\sigma)}{\partial \sigma} u(q_\sigma(\theta)) d\theta$. Finally, by combining both inequalities and dividing them by $(\sigma' - \sigma)^2$ and let $\sigma' \to \sigma$, we get $-\frac{1}{\beta_n}\int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial \Phi^n(\theta|\sigma)}{\partial \sigma}\frac{\partial q_\sigma(\theta)}{\partial \sigma} u'\left(q_\sigma(\theta)\right) d\theta \geq 0$. We obtain Lemma 4 when $\beta_n = 1$. ■

**Proof of Lemma 5:** Let sophistication $\hat{\beta} = \beta_n \in (\beta, 1]$. If the imaginary allocations are perceived ex-post incentive compatible, then the fooling constraints (6) and executability constraints (7) can be written as

$$\theta u\left(q_\sigma^I(\theta)\right) - \beta_n p_\sigma^I(\theta) \geq \max_{\theta'} \left\{\theta u\left(q_\sigma^R(\theta')\right) - \beta_n p_\sigma^R(\theta')\right\}. \tag{14}$$

$$\theta u\left(q_\sigma^R(\theta)\right) - \beta p_\sigma^R(\theta) \geq \max_{\theta'} \left\{\theta u\left(q_\sigma^I(\theta')\right) - \beta p_\sigma^I(\theta')\right\}. \tag{15}$$

Since the real allocations are ex-post incentive compatible, the non-sophisticated consumers at $t = 0$ predict their future-selves choosing to report

$$\frac{\beta}{\beta_n}\theta = \arg\max_{\theta'} \left[\theta u\left(q_\sigma^R(\theta')\right) - \beta_n p_\sigma^R(\theta')\right]$$

if they have realized type $\theta$ and chose the real allocations. Similarly, since the imaginary allocations are perceived ex-post incentive compatible, then at $t = 1$ the consumers would report $\frac{\beta_n}{\beta}\theta$ if they have realized type $\theta$ and chose the imaginary allocations. Hence, (14) and (15) can be simplified as

$$\theta u\left(q_\sigma^I(\theta)\right) - \beta_n p_\sigma^I(\theta) \geq \theta u\left(q_\sigma^R\left(\frac{\beta}{\beta_n}\theta\right)\right) - \beta_n p_\sigma^R\left(\frac{\beta}{\beta_n}\theta\right). \tag{16}$$

$$\theta u\left(q_\sigma^R(\theta)\right) - \beta p_\sigma^R(\theta) \geq \theta u\left(q_\sigma^I\left(\frac{\beta_n}{\beta}\theta\right)\right) - \beta p_\sigma^I\left(\frac{\beta_n}{\beta}\theta\right). \tag{17}$$

By Lemma 1 and revenue equivalence for belief $\beta_n$, $p_\sigma^I(\theta) = p_\sigma^I(\underline{\theta}) + \frac{1}{\beta_n}\left[\theta u\left(q_\sigma^I(\theta)\right) - \underline{\theta} u\left(q_\sigma^I(\underline{\theta})\right) - \int_{\underline{\theta}}^{\theta} u\left(q_\sigma^I(t)\right) dt\right]$, the following can be derived

$$V^R(\sigma, \theta) = V^R(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\theta} u\left(q_\sigma^R(t)\right) dt,$$

$$W^I(\sigma, \theta) = W^I(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\theta} u\left(q_\sigma^I(t)\right) dt,$$

$$\theta u\left(q_\sigma^I\left(\frac{\beta_n}{\beta}\theta\right)\right) - \beta p_\sigma^I\left(\frac{\beta_n}{\beta}\theta\right) = \frac{\beta}{\beta_n}\left[W^I(\sigma, \underline{\theta}) + \int_{\underline{\theta}}^{\frac{\beta_n}{\beta}\theta} u\left(q_\sigma^I(t)\right) dt\right],$$

$$\theta u\left(q_{\sigma}^{R}\left(\frac{\beta}{\beta_n}\theta\right)\right)-\beta_n p_{\sigma}^{R}\left(\frac{\beta}{\beta_n}\theta\right)=\frac{\beta_n}{\beta}\left[V^{R}\left(\sigma,\underline{\theta}\right)+\int_{\underline{\theta}}^{\frac{\beta}{\beta_n}\theta}u\left(q_{\sigma}^{R}\left(t\right)\right)dt\right],$$

which shows the fooling and executability constraints. Setting $\beta_n=1$ yields the results. ■

**Proof of Theorem 3:** Let sophistication $\hat{\beta}=\beta_n\in(\beta,1]$. The firm maximizes profit $\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[p_{\sigma}^{R}\left(\theta\right)-cq_{\sigma}^{R}\left(\theta\right)\right]\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)d\theta$ subject to the participation constraints: $U^{I}\left(\sigma\right)\geq 0$ for any $\sigma\in\Sigma$, the ex-ante incentive compatibility constraints: $U^{I}\left(\sigma\right)\geq U^{I}\left(\sigma'|\sigma\right)$ for any $\sigma,\sigma'\in\Sigma$, and the fooling and executability constraints: (10), (11).

Let $\beta_n\in(\beta,1]$ be the level of sophistication. By examining (10) and (11), (11) binds at the optimum. It is also clear that $U^{I}\left(\underline{\sigma}\right)=0$. The analysis proceeds by ignoring (10), which is verified at the end. With Lemma 1 and (11) binding,

$$p_{\sigma}^{R}\left(\theta\right)=\frac{1}{\beta}\theta u\left(q_{\sigma}^{R}\left(\theta\right)\right)-\frac{1}{\beta_n}\left[W^{I}\left(\sigma,\underline{\theta}\right)+\int_{\underline{\theta}}^{\frac{\beta_n\theta}{\beta}}u\left(q_{\sigma}^{I}\left(t\right)\right)dt\right].$$

Hence, the firm's objective is to maximize

$$\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[\theta u\left(q_{\sigma}^{R}\left(\theta\right)\right)-c\beta q_{\sigma}^{R}\left(\theta\right)\right]\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)d\theta \\ -\frac{\beta}{\beta_n}\int_{\underline{\sigma}}^{\overline{\sigma}}W^{I}\left(\sigma,\underline{\theta}\right)\gamma\left(\sigma\right)d\sigma-\frac{\beta}{\beta_n}\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\frac{\beta_n}{\beta}\theta}u\left(q_{\sigma}^{I}\left(t\right)\right)\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)dtd\theta d\sigma.$$

Immediately, the optimal real allocations satisfy $\theta u'\left(q_{\sigma}^{R}\left(\theta\right)\right)=c\beta$ for all $\sigma$ and $\theta$. Also, the imaginary allocations are chosen to maximize

$$-\int_{\underline{\sigma}}^{\overline{\sigma}}W^{I}\left(\sigma,\underline{\theta}\right)\gamma\left(\sigma\right)d\sigma-\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\frac{\beta_n}{\beta}\theta}u\left(q_{\sigma}^{I}\left(t\right)\right)\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)dtd\theta d\sigma.$$

Since $W^{I}\left(\sigma,\underline{\theta}\right)=\beta_n U^{I}\left(\sigma\right)-\int_{\underline{\theta}}^{\overline{\theta}}\left[1-\Phi^{n}\left(\theta|\sigma\right)\right]u\left(q_{\sigma}^{I}\left(\theta\right)\right)d\theta$, where $\Phi^{n}\left(\theta|\sigma\right)=\Pi\left(\theta|\sigma\right)+\theta\left(1-\beta_n\right)\pi\left(\theta|\sigma\right)$, then

$$\int_{\underline{\sigma}}^{\overline{\sigma}}W^{I}\left(\sigma,\underline{\theta}\right)\gamma\left(\sigma\right)d\sigma=\beta_n\int_{\underline{\sigma}}^{\overline{\sigma}}U^{I}\left(\sigma\right)\gamma\left(\sigma\right)d\sigma-\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[1-\Phi^{n}\left(\theta|\sigma\right)\right]u\left(q_{\sigma}^{I}\left(\theta\right)\right)\gamma\left(\sigma\right)d\theta d\sigma \\ =-\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[\frac{1-\Gamma\left(\sigma\right)}{\gamma\left(\sigma\right)}\frac{\partial\Phi^{n}\left(\theta|\sigma\right)}{\partial\sigma}+1-\Phi^{n}\left(\theta|\sigma\right)\right]u\left(q_{\sigma}^{I}\left(\theta\right)\right)\gamma\left(\sigma\right)d\theta d\sigma.$$

Integration by parts using Lemma 4 and $U^{I}\left(\underline{\sigma}\right)=0$ implies $\int_{\underline{\sigma}}^{\overline{\sigma}}U^{I}\left(\sigma\right)\gamma\left(\sigma\right)d\sigma$ can be rewritten as $-\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[1-\Gamma\left(\sigma\right)\right]\frac{\partial\Phi^{n}(\theta|\sigma)}{\partial\sigma}u\left(q_{\sigma}\left(\theta\right)\right)d\theta d\sigma$, which gives us the second equality above. Hence, when $\beta_n = 1$, $\Phi^{n}\left(\theta|\sigma\right)=\Pi\left(\theta|\sigma\right)$, so the optimal imaginary allocations are $q_{\sigma}^{I}\left(\theta\right)=0$ for any $\sigma\in\Sigma$ and $\theta\in\Theta$. The optimal payment scheme immediately follows from the revenue equivalence for both real and imaginary allocations.

To check the fooling constraints are satisfied, recall that (11) is binding at the optimum and if $q_{\sigma}^{I}\left(\theta\right)=0$, then $V^{R}\left(\sigma,\underline{\theta}\right)=-\int_{\underline{\theta}}^{\theta}u\left(q_{\sigma}^{R}\left(t\right)\right)dt$. From (10), for any $\sigma$, the fooling constraints are rewritten as

$$0\geq-\int_{\underline{\theta}}^{\theta}u\left(q_{\sigma}^{R}\left(t\right)\right)dt+\int_{\underline{\theta}}^{\beta\theta}u\left(q_{\sigma}^{R}\left(t\right)\right)dt.$$

Since the optimal $q_{\sigma}^{R}$ is increasing in $\theta$, the expression above is strictly negative. As a result, the fooling constraints hold. ∎

**Proof of Theorem 4:** The proof proceeds by first focusing on local ex-ante incentive constraints and $U\left(\underline{\sigma}\right)=0$, and ignoring other participation and ex-ante incentive compatibility constraints. It ends by checking that all participation constraints hold when $\beta$ is sufficiently large and $U\left(\underline{\sigma}\right)=0$ and Assumption 5 implies the ex-ante incentive compatibility constraints hold globally.

The firm maximizes expected profit $\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[p_{\sigma}\left(\theta\right)-cq_{\sigma}\left(\theta\right)\right]\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)d\theta d\sigma$, which can be rewritten as

$$\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[\theta u\left(q_{\sigma}\left(\theta\right)\right)-cq_{\sigma}\left(\theta\right)\right]\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)d\theta d\sigma-\int_{\underline{\sigma}}^{\overline{\sigma}}U\left(\sigma\right)\gamma\left(\sigma\right)d\sigma.$$

By integration by parts and Lemma 4 (from the proof, $\frac{\partial U(\sigma)}{\partial\sigma}=-\frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}}\frac{\partial\Phi(\theta|\sigma)}{\partial\sigma}u\left(q_{\sigma}\left(\theta\right)\right)d\theta$),

$$\int_{\underline{\sigma}}^{\overline{\sigma}}U\left(\sigma\right)\gamma\left(\sigma\right)d\sigma=U\left(\underline{\sigma}\right)-\frac{1}{\beta}\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[1-\Gamma\left(\sigma\right)\right]\frac{\partial\Phi\left(\theta|\sigma\right)}{\partial\sigma}u\left(q_{\sigma}\left(\theta\right)\right)d\theta d\sigma.$$

Therefore, the firm chooses $q_{\sigma}\left(\theta\right)$ to maximize

$$\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left\{\left[\theta+\frac{1-\Gamma\left(\sigma\right)}{\beta\gamma\left(\sigma\right)}\frac{\partial\Phi(\theta|\sigma)/\partial\sigma}{\pi\left(\theta|\sigma\right)}\right]u\left(q_{\sigma}\left(\theta\right)\right)-cq_{\sigma}\left(\theta\right)\right\}\pi\left(\theta|\sigma\right)\gamma\left(\sigma\right)d\theta d\sigma-U\left(\underline{\sigma}\right).$$

Clearly, at the optimum, $U\left(\underline{\sigma}\right)=0$. Also, the first order condition gives us the optimal quantity $q_{\sigma}\left(\theta\right)$ and Assumption 3 guarantees ex-post incentive compatibility.

Next, to see that Assumptions 4 and 5 imply that local ex-ante incentive compatibility is sufficient for global ex-ante incentive compatibility, notice that ex-ante incentive compatibility implies for any $\sigma, \sigma' \in \Sigma$,

$$U(\sigma) - U(\sigma'|\sigma) = U(\sigma) - U(\sigma') + U(\sigma') - U(\sigma'|\sigma) = \int_{\sigma'}^{\sigma} \left[ U'(\hat{\sigma}) - \frac{\partial U(\sigma'|\hat{\sigma})}{\partial \hat{\sigma}} \right] d\hat{\sigma} \geq 0.$$

By Lemma 2, $\frac{\partial U(\sigma'|\hat{\sigma})}{\partial \hat{\sigma}} = -\frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial \Phi(\theta|\hat{\sigma})}{\partial \hat{\sigma}} u(q_{\sigma'}(\theta)) d\sigma$, and along with Lemma 4, we have

$$U(\sigma) - U(\sigma'|\sigma) = \frac{1}{\beta} \int_{\sigma'}^{\sigma} \left[ u(q_{\sigma'}(\theta)) - u(q_{\hat{\sigma}}(\theta)) \right] \frac{\partial \Phi(\theta|\hat{\sigma})}{\partial \hat{\sigma}} d\hat{\sigma}.$$

For $\hat{\sigma} > \sigma'$ and $\theta < \theta^c$, Assumptions 4 and 5 guarantee $u(q_{\sigma'}(\theta)) \leq u(q_{\hat{\sigma}}(\theta))$, and with Assumption 3, $\frac{\partial \Phi(\theta|\hat{\sigma})}{\partial \hat{\sigma}} \leq 0$. For $\hat{\sigma} > \sigma'$ and $\theta > \theta^c$, $u(q_{\sigma'}(\theta)) \geq u(q_{\hat{\sigma}}(\theta))$, and with Assumption 3, $\frac{\partial \Phi(\theta|\hat{\sigma})}{\partial \hat{\sigma}} \geq 0$. This implies that $U(\sigma) \geq U(\sigma'|\sigma)$, so the contract is ex-ante incentive compatible for any downward deviation. A similar procedure also shows it is ex-ante incentive compatible for upward deviations.

Finally, participation constraints for any $\sigma > \underline{\sigma}$ is slack when $\beta$ is large. By Lemma 4,

$$\begin{aligned}
\frac{\partial U(\sigma)}{\partial \sigma} &= \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} -\frac{\partial \Phi(\theta|\sigma)}{\partial \sigma} u(q_{\sigma}(\theta)) d\theta \\
&= \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} -\left[ \frac{\partial \Pi(\theta|\sigma)}{\partial \sigma} + \theta(1-\beta) \frac{\partial \pi(\theta|\sigma)}{\partial \sigma} \right] u(q_{\sigma}(\theta)) d\theta \\
&= \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} -\frac{\partial \Pi(\theta|\sigma)}{\partial \sigma} u(q_{\sigma}(\theta)) d\theta \\
&\quad + \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} (1-\beta) \frac{\partial \Pi(\theta|\sigma)}{\partial \sigma} \left[ u(q_{\sigma}(\theta)) + \theta u'(q_{\sigma}(\theta)) \frac{\partial q_{\sigma}(\theta)}{\partial \theta} \right] d\theta \\
&= \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} -\frac{\partial \Pi(\theta|\sigma)}{\partial \sigma} \left[ \beta u(q_{\sigma}(\theta)) - \theta(1-\beta) \Xi(\sigma, \theta) \frac{u'(q_{\sigma}(\theta))}{u''(q_{\sigma}(\theta))} u'(q_{\sigma}(\theta)) \right] d\theta,
\end{aligned}$$

where $\Xi(\sigma, \theta) = -\left[ 1 - \frac{1-\Gamma(\sigma)}{\beta\gamma(\sigma)} \frac{\partial \Psi_{\sigma}(\theta)}{\partial \theta} \right] \left[ \theta - \frac{1-\Gamma(\sigma)}{\beta\gamma(\sigma)} \Psi_{\sigma}(\theta) \right]^{-1} \leq 0$. The third line comes from integration by parts, and the fourth line comes from implicit differentiation of the optimal quantity. It remains to show that when $\beta$ is sufficiently large, $\beta u(q_{\sigma}(\theta)) - \theta(1-\beta) \Xi(\sigma, \theta) \frac{u'(q_{\sigma}(\theta))}{u''(q_{\sigma}(\theta))} u'(q_{\sigma}(\theta)) \geq 0$. To see this, notice that by Assumption 3, $q_{\sigma}(\theta) > 0$ for any $\sigma \in \Sigma$ and $\theta \in \Theta$. Hence, there exists $M \in (0, \infty)$ such that $-\frac{u'(q_{\sigma}(\theta))}{u''(q_{\sigma}(\theta))} \frac{u'(q_{\sigma}(\theta))}{u(q_{\sigma}(\theta))} < M$. Also, $-\theta \Xi(\sigma, \theta) = \left[ \theta - \frac{1-\Gamma(\sigma)}{\beta\gamma(\sigma)} \Psi_{\sigma}(\theta) \epsilon_{\psi}(\sigma, \theta) \right] \left[ \theta - \frac{1-\Gamma(\sigma)}{\beta\gamma(\sigma)} \Psi_{\sigma}(\theta) \right]^{-1}$ is finite by Assumption 3. As a result, there exists $\bar{\beta}$ such that for any $\beta \geq \bar{\beta}$, $\frac{\partial U(\sigma)}{\partial \sigma} > 0$, so it is optimal to choose the contract such that $U(\underline{\sigma}) = 0$ and $U(\sigma) > 0$ for all $\sigma > \underline{\sigma}$. ■

# B Model of Late Fees

This appendix provides the main technical derivations for Section 4.2. Let the direct mechanism be $C = \{(p_{\sigma,1}(\theta), p_{\sigma,2}(\theta))\}_{\sigma\in\Sigma,\theta\in\Theta}$.

**Lemma 6** *For any $\sigma \in \Sigma$, $C$ is ex-post incentive compatible if and only if for any $\sigma \in \Sigma$ :*

*i. $p_{\sigma,1}(\theta)$ is non-increasing in $\theta$,*

*ii. $V(\sigma,\theta)$ is absolutely continuous in $\theta$, and $\frac{\partial V}{\partial \theta} = -\frac{1}{2}(p_{\sigma,1}(\theta))^2$ when differentiable.*

Lemma 6 yields the revenue equivalence:

$$p_{\sigma,2}(\theta) = \frac{1}{\beta}\left[-\frac{\theta}{2}(p_{\sigma,1}(\theta))^2 - V(\sigma,\underline{\theta}) + \frac{1}{2}\int_{\underline{\theta}}^{\theta}(p_{\sigma,1}(t))^2\,dt\right].$$

This gives us the following lemma.

**Lemma 7** *If $C$ is ex-post incentive compatible, then for any $\sigma, \sigma' \in \Sigma$,*

$$U(\sigma'|\sigma) = q + \frac{1}{\beta}V(\sigma',\underline{\theta}) - \frac{1}{2\beta}\int_{\underline{\theta}}^{\overline{\theta}}[1-\Phi(\theta|\sigma)](p_{\sigma,1}(\theta))^2\,d\theta,$$

*where $\Phi(\theta|\sigma) = \Pi(\theta|\sigma) + \theta(1-\beta)\pi(\theta|\sigma)$.*

The creditor is maximizing the sum of $\int_{\underline{\theta}}^{\overline{\theta}}[p_{\sigma,1}(\theta) + p_{\sigma,2}(\theta) - q]\,\pi(\theta|\sigma)\,d\theta$. From previous analysis, it is clear that at the optimum, $U(H) = 0$ and $U(L) = U(H|L) \geq 0$. The following theorem derives the optimal contract when $\beta$ is large. The proof is similar to the proof of Theorem 1, so it is omitted.

**Proposition 4** *If $\beta$ is sufficiently large, then the optimal contract has $p_{L,1}(\theta) = \frac{1}{\theta}$ and $p_{H,1}(\theta) = \left[\theta + \left(\frac{1-\gamma}{\beta\gamma}\right)\Psi(\theta)\right]^{-1}$, where $U(H) = 0$ and $U(L) = U(H|L) \geq 0$. The late payment schedule is $p_{\sigma,2}(\theta) = \frac{1}{\beta}\left[-\frac{\theta}{2}(p_{\sigma,1}(\theta))^2 - V(\sigma,\underline{\theta}) + \frac{1}{2}\int_{\underline{\theta}}^{\theta}(p_{\sigma,1}(t))^2\,dt\right]$ for all $\sigma \in \{L, H\}$.*

For exploitative contracts, the firm designs the contract so that it satisfies the fooling and executability constraints. Let $\beta_n \in (\beta, 1]$ denote the sophistication level. I continue to focus on perceived ex-post incentive compatible contracts. The fooling constraints are for any $\sigma \in \Sigma$ and $\theta \in \Theta$,

$$-\frac{\theta}{2}\left(p^I_{\sigma,1}(\theta)\right)^2 - \beta_n p^I_{\sigma,2}(\theta) \geq -\frac{\theta}{2}\left(p^R_{\sigma,1}\left(\frac{\beta}{\beta_n}\theta\right)\right)^2 - \beta_n p^R_{\sigma,2}\left(\frac{\beta}{\beta_n}\theta\right),$$

and the executability constraints are for any $\sigma \in \Sigma$ and $\theta \in \Theta$,

$$-\frac{\theta}{2}\left(p^R_{\sigma,1}(\theta)\right)^2 - \beta p^R_{\sigma,2}(\theta) \geq -\frac{\theta}{2}\left(p^I_{\sigma,1}\left(\frac{\beta_n}{\beta}\theta\right)\right)^2 - \beta p^I_{\sigma,2}\left(\frac{\beta_n}{\beta}\theta\right).$$

**Lemma 8** *If the real allocations are ex-post incentive compatible and the imaginary allocations are perceived ex-post incentive compatible, then the fooling constraints are, $\forall \sigma \in \Sigma$,*

$$W^I(\sigma, \underline{\theta}) - \frac{1}{2}\int_{\underline{\theta}}^{\theta}\left(p^I_{\sigma,1}(t)\right)^2 dt \geq \frac{\beta_n}{\beta}\left[V^R(\sigma,\underline{\theta}) - \frac{1}{2}\int_{\underline{\theta}}^{\frac{\beta}{\beta_n}\theta}\left(p^R_{\sigma,1}(t)\right)^2 dt\right], \tag{18}$$

*and the executability constraints are, $\forall \sigma \in \Sigma$ and $\forall \theta \in \Theta$,*

$$V^R(\sigma, \underline{\theta}) - \frac{1}{2}\int_{\underline{\theta}}^{\theta}\left(p^R_{\sigma,1}(t)\right)^2 dt \geq \frac{\beta}{\beta_n}\left[W^I(\sigma,\underline{\theta}) - \frac{1}{2}\int_{\underline{\theta}}^{\frac{\beta_n}{\beta}\theta}\left(p^I_{\sigma,1}(t)\right)^2 dt\right]. \tag{19}$$

Similar to previous analysis, the executability, downward ex-ante incentive compatibility and participation constraints for $L$-consumers are binding at the optimum. The firm chooses $p^R_{\sigma,1}$ and $p^I_{\sigma,1}$ to maximize

$$\int_{\underline{\sigma}}^{\overline{\sigma}}\int_{\underline{\theta}}^{\overline{\theta}}\left[p^R_{\sigma,1}(\theta) - \frac{\theta}{2\beta}\left(p^R_{\sigma,1}(\theta)\right)^2\right]\pi(\theta|\sigma)\gamma(\sigma)\,d\theta d\sigma - \int_{\underline{\sigma}}^{\overline{\sigma}}\gamma(\sigma)B^I(\sigma)\,d\sigma$$

where $B^I(\sigma) = q + \frac{1}{\beta_n}W^I(\sigma,\underline{\theta}) - \frac{1}{2\beta_n}\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\frac{\beta_n}{\beta}\theta}\left(p^I_{\sigma,1}(t)\right)^2\pi(\theta|\sigma)\,dtd\theta$. Notice that the firm wants to set $p^I_{\sigma,1}(\theta)$ as large as possible, so that consumers incorrectly believe that they will face a front-loaded payment scheme. The following proposition formalizes the exploitative contract with late fees.

**Proposition 5** *When $\beta_n = 1$, for all $\sigma \in \Sigma$, $p^R_{\sigma,1}(\theta) = \frac{\beta}{\theta}$ for any $\theta \in \Theta^R = \left[\frac{\beta}{\beta_n}\underline{\theta}, \overline{\theta}\right]$, and $p^I_{\sigma}$ is front-loaded.*