Contract Design and Self-Control: Theory and Evidence*

Stefano DellaVigna
University of California, Berkeley
sdellavi@econ.berkeley.edu

Ulrike Malmendier
Harvard University
malmendier_ulrike@gsb.stanford.edu

This version: February 3, 2003.

Abstract

How do rational firms respond to consumer biases? In this paper, we analyze the profit-maximizing contract design of firms if consumers have time-inconsistent preferences and are partially naive about it. We consider markets for two types of goods: goods with immediate costs and delayed benefits (investment goods) such as health club attendance, and goods with immediate benefits and delayed costs (leisure goods) such as credit card borrowing. We establish three features of the profit-maximizing contract design with partially naive time-inconsistent consumers. First, firms price investment goods below marginal cost. Second, firms price leisure goods above marginal cost. Third, in contracts with automatic renewal firms charge higher per-unit prices and additional fees after renewal, and make contract cancellation costly. These contractual features target the consumer misperception of future consumption and underestimation of the renewal probability. The predictions of the theory match the empirical contract design in the credit card, health club, mail order, mobile phone, newspaper, and vacation time-sharing industries. We also show that time inconsistency has adverse effects on consumer welfare only if consumers are naive.

*We are particularly grateful to Philippe Aghion, Edward Glaeser, Lawrence Katz, and David Laibson for their invaluable support and comments. We thank George Baker, Daniel Benjamin, Drew Fudenberg, Jerry Green, Oliver Hart, Caroline Hoxby, Markus Möbius, Daniele Paserman, Ben Polak, Andrei Shleifer and the participants of seminars at the Econometric Society Summer Meeting 2001 (University of Maryland), in the Economics Departments at UC Berkeley, Harvard, Northwestern, at Haas (EAP and Marketing), Harvard Business School (NOM), Kellogg (M&S), and Yale SOM for their comments. Tricia Glynn and Boris Nenchev provided excellent research assistance. For financial support, DellaVigna thanks Bank of Italy and Harvard University, Malmendier thanks Harvard University and the DAAD.

1 Introduction

A growing body of experimental evidence documents deviations from standard preferences and decision-making.[1] Many economists, however, remain unconvinced about the importance and validity of this evidence. Common objections point to artificial laboratory settings, low stakes, and the lack of repetition as reasons for the anomalous findings. Economists, after all, are interested in market interactions. In the market, high-stake incentives and repeated transactions are likely to discipline the deviations (Stigler, 1958; Becker, 1957).

A natural response to these concerns is to examine the impact of the deviations in a market setting. Odean (1998) and Genesove and Mayer (2001) among others take this route and provide seeming evidence of consumer biases in asset and housing markets. One can, however, take the logic of market interaction one step further. If consumers display systematic deviations, profit-maximizing firms should respond to them and offer contracts that target the deviations. This suggests that a key test for the relevance of the deviations is whether they explain the contractual design of the firms.

In this paper we use this contract-based approach to investigate the time consistency of consumer preferences. Since the pioneering contributions of Samuelson (1937) and Koopmans (1960), exponential time preferences have been a standard assumption in economics. These preferences are time-consistent because the discount rate between any two equidistant periods is constant.

The experimental evidence on intertemporal preferences calls this assumption into question. According to this evidence, the average discount rate for adjacent periods is higher in the near future than in the more distant future (Benzion, Rapoport and Yagil, 1989; Kirby, 1997; Kirby and Herrnstein, 1995). Individuals who display this discounting pattern in all periods have time-inconsistent preferences. Since their long-run plans may conflict with their actual decisions, these individuals have self-control problems. Collecting convincing market evidence about time preferences is crucial since time inconsistency can have large implications for intertemporal decisions in a wide range of fields such as addiction (Gruber and Koszegi, 2001), consumption (Laibson, 1997), tax policy (Krusell, Kuruscu, and Smith 2000), and economic growth (Barro, 1999).

In this paper we consider the interaction between profit-maximizing firms and consumers with time-inconsistent preferences. We derive the optimal contract design and compare it with observed features of contracts in several industries. We also consider the implications of the market interaction for consumer welfare.

Formally, we assume that the consumers have a hyperbolic discount function (Strotz, 1956; Phelps and Pollak, 1968; Laibson, 1997; O'Donoghue and Rabin, 1999a), with a (weakly) higher discount rate between the present and the next period than between any of the subsequent

[1] Camerer (1995) summarizes the evidence for individual decision making.

periods. Whenever the inequality is strict, this discount function implies time inconsistency, since the discount rate between two periods depends on the time of evaluation. We consider both agents who are sophisticated about their time inconsistency, as well as agents who are (partially) naive about it (O'Donoghue and Rabin, 2001). The latter assumption is motivated by the experimental evidence on overconfidence about positive personal attributes (Larwood and Whittaker, 1977; Svenson, 1981) and is consistent with field evidence on 401(k) investment (Madrian and Shea, 2001) and health club attendance (DellaVigna and Malmendier, 2001).

We assume that the firms can access data on consumer behavior, and therefore are aware of the time preferences of the consumers. Using this information, the firms design profit-maximizing contracts in order to sell goods to consumers. We consider firms that produce investment or leisure goods under monopoly or perfect competition. Investment goods have current costs and future benefits relative to the best alternative activity—health club attendance with current effort cost and future health benefits is one example. Leisure goods have current benefits and future costs relative to the best alternative activity—for example, credit card borrowing increases current consumption at the expense of future consumption. In Section 2 we obtain two main results about the profit-maximizing contractual design.

First, in a market for investment goods, a monopolistic firm prices below marginal cost if consumers have time-inconsistent preferences. The reason for the deviation from marginal cost pricing is twofold. Individuals that are sophisticated about their time-inconsistency look for commitment devices to increase their consumption of investment goods. The firm provides these devices in the form of low per-usage prices. Naive users, instead, overestimate their future self-control and therefore their usage of the investment goods. The firm offers contracts with discounts on the per-usage price, since the individuals overestimate the value of the discount. These results do not require monopoly power. Under perfect competition, firms price below marginal cost as well. In addition, the zero-profit condition implies that firms charge an initiation fee at sign-up.

Second, in a market for leisure activities the opposite result holds. Individuals with time-inconsistent preferences look for ways to limit their usage (if sophisticated) or they underestimate their usage (if naive). In both cases, competitive or monopolistic firms price usage above marginal cost. In addition, competitive firms offer a bonus.

In Section 3 we compare these results with the empirical features of contracts in several industries. The pattern of above- and below-marginal cost pricing predicted for time-inconsistent preferences fits well the observed features of contracts. The health club and the credit card industry are two examples. In the typical health club contract, users pay an enrollment fee and a flat fee but no price per visit, despite marginal costs per attendance to the club estimated between \$3 and \$6. This type of contract is puzzling for standard theory since it deviates from marginal cost pricing in a direction which conflicts with standard price discrimination. Frequent health club users, who presumably have the highest willingness to pay for exercising, pay the same flat fee as infrequent users and thus a lower price per visit. Price discrimination would instead predict that frequent users should pay more. Marginal cost pricing would align the incentives of the customers with the costs of the firms, and would allow firms to charge high attenders more than low attenders.

While health clubs price attendance below marginal cost, credit card companies price usage of the credit line above marginal cost, as indicated by the 20 percent premium on the resale of credit card debt (Ausubel, 1991). They also offer a bonus: users pay no annual fee and receive benefits such as car rental and luggage insurance. This pricing structure is puzzling since users with no outstanding balance pay a negative price for the credit card services.

In Section 4 we extend the model to contracts which are automatically renewed unless the consumer cancels. Automatic renewal is a common contractual feature in the newspaper, credit card, utility, and health club industries. We assume that firms choose the optimal contract design, including the transaction cost of cancellation, in a simple class of generalized two-part tariffs. For this class of contracts, the relevant distinction is not between time-consistent and time-inconsistent preferences, but rather between agents who are sophisticated about their time preferences (whether time consistent or not), and agents who are (partially) naive.

We show that, if consumers are sophisticated, the firms charge no renewal fee and create no cancellation cost, and the profit-maximizing pricing remains as in the previous model. If instead consumers are naive, the firms charge a fee upon each renewal, and, in the case of leisure goods, set a higher per-unit price in later periods. In addition, firms may introduce transaction costs of cancellation, even though these costs are a net loss from the point of view of the consumers and the firms. The intuition for these results is simple: naive individuals overestimate the cancellation probability, and the firms finds it advantageous to shift the pricing towards the later periods. The results for naive agents do not hinge on full naiveté: even if the agents are only partially naive, the firms charge fees at renewal and create additional transaction costs.

In Section 5 we collect evidence on the pricing of contracts with automatic renewal. Overall, fees at renewal and rising per-unit prices are common for contracts with automatic renewal, as predicted for naive consumers. Teaser rates for credit cards are a leading example (Laibson and Yariv, 2000). The typical credit card offer features a low initial interest rate on outstanding balances for an initial period, after which the interest rate rises to a high level. This introductory offer does not seem due to asymmetric information about quality, since the services provided by a credit card company are standard and highly observable.

The health club industry provides an example of automatic contract renewal with endogenous costs of cancellation. In the typical contract, membership is extended automatically from month to month. Customers who want to quit are typically required to mail a cancellation letter or to cancel in person, even though firms could make cancellation as easy as sending an email or giving a phone call. In other industries even the purchases are automatic. Members of book and compact disc clubs automatically receive the selection of the month and are charged for it, unless they send back a card. Efficiency reasons suggest that the consumers should choose whether and which products to purchase.

We also present evidence on the contractual design in the mobile phone, newspaper, and vacation time-sharing industries. Overall, the empirical contract design in these industries supports the view that time-inconsistency and partial naiveté are systematic, persistent, and widespread features of consumer preferences. If the share of consumers with these features were small or quickly vanishing, the firms would not deviate from marginal cost pricing, create renewal fees and cancellation costs. These latter features reduce the firm profits when consumers are time-consistent. Time-consistent agents, for example, can use teaser rate offers to borrow at a very low interest rate by switching card every six months. Credit card companies make losses on these agents.

In Section 6 we consider the welfare effects of time inconsistency in the market. We show that time inconsistency has adverse effects on consumer welfare only if consumers are naive. If the agents are (partially) naive, the firms design contracts that maximize the consumers' misperception of their future behavior, with two adverse effects on consumer welfare. First, the market outcomes are inefficient, since the firms do not maximize the actual consumer-firm surplus, but only the fictitious surplus. Second, under monopoly naiveté induces a redistribution of surplus from the naive agents to the firms. If, on the other hand, the agents are aware of their time inconsistency, the market interaction with the firms has favorable welfare effects. The firms offer a perfect commitment device, which enables the agents to achieve the efficient consumption level. In the market equilibrium, therefore, imperfect self-control has no effect on consumer welfare for sophisticated agents.

The contribution of this paper is threefold. First, we contribute to the literature on the general equilibrium analysis of near-rationality. Previous papers examined the interplay between rational and non-rational consumers (Akerlof and Yellen, 1985), particularly in asset markets (DeLong et al., 1990; Shleifer, 2000).[2] While these papers analyze the case in which rational and non-rational agents are on the same side of the market, we consider the asymmetric market interaction between rational firms and consumers with time-inconsistent preferences. Several results obtained in this paper for time-inconsistent agents are likely to apply more generally. In particular, deviations from rationality need not have any welfare effects in the market, as long as the consumers are fully aware of these deviations. As for the firm pricing, even a small amount of naiveté can generate large deviations in firms' responses.

As a second contribution, we provide field evidence on time inconsistency[3] and naiveté using the contractual design of firms to identify the time preferences of the consumers. A

[2]O'Donoghue and Rabin (1999b), Russell and Thaler (1985), and Benabou and Tirole (2000) also consider the interaction of standard and non-standard agents.

[3]Angeletos et al. (2001), DellaVigna and Malmendier (2001), DellaVigna and Paserman (2000), Gruber and Koszegi (2001), Fang and Silverman (2001) use field data to test for time consistency.

key feature of contract-based evidence is that it naturally incorporates concerns about the relevance of deviations for economics. Firms do not respond to consumer deviations that are not systematic, or are limited to small stakes. Moreover, firms have no incentives to change the contractual design if the fraction of users displaying deviations is small, or vanishing over time. The contract design, therefore, provides evidence on the prevalence of the deviation among the consumers.

Third, we employ time inconsistency and naiveté to suggest a new, unified explanation for the pricing in industries as diverse as the credit card, health club, mobile phone. We suggest that the intertemporal features of the goods are key determinants of the contract design. In addition, we characterize the pricing of contracts with automatic renewal.

The remainder of the paper is organized as follows. In Section 2 we present a simple model of market interaction between consumers and profit-maximizing firms. In Section 3 we compare the predictions of the model with stylized features of contract design in several industries. In Section 4 we extend the model to contracts with automatic renewal, and in Section 5 we compare once again the predictions of the model with observed contract design. In Section 6 we analyze the welfare implications of time-inconsistency and naiveté, and discuss the policy implications. In Section 7 we present alternative interpretations of the evidence on contract design and in Section 8 we conclude.

2 A simple model

In this Section, we set up a simple model that illustrates the basic intuitions about consumer behavior and firm pricing. We allow consumers to have time-inconsistent time preferences (with time-consistent preferences as the benchmark case) and examine the consequences on the profit-maximizing contract design.

2.1 The setting

In this model, there is a monopolistic firm and a consumer (Figure 1a). In period 0, the firm proposes a two-part tariff (L, p) to the consumer. If the consumer rejects, the firm makes zero profits and the consumer attains the reservation utility $\bar{u}$ at $t = 1$. If the consumer accepts, she pays the initiation fee L in period 1 to the firm. Upon accepting, the consumer learns her type c and then chooses her consumption in period 1. The agent can choose to consume C or not to consume (NC). If she chooses C, she pays p to the firm and incurs cost c in $t = 1$; she then receives b in $t = 2$. If she chooses NC, she attains payoff 0 in $t = 1$ and in $t = 2$.

Consumer. Good C provides payoff $-c$ when the activity is undertaken $(t = 1)$ and positive payoff $b > 0$ in the later periods $(t = 2)$. Compared to the alternative activity NC, therefore, consumption C is costly at present and provides benefits in the future. We call goods like C with a net positive future payoff relative to the alternative activity *investment goods*. Examples include activities beneficial to health (visits to a doctor, health club attendance, sports), to income (educational and vocational training, financial services and transactions) or to culture (theatre and concert attendance).

The investment activity C is an experience good. At $t = 0$, the agent knows the distribution F out of which the costs c are drawn. We assume that F has a strictly positive density f over $\mathbb{R}$. At the end of period 0 the agent learns the type c. The benefits b are deterministic and known from the outset.[4]

Intertemporal preferences. We assume that the agent has quasi-hyperbolic preferences (Phelps and Pollak 1968, Laibson 1997, O'Donoghue and Rabin 1999). The discount function for time s, when evaluated at period t, equals 1 for $s = t$ and equals $\beta\delta^{s-t}$ for $s = t+1, t+2, ...$ with $\beta \leq 1$. The present value of a flow of future utilities $(u_s)_{s \geq t}$ as of time t is

$$u_t + \beta \sum_{s=t+1}^{\infty} \delta^{s-t} u_s. \tag{1}$$

We can interpret β as the parameter of short-run discounting and δ as the parameter of long-run discounting. The standard *time-consistent exponential* model corresponds to the case where β is equal to 1. If β is smaller than 1, the individual exhibits time-varying discounting. The discount factor between the present period and the next period is $\beta\delta$, while the discount factor between any two periods in the future is simply δ. The difference between the short-run and the long-run discount factors generates *time inconsistency.*

We allow for consumers who overestimates their time consistency. A *partially naive hyperbolic* agent with parameters $(\beta, \hat{\beta}, \delta)$ (O'Donoghue and Rabin, 2001) expects (erroneously) to have the discount function $1, \hat{\beta}\delta, \hat{\beta}\delta^2, ...$ with $\beta \leq \hat{\beta} \leq 1$ in all future periods. The individual therefore correctly anticipates that she will have hyperbolic preferences in the future, but she overestimates the future parameter of short-run discounting if $\beta < \hat{\beta}$. The difference between the perceived and actual future short-run discount factor $\hat{\beta} - \beta$ reflects the *overconfidence* about future self-control.

Three special cases deserve mention. An *exponential* agent has time-consistent preferences ($\beta = 1$) and is aware of it ($\hat{\beta} = 1$). A *sophisticated* agent has time-inconsistent preferences ($\beta < 1$) and is aware of it ($\hat{\beta} = \beta$). A fully *naive* agent has time-inconsistent preferences ($\beta < 1$), but is completely unaware of it ($\hat{\beta} = 1$). She believes that she will behave like a time-consistent agent in the future.

Firm. The production costs consist of a set-up cost $K \geq 0$ that the firm incurs at $t = 1$ whenever a consumer signs the contract, and a per-unit cost $a \geq 0$, incurred at $t = 1$ whenever an agent consumes C. The monopolistic firm has all the bargaining power and offers a non-renegotiable contract to the consumer at $t = 0$. In this simple framework, the most general

[4]The results in this section do not change if we assume stochastic benefits b.

contract is a two-part pricing scheme, $(L,\, p)$ with a lump-sum initiation fee L due at $t = 1$ and a per-usage price p due at $t = 1$.

At time $t = 0$ the firm chooses the two-part tariff to maximize the discounted net present value of the expected future profits. Given that the firm can borrow and lend on the credit market, the discount factor is determined by the market interest rate r as $1/(1+r)$. We assume that this discount factor equals the long-run discount factor[5] for individuals, i.e., $1/(1+r) = \delta$. Finally, we suppose that the firm has complete knowledge of the individuals' preferences and that, as of $t = 0$, knows the cost distribution F.

2.2 Consumer behavior

In this setting, the consumer displays the following behavior. At $t = 0$, she assigns discounted utility $\beta\delta(\delta b - p - c)$ to consumption C and utility 0 to NC. Thus, she would like her future self to invest at $t = 1$, upon learning the type c, whenever $c \leq \delta b - p$.

A *time-inconsistent* agent, though, will undertake C less often than her previous self wishes. At the moment of deciding between C and NC, the net payoff from consuming C equals $\beta\delta b - p - c$. Therefore, at $t = 1$ she consumes C if $c \leq \beta\delta b - p$, i.e., with probability $F(\beta\delta b - p)$. The parameter of short-run impatience, β, determines the difference between desired and actual consumption probability $F(\delta b - p) - F(\beta\delta b - p)$. This difference is zero for individuals with time-consistent preferences ($\beta = 1$). The smaller is β, the larger is this difference and the more serious are the self-control problems.

A partially naive hyperbolic individual is not fully aware of her time-inconsistency. Therefore, as of $t = 0$, she overestimates the probability that her future self will consume C at $t = 1$. She expects that she will consume if $\hat{\beta}\delta b - p - c \geq 0$, i.e., with probability $F(\hat{\beta}\delta b - p)$. The difference between forecasted and actual consumption probability, $F(\hat{\beta}\delta b - p) - F(\beta\delta b - p) \geq 0$, is a measure of the consumer's *overconfidence*. Time-consistent ($\beta = \hat{\beta} = 1$) and sophisticated agents ($\beta = \hat{\beta} < 1$) have rational expectations about their future time preferences and display no overconfidence.

Given the above, at time 0 a consumer that signs the contract (L, p) expects to attain a net benefit $\beta\delta\left[-L + \int_{-\infty}^{\hat{\beta}\delta b - p}(\delta b - p - c)dF(c)\right]$ (notice the $\hat{\beta}$ in the integral).

2.3 Firm behavior: The profit-maximizing contract

If the consumer signs the contract, profits accrue to the monopolistic firm from the difference between the initiation fee L and the set-up cost K and from the per-usage net gain $p - a$. The firm earns this latter part of the profit whenever the user undertakes C, i.e., with

[5] In an economy populated mostly by hyperbolic agents, the market interest rate satisfies $1/(1+r) = \delta$ if either (i) the richest agents have exponential preferences with discount factor δ, or (ii) the richest agents are sophisticated hyperbolic individuals that can afford commitment devices.

probability $F(\beta\delta b - p)$. Therefore, as of $t = 0$, the expected per-usage net gain amounts to $\delta F(\beta\delta b - p)(p - a)$. The firm maximizes profits subject to the participation constraint that the discounted perceived utility equal the reservation utility $\beta\delta\overline{u}$. The maximization problem is thus[6]

$$\max_{L,p} \delta \{L - K + F(\beta\delta b - p)(p - a)\} \tag{2}$$

$$\text{s.t. } \beta\delta \left[-L + \int_{-\infty}^{\hat{\beta}\delta b - p} (\delta b - p - c)\, dF(c) \right] = \beta\delta\overline{u}. \tag{3}$$

Substituting for L in (2) yields:

$$\max_{p} \Pi(p) = \delta \left[\left(\int_{-\infty}^{\beta\delta b - p} (\delta b - a - c)\, dF(c) - K - \overline{u} \right) + \int_{\beta\delta b - p}^{\hat{\beta}\delta b - p} (\delta b - p - c)\, dF(c) \right], \tag{4}$$

where L^* is determined by (3). The first term of (4) is the actual *social* surplus from the market interaction of the two parties. The integrand consists of the net benefit to the user, $\delta b - c$, minus the cost for the provider, a. The per-usage price p does not appear in the integrand since it is a mere transfer between the two parties. The second term of (4) is the fictitious *consumer* surplus that reflects the overconfidence in consumption. It is null for time-consistent and sophisticated users ($\hat{\beta} = \beta$) and it increases with naiveté as measured by $\hat{\beta} - \beta$.

In order to guarantee existence of a profit-maximizing contract, we introduce a technical assumption that we maintain throughout the paper.

Assumption ABP. (Asymptotically bounded peaks) There is a pair $(M, z) \in I\!R^2$ such that $f(y'') \leq M f(y')$ for all y', y'' with $z < |y'| < |y''|$ and $y' \cdot y'' > 0$.

Assumption ABP rules out the anomalous case of unbounded peaks on the tails of $f(c)$ (*Figure 2b*). It is satisfied by all standard distribution functions.

Proposition 1 (Simple model, investment goods, monopoly) *Under monopoly, for $b > 0$, a profit-maximizing contract (L^*, p^*) exists. The per-usage price p^* equals marginal cost ($p^* = a$) for $\beta = 1$ and is set below marginal cost ($p^* < a$) for $\beta < 1$. The lump-sum fee L^* is set to satisfy the individual rationality constraint* (3).

Proof. We show that there exists a finite $\underline{M}$ such that the profit-maximing price p must lie in $[\underline{M}, a]$. The derivative of the profit function with respect to p is

$$\frac{\partial \Pi(p)}{\partial p} = \delta f(\beta\delta b - p) \left[(a - p) - (1 - \hat{\beta}) \frac{f(\hat{\beta}\delta b - p)}{f(\beta\delta b - p)} - \frac{F(\hat{\beta}\delta b - p) - F(\beta\delta b - p)}{f(\beta\delta b - p)} \right] \tag{5}$$

[6] We are assuming the existence of a two-part tariff (L', p') that satisfies the individual rationality constraint and that provides non-negative profits to the firm. Otherwise, there is no market for the investment good.

Given $f > 0$, $\partial\Pi(p)/\partial p < 0$ for $p > a$. Therefore any contract with $p > a$ (and L satisfying (3)) generates lower profits than the contract with $p = a$. We now use Assumption ABP to construct the lower bound $\underline{M}$. Consider a pair (z, M) satisfying Assumption ABP. For $p < \beta\delta b - z$, the second term in brackets in (5) is bounded below by $-\left(1-\hat{\beta}\right)\delta bM$ and the third term is bounded below by $-\left(\hat{\beta}-\beta\right)\delta bM$. Therefore, the expression in brackets is positive for $p^* < a + \min(-(1-\beta)\delta bM, \beta\delta b - z - a)$. The right-hand side of the inequality provides the lower bound $\underline{M}$, since any contract with p smaller than $\underline{M}$ generates lower profits than the contract with p equal to $\underline{M}$. The existence of a solution for p^* follows from continuity of the profit function on the compact set $[\underline{M}, a]$. Therefore, the profit-maximizing price p^* must satisfy $\partial\Pi(p^*)/\partial p = 0$. As for the value of p^*, $\beta < 1$ implies $\partial\Pi(p)/\partial p < 0$ at $p = a$ and therefore $p^* < a$. Finally, $\beta = \hat{\beta} = 1$ implies that $\partial\Pi(p)/\partial p = 0$ is satisfied only for $p = a$. **Q.E.D.**

The first order condition of (4) with respect to the per-usage price p can be rearranged to yield

$$p^* - a = -\left(1-\hat{\beta}\right)\delta b\frac{f\left(\hat{\beta}\delta b - p^*\right)}{f\left(\beta\delta b - p^*\right)} - \frac{F\left(\hat{\beta}\delta b - p^*\right) - F\left(\beta\delta b - p^*\right)}{f\left(\beta\delta b - p^*\right)}. \qquad (6)$$

For a consumer with standard *time consistent* preferences ($\beta = \hat{\beta} = 1$), the right-hand side of (6) is zero and the per-usage price p^* is set equal to the marginal cost a of the firm. The best policy for a firm facing an individual with perfect self-control is to align the incentives of the user with the cost for the firm. This guarantees that the agent undertakes C only if the investment generates positive surplus, that is, if $c \leq \delta b - a$.

For time-inconsistent users ($\beta < 1$), the optimal per-usage price p^* lies below a. Below marginal cost pricing occurs for two distinct reasons. The first is a commitment rationale. An individual who is at least partially aware of the time inconsistency looks for ways to increase the probability of future investment. Choosing a contract with low p is one such way. The first term in (6) expresses this rationale. The firm lowers p^* below a to the extent that the user is conscious about her future *time inconsistency*, as measured by $1-\hat{\beta}$. In fact, for a *sophisticated* individual ($\hat{\beta} = \beta$), below marginal cost pricing works as a perfect commitment device: given $p_s^* = a - (1-\beta)\delta b$, a sophisticated agent will consume if $c \leq \beta\delta b - p_s^* = \delta b - a$, i.e., as frequently as a time-consistent user with the same δ.

The second reason for pricing below marginal cost is the presence of naiveté. The firm knows that a (partially) naive user overestimates future consumption. Therefore, it offers a contract with a discount on p and an increase in L relative to the contract for a time-consistent agent. While the user would be indifferent between the two contracts if both were offered to her, the actual welfare is lower for the contract with a discount on p. The user will take advantage of the discount less often than she anticipates, and the firm will make higher profits. The firm sets p so as to increase this fictitious surplus, the second term in the profit function

(4). To a first approximation, this fictitious surplus depends on $F(\hat{\beta}\delta b - p) - F(\beta\delta b - p)$, the overestimation of future consumption. Figure 2a illustrates this for a distribution F with a peak in the right tail. The shaded area in the figure corresponds to $F(\hat{\beta}\delta b - p) - F(\beta\delta b - p)$. The higher is the density f, the larger is this shaded area, and the higher are the profits from the overestimation of consumption. By choosing p so that the shaded area coincides with the peak, therefore, the firm increases the fictitious surplus. Condition ABP rules out peaks in the extreme right tail of the distribution (Figure 2b) which might lead the firm to set an arbitrarily negative p.

Summing up, in equilibrium both types of hyperbolic agents pay a lower per-usage price than time-consistent consumers. To the extent that the individual is aware of the time-inconsistency, she prefers a low per-usage price as a commitment device that increases the usage in the future; the firm provides such device. If the user is at least partially overconfident about the strength of her future will-power, the firm exploits their misperception of the investment probabilities by tilting the pricing toward the initiation fee L.

Perfect competition. Do the previous results depend on the monopoly power of the firm? We now consider the effect of competition on the profit-maximizing contract. In problem (2)-(3), the degree of competition affects $\overline{u}$, the utility arising from the best alternative activity for the agent at time $t = 1$. Competing firms offer alternative contracts for the provision of the investment good C, and therefore raise $\overline{u}$. Even if all the other parameters in the model are constant, this may change the profit-maximizing contract. However, the solution for p^* in problem (2)-(3) does not depend on $\bar{u}$, as expression (4) shows. As a consequence, the below-marginal cost pricing of p^* does not depend on monopoly power.

The equilibrium level of L, on the other hand, does depend on $\overline{u}$ and therefore on the market structure. Under perfect competition, $\overline{u}$ is determined so as to equate expected profits to 0. For a time-consistent individual ($\beta = 1$), zero profits and $p^* = a$ imply that L^* equals the set-up cost K. For an individual with time-inconsistent preferences ($\beta < 1$), zero profits and $p^* < a$ imply that L^* is higher than K. Therefore, an individual with time-inconsistent preferences pays a larger initiation fee under the profit-maximizing contract than a time-consistent individual. The following Remark summarizes these results.

Remark 1 (Simple model, investment goods, perfect competition) *Under perfect competition, for $b > 0$, the profit-maximizing contract (L^*, p^*) exists. The per-usage price p^* equals marginal cost ($p^* = a$) for $\beta = 1$ and is set below marginal cost ($p^* < a$) for $\beta < 1$. The lump-sum fee L^* equals the set-up cost ($L^* = K$) for $\beta = 1$ and is set above the set-up cost ($L^* > K$) for $\beta < 1$.*

2.4 Leisure and neutral goods

The above model captures the case of activities with immediate benefits and delayed costs as well. Recall that C and NC are alternative ways to employ a given amount of resources (time and money). In contrast with the investment good case, we now assume that commodity C provides a higher payoff than NC at time 1 and a lower payoff at time 2. We call commodities with this front-loaded intertemporal profile *leisure goods.* Examples include goods that are harmful to future health (sweets, addictive goods), and goods that tempt the user to forego more productive activities (compact discs, cellular phone calls). With some abuse of notation, we employ the same variables as in the investment case (Figure 1a): the net payoff of C relative to NC is $-c$ in period 1 and b in period 2. Unlike in the investment case, however, $-c$ is the current benefit of C and $b < 0$ is the future cost. We otherwise maintain the same assumptions: the payoff $-c$ at time 1 (now, an immediate benefit) is stochastic, with F as the distribution of c, and the payoff b (now, a delayed cost) at time 2 is deterministic.

As for the consumption decision, at $t = 0$ the user desires to undertake C in the future with probability $F(\delta b - p)$. At $t = 1$, she ends up undertaking C with probability $F(\beta\delta b - p)$. Notice that for leisure goods time-inconsistency leads to *overconsumption:* $F(\beta\delta b - p) > F(\delta b - p)$. A partially naive user is not fully aware of the future overconsumption. She expects to undertake C with probability $F(\hat{\beta}\delta b - p) < F(\beta\delta b - p)$. She thus anticipates buying *fewer* leisure goods than they actually do. The difference between forecasted and actual consumption probability, $F(\hat{\beta}\delta b - p) - F(\beta\delta b - p) < 0$, is a measure of the *overconfidence*.

Similarly, we can deal with goods with no systematic intertemporal trade-off. We call *neutral goods* activities C with future payoff b equal to 0. Although formally this category is a knife-edge case between the case $b > 0$ (investment) and the case $b < 0$ (leisure), we can think of it as an approximation for goods that do not challenge self-control. For neutral goods desired, perceived and actual consumption coincide. These goods tempt the consumer to neither excessive consumption, as leisure goods, nor insufficient consumption, as investment goods.

Corollary 1 *(***Simple model, leisure and neutral goods***) For $b \leq 0$, the profit-maximizing contract (L^*, p^*) exists and has the following properties. (i) Under both monopoly and perfect competition, the per-usage price p^* equals marginal cost ($p^* = a$) for $\beta = 1$ or $b = 0$, and is set above marginal cost ($p^* > a$) for $\beta < 1$ and $b < 0$. (ii) Under perfect competition, the lump-sum fee L^* equals the set-up cost ($L^* = K$) for $\beta = 1$ or for $b = 0$, and is set below the set-up cost ($L^* < K$) for $\beta < 1$ and $b < 0$.*

Proof. The proof that the solution for p^* is interior follows along the lines of the proof of Proposition 1. Result (i) follows since the right-hand side of (6) is positive for $\beta < 1$ and $b < 0$. **Q.E.D.**

For a time-consistent user ($\beta = 1$), marginal cost pricing aligns incentives for the consumer and costs of the firms. For a time-inconsistent user ($\beta < 1$), above marginal cost pricing occurs for both a commitment and an overconfidence reason. To the extent that the agents is sophisticated, the high per-unit cost is a commitment device designed to solve the overconsumption problem. To the extent that the agent is naive, above marginal cost pricing is aimed at exploiting the underestimation of the number of purchases. Under perfect competition, the firms are likely to offer an initial bonus to a customer with weak will power, while they offer no bonus to a time-consistent user. Finally, the pricing of neutral goods equals the pricing for time-consistent agents: given the absence of temptation, marginal cost pricing is optimal despite the presence of limited self-control.

3 Evidence on contracts

This simple model has two testable implications about contractual features. For investment activities, it predicts an initiation fee and pricing below marginal cost. For leisure goods, it predicts a bonus (or a small initiation fee) at sign-up and above marginal cost pricing. We compare these predictions with empirical evidence on contract design in industries for investment or leisure goods. We consider cases in which firms access data on consumer behavior and therefore can price per usage. In Section 7 we consider some alternative interpretations of these facts.

3.1 Investment goods

Health club industry. The US health club industry is large and fast-growing. As of January 2001, 16,983 clubs were operating in the US. The industry revenues for the year 2000 totalled \$11.6bn, and the memberships in this same period summed to 32.8m, up from 17.4m in 1987. Fifty-one percent of the users were members in commercial health clubs, while thirty-four percent were members in non-profit facilities.[7] The upper part of Table 1 presents the top 10 commercial health clubs in the US. Only the market leader, Bally Total Fitness with \$1,007m revenues and 4m members, is publicly traded. Interestingly, few companies operate in more than 10 states. The lower part of Table 1 shows the Herfindahl and Concentration Ratio indices for the industry.[8] Ownership concentration is in the 10th percentile of US industries. In the highly competitive health club industry, the vast majority of clubs accounts for minor shares

[7]The remaining fifteen percent are members of miscellaneous for-profit facilities (corporate fitness centers, aerobics studios, resorts, spas, hotels, and country clubs.) Source: International Health, Racquet and Sportsclub Association (IHRSA), www.ihrsa.org.

[8]The Herfindahl index is computed using data on the 100 largest firms. Since the 100-th firm has a share of only 0.00031 of the industry revenue, the bias from neglecting the other firms is at most $0.00031^2 * 16,883 = 0.00162$.

in local markets.

Given the competitiveness of the industry, we expect the clubs to offer contracts that are approximately profit-maximizing. To document the types of contracts offered in a large urban market, we conducted during the months April-October 2001 a telephone survey of all the clubs in the metropolitan Boston area. The selection criterion for inclusion in the sample is the listing of the health club in the Yellow Pages for the year 2000 in the Boston area. Clubs belonging to the same company in different locations were contacted only once. After eliminating 38 companies that had gone out of business, changed telephone number, merged with another company or that serve only professional athletes, the final sample includes 67 companies, all of which provided the desired information. These companies overall operate 100 clubs in the Boston area.

The survey, whose transcript is in Appendix B, documents the menu of contracts offered by the health clubs. For each club, we divide contracts into two groups, Frequent and Infrequent contracts, to capture the frequency of adoption in the club. We classify a contract as 'Frequent' if it is mentioned initially in answer to the question "which contracts are available at the health cub", and as 'Infrequent' if it is mentioned later in the call or in response to a specific inquiry.[9]

Table 2 shows the results of the survey. Columns 1 to 3 present the unweighted results for the 67 companies. In Columns 4 to 6 we present the result by clubs. The health clubs in the Boston area offer three types of contracts:

1. *Monthly contract.* Across all companies, the user pays an average initiation fee of $128, and an average monthly fee of $57. These fees are somewhat higher when weighed by the number of clubs. The fee per visit is null.

2. *Annual contract.* Across all companies, members pay an average initiation fee of $73, a fee of $594 for one year, and no fee per visit. These numbers are 10 percent higher when weighed by clubs. We include in this category contracts with a commitment of two or three years; for these cases, we display the annual equivalent of the overall fee.

3. *Pay-per-visit contract.* Users of this contract pay a fee per visit that averages $11.30 across companies and $12 across clubs.

The middle part of Table 2 provides information on the menu of contracts. A large majority of the companies (Columns 1 to 3) offer the three contracts. There are, however, differences in the proxy for frequency of adoption. The monthly contract is a Frequent contract for a large fraction of the companies (47 out of 67), the annual contract for about half of them (33 out of 67), and the pay-per-visit contract for only 2 companies out of 67. The results are similar in the sample of all the clubs (Columns 4 to 6).

[9] We inquired about three specific types of contracts, in case they had not been mentioned: the monthly, the annual, and the pay-per-visit contract (details below).

In order to test the theory, we would like to compare the fee per visit, which is null under the two most common contracts, with the marginal cost of a visit for a club. For the clubs in DellaVigna and Malmendier (2001), we can estimate the latter as the total variable cost for a month divided by the total number of visits in the month. The resulting figure of $5 includes the cost of providing towels, personnel, and replacing broken machines, but it excludes congestion effects, which are sizeable at peak hours. Given the variation in personnel and towel provision, the marginal cost in the industry is likely to lie between $3 and $6, excluding congestion costs. Health clubs therefore price attendance substantially below marginal cost, as predicted by the model for users with time-inconsistent preferences.

Interestingly, the industry seems to have converged over time to below-marginal cost pricing. In the 50s, many health clubs operated under a pay-per-usage system with low operating costs: the clubs used coupons that the users could redeem at the entrance. This pricing scheme was largely abandoned by the 70s in favor of contracts with no per-visit price. The pay-per-visit type of contract is not common today even though the introduction since the 80s of electronic cards to track customer attendance would allow firms to charge per-visit fees at a small cost. One interpretation of this evolution of contracts is that firms in the health club industry have learned over time to design contracts that maximize profits given that the average user has limited self control.

Consumer behavior in this industry also conforms with a model of limited self-control. Using a panel data set on attendance and contract decisions of members of three US health clubs, DellaVigna and Malmendier (2001) show that consumers that pick monthly or annual contracts would on average have saved money paying per visit. This behavior can arise if consumer look for a commitment device or if they are naive about their future self-control, and therefore overestimate future attendance.

Vacation timesharing. Vacation timeshare companies such as Resort Condominiums International (RCI) and Hapimag offer members the opportunity to book one or more weeks of holiday in different resorts each year. The estimated sales of time-shares in year 2000 amount to $7.5 billion worldwide. Booking a holiday is an activity with current effort costs—planning of holiday time, location, and logistics—and future benefits, the holiday itself. Sophisticated individuals look for ways to ease the burden of this investment activity, such as reducing the monetary cost of the week. Naive individuals are overconfident about their ability to book. In both cases, firms should charge a high initial fee and below marginal cost pricing of the weeks of holiday. Indeed, users typically pay a large initial fee (on average, $11,000 in the US) to become members, and only a small fee for each week of holiday used: RCI charges an exchange fee of $140.[10]

Other examples of below-marginal cost pricing for investment goods include concert and opera tickets, and the enrollment at educational institutions. In both cases, there are immediate

[10]Amended Annual Report 10-K/A by Cendant Corporation for year 2000.

costs— foregoing more pressing or more leisurely activities to attend the event, foregoing current wages to enroll in an educational program— and future benefits in terms of culture, knowledge and skills. In both cases, usage is priced below marginal cost. Pre-paid season subscription with no fee per event are typical for concerts. Academic institutions typically require a lump-sum fee for the year and do not charge users based on the number of classes taken or attended.

3.2 Leisure goods

Credit cards. An easily accessible credit line has the intertemporal features of a leisure good: it allows credit-constrained individuals to increase current consumption (current benefit) at the expense of future consumption (future cost). Credit card ownership is an easy and widespread way to obtain credit in the US: the average credit card debt across all US households amounted to about $5,000 in 1998 (Laibson, Repetto, and Tobacman, forthcoming; Gross and Souleles, forthcoming). The model suggests that naive individuals underestimate the usage of the credit line, and that credit card companies respond by charging an interest rate on the outstanding balance above marginal cost. An additional prediction is that firms charge a low initial fee or even offer a bonus.

We test the predictions of the model using representative credit card offers from major US issuers.[11] As Table 3 shows, most credit card issuers charge an interest rate on outstanding balances that exceeds the prime rate by as much as 10 percentage points (Column 2). This high interest rate could be due to above-marginal-cost pricing of borrowing or to high default costs. In a pioneering paper, Ausubel (1991) provided clear evidence in favor of the first interpretation: credit card debt resells on the private market at a 20 percent premium. This implies that the interest rate on credit card debt, net of default and operating costs, exceeds the cost of capital by 20 percent. Interestingly, Ausubel was the first to suggest that overconfidence about future borrowing may explain the high rates of interest. Our model embeds this prediction in a general theory of leisure good pricing for individuals with time-inconsistent preferences and naiveté.

The second prediction of our model is also met. Despite sizeable costs of credit card accounts[12], firms offer a bonus to users. The issuers typically require no annual fee (Column

[11] We selected the largest 6 issuers ranked by outstandings as of 1997, excluding First Chicago NBD that merged with BancOne/FirstUSA (Evans and Schmalensee, 1999, p. 229). We included also information on Providian (12th largest issuer) and CapitalOne (8th largest issuer) because of the availability of information on the menu of cards offered. Finall, we included Discover and American Express that are the biggest issuers outside the Visa and Mastercard circle. We collected the information on the main credit card offer(s) from each issuer's website.

[12] The operating expenses, net of the revenues from the interchange fee, average 3.4 percent of outstanding balances in 1997 (calculation of the authors using data from Evans and Schmalensee, 1999, p. 249). These expenses do not include borrowing and default costs.

3), and in addition offer valuable benefits such as car rental and luggage insurance (for the Platinum offers), discounts on future purchases or cash back (Column 4).[13] As a consequence, credit card companies make losses on users with no outstanding balances, who can use the card for transactions, make use of the benefits, and borrow for up to 30 days at no cost.

Cellular phones. Cellular phones are widespread in the US. As of year 2000, the mobile phone industry had $52.5bn revenues and 109.5m subscribers. The cellular phone, a convenient communication tool as well as the latest gadget, tempts users with limited self-control to spend time on the telephone rather than in other more productive activities. Furthermore, users are tempted to neglect the future health hazards. At the moment of signing a contract, sophisticated users look for a way to commit themselves to a lower usage and naive users underestimate the number of future calls. Cellular phone companies in the US provide a pricing scheme that fits both types. In the typical contract, the consumers choose a monthly airtime allowance within a menu. Table 4 shows the annual revenue in millions of dollars (Column 1), the menu of allowances (Column 2) and the monthly fee (Column 3) for three major cellular phone companies. The allowances and rates are for daytime, weekday minutes of call for plans with domestic long distance included; additional weekend and night-time minutes (where available) are not included. The marginal price for minutes beyond the allowance (Column 5) is typically two to four times higher than the average price of a call within the limit (Column 6). The high marginal cost of calls beyond the limit conforms to the theory for leisure goods. Sophisticated users appreciate the high marginal cost over the limit as a commitment device. Naive users underestimate their usage, set too low a limit and end up paying high sums for calls above the monthly allowance.

4 Renewal decision

In the credit card, mail order, newspaper, and utilities industries firms often offer long-term contracts which are automatically renewed unless the consumer pays a transaction cost to cancel. For these types of contracts firms charge higher per-unit prices and additional fees after renewal, and often create endogenous cancellation costs. In order to explain these puzzling facts, we extend the model of Section 2 to include a renewal decision. Unlike for the simple model, the profit-maximizing contract design will allow us to tell apart the sophistication and the naiveté hypotheses.

[13] In the absence of profits from lending, credit card companies would have to charge an annual fee to recover their costs. Therefore charge cards, on which the balance has to be paid in full within 30 days, should require substantial annual fees. In fact, as the last row in Table 1 shows, the American Express charge card, the most common in the US, has an annual fee of $55 or $75 (if Gold).

4.1 The setting

Timing. Figure 1b illustrates the timing of the extended model. The choices in periods 0 and 1 are as in the simple model. In addition, at the end of period 1 the agent has the option to renew the contract for one more period. If she renews (R), she pays the lump-sum renewal fee L_R at time 2 and becomes entitled to choose either C or NC in period 2. If she quits, she pays an effort cost $k \geq 0$ at time 1 and obtains 0 utility thereafter.

Consumers. As in the simple model, the consumer chooses between activities NC and C at each period of enrollment, in this case $t = 1, 2$. Relative to activity NC, activity C provides an immediate payoff $-c_t$ at time t and delayed payoffs with a net present value of b as of time $t+1$. We analyze the cases of *investment goods* ($b > 0$), *leisure goods* ($b > 0$), and *neutral goods* ($b > 0$). We assume that the consumer in period 0 knows only the distribution F from which c_1 is drawn. In period 1, before choosing C or NC, the agent experiences the realization of cost c_1, with $c_1 = c_2 = c$. Therefore, we omit the time subscript on c. The payoff b is known ex ante.

Firms. The firm incurs a set-up cost $K \geq 0$ for enrollment at $t = 1$ and a per-unit cost $a \geq 0$ whenever an agent consumes C. The firm has all the bargaining power and offers a non-renegotiable contract to the customer at $t = 0$. The contract, a generalized two-part tariff[14] $((L, p), (L_R, p_R))$, specifies a lump-sum amount L due for sign-up at $t = 1$, and a price per usage p at $t = 1$. The new contractual elements are L_R, a lump-sum amount due at $t = 2$ if the agent *R*enews, and p_R, the per-usage price at $t = 2$ after *R*enewal. We analyze both the cases of monopoly and of perfect competition at $t = 0$. The results in this Section hold for both market assumptions, unless noted otherwise.

4.2 Consumer behavior

The decision to consume C has the same features as in Section 2. The new element of consumer behavior is the renewal decision R. At the end of period 1, the agent compares the cost of quitting $-k$ with the benefit of renewing the contract. If she renews, she pays the lump-sum fee L_R at $t = 2$ and she expects to achieve the benefits of consumption $(\delta b - p_R - c)$ if $c \leq \hat{\beta}\delta b - p_R$ (notice the $\hat{\beta}$ in this expression). Therefore, renewal R occurs if

$$\beta\delta[-L_R + \mathbf{1}_{\{c \leq \hat{\beta}\delta b - p_R\}}(\delta b - p_R - c)] \geq -k \tag{7}$$

where $\mathbf{1}_X$ is the indicator function for set X. Define $R^{\beta,\hat{\beta},\delta} = \{c | \text{expression (7) holds}\}$ to be the set of values of c for which an individual with preferences $(\beta, \hat{\beta}, \delta)$ renews at $t = 1$. It is easy to show that for a given $\hat{\beta}$ the set $R^{\beta,\hat{\beta},\delta}$ is decreasing in β, $R^{\beta_1,\hat{\beta},\delta} \supseteq R^{\beta_2,\hat{\beta},\delta}$ for $\beta_1 < \beta_2 \leq \hat{\beta}$. More limited self-control (lower β) is associated with a higher renewal probability.

[14] This contract is the most general contract under the restriction that the firm cannot condition the contract on observed consumer choices. If we allow for conditioning, the firm would use it only if the consumer is naive.

At time 0, when signing the contract, the agent forms expectations about the future renewal behavior. Conditional on a realization of c, the agent expects to renew at $t = 1$ if

$$\hat{\beta}\delta[-L_R + \mathbf{1}_{\{c \leq \hat{\beta}\delta b - p_R\}}(\delta b - p_R - c)] \geq -k. \tag{8}$$

While time-consistent and sophisticated individuals ($\beta = \hat{\beta}$) correctly anticipate the renewal probability, partially naive agents ($\beta < \hat{\beta}$) underestimate it. This is not surprising: partially naive individuals overestimate the probability of undertaking activities with current costs and delayed benefits, such as contract cancellation. Formally, an agent with preferences $(\beta, \hat{\beta}, \delta)$ expects to renew if c belongs to $R_t^{\hat{\beta},\hat{\beta},\delta}$. Therefore, the set of payoffs c for which the agent renews against her expectations is $R^{\beta,\hat{\beta},\delta} \backslash R^{\hat{\beta},\hat{\beta},\delta}$, which is well defined given that the set $R^{\beta,\hat{\beta},\delta}$ is decreasing in β. As we will see below, the underestimation of renewal plays a key role in the determination of the profit-maximizing contract.

4.3 Firm behavior: The profit-maximizing contract

At $t = 0$, the firm chooses the profit-maximizing contract $((L^*, p^*), (L_R^*, p_R^*))$ to maximize the discounted net present value of the expected future profits. In period 1, as in the simple model, profits accrue from the initiation fee net of the set-up cost, $L - K$, and from the expected net gain from consumption, $(p - a)F(\beta\delta b - p)$. In period 2, the firm earns the renewal fee L_R with probability $P(R^{\beta,\hat{\beta},\delta})$, and the net profit from consumption, $(p_R - a)$ if the agent renews and consumes C, i.e., with probability $P(R^{\beta,\hat{\beta},\delta} \cap \{c | c \leq \beta\delta b - p_R\})$.

The firm maximizes profits subject to the individual rationality constraint as perceived by the individual agent. The agent expects to attain the net utility of consumption in period 1, to pay the renewal fee L_R with probability $P(R^{\hat{\beta},\hat{\beta},\delta})$ (notice the additional $\hat{\beta}$), and to pay the cancellation cost k otherwise. Conditional on renewal, the agent expects to attain the utility of consumption in period 2. The maximization problem for the firm at time 0 is thus

$$\max_{L,L_R,p,p_R} \delta \begin{bmatrix} L - K + \int_{-\infty}^{\beta\delta b - p} (p-a)\, dF(c) \\ +\delta \left[L_R \cdot P(R^{\beta,\hat{\beta},\delta}) + \int_{R^{\beta,\hat{\beta},\delta} \cap \{c | c \leq \beta\delta b - p_R\}} (p_R - a)\, dF(c) \right] \end{bmatrix} \tag{9}$$

$$\text{s.t. } \beta\delta \begin{bmatrix} -L + \int_{-\infty}^{\hat{\beta}\delta b - p} (\delta b - p - c)\, dF(c) + \delta[-L_R \cdot P(R^{\hat{\beta},\hat{\beta},\delta}) \\ + \int_{R^{\hat{\beta},\hat{\beta},\delta} \cap \{c | c \leq \hat{\beta}\delta b - p_R\}} (\delta b - p_R - c)\, dF(c) - \frac{k}{\delta} P(c \notin R^{\hat{\beta},\hat{\beta},\delta})] \end{bmatrix} = \beta\delta\overline{u}. \tag{10}$$

As in Section 2.3, we focus on the case in which there is a market for the good. We assume the existence of a contract that satisfies (10) and guarantees non-negative profits to the firm. Program (9-10) reduces to the following surplus maximization:

$$\max_{L_R,p,p_R} \delta \left[\int_{-\infty}^{\beta\delta b - p} (p-a)\, dF(c) + \int_{-\infty}^{\hat{\beta}\delta b - p} (\delta b - p - c)\, dF(c) - \overline{u} - K \right] \tag{11}$$

$$+\delta^2 \left[\begin{array}{c} L_R \cdot P\left(R^{\beta,\hat{\beta},\delta} \backslash R^{\hat{\beta},\hat{\beta},\delta}\right) - \frac{k}{\delta} P\left(c \notin R^{\hat{\beta},\hat{\beta},\delta}\right) + \\ \int_{R^{\beta,\hat{\beta},\delta} \cap \{c | c \leq \beta\delta b - p_R\}} (p_R - a)\, dF(c) + \int_{R^{\hat{\beta},\hat{\beta},\delta} \cap \left\{c | c \leq \hat{\beta}\delta b - p_R\right\}} (\delta b - p_R - c)\, dF(c) \end{array} \right],$$

where L^* is determined by the individual rationality constraint in (10). Notice that expression (11) is additively separable in p—the first line—and (L_R, p_R)—the second line. We can therefore solve for p^* and defer the solution for (L_R^*, p_R^*) until the next Subsection. The first line in program (11) coincides with the maximization problem (4) in the simple model. Therefore, as the following Remark states, the profit-maximizing per-usage price p^* equals the one in the simple model of Section 2.

Remark 2 (Model with renewal, initial per-usage price) *The set of solutions for p^* coincides with the set of solutions in the simple model in (4). Therefore, for $\beta = 1$, $p^* = a$. For $\beta < 1$, $p^* < a$ for investment goods ($b > 0$), $p^* > a$ for leisure goods ($b < 0$), and $p^* = a$ for neutral goods ($b = 0$).*

In the period before renewal, the profit-maximization is driven only by static forces: the need to align incentives for the agent with the marginal cost of the firm (time-consistent agents), the provision of commitment (sophisticated agents), and the presence of overconfidence (naive agents). For users with time-inconsistent preferences we thus get the familiar result of below marginal cost pricing for investment goods, above marginal cost pricing for leisure goods and marginal cost pricing for neutral goods.

4.3.1 Time-consistent and sophisticated agents

We consider now the solution for (L_R^*, p_R^*) for time-consistent and sophisticated hyperbolic agents ($\beta = \hat{\beta} \leq 1$). For these agents, expected and actual cancellation and consumption probabilities coincide. Therefore, the second line of surplus maximization (11) reduces to

$$\max_{L_R, p_R} \delta^2 \left[\int_{R^{\beta,\beta,\delta} \cap \{c | c \leq \beta\delta b - p_R\}} (\delta b - a - c)\, dF(c) - \frac{k}{\delta} P(c \notin R^{\beta,\beta,\delta}) \right] \tag{12}$$

where L^* is determined by (10). Proposition 2 characterizes the solution for p_R^*, L^*, and L_R^*.

Proposition 2 *(***Model with renewal, time-consistent and sophisticated agents***) For $k > 0$, the profit-maximizing contract $((L^*, p^*)(L_R^*, p_R^*))$ for sophisticated agents ($\beta = \hat{\beta}$) has the following features:*

(i) The per-usage price after renewal p_R^* is

$$p_R^* = p^* = a - (1 - \beta)\, \delta b.$$

(ii) For $\beta = 1$, $p_R^* = p^* = a$. For $\beta < 1$, $p_R^* = p^* < a$ for investment goods, $p_R^* = p^* > a$ for leisure goods, and $p_R^* = p^* = a$ for neutral goods.

(iii) The renewal fee L_R^* is any $L_R \in I\!R$ such that

$$L_R \leq k/\beta\delta + \min\{(1-\beta)\,\delta b, 0\}\,. \tag{13}$$

(iv) Under perfect competition, the initiation fee is $L^* = K + [F(\delta b - a)(1-\beta)\,\delta b](1+\delta) - \delta L_R^*$.

Proof. Consider the expression (12). The set $\{c|c \leq \delta b - a\}$ is the essentially unique (up to sets of measure 0) maximizer of $\int_S (\delta b - a - c)\,dF(c)$ with respect to S. Similarly, the empty set is the essentially unique maximizer of $-\frac{k}{\delta}P(N)$ with respect to N. Therefore, all pairs (L_R, p_R) that satisfy both $R^{\beta,\beta,\delta} \cap \{c|c \leq \beta\delta b - p_R\} = \{c|c \leq \delta b - a\}$ and $R^{\beta,\beta,\delta} = I\!R$ are the maximizers of (12), if any such pair exists. Denote by M the set of these maximizers. For given L_R and p_R, these conditions are equivalent to $\{c|c \leq \beta\delta b - p_R\} = \{c|c \leq \delta b - a\}$ and $R^{\beta,\beta,\delta} = I\!R$. As a consequence, $M = \{(p_R, L_R) \mid p_R = a - (1-\beta)\,\delta b,\ L_R \leq k/\beta\delta + \min[(1-\beta)\,\delta b, 0]\}$, where the last inequality guarantees $R^{\beta,\beta,\delta} = I\!R$. This proves (i), (ii), and (iii). Statement (iv) is a consequence of equalizing profits to zero and setting $p_R^* = p^* = a - (1-\beta)\,\delta b$. **Q.E.D.**

For time-consistent and sophisticated agents the per-usage price equals $a - (1-\beta)\,\delta b$ both before and after renewal. This choice of p^* and p_R^* ensures that at $t = 1$ and at $t = 2$ the individuals with positive firm-agent surplus (such that $c \leq \delta b - a$) consume the good. In addition, L_R^* is set low enough so that all the individuals renew (see expression (7)) and no agent bears the effort cost k, a net waste from the point of view of the two parties. The equilibrium initiation fee L^* depends on the level of the renewal fee L_R^*. For a given L_R^*, the initiation fee is increasing in b for $\beta < 1$. Therefore, it is highest for investment goods, and lowest (possibly a bonus) for leisure goods.

Overall, the profit-maximizing contract for time-consistent and sophisticated individuals is as in the simple model. The per-visit prices are the same, and the renewal fee in the model with renewal is low enough that the renewal decision becomes trivial. Given that both agents have rational expectations about the renewal probability, it is in the interest of the firm to minimize the probability that the users incur the cancellation costs k. As we will see, it is the lack of rational expectations about quitting that leads to a distinctly different optimal contract for naive agents.

4.3.2 Naive agents

For (fully) naive hyperbolic consumers ($\beta < \hat{\beta} = 1$), the maximization problem of the firm with respect to p_R and L_R is

$$\max_{L_R, p_R} \delta^2 L_R \cdot P\left(R^{\beta,1,\delta} \backslash R^{1,1,\delta}\right) - \delta^2 \frac{k}{\delta} P\left(c \notin R^{1,1,\delta}\right) \tag{14}$$
$$+\delta^2 \left[\int_{R^{\beta,1,\delta} \cap \{c|c \leq \beta\delta b - p_R\}} (p_R - a)\,dF(c) + \int_{R^{1,1,\delta} \cap \left\{c|c \leq \hat{\beta}\delta b - p_R\right\}} (\delta b - p_R - c)\,dF(c)\right].$$

Compared to the case of time-consistent and sophisticated agents, the firm earns profits from two additional sources. First, naive agents underestimate the renewal probability, and therefore the payment of $L_{R,}$ by $P(R^{\beta,1,\delta}\backslash R^{1,1,\delta})$. The firm maximizes the revenue from the underestimation, captured in the first line of (14). Second, the firm extracts the fictitious surplus from the consumer misperception about consumption at $t = 2$, as in the second line of (14). The following Proposition characterizes the profit-maximizing contract.

Proposition 3 (Model with renewal, naive agents) *For $k > 0$, the profit-maximizing contract $((L^*, p^*), (L_R^*, p_R^*))$ for naive agents ($\beta < \hat{\beta} = 1$) has the following features:*
(i) The renewal fee L_R^* satisfies

$$L_R^* = k/\beta\delta \tag{15}$$

for leisure and neutral goods ($b \leq 0$) and

$$L_R^* \geq k/\beta\delta \tag{16}$$

for investment goods ($b > 0$).
(ii) A solution for p_R^* exists. For leisure and neutral goods ($b \leq 0$), p_R^* is a solution of the equation

$$p_R^* - a = -\frac{F\left(\delta b - p_R^* - \frac{1-\beta}{\beta\delta}k\right) - F\left(\beta\delta b - p_R^*\right)}{f(\beta\delta b - p_R^*)}. \tag{17}$$

The per-usage price after renewal is higher than before renewal: $p_R^* > p^* \geq a$. For investment goods ($b > 0$), under Condition A1 in Appendix A, we observe below-marginal cost pricing after renewal: $p_R^* < a$.
(iii) Under perfect competition, for leisure and neutral goods ($b \leq 0$) the initiation fee L satisfies $L < K$.

Proof. The expressions for L_R^* and p_R^* are derived in Proposition A.1 in Appendix A. To prove $p_R^* > p^*$ for leisure and neutral goods ($b \leq 0$) we employ the concept of monotone comparative statics (Milgrom and Roberts, 1990). Consider the profit function Π after imposing $L_R^* = k/\beta\delta$, the optimal value for L_R. The profit function Π depends on one maximization variable, p, and one parameter, k. For $k = 0$, Π coincides with the static profit function that generates p^*, and therefore, p_R^* equals p^*. For $k \geq 0$, Π satisfies the sufficient condition for monotone comparative statics of p_R as a function of k since $(\partial^2\Pi/\partial p_R \partial k) = f(\delta b - p_R^* - \frac{1-\beta}{\beta\delta}k) > 0$. By Theorem 5 from Milgrom and Roberts (1990), p_R^* is an increasing function of k for all $k \geq 0$. We have therefore proven $p_R^* > p^*$ for $k > 0$. **Q.E.D.**

Proposition 3 states three results. First, the firm chooses L_R^* to extract maximal revenue from the underestimation of renewal, as captured by $L_R \cdot P(R^{\beta,1,\delta}\backslash R^{1,1,\delta})$ in (14). For leisure and neutral goods ($b \leq 0$), this involves setting L_R^* equal to $k/\beta\delta$. For this level fully naive users always renew (expression (7)) but they expect to renew only if $c \leq \delta b - p_R^* - (L_R^* - k/\delta)$ (expression (8)). For $c > \delta b - p_R^* - (L_R^* - k/\delta)$, therefore, the users anticipate quitting and require ex ante a compensation for the cost k; in reality, however, they renew and pay L_R^* to the firm. For this level of c the firm makes a net gain as of time 0 of $\delta [\delta L_R^* - k]$, which is positive since $L_R^* = k/\beta\delta > k/\delta$.

For investment goods ($b > 0$), the optimal renewal fee L_R^* may be higher than $k/\beta\delta$. When increasing L_R past $k/\beta\delta$, the firm can extract more revenue, but it also faces a cost: the higher L_R induces the users with $c \geq \delta b - p_R$ to quit and to pay cancellation cost k. For investment goods the loss from the quitters is limited: the users dropping out ($c \geq \delta b - p_R$) would not have consumed C since $\beta\delta b - p_R < \delta b - p_R$. The firm may find it profitable to set $L_R^* > k/\beta\delta$. For leisure goods, instead, the loss is large, since some of the users that quit—those with $\delta b - p_R \leq c \leq \beta\delta b - p_R$—would have consumed C at the high marginal price $p_R^* > a$. The firm therefore sets $L_R^* = k/\beta\delta$.

Second, for leisure and neutral goods the firm increases the per-usage price after renewal ($p_R^* > p^*$) to maximize the fictitious surplus from the consumption at $t = 2$. Compare the fictitious surplus for $t = 1$ and for $t = 2$. At $t = 1$ the agent simply underestimates the probability of consumption of the leisure good. In contrast, at $t = 2$ the agent underestimates *both* the renewal probability and the probability of consumption conditional on renewal. The numerator of expression (17) captures this double underestimation. At $t = 2$ the agent expects to consume C whenever she renews, i.e., if $c \leq \delta b - p_R^* - (L_R^* - k/\delta)$, and she truly consumes C if $c \leq \beta\delta b - p_R^*$, since she always renews. Since $L_R^* > k/\delta$, the overall underestimation of future consumption is higher at $t = 2$ than at $t = 1$. The firm responds to the increased underestimation after renewal by raising the price per usage at $t = 2$ even more above marginal cost. Hence, $p_R^* > p^* > a$. For investment goods, two opposing forces drive the pricing of the per-usage price p_R: overestimation of future consumption conditional on renewal and underestimation of renewal. The firms respond to overestimation of consumption by reducing p_R, and to underestimation of renewal by increasing p_R. Condition A1 guarantees that the latter force does not dominate; therefore, below marginal cost pricing ensues.

Third, under perfect competition, firms offer a bonus at sign-up for leisure and neutral goods, i.e., $L^* < K$. The firms compensate users ex ante for the above marginal cost pricing of p_R^* and p^*, and for the high renewal fee L_R^*. For investment goods the sign of $L^* - K$ is ambiguous. The below marginal cost pricing of p^* and p_R^* (under Condition A1) drives the initial fee up, while the high renewal fee L_R has the opposite effect.

To sum up, the model with renewal replicates most results of the simple model for time-inconsistent agents—above marginal cost pricing and bonus for leisure goods, below marginal cost pricing for investment goods. More important, two new features of the profit-maximizing contract distinguish naive from sophisticated agents—high renewal fee L_R, and scaling up of the per-usage price p_R for leisure goods. Naive agents underestimate the likelihood of renewal, and the firms charge higher per-unit prices and additional fees after renewal. Interestingly, these latter features of the pricing extend to neutral goods. Firms tailor contracts to respond to time-inconsistency, even if the consumption good has a flat intertemporal profile. Given that cancellation is an activity with current costs and future benefits, firms set contracts so as to respond to the overestimation of cancellation.

4.3.3 Endogenous cancellation cost

In Section 4.3 we derived the profit-maximizing contract for a given level of the cost k. Denote by $\Pi(k)$ the corresponding profit level as a function of k. In this Section, we allow the firm to choose the profit-maximizing transaction cost k and we examine the effect on $\Pi(k)$. We assume that the firm can increase the transaction cost k of cancellation from a minimum level $\underline{k} \geq 0$ up to a maximum level $\overline{k}$; formally, $k \in [\underline{k}, \overline{k}]$. Examples of endogenous cancellation costs are the request of additional documents or of in-person cancellation. The following Proposition (proved in Appendix A) summarizes the results.

Proposition 4 (Model with renewal, endogenous cancellation cost) *(i) For time-consistent and sophisticated agents ($\beta = \hat{\beta} \leq 1$), the optimal profit $\Pi(k)$ is independent of k. (ii) For naive agents ($\beta < \hat{\beta} = 1$), the optimal profit $\Pi(k)$ is bounded below by a function $\underline{\Pi}(k)$ which satisfies $\Pi(0) = \underline{\Pi}(0)$ and is strictly increasing in k; in addition, for leisure ($b < 0$) and neutral goods ($b = 0$) $\Pi(k)$ is strictly increasing in k, so $k^* = \overline{k}$.*

In the optimal contract for time-consistent and sophisticated agents the firm is indifferent regarding the level of k because in equilibrium no cancellation takes place.[15] In the optimal contract for naive agents, instead, the firm creates additional cancellation costs k, even though these transaction costs may decrease the actual surplus. A higher k allows the firm to charge a higher L_R, and therefore to obtain more revenue from the underestimation of renewal. For leisure goods, the firm profits $\Pi(k)$ are increasing in k, so $k^* = \overline{k}$. For investment goods, the relationship between k and the profits is not necessarily monotonic. Nevertheless, Proposition 4(ii) implies that the firm wants to increase k beyond 0 whenever $\underline{k} = 0$.

4.3.4 Profit-maximizing contract with partial naiveté

So far we have characterized the polar cases of time-consistent, sophisticated, and naive agents. We now consider the general features of the profit-maximizing contract for partially naive hyperbolic agents ($\beta < \hat{\beta} \leq 1$). We show that the optimal contract for partial naive agents

[15] In a more general setting where contract renewal is somewhat costly for the firm, some users cancel under the optimal contract and therefore the firm strictly prefers to set $k^* = \underline{k}$.

resembles closely the one for fully naive agents. For simplicity, we focus on the case of investment and neutral goods ($b \geq 0$).

Proposition 5 (Model with renewal, partially naive agents) *For partially naive hyperbolic agents ($\beta < \hat{\beta} \leq 1$) in a market for investment or neutral goods ($b \geq 0$), the renewal fee L_R^* satisfies*

$$L_R^* \geq k/\beta\delta. \tag{18}$$

Moreover, the optimal profit $\Pi(k)$ is bounded below by a function $\underline{\Pi}(k)$ which satisfies $\Pi(0) = \underline{\Pi}(0)$ and is strictly increasing in k.

Proof. In Appendix A. **Q.E.D.**

The pricing of both the renewal fee L_R^* and the cancellation cost k^* is essentially the same as for fully naive agents. First, L_R^* is set to take maximal advantage of consumer underestimation of renewal, as captured by $L_R P\left(R^{\beta,\hat{\beta},\delta} \backslash R^{\hat{\beta},\hat{\beta},\delta}\right)$ in (11). For $L_R^* \geq k/\beta\delta$ partially naive users renew more than they expect to (expressions (7) and (8)). With probability $P\left(R^{\beta,\hat{\beta},\delta} \backslash R^{\hat{\beta},\hat{\beta},\delta}\right) > 0$, therefore, the firm makes a net gain of $\delta\left[\delta L_R^* - k\right]$, which is positive since $L_R^* > k/\delta$. Second, the profit function $\Pi(k)$ is increasing in k at least for k close enough to 0. The firm therefore creates some additional transaction costs if $\underline{k} = 0$.

Interestingly, the profit-maximizing contract is discontinuous with respect to the introduction of an arbitrarily small amount of naiveté. While the renewal fee for sophisticated agents is essentially 0, the renewal fee for partially naive agents is high enough to maximize the underestimation of cancellation. While the firm has no incentive to generate transaction costs for sophisticated agents, it does so for partially naive agents. This qualitative difference does not depend on the size of the naiveté, and it does not go away as $\hat{\beta}$ converges to β.

The discontinuity in the optimal contract feeds back into the consumer behavior, and may induce a discontinuity in the consumer behavior. As Proposition 5 states, even a small amount of naiveté may induce firms to set $L_R^* > k/\beta\delta$—for this level of L_R^*, all the partially naive consumers with $c < \delta b - p_R$ quit. No cancellation would instead take place in a market with fully sophisticated agents.

This discontinuity in behavior illustrates a general difference between non-market and market behavior. In the absence of firms, small deviations from rational expectations generate unnoticeable differences in the behavior of the agents. So long as the agents face 'random' tasks, in very few cases the deviation makes a difference. In a market setting this conclusion need not hold, since the tasks are highly non-random. The firms, aware of the consumer weakness, offer contracts which are explicitly designed to target the deviation, no matter how small. The consumer, therefore, faces tasks which systematically magnify the effect of the deviation.

4.4 Long-term contracts

So far we have presented a simple model with two periods of consumption and one renewal option. We now sketch an extension to multiple periods of consumption, and multiple renewals. We explore some new predictions for the profit-maximizing contract design.

Consider the timing in Figure 1c. The agent at $t = 0$ signs a contract for T periods, and at time T has the option to renew the contract for additional T periods, or to cancel at an immediate transaction cost $k \geq 0$. Conditional on renewal, the agent faces a similar choice at $t = 2T$, and so on for infinite periods. In each of the T periods of enrollment, the agent decides whether to consume or not. For simplicity, we assume that the firm offers a generalized two-part tariff $((L, p), (L_R, p_R))$, where L is the initiation fee, and p is the per-usage price in the first T periods. L_R is the fee paid by the agent at the period after each renewal, and p_R is the per-usage price in the T periods after each renewal.

The contract design for time-consistent, sophisticated, and naive agents has the same qualitative features as in the baseline model with renewal, but the magnitudes differ. The next Proposition characterizes the renewal fee. The detailed solution is in Appendix A.

Proposition 6 (Model with renewal, long-term contracts) *For time-consistent ($\beta = \hat{\beta} = 1$) or sophisticated agents ($\beta = \hat{\beta} \leq 1$), the renewal fee L_R^* is any $L_R \in I\!R$ such that*

$$L_R \leq \frac{1 - \delta^T}{\beta\delta} k + \min\left\{\frac{1 - \delta^T}{1 - \delta}(1 - \beta)\delta b, 0\right\}. \quad (19)$$

For naive agents ($\beta < \hat{\beta} = 1$), the renewal fee L_R^* satisfies

$$L_R^* \geq \frac{1 - \beta\delta^T}{\beta\delta} k. \quad (20)$$

Inequality (19) for sophisticated agents corresponds to (13), and inequality (20) for naive agents corresponds to (15) and (16) in the simple model with renewal. This extended model lends itself to a simple calibration exercise on the renewal fee L_R. Consider a contract with monthly renewal (time unit = 1 day, $T = 30$), a daily discount factor $\delta = .99985$ corresponding to an annual discount factor of .947, and a short-run discount factor $\beta = .8$. Inequality (19) implies that the renewal fee for sophisticated agents is at most as high as $\left(1 - \delta^T\right) k/\beta\delta \approx .0056k$. Inequality (20) implies that the renewal fee for naive agents is at least as high as $\left(1 - \beta\delta^T\right) k/\beta\delta \approx .254k$. The renewal fee for naive agents is at least one order of magnitude higher than the one for sophisticated agents.

This model with long-term contract, therefore, adds a quantitative prediction: we expect a 'high' renewal fee with naive agents, and an essentially null renewal fee for time-consistent and sophisticated agents.

5 Evidence on contracts with automatic renewal

Summary. The profit-maximizing contract design of the simple model generalizes to the introduction of a renewal option. In the presence of time-inconsistent preferences we expect above marginal cost pricing for leisure goods and below marginal cost pricing for investment goods. Furthermore, initiation fees are more likely for investment than for leisure goods. If consumers have perfect self-control, instead, marginal cost pricing and initial fees equal to the set-up costs prevail.

In addition, the above Section provides an array of new predictions that distinguish between agents with rational expectations (whether time-consistent or sophisticated) and naive agents. Naive agents underestimate the probability that they will renew the contract. The firms respond to this underestimation by charging higher per-unit prices and additional high fees after renewal. Firms also create additional transaction costs for cancellation, even though these costs are a burden for the consumer. Sophisticated and time-consistent agents, instead, are fully aware of their future renewal behavior. As a consequence, the firms have no interest in charging higher per-unit prices and additional fees after renewal that would induce some individuals with positive surplus to quit. Moreover, firms do not generate cancellation costs: individuals anticipate the possibility of paying these costs in the future, and require an ex ante compensation. Importantly, the predictions for naive agents do not hinge on full naiveté: even if the agents are only partially naive, the firms charge fees at renewal and create additional transaction costs.

This allows us to test for the prevalence of time-inconsistency and naiveté. If only a small portion of the individuals displays naiveté, or if naiveté is a short-lived phenomenon, we should not observe pricing tilted toward the later periods and endogenous cancellation costs. The firms would make higher profits offering the contract for time-consistent or sophisticated users. Therefore, if firms do choose these contractual features, the contract design provides qualitative evidence that time-inconsistency and naiveté are common and persistent features of consumers preferences. We collect evidence from several industries on the design of contracts with automatic renewal. In Section 7 we discuss some alternative interpretations of these facts.

Credit cards. The interest rate on outstanding balances in the credit card industry rises after an initial period. As Table 3 shows, the typical credit card features a low interest rate for borrowing (Column 5) for an introductory period of typically six months (Column 6), followed by a high interest rate for the subsequent period (Column 2). The renewal after the introductory period is automatic, and the consumer can quit at any time. Presumably, these cards target naive users who expect to cancel once the interest rate spikes but delay doing so. Laibson and Yariv (2000) were the first to provide an explanation of this phenomenon based on naiveté and to model specific features of these offers. Alternative explanations of teaser rate offers based on asymmetric information are less plausible: the services provided by the company are standard and highly observable. Also, teaser rate offers generate an adverse selection of consumers: rational consumers use the credit card to borrow at the initial low rate, and switch to another card at the end of the introductory period. In a world without overestimation of quitting, the prevalence of these offers would constitute a puzzle. An additional feature of the pricing is that credit card issuers typically charge sizeable fees to consumers who either delay the monthly payments (Column 7), or run over the credit limit (Column 8).

Mail order industry. Compact disc, book and DVD clubs keep members updated about new offers and mail the purchases to them. Sales of records through music clubs accounted for 14.7% of all albums sold in the US in the mid 1990s. The two main companies, BMG and ColumbiaHouse, had $1.6bn revenues in year 2000. Members of these clubs receive automatically the periodic selection of the club and are charged for it, unless they return a card to turn down the offer.[16] Given that the purchase of these leisure goods happens by default, the theory for naive users predicts an initial bonus and a high marginal price for the selection. In fact, these clubs offer free goods (4 books, 11 compact discs) as an initial gift to new users, and market additional purchases at high prices: the CDs of the month for October and November 2001 at BMG sold at a $1.50 premium over the Amazon price. These prices are well above marginal cost, given the cost savings on limited royalties and cheap production methods. This pricing scheme with initial bonus and high price of subsequent purchases allows for a buying strategy that could be very costly for the firms: users can enroll, get the free goods, do the minimum required purchases, quit and (possibly) reenroll. The profitability of the contractual design of the clubs depends on the prevalence among the users of naive agents that remain enrolled and purchase several high-priced selections.

Health club industry. As Column 4 in Table 2 shows, the monthly contract is a frequent contract in 73 health clubs in the Boston area out of 100. This contract is automatically renewed from month to month, with an average monthly fee of $61 debited on the credit card or back account. Users can quit the club at any month. Out of the 83 clubs that offer this contract, all the clubs allow in-person cancellation, and 48 clubs allow cancellation by letter (certified letter for 27 clubs). Only 6 clubs accept cancellation by phone. This contract matches the features of the optimal contract for naive agents. The contract has a sizeable renewal fee (the monthly membership fee) which is automatically debited unless the consumer cancels, and endogenous cancellation costs (no phone cancellation, and frequently no cancellation by letter).

The prevalence of the monthly contract raises a question. Why do firms offer contracts with automatic renewal in the first place? While the model does not explicitly address this point, it is clear that, if consumers are naive, firms prefer automatic renewal to automatic expiration. Users that are not attending may delay cancellation and keep paying fees to the club. DellaVigna and Malmendier (2001) indeed find that the average delay between the last

[16] Recently, these companies have allowed consumers to use their website to decline the periodic selection.

attendance and contract termination for a monthly contract is two full months.

An alternative, perhaps simpler explanation for the automatic renewal is the quest for efficiency. Automatic renewal is likely the efficient default for a contract with a monthly duration. Efficiency reasons, however, can account only partially for the prevalence of automatic renewal. Health clubs could easily devise a more efficient contract. They could automatically renew the membership of attenders, but stop charging users who have not attended the health club for more than, say, three months. Users who wanted to restart could do so, by paying the membership fee again. Health clubs have the information to implement this contract, since they typically require consumers to swipe an electronic card at the entrance. Despite the gain in efficiency, to our knowledge no US club offers this contract as an option.

As Column 5 in Table 2 shows, the second most common contract is the annual contract, offered by 47 clubs as a typical contract. Although in four fifths of the cases this contract is terminated at the end of the year, the trend in the industry is to offer annual contracts with automatic renewal into a monthly contract after 12 months.[17]

Newspaper subscriptions. Major US newspapers such as the New York Times, the Washington Post, and the Boston Globe offer subscriptions with automatic renewal from week to week, with an option to quit at any time. As in the credit card example, the pricing is consistent with the theory for partially naive agents: the price is low in the first 8 to 12 weeks (50 percent off—New York Times, Boston Globe—or 90 percent off—Washington Post), after which the full subscription price applies. Asymmetric information about quality is a less likely explanation of the initial offer: journals are commonly available in libraries, cafes, and public places. An additional interesting feature is that consumers get the low price for four additional weeks if they pay by credit card instead of by check. Arguably, the transaction costs of cancellation relative to renewal are higher if renewal is automatic (credit card billing) than if it requires sending a check.

Video rental. Consumers who rent a movie have to take an action to return it. Automatic renewal of the video is the default. If users are naive companies should charge a high fee for late returns. Until February 2000, Blockbuster, the market leader with $4.96bn revenues as of year 2000, charged for each day of delay as much as for the previous five-day rental.[18] From February 2000, the fee for any delay (up to five days) is the same as the fee for the previous five-day rental.

One can extend the model of this section to encompass other examples of contract design directed toward naive users. In general, firms target the overestimation of the probability of undertaking a costly transaction. In vacation timesharing and frequent flyer programs, members that miss the deadline to book the holiday resort[19] or to redeem the miles lose

[17] Personal communication, Bill Howland, IHRSA Director of Public Relations.

[18] Late fees accounted for 16.7 percent of the company revenues in the first quarter of year 2000.

[19] Interestingly, vacation timesharing companies require booking well in advance of the actual holiday, eleven

their benefits. Naive users, while signing up in these programs, overestimate the likelihood of using the services, and the firms find this contract design profitable. Similarly, money-back return guarantees, in addition to having a signaling function, may be targeted at naive users who overestimate the probability of returning the good in case of dissatisfaction. Finally, the diffusion of mail-in coupons may depend on the prevalence of naive users that expect to send the coupon but end up not doing so.[20]

6 Profits and welfare

In this Section we address two questions on welfare within the simple model of Section 2. First, what is the effect of time-inconsistency and naiveté on profits and welfare? Second, are there feasible policies to increase consumer welfare? We also hint at the welfare effects of the general market interaction of profit-maximizing firms and consumers with non-standard features.

6.1 Time-inconsistency and naiveté

Definitions. The monopoly profits $\Pi_M^{\beta,\hat{\beta},\delta}$ are given by expression (4) evaluated at $p = p^*$, and the perfect competition profits satisfy $\Pi_{PC}^{\beta,\hat{\beta},\delta} = 0$. The consumer welfare is the actual discounted utility of the agents from the perspective of time 0, $\beta\delta\left[-L^* + \int_{-\infty}^{\beta\delta b - p^*}(\delta b - p^* - c)\,dF(c)\right]$. For partially naive agents this utility differs from the one that they (mistakenly) expect to experience (expression (3)). After substituting for the profit-maximizing L^*, one obtains $U_j^{\beta,\hat{\beta},\delta} = \beta\delta\left[\int_{-\infty}^{\beta\delta b - p^*}(\delta b - a - c)\,dF(c) - K\right] - \beta\Pi_j^{\beta,\hat{\beta},\delta}$, with $j \in \{M, PC\}$ for the cases of Monopoly and Perfect Competition. Finally, we define the joint consumer-firm surplus as the sum of firm profits and the consumer welfare adjusted for discounting[21], measured as $U_j^{\beta,\hat{\beta},\delta}/\beta$. Formally,

$$S_j^{\beta,\hat{\beta},\delta} \equiv \Pi_j^{\beta,\hat{\beta},\delta} + U_j^{\beta,\hat{\beta},\delta}/\beta = \delta\left[\int_{-\infty}^{\beta\delta b - p^*}(\delta b - a - c)\,dF(c) - K\right]. \tag{21}$$

Given that p^* coincides under monopoly and perfect competition (Proposition 1), $S_M^{\beta,\hat{\beta},\delta}$ and $S_{PC}^{\beta,\hat{\beta},\delta}$ coincide. We therefore drop the subscript on $S^{\beta,\hat{\beta},\delta}$. Finally, we consider investment and leisure goods and neglect neutral goods ($b = 0$). Within the simple model, time-inconsistency and naiveté have no welfare effect for goods with no significant intertemporal profile.

to twelve months for RCI and five months for Hapimag.

[20] A cute example that supports this interpretation is Cyberrebates.com. This company operated until May 2001 by offering goods with a 100 percent mail-in rebate. Users who sent in the rebate received all the money back within three months.

[21] This alternative measure of welfare coincides with the long-run utility of the individual, that is, the utility that the individual would experience if he had exponential preferences. See O'Donoghue and Rabin (2001) for a discussion of this welfare measure.

Proposition 7 (Welfare in the simple model)
(i) For sophisticated agents ($\beta = \hat{\beta} \leq 1$) the consumer welfare adjusted for discounting $U_j^{\beta,\beta,\delta}/\beta$, the firm profits $\Pi_M^{\beta,\beta,\delta}$ and the joint consumer-firm surplus $S^{\beta,\hat{\beta},\delta}$ are unaffected by the short-run discount β for $j \in \{M, PC\}$.
(ii) For partially naive agents ($\beta < \hat{\beta} \leq 1$) with $b \neq 0$, the profits $\Pi_M^{\beta,\beta,\delta}$ are strictly increasing in $1-\beta$ for fixed $\hat{\beta}-\beta$.
(iii) For partially naive agents ($\beta < \hat{\beta} \leq 1$) with $b \neq 0$, the profits $\Pi_M^{\beta,\hat{\beta},\delta}$ are strictly increasing in $\hat{\beta}-\beta$ for fixed β and $\hat{\beta} < 1$. In addition, surplus and consumer welfare are lower for partially naive agents ($\beta < \hat{\beta} \leq 1$) than for sophisticated agents ($\beta = \hat{\beta} \leq 1$): $S^{\beta,\hat{\beta},\delta} \leq S^{\beta,\beta,\delta}$ and $U_j^{\beta,\hat{\beta},\delta} \leq U_j^{\beta,\beta,\delta}$ for $1 \geq \hat{\beta} > \beta$, $j \in \{M, PC\}$.
(iv) For $b \neq 0$, increases in naiveté $\hat{\beta}-\beta$ for fixed β decrease the welfare of the agents more under monopoly than under perfect competition: $dU_M^{\beta,\hat{\beta},\delta}/d\hat{\beta} < dU_{PC}^{\beta,\hat{\beta},\delta}/d\hat{\beta}$ for $\hat{\beta} < 1$.

Proof. In Appendix A. **Q.E.D.**

Time-inconsistency. We consider first the welfare effect of time-inconsistency, as measured by $(1-\beta)$. If the agents are sophisticated ($\beta = \hat{\beta} \leq 1$), the profit-maximizing level of the per-usage price from (6) satisfies $p^* = a - (1-\beta)\delta b$. Interestingly, this is also the level of p that maximizes the surplus $S^{\beta,\hat{\beta},\delta}$ in (21) for any given β. The firm provides a contract which works as a perfect commitment device, and allows a time-inconsistent agent to consume as much as a time-consistent agent with the same δ: the agents consumes if $c < \beta\delta b - p^* = \delta b - a$. Therefore, in the market limitations to self-control do not affect the total surplus. Given that the reservation utility $\overline{u}$ does not vary with β, the division of the surplus into welfare and profits does not depend on the time-inconsistency either. The welfare measures are therefore orthogonal to self-control for sophisticated agents (Proposition 7(i)).

Do these results extend to partially naive agents? Proposition 7(ii) provides a partial answer, in the negative. An increase in $1-\beta$ for fixed naiveté $\hat{\beta}-\beta > 0$ reduces the monopoly profits $\Pi_M^{\beta,\hat{\beta},\delta}$. When self-control is more bounded (higher $1-\beta$), the firm has more leeway to extract profits.

Naiveté. We now address the welfare effects of increases in naiveté, measured as increases in $\hat{\beta}-\beta$ for fixed β. As naiveté increases, the overconfidence about future consumption increases, and the fictitious surplus $\int_{\beta\delta b-p^*}^{\hat{\beta}\delta b-p^*}(\delta b - p^* - c)\,dF(c)$ in (4) increases. It is not surprising, therefore, that a monopolistic firm can extract more profits when the naiveté is higher (Proposition 7(iii)). Partially naive agents are like people with a 'fictitious demand curve'. A monopolistic firm can get them to pay for a fictitious product that they will never actually consume.

While naiveté per se generates high profits for a monopolistic firm, the firm can increase profits further by setting the per-usage price p to maximize the overconfidence about attendance. The profit-maximizing p^* for partially naive agents maximizes the fictitious surplus, and not the real surplus $S^{\beta,\hat{\beta},\delta}$. Naiveté, therefore, has an adverse efficiency effect (Proposition 7(iii)): the real surplus is lower for a partially naive than for a fully sophisticated agent with the same β, since the firm has no interest in offering the contract that maximizes the real surplus. A second implication of this adverse efficiency effect is that the welfare of a partially naive agent is lower than the welfare of a fully sophisticated agent.

Perfect competition and naiveté. While the qualitative effects of naiveté on consumer welfare do not depend on market power, the magnitudes do, as Proposition 7(iv) states. Under monopoly, in addition to the adverse efficiency effect, naiveté has a distributional effect. As $\hat{\beta} - \beta$ increases, the fictitious surplus increases and the monopolistic firm drains more revenue from the agent. Under perfect competition this distributive effect is null since the individual appropriates the whole surplus. Thus the detrimental effect of increases in naiveté on consumer welfare, captured by $dU^{\beta,\hat{\beta},\delta}/d\hat{\beta}$, is higher under monopoly. Stated differently, the loss in consumer welfare due to monopoly power, $U_M^{\beta,\hat{\beta},\delta} - U_{PC}^{\beta,\hat{\beta},\delta}$, becomes larger in absolute value as naiveté increases.

Do these results extend to a more general market interaction of profit-maximizing firms and non-standard consumers? The above results suggest a clear distinction between non-standard preferences (time-inconsistency in the paper) and non-rational expectations (naiveté). The first type of deviation does not necessarily affect surplus, profits and welfare. The firms offer the contract that maximizes the joint surplus, and this contract may neutralize the behavioral effect of the non-standard feature. For example, if consumers have limited computational abilities and are aware of it, firms may offer simple contracts, or devices that help consumers perform the computations. In either case, in equilibrium the limited computational power is likely to be non-binding. Similarly, the limitation of self-control is not binding for sophisticated agents, since the firms offer perfect commitment devices.

This conclusion changes if consumers have non-rational expectations. Because consumers misperceive their objective function, the firms do not in general offer the contract that maximizes the joint surplus, but rather a contract that accentuates the effect of the non-rational expectations so as to increase the profits. As an example, consider again consumers with limited computational power, but now assume that they are unaware of their limits. The firms are likely to offer computationally complex contracts that induce the consumers to make choices that are profitable for the firms. Under monopoly the firms keep these profits, while under perfect competition consumers themselves receive the "returns" to naiveté. Competition therefore tempers the adverse effects of naiveté on consumer welfare.

6.2 Contract regulation

Consider now the second question. Suppose that a government can regulate the contracts by imposing a choice of either p or L but not of both. What kind of regulations, if any, would be welfare-improving for the consumers? Needless to say, the effect of any such regulation on firm profits is necessarily adverse, since the firm faces a restricted choice of contracts, while still having to satisfy the individual rationality constraint in (3).

If consumer are sophisticated ($\beta = \hat{\beta} \leq 1$), there is no scope for government intervention since the market is already taking care of the efficient pricing. Under monopoly, the consumer achieves the reservation utility $\beta\delta\overline{u}$ regardless. Under perfect competition, any choice of p different from the first best choice, $p_{FB} \equiv a - (1 - \beta)\delta b$, which characterizes the profit-maximizing contract, reduces the consumer welfare.

For partially naive individuals ($\beta < \hat{\beta} \leq 1$), instead, government intervention could be beneficial, at least in principle. The features of a welfare-enhancing intervention, however, depend crucially on the market structure. Under perfect competition, the consumer welfare coincides with the joint surplus in (21). The derivative of $S^{\beta,\hat{\beta},\delta}$ with respect to p equals $-(p - p_{FB})f(\beta\delta b - p)$. Therefore, the government can increases consumer surplus to the extent that it brings the per-usage price p closer to the first-best price p_{FB}, that is, the perfect commitment device that solves the self-control problem. Unfortunately, this policy requires information on both consumer preferences (the parameters β, δ, b) and the production function (the cost a). Simpler policies may harm consumer welfare. For example, consider the restriction to no initial fees ($L = 0$) for investment goods ($b > 0$): while the first-best pricing p_{FB} involves below-marginal-cost pricing, the restriction $L = 0$ induces above-marginal-cost pricing in the presence of set-up costs.

Government intervention requires even more information under monopoly. The consumer welfare is

$$U_M^{\beta,\hat{\beta},\delta} = \beta\delta\left[\overline{u} - \int_{\beta\delta b - p}^{\hat{\beta}\delta b - p} (\delta b - p - c)\, dF(c)\right]. \tag{22}$$

The agent achieves the reservation utility ($\beta\delta\overline{u}$) minus the fictitious surplus. To a good approximation, this fictitious surplus increases with $F(\hat{\beta}\delta b - p) - F(\beta\delta b - p)$, the overconfidence about future consumption. Since this expression depends on the per-usage price p, a monopolistic firm or a government can manipulate the overconfidence by choosing an appropriate value of p. Since the fictitious surplus contributes positively to the profits in (4), a profit-maximizing firm attempts to maximize it. Unlike a firm, a government that seeks to maximize the welfare in (22) chooses p so as to *minimize* the fictitious surplus. In doing so, the government faces serious informational constraints. While the firm may be able to infer the distribution F from data on consumer behavior, the government has no easy access to such information. Moreover, even if the government knows F, it cannot fully offset the effects of naiveté. No matter what the per-usage fee is, a partially naive agent ($\beta < \hat{\beta}$) never attains the reservation utility $\beta\delta\overline{u}$, given the assumption of $f > 0$ over $\mathbb{R}$.

To sum up, we have explored the scope and the limits of the intervention of a benevolent government that attempts to maximize consumer welfare. There is no scope for such intervention if consumers are sophisticated, in which case the market interaction leads to the socially optimal outcome. If consumers do not have rational expectations, instead, government intervention can in principle be beneficial. The intervention, however, is subject to several limits: it requires extensive information that the government is unlikely to have; it depends heavily on the market structure; it may not fully remedy the adverse effects of naiveté. A better policy for the government, in general, would be to make the partially naive users aware of their naiveté.

7 Alternative interpretations

We now comment on some alternative explanations of the results on the profit-maximizing contract design.

Transaction costs and distaste for payment per usage. Transaction costs, or dislike of per-usage charges, could explain the lack of payment per visit in the health club industry. The transaction costs in this industry, however, are likely to be small, since most clubs keep track of attendance using electronic cards, and could use this record to charge users.

Price discrimination. Price discrimination can explain the diffusion of above-marginal cost pricing. This explanation, however, makes the observed cases of below-marginal-cost pricing even more puzzling.

Temptation. Other models of self-control, such as Gul and Pesendorfer (2001), are likely to generate the results on below- and above-marginal cost pricing. These models, however, to the extent that they assume sophistication about the self-control, cannot predict the rising per-unit prices, or the endogenous cancellation costs.

Over- and underestimation of consumption. An alternative explanation of the deviations from marginal cost pricing is overconfidence about future consumption. This story, however, does not easily explain why the average user seems to *over*estimate usage at the health club but to *under*estimate usage of a credit line, or of a mobile phone.

8 Conclusion

In this paper we have analyzed the contract design of rational, profit-maximizing firms that sell goods to consumers with time-inconsistent preferences and partially naive expectations. In the profit-maximizing contract firms deviate from marginal cost pricing, charge additional fees and higher prices after renewal, and generate cancellation costs. We find that these predictions of the model match the empirical contract design in industries such as the credit card, health club, mail order, mobile phone, newspaper, and vacation time-sharing industries. The empirical findings on contract design are consistent with the view that time-inconsistent preferences and naiveté are widespread features of consumer preferences.

We also discuss the welfare effect of the profit-maximizing contracts for both sophisticated and naive agents with time-inconsistent preferences. If the agents are sophisticated, the market interaction with the firms enables the individuals to achieve the efficient consumption level. If the agents are naive, instead, the firms design the pricing so as to take maximal advantage of the consumer overconfidence and delay in cancellation. As a consequence, the interaction with the firms generates inefficient outcomes and, under monopoly, a redistribution of surplus from the agent to the firm.

The empirical findings on contracts suggest that naiveté is a common and persistent aspect of consumers' preferences. This conclusion is puzzling. How can individuals remain naive about their own preferences after a lifetime of experiences? In work in progress we try to explain this puzzle. We collect evidence suggesting that individuals are often unaware of the underlying forces driving their behavior. We model a learning process and show that the agents' beliefs about their self-control need not converge to the underlying preference parameters. Moreover, we show that firms have limited incentives to educate naive consumers.

A Appendix A. Mathematical Section

Derivation of Propositions 3. To derive the solutions for p_R^* and L_R^* in program (14), we first divide the domain of L_R into three regions and find the candidate solutions within these regions.

Region I ($L_R \leq k/\beta\delta$). In this region, $R^{\beta,1,\delta} = I\!R$. Thus, program (14) becomes

$$\begin{aligned}\max_{L_R,p_R} \Pi_I(L_R,p_R) &= \delta^2 \int_{\left(R^{1,1,\delta}\right)^c} \left(L_R - \frac{k}{\delta}\right) dF(c) \\ &+\delta^2 \left[\int_{-\infty}^{\beta\delta b-p_R} (p_R - a)\, dF(c) + \int_{R^{1,1,\delta}\cap\{c|c\leq\delta b-p_R\}} (\delta b - p_R - c)\, dF(c)\right]\end{aligned} \quad (23)$$

where $R^{1,1,\delta} = \{c|c \leq \delta b - p_R - (L_R - k/\delta)\}$ for $L_R \geq k/\delta$ and $R^{1,1,\delta} = I\!R$ otherwise, and $\left(R^{1,1,\delta}\right)^c$ is the complement of $R^{1,1,\delta}$.

The derivative of Π_I with respect to L_R equals $\delta^2 P\left[(R^{1,1,\delta})^c\right] \geq 0$. Therefore $L_{R,I}^*$, the (weakly) optimal L_R within Region I, lies at the right border of Region I: $L_{R,I}^* = k/\beta\delta$. An interior solution for $p_{R,I}^*$ satisfies

$$p_{R,I}^* - a = -\frac{P\left(R^{1,1,\delta}\right) - F\left(\beta\delta b - p_{R,I}^*\right)}{f\left(\beta\delta b - p_{R,I}^*\right)} = -\frac{F\left(\delta b - p_{R,I}^* - \frac{1-\beta}{\beta\delta}k\right) - F\left(\beta\delta b - p_{R,I}^*\right)}{f\left(\beta\delta b - p_{R,I}^*\right)} \quad (24)$$

where the last expression substitutes $L_R^* = k/\beta\delta$. Assumption ABP guarantees that the optimum level of $p_{R,I}^*$ is interior and therefore satisfies the first order condition above (see the proof of Proposition 1). This in turn implies, using the assumption of f positive over $I\!R$, that $P\left[(R^{1,1,\delta})^c\right] > 0$ and therefore that $L_{R,I}^* = k/\beta\delta$ is the strictly optimal solution in Region I.

Region II ($k/\beta\delta \leq L_R \leq k/\beta\delta + (1-\beta)\,\delta b$). The left inequality implies that users with low consumption utility, including those with $c > \delta b - p_R$, cancel. The right inequality guarantees $R^{\beta,1,\delta} \cap \{c|c \leq \beta\delta b - p_R\} = \{c|c \leq \beta\delta b - p_R\}$, i.e., all the agents that would undertake C at $t = 2$ conditional on renewing actually renew. Notice $R^{\beta,1,\delta} = \{c|c \leq \delta b - p_R - [L_R - k/\beta\delta]\}$. Program (14) can be written as

$$\begin{aligned}\max_{L_R,p_R} \Pi_{II}(L_R,p_R) &= \delta^2 \int_{R^{\beta,1,\delta}\setminus R^{1,1,\delta}} L_R dF(c) - \delta^2 \frac{k}{\delta} P\left[\left(R^{1,1,\delta}\right)^c\right] \\ &+\delta^2 \left[\int_{-\infty}^{\beta\delta b-p_R} (p_R - a)\, dF(c) + \int_{R^{1,1,\delta}\cap\left\{c|c\leq\hat{\beta}\delta b-p_R\right\}} (\delta b - p_R - c)\, dF(c)\right].\end{aligned}$$

A comparison of Π_I and Π_{II} for given p_R and L_R yields $\Pi_I - \Pi_{II} = \delta^2 \int_{(R^{\beta,1,\delta})^c} L_R dF(c)$ which is positive for $L_R > 0$. This implies a downward discontinuity of profits at $L_R = k/\beta\delta$ between Region I and II. Therefore, the global optimum for L_R cannot be a corner solution at the left boundary of Region II with $L_R = k/\beta\delta$. As for stationary points, the f.o.c. with respect to L_R yields

$$L_{R,II}^* = P\left(R^{\beta,1,\delta}\setminus R^{1,1,\delta}\right)/f\left[\delta b - p_R - \left(L_{R,II}^* - k/\beta\delta\right)\right] > 0. \quad (25)$$

Assumption ABP guarantees that for given p_R a solution for L_R exists and satisfies this f.o.c.. The f.o.c. with respect to p_R gives $p^*_{R,II} - a = -\left[P\left(R^{1,1,\delta}\right) - F\left(\beta\delta b - p^*_{R,II}\right)\right]/f\left(\beta\delta b - p^*_{R,II}\right) - L^*_{R,II} f\left[\delta b - p^*_{R,II} - \left(L^*_{R,II} - k/\beta\delta\right)\right]/f\left(\beta\delta b - p^*_{R,II}\right)$. Therefore, in an interior solution for $L^*_{R,II}$, $p^*_{R,II}$ satisfies

$$p^*_{R,II} - a = -\left[P\left(R^{\beta,1,\delta}\right) - F\left(\beta\delta b - p^*_{R,II}\right)\right]/f\left(\beta\delta b - p^*_{R,II}\right) < 0. \tag{26}$$

Once again, Assumption ABP guarantees that the optimum level of $p^*_{R,II}$ satisfies this f.o.c..

Region III ($L_R \geq \max[k/\beta\delta, k/\beta\delta + (1-\beta)\,\delta b]$). The inequality defining Region III implies $R^{\beta,1,\delta} \cap \{c | c \leq \beta\delta b - p_R\} = R^{\beta,1,\delta} = \{c | c \leq \delta b - p_R - [L_R - k/\beta\delta]\}$. The profit function (14) becomes

$$\begin{aligned} \max_{L_R, p_R} \Pi_{III}(L_R, p_R) \;\; &= \;\; \delta^2 \int_{R^{1,1,\delta} \cap \{c | c \leq \delta b - p_R\}} (\delta b - a - c)\, dF(c) + \\ & \;\;\;\;\; \delta^2 \int_{R^{\beta,1,\delta} \backslash R^{1,1,\delta}} (p_R - a + L_R)\, dF(c) - \delta^2 \frac{k}{\delta} P\left[\left(R^{1,1,\delta}\right)^c\right]. \end{aligned} \tag{27}$$

It is easy to check that Π_{III} is a function of p_R and L_R only through their linear combination $p_R + L_R$. The f.o.c. gives

$$p^*_{R,III} + L^*_{R,III} - a = \frac{P(R^{\beta,1,\delta} \backslash R^{1,1,\delta})}{f\left[\delta b - p^*_{R,III} - \left(L^*_{R,III} - k/\beta\delta\right)\right]} > 0. \tag{28}$$

for any $L^*_{R,III} \geq k/\beta\delta + (1-\beta)\,\delta b$. Assumption ABP guarantees that the solution for $p_R + L_R$ is interior. Further, a comparison between Π_{II} and Π_{III} for $b > 0$ shows that for $L_R = k/\beta\delta + (1-\beta)\,\delta b$ the two profit functions coincide and have the same solution for p^*_R. Therefore, a corner solution for $L^*_{R,II}$ at the right border is a special case of the solution in Region III.

We have found three candidates for solution: $(p^*_{R,I}, L^*_{RI})$ with $p^*_{R,I}$ as in (24) and $L^*_{R,I} = k/\beta\delta$; $(p^*_{R,II}, L^*_{R,II})$ with $p^*_{R,II}$ as in (26), $L^*_{R,II}$ as in (25), and $k/\beta\delta \leq L^*_{R,II} \leq k/\beta\delta + (1-\beta)\,\delta b$; $(p^*_{R,III}, L^*_{R,III})$ with $p^*_{R,III} + L^*_{R,III}$ as in (28) and $L^*_{R,III} \geq k/\beta\delta + (1-\beta)\,\delta b$. The following Proposition characterizes the solution for leisure and neutral goods ($b \leq 0$).

Proposition A.1 *For leisure and neutral goods ($b \leq 0$), the optimal solution lies in Region I. L^*_R equals $k/\beta\delta$ and p^*_R is implicitly defined by expression (24).*

Proof. The strategy of the proof is to rule out the existence of an optimum of program (9-10) in Regions II and III. It is easy to check that Region II is not defined for $b \leq 0$ (leisure and neutral goods). It remains to show $\Pi_I(L^*_{R,I}, p^*_{R,I}) > \Pi_{III}(L^*_{R,III}, p^*_{R,III})$. The optimum in Region III can be obtained for any L_R setting p_R equal to $p_R(L_R)$, the level that satisfies (28) for the optimal $p^*_{R,III} + L^*_{R,III}$. In particular, we can set (L_R, p_R) equal to $(k/\beta\delta, p_R(k/\beta\delta))$; for these values, $\Pi_I(\tilde{L}_R, \tilde{p}_R) - \Pi_{III}(\tilde{L}_R, \tilde{p}_R) = \delta^2 \int_{\delta b - \tilde{p}_R - (\tilde{L}_R - k/\beta\delta)}^{\beta\delta b - \tilde{p}_R} \left(\tilde{p}_R - a + \tilde{L}_R\right) dF(c) + \delta^2 \int_{\beta\delta b - \tilde{p}_R}^{\infty} \tilde{L}_R dF(c)$. This first term is nonnegative since $\beta\delta b - \tilde{p}_R \geq \delta b - \tilde{p}_R$ for leisure and neutral goods ($b \leq 0$) and $\tilde{p}_R - a + \tilde{L}_R > 0$ by (28); the second term is positive since $\tilde{L} > 0$. By the optimality of $(L^*_{R,I}, p^*_{R,I})$, $\Pi_I(L^*_{R,I}, p^*_{R,I}) \geq \Pi_I(\tilde{L}_R, \tilde{p}_R)$ and therefore $\Pi_I(L^*_{R,I}, p^*_{R,I}) \geq \Pi_I(\tilde{L}_R, \tilde{p}_R) > \Pi_{III}(\tilde{L}_R, \tilde{p}_R) = \Pi_{III}(L^*_{R,III}, p^*_{R,III})$. **Q.E.D.**

In order to qualify the solution for investment goods, we introduce the following condition.

Condition A1.
$(1-\beta)\,\delta b > \max\left(k/\beta\delta, -k/\beta\delta + [F(\beta\delta b - p_R) - F(\beta\delta b - p_R - k/\beta\delta)]/f\,[\beta\delta b - p_R]\right)$ for all $p_R \in I\!R$.

Proposition 3(i) for investment goods follows from inspection of the three candidates for optimum. The inequality $p^*_R < a$ in (ii) is satisfied whenever $(1-\beta)\,\delta b > k/\beta\delta$ for Region I, always for Region II, and if $p^*_{R,III} - a + L^*_{R,III} < k/\beta\delta + (1-\beta)\,\delta b$ for Region III. Condition A1 guarantees $p^*_R < a$ in the three Regions. **Q.E.D.**

Proof of Proposition 4. For sophisticated and time-consistent agents ($\beta = \hat{\beta} \leq 1$), Proposition 2 implies that the equilibrium profits $\Pi(k) = \delta(1+\delta)\int_{-\infty}^{\delta b - a}(\delta b - a - c)\,dF(c) - \delta K - \delta\overline{u}$ are independent of k. For naive agents ($\beta < \hat{\beta} = 1$), the derivative of the profit function (11) with respect to k equals $\frac{1-\beta}{\beta\delta}P(c \notin R^{1,1,\delta}) \geq 0$ whenever (L^*_R, p^*_R) is in Region I. Notice that, for $k > 0$, $P(c \notin R^{1,1,\delta}) > 0$. For leisure and neutral goods, the solution lies in Region I (Proposition A.1) for all k and therefore the equilibrium profits $\Pi(k)$ are strictly increasing in k. In general, it is easy to show that for $k = 0$ the solutions of Region I and II coincide and dominate the solution in Region III. Therefore the optimal profits in region I provide the desired function $\underline{\Pi}$ with the properties $\Pi(0) = \underline{\Pi}(0)$ and $\underline{\Pi}$ strictly increasing in k. **Q.E.D.**

Proof of Proposition 5. To prove $L^*_R \geq k/\beta\delta$ for investment and neutral goods ($b \geq 0$), we show that any level of the renewal fee smaller than $k/\beta\delta$ is dominated by $k/\beta\delta$. For $L_R \leq k/\beta\delta$, inequality (7) implies $R^{\beta,\hat{\beta},\delta} = I\!R$ an therefore the maximization of program (11) with respect to L_R can be simplified to

$$\begin{aligned}\max_{L_R, p_R} \Pi \;=\;& \delta^2 \int_{\left(R^{\hat{\beta},\hat{\beta},\delta}\right)^c} \left(L_R - \frac{k}{\delta}\right) dF(c) \\ &+\delta^2 \left[\int_{-\infty}^{\beta\delta b - p_R} (p_R - a)\,dF(c) + \int_{R^{\hat{\beta},\hat{\beta},\delta} \cap \left\{c | c \leq \hat{\beta}\delta b - p_R\right\}} (\delta b - p_R - c)\,dF(c)\right]\end{aligned} \tag{29}$$

where $R^{\hat{\beta},\hat{\beta},\delta} = \left\{c | c \leq \delta b - p_R - \left(L_R - k/\hat{\beta}\delta\right)\right\}$ for $L_R \geq k/\hat{\beta}\delta$ and $R^{\hat{\beta},\hat{\beta},\delta} = I\!R$ otherwise, and $\left(R^{\hat{\beta},\hat{\beta},\delta}\right)^c$ is the complement of $R^{\hat{\beta},\hat{\beta},\delta}$. For $L_R \leq k/\beta\delta$, the derivative of Π with respect to L_R is bounded below by $\delta^2 P\left[(R^{\hat{\beta},\hat{\beta},\delta})^c\right]$. The probability $P\left[(R^{\hat{\beta},\hat{\beta},\delta})^c\right]$ equals 0 for $L_R < k/\hat{\beta}\delta$ and equals $1 - F(\delta b - p_R - \left(L_R - k/\hat{\beta}\delta\right)) > 0$ for $k/\hat{\beta}\delta \leq L_R \leq k/\beta\delta$. If follows that any renewal fee smaller than $k/\beta\delta$ is strictly dominated by $L_R = k/\beta\delta$, as we intended to prove.

To prove that the optimal profit $\Pi(k)$ is bounded below by a function $\underline{\Pi}(k)$, consider the case $k = 0$. For this simpler case, it is not difficult to prove that the optimal solution is to set L_R^* at an arbitrary value L_R that satisfies $L_R \leq (1-\hat{\beta})\delta b$, and to set $p_R^* = p^*$ at the level defined by (6). In particular, setting $L_R = k/\beta\delta$ and maximizing expression (29) with respect to p_R provides the same solution for $k = 0$. Define by $\underline{\Pi}(k)$ the profit level attained by setting $L_R = k/\beta\delta$ and maximizing expression (29) with respect to p_R. We have just proved $\Pi(0) = \underline{\Pi}(0)$. All that is left to prove is that $\underline{\Pi}(k)$ is strictly increasing in k. We can compute $\partial\underline{\Pi}(k)/\partial k$ as the derivative of (29) with respect to k, after setting $L_R = k/\beta\delta$. The envelope theorem guarantees that we can disregard the indirect effect of k on p_R, since the solution for p_R is interior (as is easily proven using Assumption ABP). This derivative is bounded below by $\delta^2 P\left[(R^{\hat{\beta},\hat{\beta},\delta})^c\right](1-\beta)/\beta\delta$. This expression is strictly positive since $\beta < 1$ and $P\left[(R^{\hat{\beta},\hat{\beta},\delta})^c\right]$ equals $1 - F(\delta b - p_R - \left(k/\beta\delta - k/\hat{\beta}\delta\right)) > 0$. This completes the proof. **Q.E.D.**

Proof of Proposition 6. First, we introduce some necessary notation. Denote with $t_Q^{\beta,\hat{\beta},\delta}(c)$ the time at which an agent with preferences $(\beta,\hat{\beta},\delta)$ and immediate payoff $-c$ performs the one-time cancellation activity. The quitting time $t_Q^{\beta,\hat{\beta},\delta}(c)$ equals $T+nT$ for some $n \in I\!N$ if the agent ever quits. Define $t_Q^{\beta,\hat{\beta},\delta}(c) = \infty$ if the agent never quits. Further, denote with $\hat{U}^{\beta,\hat{\beta},\delta}(L_R, p_R|c)$ the perceived utility that an agent with preferences $(\beta,\hat{\beta},\delta)$ and immediate payoff $-c$ expects to attain from contract (L_R, p_R): $\hat{U}^{\beta,\hat{\beta},\delta}(L_R,p_R|c) \equiv \beta\delta[-L_R + \frac{1-\delta^T}{1-\delta}\mathbf{1}_{\{c\leq\hat{\beta}\delta b - p_R\}}(\delta b - p_R - c)]$ where $\mathbf{1}_X$ is the indicator function for set X. Lemma A.1 (adapted from O'Donoghue and Rabin, 2001) characterizes the time of cancellation $t_Q^{\beta,\hat{\beta},\delta}(c)$ for time-consistent, sophisticated and naive individuals.

Lemma A.1 *(***Quitting decision***) (i) For time-consistent agents ($\beta = \hat{\beta} = 1$), $t_Q^{1,1,\delta}(c) = T$ if $\hat{U}^{1,1,\delta}(L_R,p_R|c) < -\left(1-\delta^T\right)k$, and $t_Q^{1,1,\delta}(c) = \infty$ otherwise. (ii) For hyperbolic sophisticated agents ($\beta = \hat{\beta} < 1$), $t_Q^{\beta,\beta,\delta}(c) \leq T+T'$ where $T' \geq 0$ is the largest integer multiple of T satisfying $\frac{1-\delta^{T'}}{1-\delta^T}\hat{U}^{\beta,\beta,\delta}(L_R,p_R|c) > -(1-\beta\delta^{T'})k$ if $\hat{U}^{\beta,\beta,\delta}(c) < -\left(1-\delta^T\right)k$ and $t_Q^{\beta,\beta,\delta}(c) = \infty$ otherwise. (iii) For naive hyperbolic agents ($\beta < \hat{\beta} = 1$), $t_Q^{\beta,1,\delta}(c) = T$ if $\hat{U}^{\beta,1,\delta}(L_R,p_R|c) < -\left(1-\beta\delta^T\right)k$ and $t_Q^{\beta,1,\delta}(c) = \infty$ otherwise.*

Proof. (i) Given the stationary setting after T, a time-consistent agent quits either immediately or never. The solution is $t_Q^{1,1,\delta}(c) = T$ if quitting is better than renewing forever: $-k > \frac{1}{1-\delta^T}\hat{U}^{1,1,\delta}(L_R,p_R|c)$. (ii) A sophisticated hyperbolic agent cancels at time T if the next self that will cancel is too far off into the future. The agent is willing to wait at most T' periods, where T' is the highest multiple of T such that $\frac{1-\delta^{T'}}{1-\delta^T}\hat{U}^{\beta,\beta,\delta}(L_R,p_R|c) > -(1-\beta\delta^{T'})k$. Existence of a finite T' is guaranteed if $\hat{U}^{\beta,\beta,\delta}(L_R,p_R|c) < -\left(1-\delta^T\right)k$. If $\hat{U}^{\beta,\beta,\delta}(L_R,p_R|c) \geq -\left(1-\delta^T\right)k$ holds, the agent prefers renewing forever to cancelling, and therefore $t_Q^{\beta,\beta,\delta}(c) = \infty$. (iii) A naive hyperbolic agent in period T believes that the future self in T periods will undertake the cancellation activity at $2T$ if $\hat{U}^{1,1,\delta}(L_R,p_R|c) < -\left(1-\delta^T\right)k$, and never otherwise (see part (i) of this Proposition). If $\hat{U}^{1,1,\delta}(L_R,p_R|c) \geq -\left(1-\delta^T\right)k$, the future self never cancels and the present self (that discounts the future more heavily) does not cancel either: $t_Q^{\beta,1,\delta}(c) = \infty$. Under the opposite inequality, the agent chooses between canceling at T and at $2T$. She prefers canceling at $2T$ if $\hat{U}^{\beta,1,\delta}(L_R, p_R|c) - \beta\delta^T k \geq -k$ or equivalently if $\hat{U}^{\beta,1,\delta}(L_R, p_R|c) \geq -\left(1 - \beta\delta^T\right) k$. Therefore, if $-\beta\left(1 - \delta^T\right) k > \hat{U}^{\beta,1,\delta}(L_R, p_R|c) \geq -(1 - \beta\delta^T)k$, then $t_Q^{\beta,1,\delta}(c) = \infty$. If $\hat{U}^{\beta,1,\delta}(L_R, p_R|c) < -\left(1 - \beta\delta^T\right) k$, then $t_Q^{\beta,1,\delta}(c) = T$. **Q.E.D.**

For a given contract $((L,p),(L_R,p_R))$, we define $R_t^{\beta,\hat{\beta},\delta} = \left\{c | t_Q^{\beta,\hat{\beta},\delta}(c) > t\right\}$ to be the set of values of c for which an individual with preferences $(\beta,\hat{\beta},\delta)$ has not yet cancelled at time t. The set $R_t^{\beta,\hat{\beta},\delta}$ is decreasing in t, $R_s^{\beta,\hat{\beta},\delta} \supseteq R_u^{\beta,\hat{\beta},\delta}$ for $s < u$. It is easy to see, along the lines of Section 4.3, that we can decompose the maximization problem for the firm into the maximization problem for the first T periods, and for the subsequent periods. The first problem admits the same solution for p^* as in Proposition 2. The maximization problem for the later periods can be written as

$$\max_{L_R,p,p_R} \sum_{i=0}^{\infty} \delta^{T+1+iT} \left\{ \begin{array}{c} L_R \cdot P\left(R_{T+iT}^{\beta,\hat{\beta},\delta} \backslash R_{T+iT}^{\hat{\beta},\hat{\beta},\delta}\right) - \frac{k}{\delta} P\left(R_{T+(i-1)T}^{\hat{\beta},\hat{\beta},\delta} \backslash R_{T+iT}^{\hat{\beta},\hat{\beta},\delta}\right) + \\ \frac{1-\delta^T}{1-\delta} \int_{R_{T+iT}^{\beta,\hat{\beta},\delta} \cap \{c | c \leq \beta\delta b - p_R\}} (p_R - a)\, dF(c) + \\ \frac{1-\delta^T}{1-\delta} \int_{R_{T+iT}^{\beta,\hat{\beta},\delta} \cap \left\{c | c \leq \hat{\beta}\delta b - p_R\right\}} (\delta b - p_R - c)\, dF(c) \end{array} \right\}.$$

For sophisticated and time-consistent agents, we can prove the following Lemma. The proof is almost identical to the proof of Proposition 2.

Lemma A.2 *For $k > 0$ the profit maximizing contract for sophisticated and time-consistent agents ($\beta = \hat{\beta} \leq 1$) satisfies: (i) The per-usage fee after renewal is $p_R^* = p^* = a - (1-\beta)\delta b$. (ii) The renewal fee L_R^* is any $L_R \in I\!R$ such that $L_R \leq \frac{1-\delta^T}{\beta\delta} k + \min\left\{\frac{1-\delta^T}{1-\delta}(1-\beta)\delta b, 0\right\}$.*

This provides us with the first desired result. The next Lemma provides the second result for naive agents. The proof follows along the lines of the proof for Proposition 3, with the due substitutions to take into account the cancellation behavior outlined in Lemma A.1.

Lemma A.3 *For $k > 0$ the profit maximizing contract for naive agents ($\beta < \hat{\beta} = 1$) satisfies: (i) For leisure goods ($b > 0$), the renewal fee L_R^* satisfies $L_R^* = \frac{1-\beta\delta^T}{\beta\delta} k$ and the per-usage fee satisfies $p_R^* = p^* = a$ for $\beta = 1$ and $p_R^* > p^* \geq a$ for $\beta < 1$. (ii) For investment goods ($b < 0$), the renewal fee L_R^* satisfies $L_R^* \geq \frac{1-\beta\delta^T}{\beta\delta} k$.*

This completes the proof. **Q.E.D.**

Proof of Proposition 7. After substituting for $p^* = a - (1-\beta)\delta b$ in the expressions for $U_j^{\beta,\beta,\delta}/\beta$, for $\Pi_{PC}^{\beta,\beta,\delta}$ and for $S^{\beta,\hat{\beta},\delta}$, these variables do not depend on β. This proves (i). To prove (ii), we differentiate the profit function (4) and make use of the envelope theorem. We get $\partial\Pi/\partial\hat{\beta} = \delta(1-\hat{\beta})(\delta b)^2 f(\hat{\beta}\delta b - p^*)$ and $\partial\Pi/\partial\beta = \delta^2 b(p^* - a) f(\beta\delta b - p^*)$. The effect of a increase of β for fixed $\hat{\beta}-\beta$ equals $\partial\Pi/\partial\hat{\beta}+\partial\Pi/\partial\beta=-\delta^2 b\left[F(\hat{\beta}\delta b-p^*)-F(\beta\delta b-p^*)\right]$ where we substituted for p^*-a the expression in (6). By the assumption that $f>0$ on $\mathbb{R}$, $\partial\Pi/\partial\hat{\beta}+\partial\Pi/\partial\beta$ is strictly positive for $b\neq 0$ and $\beta<\hat{\beta}$. To prove (iii), one can use $\partial\Pi/\partial\hat{\beta}=\delta(1-\hat{\beta})(\delta b)^2 f(\hat{\beta}\delta b-p^*)$. The latter expression is strictly positive for $b\neq 0$ and $\hat{\beta}<1$. As for $S^{\beta,\hat{\beta},\delta}$, it is easy to see that $S^{\beta,\hat{\beta},\delta}$ is maximized for $p^*=a-(1-\beta)\delta b$, which is the solution for sophisticated users. The inequality $S^{\beta,\hat{\beta},\delta}\leq S^{\beta,\beta,\delta}$ follows. Since surplus and consumer welfare coincide for the case of perfect competition, $U_{PC}^{\beta,\hat{\beta},\delta}\leq U_{PC}^{\beta,\beta,\delta}$ follows as well. Finally, to prove $U_M^{\beta,\hat{\beta},\delta}\leq U_M^{\beta,\beta,\delta}$ note that the individual rationality constraint (3) implies $U_M^{\beta,\hat{\beta},\delta}<\beta\delta\overline{u}=U_M^{\beta,\beta,\delta}$ for $1\geq\hat{\beta}>\beta$. Finally, to prove (iv), consider that $dU_j^{\beta,\hat{\beta},\delta}/d\hat{\beta}=\partial U_j^{\beta,\hat{\beta},\delta}/\partial\hat{\beta}+\partial U_j^{\beta,\hat{\beta},\delta}/\partial p\cdot\partial p_j^*/\partial\hat{\beta}$ for $j\in\{M,PC\}$. Using (6), $\partial U_M^{\beta,\hat{\beta},\delta}/\partial p=\partial U_{PC}^{\beta,\hat{\beta},\delta}/\partial p=-\{p-[a-(1-\beta)\delta b]\}f(\beta\delta b-p)$ follows; we also know $p_M^*=p_{PC}^*$. Therefore, $dU_M^{\beta,\hat{\beta},\delta}/d\hat{\beta}-dU_{PC}^{\beta,\hat{\beta},\delta}/d\hat{\beta}=\partial U_M^{\beta,\hat{\beta},\delta}/\partial\hat{\beta}-\partial U_{PC}^{\beta,\hat{\beta},\delta}/\partial\hat{\beta}=-(\delta b)^2(1-\hat{\beta})f(\hat{\beta}\delta b-p^*)<0$ for $\hat{\beta}<1$ and $b\neq 0$. **Q.E.D.**

B Appendix B. Survey transcript

1. "Hi. My name is ***. I heard good things about your gym and I am considering joining. Could you please tell me which contracts you offer?" If the respondent is not willing to answer, we insist: "I would really appreciate if you could tell me. I would like to have some more information before I check out your gym."

For each contract we ask questions 2, 3 and 4 if applicable: 2. "Is there an initiation fee?" 3. "Do I have to renew the contract at the end of the period or is it automatically renewed?" 4. (if expiration is not automatic) "If I want to quit the gym, how do I cancel?"

After these questions, we ask: 5. "Are there other types of contracts?"

If the person has not mentioned them yet, we ask: 6. "Do you also offer monthly contracts?" 7. "Do you also offer contracts for a longer time period, such as a year?" 8. "Could I just come and pay each time I use the gym?"

Finally, we inquire about the type of club: 9. "One last thing: What facilities do you offer? Do you also have racket courts?" We conclude the survey. 10. "Thank you very much. I appreciated your help."

References

[1] **Akerlof, George A; Yellen, Janet L.** "Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?". *American Economic Review*, September 1985, *75*(4), pp. 708–20.

[2] **Angeletos, Marios; Laibson, David I.; Repetto, Andrea; Tobacman, Jeremy and Weinberg, Stephen.** "The Hyperbolic Buffer Stock Model: Calibration, Simulation, and Empirical Evaluation." *Journal of Economic Perspectives,* Summer 2001, *15*(3), pp. 47–68.

[3] **Ausubel, Lawrence M.** "The Failure of Competition in the Credit Card Market." *American Economic Review*, March 1991, *81*(1), pp. 50-81.

[4] **Barro, Robert J.** "Ramsey Meets Laibson in the Neoclassical Growth Model." *Quarterly Journal of Economics,* November 1999, *114*(4), pp. 1125–1152.

[5] **Becker, Gary.** *The economics of discrimination.* Chicago: University of Chicago Press, 1957.

[6] **Benabou, Roland, and Tirole, Jean.** "Self-Confidence and Social Interactions". NBER working paper W7585, March 2000.

[7] **Benzion, Uri; Rapoport, Amnon and Yagil, Joseph.** "Discount Rates Inferred from Decisions: an Experimental Study." *Management Science,* March 1989, *35*(3), pp. 270–84.

[8] **Camerer, Colin.** "Individual Decision Making." In: John H. Kagel and Alvin E. Roth (eds.), *Handbook of Experimental Economics*, Princeton University Press, 1995.

[9] **DellaVigna, Stefano and Malmendier, Ulrike.** "Self-Control in the Market: Evidence from the Health Club Industry", 2001, mimeographed.

[10] **DellaVigna, Stefano and Paserman, M. Daniele.** "Job Search and Hyperbolic Discounting." The Maurice Falk Institute for Economic Research in Israel (Jerusalem, Israel), Discussion Paper No. 00.15, December 2000.

[11] **De Long, J. Bradford; Shleifer, Andrei; Summers, Lawrence H. and Waldmann, Robert J.** "Noise Trader Risk in Financial Markets." *Journal of Political Economy*, August 1990, *98*(4), pp. 703–738.

[12] **Evans, David and Schmalensee, Richard.** *Paying with Plastic. The Digital Revolution in Buying and Borrowing*. MIT Press, 1999.

[13] **Fang, Hanming, and Silverman, Daniel.** "Time Inconsistency and Welfare Program Participation: Evidence from the NLSY", mimeographed, 2001.

[14] **Genesove, David and Mayer, Christopher.** "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics,* November 2001, *116*(4), pp. 1233–1260.

[15] **Gross, David B. and Souleles, Nicholas S.** "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data." *Quarterly Journal of Economics*, forthcoming.

[16] **Gruber, Jonathan and Koszegi, Botond.** "Is Addiction 'Rational'? Theory and Evidence." *Quarterly Journal of Economics,* November 2001, *116*(4), pp. 1261–1303.

[17] **Gul, Faruk and Pesendorfer, Wolfgang.** "Temptation and Self-Control." *Econometrica,* November 2001, *69*(6), pp. 1403–1436.

[18] **Kirby, Kris N**. "Bidding on the Future: Evidence Against Normative Discounting of Delayed Rewards." *Journal of Experimental Psychology: General,* March 1997, *126*(1), pp. 54–70.

[19] **Krusell, Per; Kuruscu, Burhanettin and Smith, Anthony A.** "Tax Policy with Quasi-geometric Discounting." *International Economic Journal*, Autumn 2000, *14*(3), pp. 1-40.

[20] **Laibson, David I.** "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, May 1997, *112*(2), pp. 443–77.

[21] **Laibson, David I.; Repetto, Andrea and Tobacman, Jeremy.** "A Debt Puzzle." In: Philippe Aghion, Roman Frydman, Joseph Stiglitz, Michael Woodford (eds.), *Knowledge, Information, and Expectations in Modern Economics: In Honor of Edmund S. Phelps*, forthcoming.

[22] **Laibson, David I. and Yariv, Leeat.** "Teaser Rates in the Credit Card Market", mimeographed, 2000.

[23] **Larwood, Laurie and Whittaker, William.** "Managerial Myopia: Self-serving Biases in Organizational Planning." *Journal of Applied Psyhology,* April 1977, *62*(2), pp. 194–198.

[24] **Loewenstein, George and Prelec, Drazen**. "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *Quarterly Journal of Economics*, May 1992, *107*(2), pp. 573–597.

[25] **Madrian, Brigitte C. and Shea, Dennis.** "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior" *Quarterly Journal of Economics*, November 2001, *116*(4), pp. 1149–1187.

[26] **Milgrom, Paul and Robert, John.** "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economics Review*, June 1990, *80*(3), pp. 511-528.

[27] **Odean, Terrance.** "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance*, October 1998, 53(5), pp. 1775-1798.

[28] **O'Donoghue, Ted D. and Rabin, Matthew**. "Doing It Now or Later." *American Economics Review*, March 1999a, *89*(1), pp. 103-124.

[29] ——. “Incentives for Procrastinators.” *Quarterly Journal of Economics*, August 1999b, *114*(3), pp. 769–816.

[30] ——. “Choice and Procrastination.” *Quarterly Journal of Economics*, February 2001, *116*(1), pp. 121–160.

[31] **Phelps, Edmund S. and Pollak Robert A**. “On Second-Best National Saving and Game-Equilibrium Growth.” *Review of Economic Studies,* April 1968, *35*(2), pp. 85–199.

[32] **Rubinstein, Ariel.** “Is It ‘Economics and Psychology’?: The Case of Hyperbolic Discounting”, mimeographed, 2000.

[33] **Russell, Thomas and Thaler, Richard H.** “The Relevance of Quasi Rationality in Competitive Markets.” *American Economic Review*, December 1985, *75*(5), pp. 1071–1082.

[34] **Shleifer, Andrei.** *Inefficient markets: An introduction to behavioral finance. Clarendon Lectures in Economics.* Oxford and New York: Oxford University Press, 2000.

[35] **Stigler, George J.** “The economies of scale”, *Journal of Law and Economics,* October 1958, *1*, pp. 54–71.

[36] **Strotz, Robert H**. “Myopia and Inconsistency in Dynamic Utility Maximization.” *Review of Economic Studies,* 1956, *23*(3), pp. 165–180.

[37] **Svenson, Ola.** “Are We All Less Risky and More Skillful than Our Fellow Drivers?” *Acta Psychologica*, February 1981, *47*(2), pp. 143–148.

[38] **Thaler, Richard H**. “Some Empirical Evidence on Dynamic Inconsistency.” *Economics Letters,* 1981, *8*, pp. 201–207.

Figure 1a. Timing of simple model.

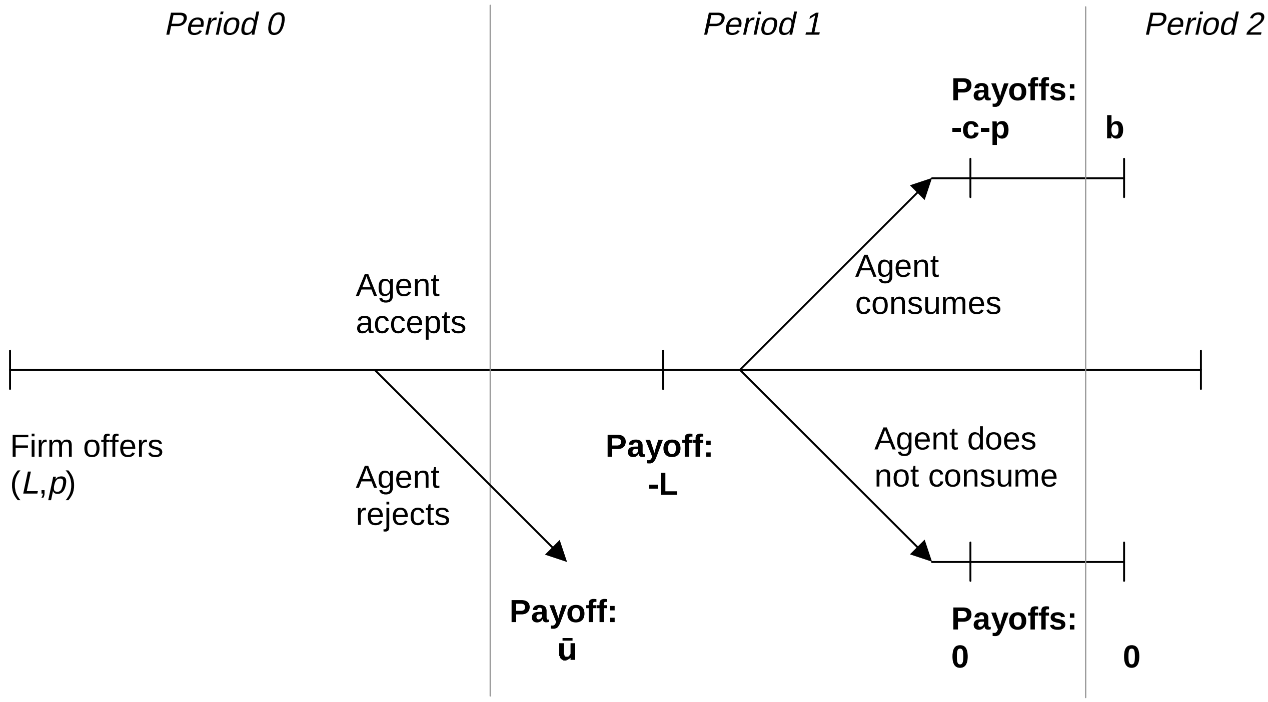

Figure 1b. Timing of model with renewal.

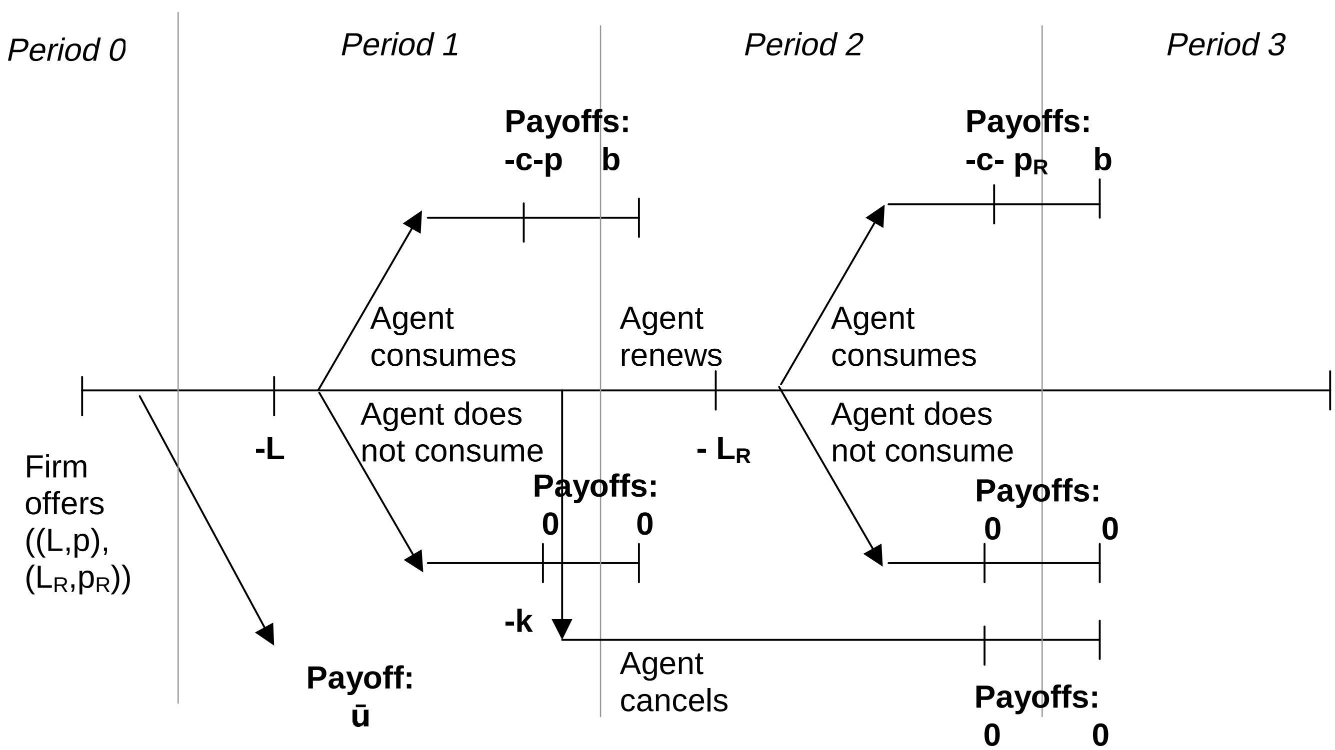

Figure 1c. Timing of model with long-term contracts.

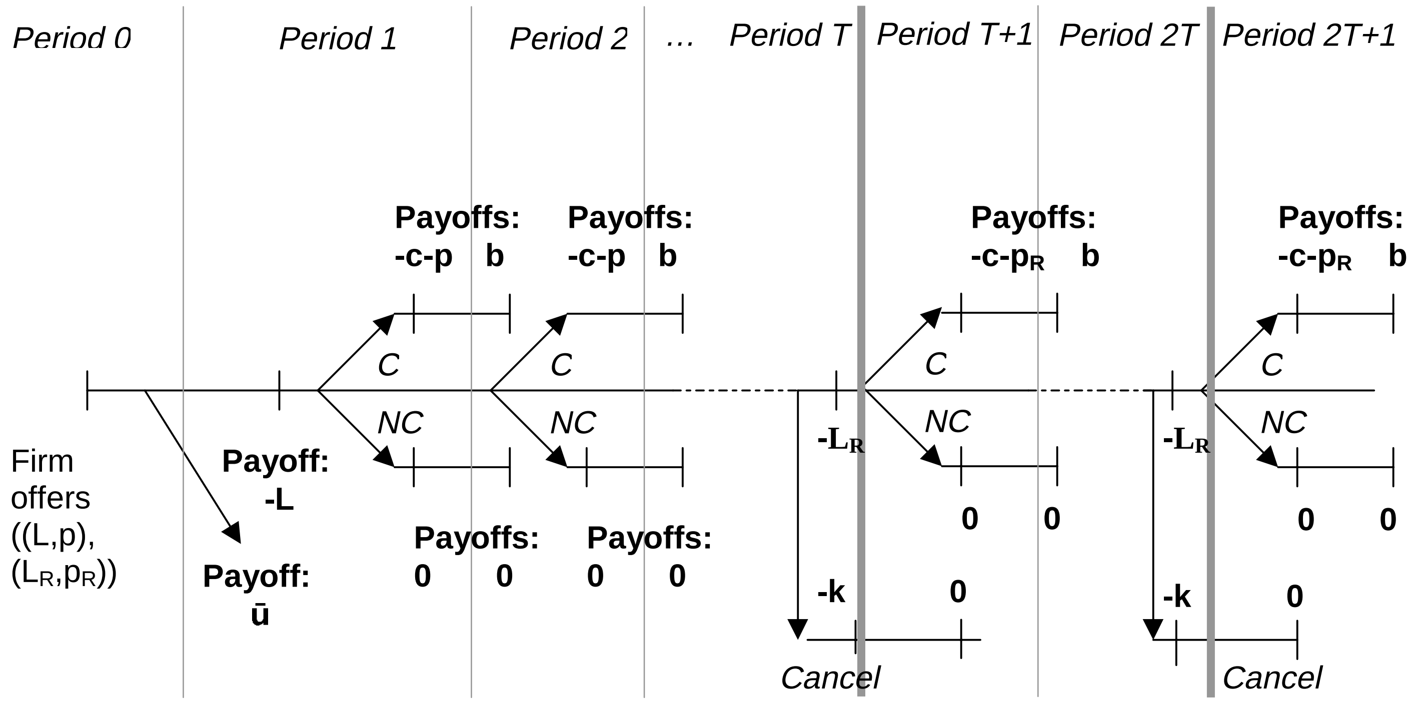

Figure 2a. Optimal level of *p*: pricing of overconfidence.

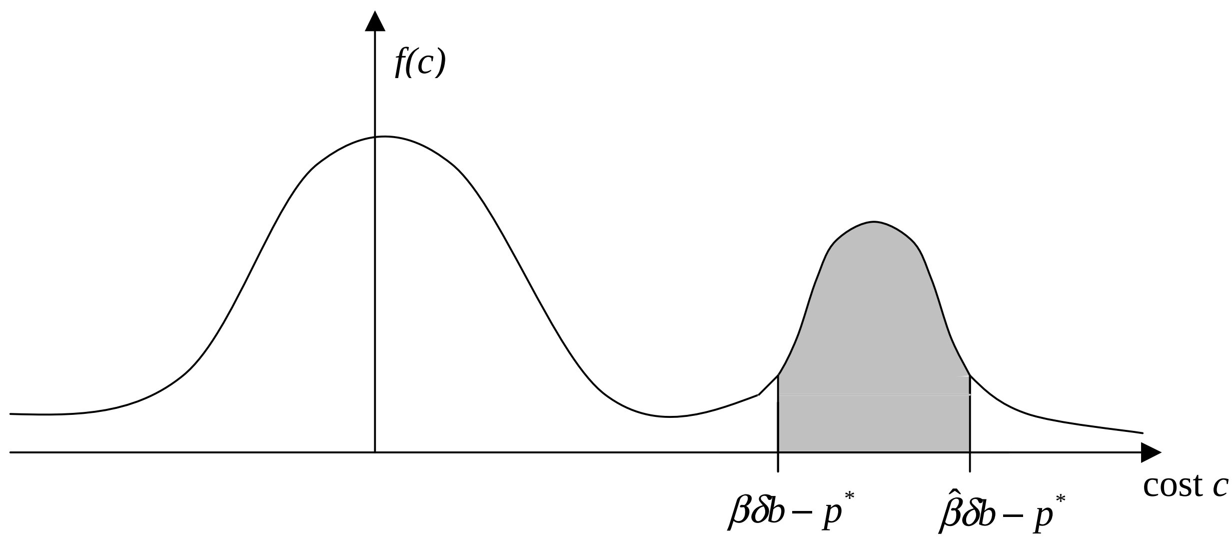

Figure 2b. Violation of asymptotically bounded peaks (ABP) condition.

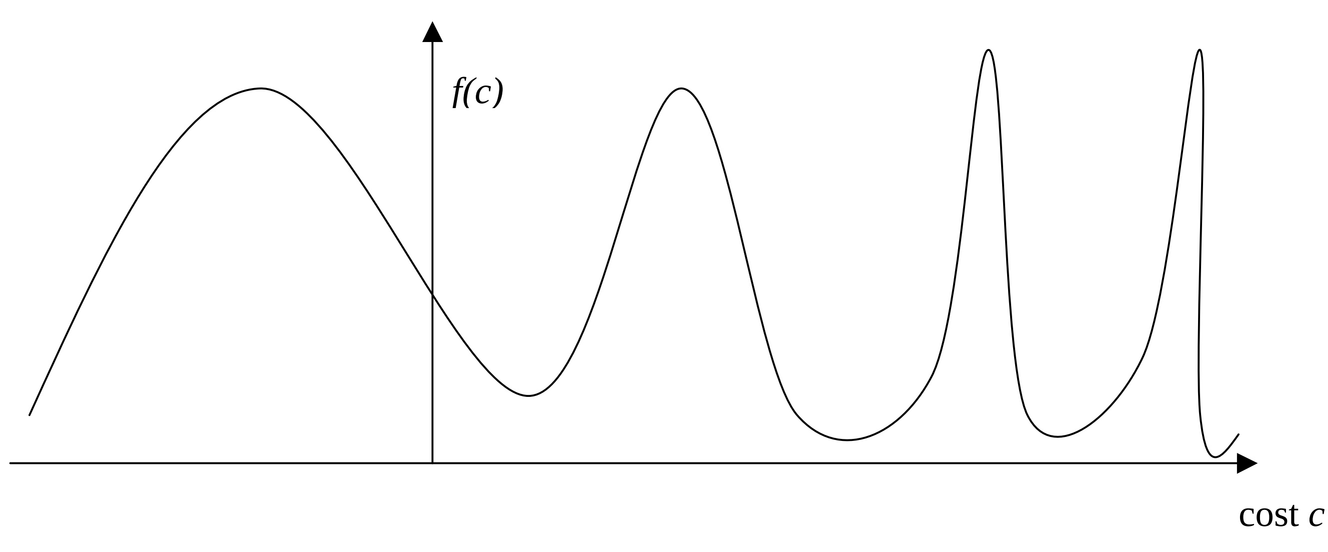

Table 1. Health Club Industry in the US. Revenue and concentration indices †

Top 10 Clubs for Revenue in the US, year 2000

	Corporate Revenue in 2000 (in m$)	Number of Clubs in 2000[1]	Number of Employees in 2000	Number of States Operates in	Year Founded
	(1)	(2)	(3)	(4)	(5)
1. Bally Total Fitness Chicago, IL	1,007	380(O), 5(F)	20,000	28 (and Canada)	1962
2. 24 Hour Fitness Worldwide San Francisco, CA	943	430(O)	26,300	15 (and 10 countries)	1983
3. Town Sports International New York, NY	225	1(O), 4(M), 104(L)	6,400	9	1973
4. Wellbridge Denver, CO	175	23(O), 14(M), 10(L)	5,000	12	1983
5. Life Time Fitness Inc. Eden Prairie, MN	101	20(O)	3,000	6 (MN,MI,OH,IN,VA,IL)	1992
6. TCA Club Management Chicago, IL	85	24(O), 22(M)	2,200	16 (and Canada)	1969
7. The Sports Club Co. Inc. Los Angeles, CA	77	7(O)	2,400	4 (CA,NY,NV,D.C.)	1978
8. Crunch Fitness International New York, NY	73	19(L), 1(LS)	2,200	5 (NY,CA,FL,IL,GA)	1989
9. Western Athletic Clubs San Francisco, CA	70	4(O), 5(L)	1,600	2 (CA,WA)	1977
10. Sport &Health Clubs McLean, VA	61.5	24(O), 1(M)	1,850	3 (VA,MD,D.C.)	1973

Concentration Indices for year 2000

Herfindahl Index (*10,000)	152.65
Concentration Ratio 4 (%)	20.25
Concentration Ratio 8 (%)	23.15
Concentration Ratio 20 (%)	27.84
Concentration Ratio 50 (%)	32.55

† **Notes:** Source: "Understanding the Top 100", Jerry Janda, Club Industry. The Corporate Revenue includes also revenue from international clubs, whenever applicable. Information comes from a survey of the companies. Since all of the companies but one are private, accuracy of the information is not guaranteed. The Concentration Indexes are from calculations of the authors using data from the publication by the Club Industry. The Herfindahl Index is the sum of the squared shares of company revenues over total industry revenue, summed across the biggest 100 companies (the contribution to the Herfindahl index of the companies past the 100^{th} is bounded above by .00162). The Concentration Ratio *n* is the ratio of the revenues by the *n* biggest companies over industry revenues.

[1] (O)=Owned, (M)=Managed, (L)=Leased, (LS)=Licensed, (F)=Franchised.

Table 2. Health club industry in Boston area. Menu of contracts [†]

	Sample: One club per company			**Sample:** All clubs		
	Monthly contract	Annual contract	Pay-per-visit	Monthly contract	Annual contract	Pay-per-visit
	(1)	(2)	(3)	(4)	(5)	(6)
Average fee [in $]:						
per visit			11.30 (4.17)			12.00 (4.75)
per month	57.59 (29.24)			61.04 (28.82)		
per year		594.23 (309.34)			649.57 (309.43)	
initiation fee	128.40 (119.49)	73.32 (121.30)		154.09 (105.05)	82.93 (115.25)	
Menu of contracts:						
No. of health clubs offering contract	53	57	51	85	90	82
No. of health clubs – Frequent Contract	47	33	2	73	47	2
No. of health clubs – Infrequent Contract	6	24	49	12	43	80
Cancellation procedure:						
Automatic renewal	51	11		83	18	
- cancel in person	51	11		83	18	
- cancel by letter	23 (12 certified)	9 (6 certified)		48 (27 certified)	16 (9 certified)	
- cancel by phone	6	1		6	1	
Automatic expiration	2	38		2	64	
Information not available	0	8		0	8	
		0				
Number of observations	N = 67	N = 67	N = 67	*N* = 100	*N* = 100	*N* = 100

[†] **Notes:** Standard deviations in parentheses. This table summarizes the features of contracts offered by health clubs in the Boston metropolitan area. Information from a survey conducted by the authors (more details in Section 3.1, transcript in Appendix B). The sample of companies contacted that fulfilled the requirements is 67. Accounting for companies with multiple clubs, the survey covers 100 different clubs. The sample "One club per company" includes only one observation from each company. The sample "All clubs" includes one observation per club. A contract is a Frequent Contract if the staff in the health club mentions the contract at the beginning of the phone interview. A contract is an Infrequent Contract if the staff mentions the contract later in the conversation or in response to specific inquiries. Fee per visit is amount due at each visit. Initiation fee is amount due at sign-up for monthly and annual contract.

Table 3. Credit card industry. Representative contracts [†]

	Type of credit card offer (1)	Regular interest rate (APR) (2)	Annual fee in $ (3)	Benefits (4)	Introductory interest rate (APR) (5)	Length of introductory offer (6)	Late Fee in $ (7)	Overlimit Fee in $ (8)
Citibank	Platinum Select Visa	Prime + 12.99%	0		2.90%*	9 months	15-25-35	29
MBNA	Platinum Plus Visa	12.99%	0		3.90%*	6 months	29	29
First USA	Platinum Visa	Prime + 6.50%	0		9.90%*	9 months	29	29
Chase Manhattan	Wal-Mart Mastercard	Prime + 3.98% to Prime + 11.98%	0		0%	6 months	29	29
Bank of America	Visa Gold	Prime + 7.99% to Prime + 12.99%	0		3.90%	6 months	29	29
Household Bank	GM Mastercard	Prime + 9.99%	0	5% toward GM	2.90%	6 months		
Providian	Visa Platinum	Prime + 3.24%	0		0%	3 months	29	29
	Visa Gold Prestige	Prime + 10.24%	0		0%	2 months	29	29
	Visa Gold Preferred	Prime + 13.24%	0		0%	2 months	29	29
	Visa Classic	Prime + 17.24%	0-59-89		0%	2 months	29	29
Capital One	Platinum Visa	9.90%	0		N/A	N/A	29	29
	Gold Visa	14.90%	0		2.90%*	6 months	29	29
	Classic Visa	19.80%	49		N/A	N/A	29	29
Discover	Platinum Card	13.99%	0	1% Cashback	1.70%*	6 months	29	29
American Express	Blue Credit Card	9.99%	0		0%	6 months	29	29
	Optima Credit Card	Prime + 7.99%	0		7.90%	6 months	29	29
	(Gold) Charge Card	N/A	55-75		N/A	N/A	29	N/A

[†] **Notes:** Information about typical credit card offers from the issuers' websites. The displayed issuers are the largest 6 issuers ranked by outstandings as of 1997, excluding First Chicago NBD that merged with BancOne/FirstUSA (Evans and Schmalensee, 1999, p. 229). We also include information on Providian (12th largest issuer) and CapitalOne (8th largest issuer) because of the availability of information on the menu of cards offered. Finally, we included Discover and American Express that are the biggest issuers outside the Visa and Mastercard circle. We include also, for comparison, a charge card (last row). The regular interest rate is the APR on the outstanding balance for customers that pay regularly. Interest rate may depend on the credit history. Introductory APR (Column 5) is the interest rate on the outstanding balance for the introductory period (Column 6), after which the relevant interest rate becomes the one in Column 2. Values with a star (*) are credit cards for which the introductory offer applies only for Balance Transfers.

Table 4. Mobile phone industry. Menu of contracts †

	Revenues in year 2000 [$m] (1)	Monthly allowance (2)	Monthly fee in $ (3)	Average price per minute in ¢ (4)	Price of additional minutes in ¢ (5)	(5)/(4) (6)
AT&T	7,627[1]	450	59.99	13.3	35	**2.63**
		650	79.99	12.3	35	**2.84**
		900	99.99	11.1	25	**2.25**
		1,100	119.99	10.9	25	**2.29**
		1,500	149.99	10.0	25	**2.5**
		2,000	199.99	10.0	25	**2.5**
Sprint PCS	6,341	20	19.99	100.0	40	**.40**
		200	34.99	17.5	40	**2.28**
		350	39.99	11.4	40	**3.50**
		450	49.99	11.1	40	**3.60**
		1,000	74.99	7.5	40	**5.33**
Verizon	14,236	150	35	23.3	40	**1.71**
		400	55	13.7	35	**2.55**
		600	75	12.5	35	**2.80**
		900	100	11.1	25	**2.25**
		1500	150	10.0	25	**2.50**
		2000	200	10.0	20	**2.00**
		3000	300	10.0	20	**2.00**

† **Notes:** Information from the website of the companies. Allowances and rates are for calls for plans with domestic long distance included. For Sprint, the allowance applies only to daytime, weekday calls. For Sprint, additional minutes for evening and weekend calls are not included in the computations. Annual Revenue is from 10K filing and refers to the cellular phone business exclusively. Monthly allowance is the total number of monthly minutes that the consumer can use without incurring any extra charge. Average price per minute is ratio of Columns 3 and 2. Price of additional minutes is cost per minute of each call beyond the monthly allowance.

[1] This is the revenue in year 1999, since the data from year 2,000 was not available.