Optimal Screening of Time Inconsistency

Simone Galperti*

November 9, 2013

Abstract

This paper develops a theory of optimal provision of commitment devices to people who value both commitment and flexibility, and whose preferences differ in the degree of time inconsistency. If time inconsistency is observable, then both a planner and a monopolist provide devices that help each person commit to the efficient level of flexibility. But the combination of unobservable time inconsistency and preference for flexibility creates an adverse-selection problem. To solve it, the monopolist and (possibly) the planner inefficiently curtail flexibility in the device for a more inconsistent person, and may have to add unused options to, or even distort, the device for a less inconsistent person. Flexibility is curtailed in a particular way that is evocative of existing commitment devices. This theory has normative as well as positive implications for private and public provision of these devices.

KEYWORDS: Time inconsistency, self-control, commitment, flexibility, screening, unused options.

JEL CLASSIFICATION: D42, D62, D82, D86, D91, G21, G23.

*I am indebted to Eddie Dekel and Alessandro Pavan for many long, fruitful discussions that greatly improved the paper. I also thank S. Nageeb Ali, Gene Amromin, Stefano DellaVigna, Jeffrey Ely, William Fuchs, Garrett Johnson, Botond Koszegi, David Laibson, Santiago Oliveros, Jonathan Parker, Nicola Persico, Henrique Roscoe de Oliveira, Todd Sarver, Ron Siegel, Bruno Strulovici, Balazs Szentes, Rakesh Vohra, Asher Wolinsky, Michael Whinston, Leeat Yariv, and seminar participants at Northwestern, NYU, Berkeley, Duke, UCSD, Michigan, NYU Stern, LSE, EEA-ESEM 2012, ESSET 2013, SAET 2013. I gratefully acknowledge financial support from the Center of Economic Theory of the Weinberg College of Arts and Sciences of Northwestern University. All remaining errors are mine.

1 Introduction

Evidence suggests that many people have self-control problems (see DellaVigna's (2009) survey). Often aware of these problems, people demand commitment devices. This demand has received the attention of different institutions: Firms, like StickK and GymPact, sell devices that help people commit to their goals, and some governments set up tax incentives to help people adequately save for retirement—in the US, through devices like individual retirement accounts (IRAs) and 401(k) plans.[1] But, uncertain about the future, people demand commitment devices that also allow for flexibility. Can firms or governments satisfy these opposite desires for commitment and flexibility? Moreover, the degree of self-control—and hence the demand for commitment—varies across people and is not immediately detectable. How does this affect the provision of commitment devices?

This paper answers these questions by developing a theory of optimal provision of flexible commitment devices—from the point of view both of a profit-maximizing firm and of a welfare-maximizing planner. The paper first shows that the *combination* of people's demand for flexibility and superior information on their self-control creates an adverse-selection problem. Then, it shows that this problem leads to a trade-off between commitment and flexibility, and characterizes how, as a result, the firm and (possibly) the planner optimally curtail flexibility for people with self-control problems. This theory could be a basis to guide private and public provision (or regulation) of commitment devices, to predict its possible inefficiencies, and to explain some broad features of existing commitment devices.

The model features a provider (she), an agent (he), and two periods. In period 1, the provider offers the agent a device that, in period 2, allows him to choose among several actions and, for each action, charges a payment. An example would be a savings device that allows the agent to make deposits and withdrawals, whose amount determines a fee or a tax. In period 1, the agent desires flexibility (in the sense of Kreps (1979)) because his preference over period-2 actions depends on an uncertain state; importantly, the state is not contractible. Moreover, in period 1, the agent can desire commitment because his preference can be time inconsistent (Strotz (1956)). In line with most of the literature, this paper uses the period-1

[1] Other typical examples of commitment devices include automatic drafts from checking to investment accounts, Christmas clubs, rotating savings and credit associations, microcredit savings accounts in developing countries, fat farms, and programs to reduce consumption of cigarettes, alcohol or drugs. (See Ashraf et al. (2003), Ashraf at al. (2006), DellaVigna and Malmendier (2004, 2006), Bryan et al. (2010)).

preference to measure efficiency. Finally, only the agent knows his degree of time inconsistency—his type—from period 1. For illustration, suppose the agent can be either consistent (type C) or inconsistent (type I).

It is common to think that coexisting preferences for commitment and for flexibility must involve a trade-off, which may be reflected in the design of commitment devices, even when the provider observes the agent's degree of inconsistency (Amador et al. (2006); Ambrus and Egorov (2013); Bond and Sigurdsson (2013)). In contrast, Section 3 shows that, without restrictions on monetary incentives (the device payments), it is possible to design a device that commits type I to a flexible plan of action that is efficient in each state. Moreover, if types were observable, the provider would always offer I an efficient device—even if she cares only about profits. This is because, in period 1, type I knows his self-control problems and therefore is willing to pay more for an efficient device.[2] Section 3 also shows how efficient devices must feature different payments as the agent's degree of inconsistency changes; nonetheless, they all induce a behavior with the same level of flexibility. Concretely, in a savings application, an efficient device for I—but not that for C—penalizes withdrawals and rewards deposits; in an exercising application, it penalizes missed workouts and rewards attended ones. Nonetheless, in each state, savings and workouts are the same across types.

With unobservable types, however, the question is whether it is possible and, in particular, optimal to make each type self-select an efficient device. Section 4, the core of the paper, first shows that a less inconsistent agent values any flexible device strictly more than a more inconsistent agent, creating the adverse-selection problem. This is because, for example, an efficient savings device for I rewards deposits and penalizes withdrawals. But if C takes this flexible device, thanks to his higher self-control, he expects to incur less penalties and enjoy more rewards than I; as a result, C expects a higher payoff. As in usual screening models, for C not to take the device for I, the device for C must then grant him an information rent. But in contrast to those models, unfortunately, the rent that makes C just prefer the device for himself can suddenly make I strictly prefer this device too. Intuitively, a device for C does not feature payments that will solve I's self-control problems, yet C's minimal rent can still be enough to lure I. This possibility creates an unusual situation: When designing each device, the provider has to worry about *both* types' incentives to mimic one another.

Section 4 then derives the screening devices that solve this unusual adverse-

[2] This result generalizes a result of DellaVigna and Malmendier (2004).

selection problem. This derivation involves designing, for each type, a *menu* of actions and associated payments and requires some nonstandard techniques (see below).

The screening device for I curtails flexibility below efficiency. Although in principle it can do so in many ways, this paper tightly characterizes the optimal way: Flexibility is curtailed at both ends of the efficient choice range. Specifically, type I reacts to both high and low states less than efficiently—resulting in a narrower choice range—and does not react at all to extreme states, again both high and low. These inefficiencies are induced by modifying the payments I faces in an efficient device. For example, suppose there are four states, $s_2 > s_1 > s_{-1} > s_{-2}$, with corresponding efficient savings 2, 1, -1, and -2 (the last two are withdrawals). Then, the device for I makes him save, say, 0.9 in s_1 and s_2 and -0.5 in s_{-1} and s_{-2}, by making the withdrawal penalties stiffer and the deposit rewards weaker beyond certain amounts. Intuitively, the device for I curtails flexibility, because by committing I to a less flexible savings plan, it also gives C fewer chances to benefit from his higher self-control and thus cuts the rent that is necessary to make C choose the device for himself.

The screening device for C may also have to depart from the optimal device under observable types, in order to ward off I. The first strategy the provider adopts to do so is unconventional. In standard models with only time-consistent agents, the provider would have to distort C's choices to deter I from mimicking C. Here, instead, she first adds options C never uses but I views as temptations, to an otherwise efficient device for C, thus making it less attractive for I. But to ward off I, the unused options must be tempting enough; otherwise, the provider will also have to distort C's choices. Therefore, screening time inconsistency violates the usual 'no distortion at the top' property that marks previous screening models, both static and dynamic (e.g., Mussa and Rosen (1978); Courty and Li (2000); Battaglini (2005)).

Some features that distinguish the screening devices from the efficient ones are broadly consistent with existing commitment devices. For example, some devices that incentivize savings—like IRAs and 401(k) plans—also restrict both withdrawals and deposits through dear monetary penalties; and there is evidence consistent with the principle that such restrictions dissuade people who value flexibility more than commitment from using those devices (Amromin (2002, 2003)). As another example, some devices that provide monetary incentives to work out regularly—like GymPact—also limit the maximum and minimum number of work-

outs they incentivize. Section 4.4 discusses this evidence and alternative explanations.

This paper relates to the literature on the trade-offs between preferences for commitment and for flexibility. Amador et al. (AWA) (2006), Ambrus and Egorov (AE) (2013), and Bond and Sigurdsson (BS) (2013) characterize commitment policies an inconsistent agent would impose on himself, when he also values reacting to future information. These papers assume technology constraints on the feasible policies—they rule out monetary transfers across states—which, as noted, can lead to trade-offs between commitment and flexibility. In contrast, in the present paper without restrictions on transfers, a tension between commitment and flexibility arises from a quite different, perhaps unexpected, source: information constraints. AWA, AE, and BS also discuss how the agent can implement his policies with contracts already available in the market (like illiquid assets), even though they are not designed to serve as commitment devices. In contrast, this paper explicitly studies the provision of such devices by third parties like banks, gyms, or governments.[3] Section 4.4 compares the policy implications of AWA, AE, BS with those of this paper.

Other papers have studied the problem of designing incentives, or contracts, for agents with low self-control. This paper, however, is the first to characterize the optimal provision of commitment devices to agents who value both commitment and flexibility, and privately know their degree of time inconsistency. In some earlier papers (e.g., O'Donoghue and Rabin (1999b); DellaVigna and Malmendier (2004)), the agent has no private information when he is offered a contract.[4] In other papers (e.g., Eliaz and Spiegler (ES) (2006); Esteban and Miyagawa (EM) (2006b); Heidhues and Koszegi (HK) (2010)), the agent has some private information from the outset, but has no preference for flexibility. Moreover, in ES, the agent privately knows whether he is aware of his time inconsistency, but its actual degree is known to the mechanism designer. ES therefore study a different screening problem: For instance, if the agent can be of two types—sophisticated or naive—then the designer can always screen them without losing any profit (see Spiegler (2011)). In contrast, here she can never screen `C` and `I` without losing

[3] O'Donoghue and Rabin (2007) as well as Bryan et al. (2010) ask whether markets can provide products that solve people's commitment problems, but leave the answer to future research.

[4] Solving specific cases of DellaVigna and Malmendier's (2004) model, Jianye (2012) shows that their results may change if the agent has private information on different aspects of his preference, including the degree of time inconsistency.

profits. In EM, the designer has to screen the agent's valuation for her good—not his degree of inconsistency. However, EM show that the agent's low self-control can help the designer extract more profits than in standard monopolistic-screening models. Finally, even when allowing for asymmetric information from the outset, HK focus on settings in which self-selection is guaranteed by the contracts that are optimal under symmetric information.[5]

The result that the device for type `C` may have to contain unused options is based on a key insight of Gul and Pesendorfer (2001): Agents who are prone to temptations (or are time inconsistent) dislike menus with more options, as such menus make the woes of temptation more likely. This insight is also behind related results in BS and EM. The present model, however, differs in ways that lead to substantive differences in when and how the provider has to rely on unused options, as explained in Section 4.4.

By examining welfare maximization, this paper also speaks to the literature on optimal paternalism (e.g., O'Donoghue and Rabin (2003)). It indicates a reason for some public provision of commitment devices: The inability to observe people's self-control problems leads profit-maximizing firms to create inefficiencies for those who most need commitment.[6] In general, a paternalistic planner can achieve higher efficiency. She may, however, not be able to reach full efficiency, because of the adverse-selection problem identified in this paper. In the specific case of devices that rely on taxation, like IRAs and 401(k) plans, this paper also shows that the planner will face a trade-off between her corrective goal—helping `I` save adequately for retirement—and her redistributive goal—collecting tax revenues from `C`.

Finally, this paper relates to the literature on dynamic mechanism design. On the methodological side, it highlights the following point about direct mechanisms. We know from Myerson (1986) that, in models with only time-consistent agents, one can always restrict attention to mechanisms that describe only options on the truthful path of play, thus preventing the agents from revealing lies. But with time-inconsistent agents, this is not true: To find the optimal mechanisms, one *must* allow for off-path options and thus allow the agents to reveal if they lie—even though lies never occur in equilibrium (see Section 4.1). On the technical side, this paper uses nonstandard and recent methods to find the mechanisms

[5] The literature on contracting with behaviorally biased agents also includes, among others, DellaVigna and Malmendier (2006); Esteban and Miyagawa (2006a, 2007); Eliaz and Spiegler (2008); Grubb (2009); and Spiegler (2011).

[6] Section 5.2 argues that competition alone may fail to remove these inefficiencies.

that optimally screen time inconsistency. For reasons that will be explained later, to handle the incentive constraints involving the agent's degree of inconsistency, it uses Lagrangian methods from Luenberger (1969). These methods differ from standard optimal-control methods and the standard dynamic-mechanism-design approach (Courty and Li (2000); Pavan, Segal, and Toikka (2012)). Finally, to ensure that the mechanisms satisfy certain monotonicity properties, it adapts Toikka's (2011) generalization of Myerson's (1981) ironing method to allow for off-path options.

Section 5 shows that the insights of the paper generalize to settings with naive agents, competition among providers, and more than two types. Section 6 concludes. All proofs are in the appendix.

2 The Model

This section sets up a simple two-period model in which a party (the provider) supplies commitment devices to another party (the agent). As noted, Section 5 will extend the model and its analysis in several directions.

The provider offers her devices in period 1. Each device allows the agent to choose in period 2 among several options, each consisting of a contractible action a and a payment p to the provider. The set of feasible actions is $[\underline{a}, \overline{a}] \subset \mathbb{R}$ with $-\infty < \underline{a} < \overline{a} < +\infty$. The provider incurs a cost to produce a in period 2, given by the twice-differentiable function $c : [\underline{a}, \overline{a}] \to \mathbb{R}$ with $c' \geq 0$ and $c'' \geq 0$. By assumption, the provider can fully commit to any device—which, if chosen in period 1, becomes binding for the agent—and no third party can, in period 2, offer the agent contracts that may interfere with the provider's devices. Relaxing either of these assumptions can undermine the provider's ability to supply commitment devices and raises issues that, though important, are beyond the scope of this paper.

The agent may have time-inconsistent preferences and is fully aware of it (sophistication). To model time inconsistency, this paper follows Strotz (1956). The agent has two selves: *self-1* lives in period 1 and chooses a device; *self-2* lives in period 2 and picks an option from the device chosen by self-1. Both selves' preferences depend on state s, which occurs in period 2 and can reflect taste, income, or health shocks. These shocks induce self-1 to desire flexibility; moreover, they are not contractible, for example because only the agent observes them.[7] Conditional

[7] If s were contractible, the agent could simply delegate his future choices to the provider and

on s, self-1's and self-2's *direct* utilities from action a are

$$u_1(a;s) = sb(a) - a \quad \text{and} \quad u_2(a;s,t) = tsb(a) - a,$$

where $b : [\underline{a}, \overline{a}] \to \mathbb{R}$ is twice differentiable with $b' > 0$ and $b'' < 0$. The distribution of s is F, which is commonly known in period 1 and has continuous and strictly positive density f over $[\underline{s}, \overline{s}] \subset \mathbb{R}$ with $0 < \underline{s} < \overline{s} < +\infty$. This simply says that, in each state, the agent assigns a weight to the benefit and cost of action a that is bounded away from zero. Finally, self-1's and self-2's total utilities are $u_1(a;s) - p$ and $u_2(a;s,t) - p$. These functions have two properties that help, as we will see, to keep the model tractable: (1) s, t, and the function b enter multiplicatively; (2) p interacts neither with s nor with t.

The positive parameter t determines the preferences' degree of (time) inconsistency and can lead self-1 to desire commitment. When $t \neq 1$, self-1 foresees that, in each state, self-2 trades off the benefit and cost of a in a systematically different way. This modeling assumption is based on the idea, proposed by cognitive psychologists, of 'salience:' Decision-makers seem to perceive the cost of their actions ($-a$) as more (or less) salient than the benefit ($sb(a)$), depending on whether they are considering an immediate or a future decision (see, e.g., Akerlof (1991) and the references therein).

The model captures, in a stylized way, some key common aspects of different settings.

Example 1 *In period 1, the government (or a bank) offers savings devices to the agent, who is planning his future savings. A device allows him to make deposits and withdrawals in period 2 (his working life), whose amount, a, determines a tax or fee, p; the device gives a return at retirement, based on an exogenous rate. The government can design different devices by changing how p depends on a.*[8] *In period 1, the agent knows his period-2 income, y, but is uncertain about his rate, s, of intertemporal substitution between period 2 and retirement—for instance, because it depends on his period-2 health. So, self-1 assigns utility $y-a-p+sb(a)$ to saving a and paying p in state s, where $y-a-p$ is self-2's immediate consumption and $b(a)$ is the expected utility at retirement. However, self-2's utility function is $y-a-p+tsb(a)$ with $0 < t \leq 1$, since period-2 consumption is then more salient*

thus bypass his self-control problems.

[8] Note that p can have two parts, p_1 and p_2, where only p_2 depends on a and p_1 is possibly paid by the agent in period 1, as an initial deposit.

(Phelps and Pollack (1968); Laibson (1997)).[9]

Example 2 *In period 1, a gym offers memberships, which allow the agent to work out in period 2 (the following month) at its facility; the time spent there determines a fee, p. A workout of e hours causes immediate discomfort, but improves future health by $\hat{b}(e)$ (with $\hat{b}' > 0$ and $\hat{b}'' < 0$). The agent's self-1 knows that his marginal disutility from e will depend, for example, on whether he is sick (s). Self-1 may also foresee that self-2 will always overweigh the discomfort of e and thus will tend to work out less than self-1 wants ex ante. To model the agent's preferences, we can use $\hat{b}(e) - se$ for self-1 and $\hat{b}(e) - tse$ for self-2, with $t > 1$. Letting $a = -\hat{b}(e)$ and assuming quasi-linearity in p, we can write these preferences as $u_1(a; s) - p$ and $u_2(a; s, t) - p$, with the properties assumed above.*[10]

For clarity's sake, in most of the paper the agent can be one of two types: type $\mathtt{C}$ has $t^{\mathtt{C}} = 1$ and is consistent; type $\mathtt{I}$ has $0 < t^{\mathtt{I}} < 1$ (as in the savings example) and is inconsistent. Being sophisticated, the agent knows t in period 1. In contrast, the provider cannot observe t; she, however, knows the possible types and the probability $\gamma \in (0, 1)$ of type $\mathtt{C}$.

The overall utility that type j gets from a device in period 1 depends on what j's self-2 does in period 2. Type j's utility from a device is then just the expected utility of self-2's ensuing decisions, computed using self-1's preference. If the agent rejects all the provider's devices, he gets the outside option whose value is normalized to zero.

As in other models with time inconsistency, the choice of a welfare criterion is delicate. This paper adopts the following criterion, which adheres to the usual interpretation of self-1's preference as the agent's long-run preference (see, e.g., O'Donoghue and Rabin (1999a); DellaVigna and Malmendier (2004)).

Definition 1 (Efficiency) *In state s, the* social surplus *of action a is $u_1(a; s) - c(a)$, and the* efficient outcome *is*

$$\mathbf{a}^*(s) = \arg\max_{a \in [\underline{a}, \overline{a}]} u_1(a; s) - c(a).$$

By the properties of u_1 and c, the function $\mathbf{a}^*$ is strictly increasing if it always takes interior values in $[\underline{a}, \overline{a}]$. A strictly increasing $\mathbf{a}^*$ sets a clearer benchmark

[9] Section 5.6 discusses the case with utility functions that are not linear in p.

[10] The interpretation of some gym memberships as examples of commitment devices already appears in DellaVigna and Malmendier (2004, 2006).

in terms of the efficient level of flexibility—in this case, efficiency always calls for different actions in different states. For this reason, this paper assumes that the smallest and largest feasible actions ($\underline{a}$ and $\overline{a}$) are never efficient. It also assumes that the maximum social surplus is positive in every state.

Assumption 1 *For every s, $\mathbf{a}^*(s)$ is interior, and $u_1(\mathbf{a}^*(s); s) - c(\mathbf{a}^*(s)) > 0$.*

Finally, when designing her devices in period 1, the provider maximizes a weighted sum of expected profits and expected social surplus, with respective weights $\pi \in [0,1]$ and $1-\pi$. This is a convenient way to allow for a monopolist ($\pi = 1$), as well as for a paternalistic planner who may have to worry about the profitability of her devices ($\pi < 1$). This situation may arise if the planner regulates the market of commitment devices and has to ensure that third-party providers expect enough profits to enter the market, or if she exclusively provides such devices and has limited funds to finance them—consider, for example, a government providing tax incentives for savings. In these cases, one could let the planner maximize the expected social surplus subject to making some minimum (possibly negative) profit. This alternative setup would only make π endogenous, without changing the thrust of the paper (see Online Appendix C).

3 Observable Time Inconsistency

To better understand how the inability to observe the agent's degree of inconsistency t affects the provider's supply of commitment devices, this section characterizes her optimal devices when she can observe t.

Such devices can be characterized using direct mechanisms (DMs) that make the agent report truthfully state s in period 2. Formally, each DM consists of two functions, $\mathbf{a} : [\underline{s}, \overline{s}] \to [\underline{a}, \overline{a}]$ and $\mathbf{p} : [\underline{s}, \overline{s}] \to \mathbb{R}$, and must satisfy the constraints

$$u_2(\mathbf{a}(s); s, t) - \mathbf{p}(s) \geq u_2(\mathbf{a}(s'); s, t) - \mathbf{p}(s') \tag{IC}$$

for all s, s' and

$$\int_{\underline{s}}^{\overline{s}} [u_1(\mathbf{a}(s); s) - \mathbf{p}(s)] f(s)\, ds \geq 0. \tag{IR}$$

Note that (IR) depends on self-1's preference, but (IC) depends on self-2's preference. In period 1, the provider maximizes

$$(1-\pi)\int_{\underline{s}}^{\overline{s}} [u_1(\mathbf{a}(s); s) - c(\mathbf{a}(s))] f(s)\, ds + \pi \int_{\underline{s}}^{\overline{s}} [\mathbf{p}(s) - c(\mathbf{a}(s))] f(s)\, ds.$$

Although varying π changes the provider's goal, it turns out that she always finds it optimal to offer an inconsistent agent a device that solves his self-control problems.

Lemma 1 (First Best) *If the agent's degree of inconsistency t is observable, then for any $\pi \in [0,1]$ and $t > 0$, the optimal device sustains $\mathbf{a}^*$ and yields the same expected profits. Moreover, let $\mathbf{p}_t^*$ be the payment scheme that sustains $\mathbf{a}^*$ with t. Then, $\frac{d\mathbf{p}_1^*(s)}{ds} = c'(\mathbf{a}^*(s))\frac{d\mathbf{a}^*(s)}{ds}$ and, for $t \neq 1$,*

$$\frac{d\mathbf{p}_t^*(s)}{ds} = \frac{d\mathbf{p}_1^*(s)}{ds} + s(t-1)b'(\mathbf{a}^*(s))\frac{d\mathbf{a}^*(s)}{ds}.$$

The intuition for Lemma 1 follows. As usual, since the agent has no private information in period 1, the provider can extract the whole utility the agent expects when choosing a device, i.e., the whole expected utility of *self-1*. For any π, the provider then wants to maximize the expected social surplus, which requires sustaining $\mathbf{a}^*$. Compared to models with only consistent agents, however, there is a twist: If the agent is inconsistent, the provider has to offer him tailored incentives (the function $\mathbf{p}_t^*$) so that his *self-2* will comply with plan $\mathbf{a}^*$. In general, such incentives may not exist (e.g., if $t < 0$ in this model). But they always exist when $t > 0$, because self-2 prefers higher actions in higher states as prescribed by $\mathbf{a}^*$.

Lemma 1 generalizes a similar result of DellaVigna and Malmendier (2004)—by allowing for more than two values of action a—and has several implications. First, profit maximization alone leads a firm to provide a full solution to the agent's time inconsistency.[11] This result, however, relies on both self-1 and self-2 preferring higher actions in higher states ($t > 1$); it may therefore not hold, even with symmetric information, for other forms of time inconsistency. Second, the provider is indifferent between trading with a consistent agent and trading with an inconsistent agent of any degree t, given their common period-1 preference. Finally, the first-best devices induce behaviors whose level of flexibility is invariant across types, even though they provide incentives ($\mathbf{p}_t^*$) that vary with t: As the agent becomes more inconsistent (i.e., t moves farther away from 1), when a increases $\mathbf{p}_t^*$ increases faster when $t > 1$ and slower then $t < 1$.

Concretely, the properties of $\mathbf{p}_t^*$ can be interpreted as follows. The first-best device for `C` involves a payment scheme, $\mathbf{p}_1^*$, that makes `C` internalize the produc-

[11] It is easy to see that, with perfect competition among providers, each device continues to be efficient, but the agent enjoys the entire surplus from it.

tion cost and hence choose efficiently. On the other hand, in the savings example with $t^{\mathtt{I}} < 1$, the first-best device for $\mathtt{I}$ involves $\mathbf{p}_1^*$ combined with payments that increase as the agent deposits less or withdraws more (e.g., strictly increasing rewards for deposits and analogous penalties for withdrawals). Similarly, in the gym example with $t^{\mathtt{I}} > 1$, the first-best device for $\mathtt{I}$ involves $\mathbf{p}_1^*$, combined with strictly increasing rewards for attended workouts and analogous penalties for missed workouts (recall that $a = -\hat{b}(e)$ where e are workout hours). To see this, for any s and $t > 0$, let $\mathbf{d}_t^*(s) = \mathbf{p}_t^*(s) - \mathbf{p}_1^*(s)$. Note that $\mathbf{d}_t^*$ is continuous in s and $\frac{d\mathbf{d}_t^*(s)}{ds} \gtrless 0$ if and only if $t \gtrless 1$. Moreover, since the expected profit and the sustained allocation are the same across t's, the expected revenue must be the same across t's; that is, $\int_{\underline{s}}^{\overline{s}} \mathbf{d}_t^*(s) f(s) ds = 0$ for all t. So, for $t < 1$, there is $\hat{s} \in (\underline{s}, \overline{s})$ such that $\mathbf{d}_t^*(s) > 0$ for $s < \hat{s}$ and $\mathbf{d}_t^*(s) < 0$ for $s > \hat{s}$. Similarly, for $t > 1$, there is $s' \in (\underline{s}, \overline{s})$ such that $\mathbf{d}_t^*(s) < 0$ for $s < s'$ and $\mathbf{d}_t^*(s) > 0$ for $s > s'$.

This paper aims to explain why and how the agent's private information on his time inconsistency alters the provider's supply of commitment devices, relative to the first best. One can show that if the number of states is finite, then for $t^{\mathtt{I}}$ close to $t^{\mathtt{C}}$ the provider can (and will) sustain $\mathbf{a}^*$ without worrying about the agent's private information. Intuitively, with discrete states, many incentive schemes (i.e., functions $\mathbf{p}$) can sustain $\mathbf{a}^*$ with each type. Moreover, for $t^{\mathtt{I}}$ close to $t^{\mathtt{C}}$, there is a single $\mathbf{p}$ that sustains $\mathbf{a}^*$ with both types, so private information does not matter (see Online Appendix B). This never happens, however, with a continuum of states. For this reason, the paper focuses on this case.

4 Unobservable Time Inconsistency

4.1 The Screening Problem

When the provider cannot observe the agent's degree of inconsistency, she has to design commitment devices—one for each type—that satisfy two kinds of incentive-compatibility conditions. First, type j's self-1 must select the device designed for j, hereafter called 'j-*device.*' Second, after selecting a j-device, type j's self-2 must choose, at each state, the option that the provider designed for j to choose in that state. As usual, this design problem can be analyzed using direct mechanisms (DMs) that make the agent reveal, sequentially, his period-1 and period-2 information.

However, since in this model the agent can be time inconsistent, one has to

specify carefully what information self-2 can reveal to the DMs. In contrast to models with only consistent agents (Myerson (1986)), DMs that let self-1 report t and self-2 report only s—his 'incremental' information—entail a loss of generality. Richer mechanisms allow the provider to offer j-devices with more options than j's self-2 will ever use. These unused options can never help her screen consistent agents, but they can help her screen inconsistent agents: By representing temptations, unused options may deter self-1 of an inconsistent agent other than j from choosing the j-device (see Proposition 4). So, to describe unused but tempting options, DMs must allow self-2 to report more than just s: They must allow him to report how his preference depends on all his information, both s and t. Note that this dependence is summarized by the product ts, which pins down self-2's marginal valuation of a. Therefore, let $\underline{v}^j = t^j \underline{s}$, $\overline{v}^j = t^j \overline{s}$, and $[\underline{v}, \overline{v}] = [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$.

Without loss of generality, we can focus on DMs that satisfy two properties. First, each DM must assign a pair (a, p) to each sequential report of t and $v \in [\underline{v}, \overline{v}]$—these reports correspond to choosing a device in period 1 and one of its options in period 2, respectively. Second, each DM must ensure that truthfully reporting t is optimal in period 1, and that truthfully reporting v is optimal in period 2 for *any* report about t.[12] Formally, each DM is then an array $\{\mathbf{A}, \mathbf{P}\} = (\mathbf{a}^j, \mathbf{p}^j)_{j=\mathtt{C},\mathtt{I}}$, where $\mathbf{a}^j : [\underline{v}, \overline{v}] \to [\underline{a}, \overline{a}]$ is an *allocation* and $\mathbf{p}^j : [\underline{v}, \overline{v}] \to \mathbb{R}$ is a *payment scheme*. Slightly abusing notation, let $u_2(a; v) = u_2(a; s, t)$ for $v = ts$. Given the device $(\mathbf{a}^j, \mathbf{p}^j)$, let $U^j(\mathbf{a}^j, \mathbf{p}^j)$ be j's expected utility in period 1, $\Pi^j(\mathbf{a}^j, \mathbf{p}^j)$ be the expected profits if j chooses it, and $W^j(\mathbf{a}^j)$ be the expected social surplus if j chooses it. The provider's problem is

$$\mathcal{P} = \begin{cases} \max_{\{\mathbf{A},\mathbf{P}\}} \pi\Pi(\mathbf{A}, \mathbf{P}) + (1-\pi)W(\mathbf{A}) & \\ \text{s.t., for } j = \mathtt{C},\mathtt{I} \text{ and } v, v' \in [\underline{v}, \overline{v}], & \\ u_2(\mathbf{a}^j(v); v) - \mathbf{p}^j(v) \geq u_2(\mathbf{a}^j(v'); v) - \mathbf{p}^j(v'), & (\mathrm{IC}_2^j) \\ U^j(\mathbf{a}^j, \mathbf{p}^j) \geq U^j(\mathbf{a}^{-j}, \mathbf{p}^{-j}), & (\mathrm{IC}_1^j) \\ U^j(\mathbf{a}^j, \mathbf{p}^j) \geq 0, & (\mathrm{IR}^j) \end{cases}$$

where

$$W(\mathbf{A}) = \gamma W^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}}) + (1-\gamma) W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}})$$

[12] This second property is stronger than requiring that truthfully reporting v be optimal only *conditional* on a truthful report about t. However, it entails no loss of generality. See, e.g., Pavan (2007).

and

$$\Pi\left(\mathbf{A},\mathbf{P}\right)=\gamma\Pi^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}},\mathbf{p}^{\mathtt{C}})+(1-\gamma)\Pi^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}},\mathbf{p}^{\mathtt{I}}).$$

The key to understanding the adverse-selection problem at the heart of screening time inconsistency is the constraint (IC_1^j), which captures j's incentives to choose between devices in period 1. To better understand these incentives, rewrite (IC_1^j) as

$$U^j(\mathbf{a}^j,\mathbf{p}^j)-U^{-j}(\mathbf{a}^{-j},\mathbf{p}^{-j})\geq U^j(\mathbf{a}^{-j},\mathbf{p}^{-j})-U^{-j}(\mathbf{a}^{-j},\mathbf{p}^{-j}).$$

If the right-hand side of this inequality is strictly positive, then j's expected payoff from partaking in the mechanism must exceed $-j$'s; in other words, j must enjoy some information rent. Note that the difference on the right-hand side simply captures whether, in period 1, j expects to get a larger payoff than $-j$'s, if he mimics $-j$. Therefore, by studying the payoffs C and I expect in period 1 when choosing the *same* device, we can understand the nature of the adverse-selection problem created by the unobservability of time inconsistency.

As the next proposition shows, in period 1, C expects a larger payoff than I for any device, and a strictly larger one if and only if the device provides some flexibility. First, this result shows that C is the 'high' type in this model, because (in expectation) he values any offer of the provider more than I. Second, it highlights that I's demand for flexibility is a key determinant of the adverse-selection problem: If I demanded a device with only one option (no flexibility), then C could not get any payoff surplus by mimicking I, and so the provider would not have to grant C any information rent.

Proposition 1 (Adverse-Selection Problem) *If the mechanism* $\{\mathbf{A},\mathbf{P}\}$ *satisfies* (IC_2^j)*, then*

$$U^{\mathtt{C}}(\mathbf{a}^j,\mathbf{p}^j)\geq U^{\mathtt{I}}(\mathbf{a}^j,\mathbf{p}^j)$$

for $j=$C,I*, with equality if and only if* $\mathbf{a}^j$ *is constant over* $(\underline{v},\overline{v})$.[13]

The intuition for Proposition 1 is simple. Fix a device $(\mathbf{a}^j,\mathbf{p}^j)$ that satisfies (IC_2^j), so that we know exactly what self-2 chooses in each state. Recall that C's self-2 always chooses the best option from self-1's point of view, but I's self-2 may

[13]The function $\mathbf{a}^j$ can jump at $\underline{v}$ and $\overline{v}$ without making $U^{\mathtt{C}}(\mathbf{a}^j,\mathbf{p}^j)>U^{\mathtt{I}}(\mathbf{a}^j,\mathbf{p}^j)$, simply because the distribution F is atomless. Since this indeterminacy has no economic content, hereafter the paper focuses on the extension of $\mathbf{a}^j$ to $[\underline{v},\overline{v}]$ by continuity, whenever possible.

not. Therefore, C's self-1 must be at least as well off, choosing $(\mathbf{a}^j, \mathbf{p}^j)$, as I's self-1—recall that both selves-1 have the same preference. Moreover, C's self-1 must be strictly better off than I's, if in period 2 C and I prefer and choose different options from $(\mathbf{a}^j, \mathbf{p}^j)$ with strictly positive probability; this always happens—as shown in the proof—unless $\mathbf{a}^j$ defines a device with only one option.[14]

To further understand Proposition 1, it is helpful to consider what changes if the agent can be of different types, but is always time consistent. Specifically, consider a model that is identical to that of Section 2, except that the agent's utility function is $u_2(a; s, t) - p$ in *both* periods;[15] also, call the type with $0 < t < 1$ L, and that with $t = 1$ H. This model shares two features with that of Section 2: In period 1, the agent values flexibility, and H expects to have a systematically higher valuation of a than L, as does C relative to I. The two models, however, differ in another key feature: H enjoys a more than L already in period 1. This implies that H can obtain a payoff surplus by mimicking L, even if their future choices coincide; in particular, H's surplus can be strictly positive even if the device for L has only one option. This also implies that, to give H no rent, the provider must forgo trading with L; in contrast, the provider can give C no rent and, at the same time, trade with I.

We can now reduce problem $\mathcal{P}$ to simple conditions that characterize the optimal screening devices in terms of the allocations $\mathbf{a}^{\mathtt{C}}$ and $\mathbf{a}^{\mathtt{I}}$ only. First, by standard arguments, each payment scheme $\mathbf{p}^j$ depends only on $\mathbf{a}^j$ up to a scalar. Indeed, (IC_2^j) holds if and only if $\mathbf{a}^j$ is increasing and, for every $v \in [\underline{v}, \overline{v}]$,

$$\mathbf{p}^j(v) = u_2(\mathbf{a}^j(v); v) + \int_v^{\overline{v}} b(\mathbf{a}^j(x))dx - k^j. \tag{E}$$

As self-2 values a more, he cannot choose a smaller one; and the price of $\mathbf{a}^j(v)$ must deter self-2 from choosing it when his valuation differs from (in particular, exceeds) v, explaining the integral in (E). Second, since C's expected payoff must exceed I's, $(\mathrm{IR}^{\mathtt{C}})$ always holds. And since payoffs decrease and profits increase in the payments, as usual, at the optimum both $(\mathrm{IR}^{\mathtt{I}})$ and $(\mathrm{IC}_1^{\mathtt{C}})$ must bind—recall that I is the 'low' type and C the 'high' type. These constraints then pin down $k^{\mathtt{C}}$ and $k^{\mathtt{I}}$, for every $\mathbf{a}^{\mathtt{C}}$ and $\mathbf{a}^{\mathtt{I}}$.[16]

[14] The proof assumes only that $0 < t^{\mathtt{I}} < t^{\mathtt{C}} \leq 1$. So what really matters is that C's ex-post preference is 'closer' to the common ex-ante preference than I's ex-post preference. For this reason, one can show that C is the 'high' type, in this model, also when $t^{\mathtt{I}} > t^{\mathtt{C}} \geq 1$.

[15] A similar model appears in Courty and Li (2000).

[16] When $\pi = 0$, $(\mathrm{IR}^{\mathtt{I}})$ and $(\mathrm{IC}_1^{\mathtt{C}})$ can be slack at the optimum. However, assuming that they hold with equality is without loss of generality, as far as characterizing the optimal $\mathbf{a}^{\mathtt{I}}$ and $\mathbf{a}^{\mathtt{C}}$ is

However, nothing guarantees that the remaining constraint ($\text{IC}_1^{\mathtt{I}}$) is always slack. In contrast to standard screening models (e.g., Mussa and Rosen (1978)), here $\mathtt{I}$—the 'low' type—may prefer the $\mathtt{C}$-device to the $\mathtt{I}$-device, even though $\mathtt{C}$—the 'high' type—is indifferent between them (see Section 4.3). By the previous observations, ($\text{IC}_1^{\mathtt{I}}$) holds if and only if

$$-R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}}) \geq R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}), \tag{R}$$

where[17]

$$R^{-j}(\mathbf{a}^j) = U^{-j}(\mathbf{a}^j, \mathbf{p}^j) - U^j(\mathbf{a}^j, \mathbf{p}^j).$$

Intuitively, (R) says that a dishonest $\mathtt{I}$ must expect to lose, relative to $\mathtt{C}$'s payoff, at least as much as a dishonest $\mathtt{C}$ expects to gain, relative to $\mathtt{I}$'s payoff.

Finally, using $\Pi^j(\mathbf{a}^j, \mathbf{p}^j) = W^j(\mathbf{a}^j) - U^j(\mathbf{a}^j, \mathbf{p}^j)$, we obtain the following.

Corollary 1 *A mechanism* $\{\mathbf{A}, \mathbf{P}\}$ *solves* $\mathcal{P}$ *if and only if* $\mathbf{a}^{\mathtt{C}}$ *and* $\mathbf{a}^{\mathtt{I}}$ *solve*

$$\mathcal{P}' = \begin{cases} \max_{\mathbf{A}} \gamma W^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}}) + (1-\gamma)\left[W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) - \pi \frac{\gamma}{1-\gamma} R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})\right] \\ \textit{s.t. } \mathbf{a}^{\mathtt{C}}, \mathbf{a}^{\mathtt{I}} \textit{ increasing and (R)} \end{cases}.$$

Due to condition (R), $\mathcal{P}'$ is not separable across allocations. So we cannot solve for the optimal allocations independently of one another. Nonetheless, the next lemma gives necessary and sufficient conditions for $\mathbf{a}^{\mathtt{C}}$ and $\mathbf{a}^{\mathtt{I}}$ to solve $\mathcal{P}'$, which will allow us to characterize their properties.

Lemma 2 (Optimality) *The allocations* $\mathbf{a}^{\mathtt{C}}$ *and* $\mathbf{a}^{\mathtt{I}}$ *solve* $\mathcal{P}'$ *if and only if, for* $\mu \geq 0$, $(\mathbf{a}^{\mathtt{C}}, \mathbf{a}^{\mathtt{I}}, \mu)$ *satisfies for* $j = \mathtt{C}, \mathtt{I}$

$$\mathbf{a}^j \in \arg \max_{\hat{\mathbf{a}}^j \textit{ increasing}} W^j\left(\hat{\mathbf{a}}^j\right) - r^{-j} R^{-j}\left(\hat{\mathbf{a}}^j\right),$$

$$R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}) + R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}}) \leq 0, \quad \textit{and} \quad \mu\left[R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}) + R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}})\right] = 0,$$

where $r^{\mathtt{C}} = \pi \frac{\gamma}{1-\gamma} + \frac{\mu}{1-\gamma}$ *and* $r^{\mathtt{I}} = \frac{\mu}{\gamma}$.

Lemma 2 relies on infinite-dimensional, global, Lagrangian methods that don't require assuming any property about $\mathbf{a}^{\mathtt{C}}$ and $\mathbf{a}^{\mathtt{I}}$ beyond the necessary monotonicity.

concerned.

[17] The functional R^{-j} depends only on $\mathbf{a}^j$ because k^j enters additively in $U^i(\mathbf{a}^j, \mathbf{p}^j)$ (see the proof of Proposition 1).

Its key implication is that we can characterize $\mathbf{a}^{\mathtt{C}}$ and $\mathbf{a}^{\mathtt{I}}$ by solving two distinct maximizations, each as a function of the weight r^{-j}. As usual, $r^{\mathtt{C}}$ depends on the hazard rate between $\mathtt{C}$ and $\mathtt{I}$ (the 'high' and 'low' type), scaled by how much the provider weighs profits (π). However, both $r^{\mathtt{C}}$ and $r^{\mathtt{I}}$ also depend on a new term that takes into account whether (R) binds. This term links the maximizations defining $\mathbf{a}^{\mathtt{C}}$ and $\mathbf{a}^{\mathtt{I}}$, thereby reducing the nonseparability of $\mathcal{P}'$ to the scalar μ. Finally, even if the provider cares only about welfare ($\pi = 0$), both $r^{\mathtt{C}}$ and $r^{\mathtt{I}}$ can be strictly positive; in this case, both devices will feature distortions at the optimum (see Section 4.3).

4.2 The Optimal Device for the Inconsistent Agent

This section characterizes the optimal screening device for type $\mathtt{I}$. It shows how the provider distorts the $\mathtt{I}$-device, so as to limit the rent of type $\mathtt{C}$, by curtailing its flexibility at both ends of the efficient choice range.

By Lemma 2, when designing the $\mathtt{I}$-device, the provider faces a standard trade-off—captured by $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) - r^{\mathtt{C}} R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$—which results in a curtailment of flexibility. On the one hand, she wants to maximize the expected social surplus with $\mathtt{I}$, which calls for a device that sustains the efficient level of flexibility ($\mathbf{a}^*$)—for the same reason as in the first best. On the other hand, she wants to minimize the rent that keeps $\mathtt{C}$ from mimicking $\mathtt{I}$, which calls for a device with no flexibility (Proposition 1); this is because $\mathtt{C}$'s rent reduces the profit from the $\mathtt{C}$-device and can make $\mathtt{I}$ mimic $\mathtt{C}$ (recall (R)). Intuitively, the optimal $\mathtt{I}$-device should then strike a balance between these two polar cases and therefore its flexibility should be curtailed below efficiency. To gain more intuition, recall that $\mathtt{C}$'s rent arises because, in each state, $\mathtt{C}$ and $\mathtt{I}$ have different valuations v and consequently behave differently under the $\mathtt{I}$-device (unless $\mathbf{a}^{\mathtt{I}}$ is constant). Curbing $\mathtt{C}$'s rent then requires curtailing this difference between $\mathtt{C}$'s and $\mathtt{I}$'s behavior, which depends on how $\mathbf{a}^{\mathtt{I}}$ responds to changes in v—i.e., its flexibility.

Although the basic trade-off here is the same as in standard screening models, those models offer no guidelines on how to optimally curtail the flexibility of the $\mathtt{I}$-device. In Mussa and Rosen (1978), for instance, a seller faces the same trade-off when designing the quality offered to low-valuation buyers, in a market with also high-valuation buyers. In that model, the seller has to lower quality below efficiency; similarly here the provider has to curtail flexibility below efficiency. But curtailing flexibility, in contrast to lowering quality, can be done in many ways

and the optimal one is not obvious.

The next proposition characterizes it. To state the result, let $\mathbf{a}^{\mathtt{I}*}$ be the allocation that defines (up to $k^{\mathtt{I}}$) the first-best $\mathtt{I}$-device: $\mathbf{a}^{\mathtt{I}*}(v) = \mathbf{a}^*(v/t^{\mathtt{I}})$ for $v \in [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$ and $\mathbf{a}^{\mathtt{I}*}(v) = \mathbf{a}^{\mathtt{I}*}(\overline{v}^{\mathtt{I}})$ otherwise.

Proposition 2 (Curtailment of Flexibility: Optimal Form) *An increasing* $\mathbf{a}^{\mathtt{I}}$ *that maximizes* $W^{\mathtt{I}}(\hat{\mathbf{a}}^{\mathtt{I}}) - r^{\mathtt{C}}R^{\mathtt{C}}(\hat{\mathbf{a}}^{\mathtt{I}})$ *exists, is unique, and is continuous in* v *and* $r^{\mathtt{C}}$*. If* $r^{\mathtt{C}} > 0$*,* $\mathbf{a}^{\mathtt{I}}$ *features:*
(a) Range Reduction with 'Overconsumption' at the Bottom and 'Underconsumption' at the Top: *there are* $v_* > \underline{v}^{\mathtt{I}}$ *and* $v^* < \overline{v}^{\mathtt{C}}$ *such that* $\mathbf{a}^{\mathtt{I}}(v) = \mathbf{a}^{\mathtt{I}*}(v)$ *at* v_* *and* v^**,* $\mathbf{a}^{\mathtt{I}}(v) > \mathbf{a}^{\mathtt{I}*}(v)$ *for* $v < v_*$*, and* $\mathbf{a}^{\mathtt{I}}(v) < \mathbf{a}^{\mathtt{I}*}(v)$ *for* $v > v^*$*;*
(b) No Flexibility at the Top: $\mathbf{a}^{\mathtt{I}}$ *is constant over* $[v^{\mathrm{b}}, \overline{v}^{\mathtt{C}}]$ *with* $v^{\mathrm{b}} < \overline{v}^{\mathtt{I}}$*.*
In addition, $\mathbf{a}^{\mathtt{I}}$ *can feature*
(c) No Flexibility at the Bottom: $\mathbf{a}^{\mathtt{I}}$ *can be constant over* $[\underline{v}^{\mathtt{I}}, v_{\mathrm{b}}]$ *with* $v_{\mathrm{b}} > \underline{v}^{\mathtt{I}}$*. A sufficient condition for (c) is that for all* $s' > s$ *in* $[\underline{s}, s^\dagger] \neq \varnothing$

$$\frac{F(s')/f(s') - F(s)/f(s)}{s' - s} \geq \frac{1 - t^{\mathtt{I}}}{t^{\mathtt{I}}} + \frac{1}{t^{\mathtt{I}} r^{\mathtt{C}}}.$$

The proof is constructive: First, it builds $\mathbf{a}^{\mathtt{I}}$ on path (over $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$), relying on Toikka's (2011) generalization of Myerson's (1981) ironing method, and then it explicitly builds the best extension of $\mathbf{a}^{\mathtt{I}}$ off path. Uniqueness follows from strict concavity of the function b.

The optimal $\mathtt{I}$-device curtails the flexibility of $\mathtt{I}$'s behavior, relative to the first best, as follows. First, $\mathtt{I}$ reacts to extreme states (both high and low) less than in the first best, so that his choice range is a strict subset of the efficient one given by $\mathbf{a}^{\mathtt{I}*}$. In general, $\mathtt{I}$ retains some flexibility to act on ex-post information, and his choice range remains rich, involving all options in a connected interval.[18] Second, $\mathtt{I}$ does not react at all to high enough states and, in some environments, to low enough states as well; over such states, $\mathtt{I}$'s behavior features no flexibility.

Of course, the $\mathtt{I}$-device sustains all these features with an appropriately tailored payment scheme $\mathbf{p}^{\mathtt{I}}$ (recall (E)). This observation is the key to Proposition 2. As noted, to curb $\mathtt{C}$'s rent the $\mathtt{I}$-device must curtail how differently $\mathtt{I}$ and $\mathtt{C}$ behave after choosing it, which depends on how $\mathbf{a}^{\mathtt{I}}$ varies across v's. So, to explain how the optimal $\mathbf{a}^{\mathtt{I}}$ departs from $\mathbf{a}^{\mathtt{I}*}$, it is crucial to understand the global effects on the

[18] For example, $\mathtt{I}$ retains some flexibility if $\mathtt{C}$'s probability γ is small (so that $r^{\mathtt{C}}$ is small) and the utility $b(\underline{a})$ of the smallest action is low (see Proposition 3 and Corollary 3).

differences in actions *and payments* across v's of changing $\mathbf{a}^{\mathtt{I}}$ at one v. Intuitively, if we raise $\mathbf{a}^{\mathtt{I}}(v)$, then $\mathbf{p}^{\mathtt{I}}(v)$ has to rise and $\mathbf{p}^{\mathtt{I}}(v')$ has to fall for every v' above v—so that self-2's reports remain truthful. The differences in actions and payments then shrink between v and every v' above v, but grow between v and every v' below v (recall that $\mathbf{a}^{\mathtt{I}}$ is increasing). The opposite effects arise, if we lower $\mathbf{a}^{\mathtt{I}}(v)$. So the relative strength of these global effects above and below v ultimately determines whether $\mathbf{a}^{\mathtt{I}*}(v)$ is distorted up or down, leading to properties (a), (b), and (c).

Specifically, to see why 'overconsumption' and possibly no flexibility arise at the bottom, start from the lowest v (see Figure 1). At $\underline{v}^{\mathtt{I}}$, it is clearly optimal to distort $\mathbf{a}^{\mathtt{I}*}$ up, as only the global effect above $\underline{v}^{\mathtt{I}}$ matters. At v close to $\underline{v}^{\mathtt{I}}$, however, the global effect below v also matters, but that above v prevails because, intuitively, the mass of $v' > v$ prevails on that of $v' < v$; it is therefore still optimal to distort $\mathbf{a}^{\mathtt{I}*}$ up. As v grows, the effect below v gains strength, shrinking the upward distortions until some v_*, where the two effects balance each other and $\mathbf{a}^{\mathtt{I}}(v_*) = \mathbf{a}^{\mathtt{I}*}(v_*)$ (e.g., curve 1). Intuitively, how fast this happens depends on how fast the ratio $F^{\mathtt{I}}(v)/f^{\mathtt{I}}(v)$ grows. So, if this ratio—which equals $t^{\mathtt{I}}F(v/t^{\mathtt{I}})/f(v/t^{\mathtt{I}})$—grows fast enough for v close to $\underline{v}^{\mathtt{I}}$, as formally stated in Proposition 2, then the upward distortions can shrink so quickly that the provider may wish that a larger v chose a smaller action than a smaller v' (e.g., curve 2). But $\mathbf{a}^{\mathtt{I}}$ must be increasing; so the provider can only pool v's close to $\underline{v}^{\mathtt{I}}$ at the same action (e.g., curve 3). Note that the condition on F that ensures bunching at the bottom is more likely to hold when $\mathtt{I}$ is less inconsistent (i.e., $t^{\mathtt{I}}$ is closer to 1), and when the provider cares more about $\mathtt{C}$'s rent (i.e., $r^{\mathtt{C}}$ is higher).

To see why 'underconsumption' and no flexibility arise at the top, first observe that $\mathbf{a}^{\mathtt{I}}$ must be constant above $\overline{v}^{\mathtt{I}}$. Otherwise, tossing any option with $a > \mathbf{a}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ curbs $\mathtt{C}$'s rent without harming $\mathtt{I}$: For $s > \overline{v}^{\mathtt{I}}$ (recall that $v = s$ for type $\mathtt{C}$), a dishonest $\mathtt{C}$ now chooses $\mathbf{a}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$, which is closer to $\mathtt{I}$'s choices. So, at $\overline{v}^{\mathtt{I}}$, only the global effect below it matters, and therefore it is optimal to distort $\mathbf{a}^{\mathtt{I}}$ down. By the same argument as before, the downward distortions persist until some v^*, where $\mathbf{a}^{\mathtt{I}}(v^*) = \mathbf{a}^{\mathtt{I}*}(v^*)$. Now suppose $\mathbf{a}^{\mathtt{I}}$ were strictly increasing near $\overline{v}^{\mathtt{I}}$. If we lower $\mathbf{a}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ a bit, a dishonest $\mathtt{C}$'s choices move closer to $\mathtt{I}$'s for almost all $s > \overline{v}^{\mathtt{I}}$, thus curbing $\mathtt{C}$'s rent even more. To do so, however, we also have to lower $\mathbf{a}^{\mathtt{I}}$ for some v's below $\overline{v}^{\mathtt{I}}$, to satisfy monotonicity. The smaller the set of v's affected by this extra distortion, the smaller the extra welfare loss with $\mathtt{I}$. So the best thing to do is to bunch at some a^{b} every v whose action was above a^{b}—formally, every $v > v^{\mathrm{b}}$—and leave $\mathbf{a}^{\mathtt{I}}$ unchanged for all other v's. Note that, for every $v^{\mathrm{b}} < \overline{v}^{\mathtt{I}}$, a

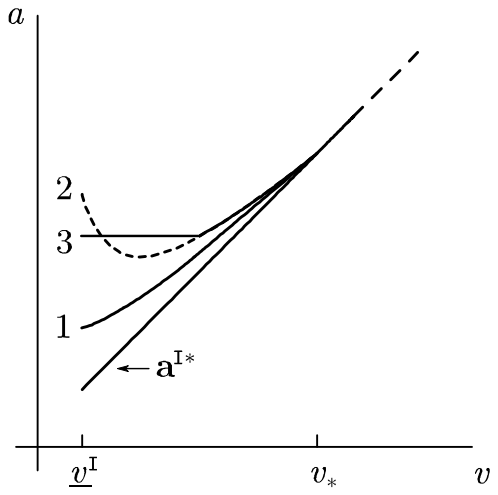

Figure 1: 'Overconsumption' and No Flexibility at the Bottom

dishonest C chooses more similarly to I for $s \in [v^{\mathrm{b}}, \overline{s}]$, whereas I must choose a^{b} only for $s \in [v^{\mathrm{b}}/t^{\mathtt{I}}, \overline{s}]$. So some bunching is always optimal, because for v^{b} close to $\overline{v}^{\mathtt{I}} = t^{\mathtt{I}}\overline{s}$, it reduces more C's rent than the welfare with I.

To better understand how the provider sustains I's distorted behavior, it is helpful to compare the payment schemes associated with $\mathbf{a}^{\mathtt{I}}$ and $\mathbf{a}^{\mathtt{I}*}$. As the next corollary shows, at the bottom of $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, $\mathbf{p}^{\mathtt{I}}$ either falls more slowly or rises more rapidly than $\mathbf{p}^{\mathtt{I}*}$ for an equal decrease in a; this reduces I's willingness to decrease a and makes him 'overconsume.' On the other hand, at the top $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, $\mathbf{p}^{\mathtt{I}}$ either falls more slowly or rises more rapidly than $\mathbf{p}^{\mathtt{I}*}$ for an equal increase in a; this reduces I's willingness to increase a and makes him 'underconsume.'

Corollary 2 *If $d\mathbf{a}^{\mathtt{I}}/dv > 0$ at v, then $\frac{d\mathbf{p}^{\mathtt{I}}/dv}{d\mathbf{a}^{\mathtt{I}}/dv}$ is strictly smaller than $\frac{d\mathbf{p}^{\mathtt{I}*}/dv}{d\mathbf{a}^{\mathtt{I}*}/dv}$ for $v \in [\underline{v}^{\mathtt{I}}, v_*)$, and strictly larger for $v \in (v^*, \overline{v}^{\mathtt{I}}]$.*

The next proposition shows how the optimal I-device changes as the weight $r^{\mathtt{C}}$ becomes very large or small. Recall that, by Lemma 2, $r^{\mathtt{C}}$ increases in C's probability γ and the weight π on profits. Also, let

$$a^{\mathrm{nf}} = \arg\max_{a \in [\underline{a}, \overline{a}]} \mathbb{E}[u_1(a; s)] - c(a),$$

which is the ex-ante efficient action if the agent is not allowed to act on ex-post information at all—'nf' means 'no flexibility.'

Proposition 3 (Comparative Statics) *Let* $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ *be the optimal allocation for* $\mathtt{I}$*. Then* $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ *converges pointwise to* $\mathbf{a}^{\mathtt{I}*}$ *as* $r^{\mathtt{C}} \to 0$ *and uniformly to* a^{nf} *as* $r^{\mathtt{C}} \to +\infty$*.*

If the provider cares very little about C's rent, then she tends to offer an efficient I-device. If instead she cares a lot about C's rent, then she tends to disregard I's desire for flexibility, in the limit offering an I-device with only the option a^{nf}—a radical reduction of flexibility vis-à-vis the first best. This happens, for example, whenever the provider cares about profits and the agent's type is very likely to be C (if $\pi > 0$, $r^{\mathtt{C}} \to +\infty$ as $\gamma \to 1$).

With more information on the distribution F, it is possible to say more about how the screening I-device differs from an efficient one. This is because the virtual valuation defining $\mathbf{a}^{\mathtt{I}}$ is in general complex. For illustrative purposes, the next lemma looks at the case with uniform F. In this case, a simple monotonic relationship also emerges between the weight $r^{\mathtt{C}}$ and the choice range of the I-device, as well as the regions involving no flexibility ($[\underline{v}^{\mathtt{I}}, v_{\mathrm{b}}]$ and $[v^{\mathrm{b}}, \overline{v}^{\mathtt{C}}]$).[19]

Lemma 3 *Suppose* s *is uniformly distributed and* $t^{\mathtt{I}} > \underline{s}/\overline{s}$*. Then, the optimal* $\mathbf{a}^{\mathtt{I}}$ *crosses* $\mathbf{a}^{\mathtt{I}*}$ *only once and is strictly increasing over* $[v_{\mathrm{b}}, v^{\mathrm{b}}]$*. As* $r^{\mathtt{C}}$ *rises,* $\mathbf{a}^{\mathtt{I}}$ *changes as follows:* v^{b} *and* $\mathbf{a}^{\mathtt{I}}(v^{\mathrm{b}})$ *decrease and* $\mathbf{a}^{\mathtt{I}}(v_{\mathrm{b}})$ *increases; when* $v_{\mathrm{b}} > \underline{v}^{\mathtt{I}}$*, then* v_{b} *increases as well.*

In general, more intricate patterns can occur, including bunching also at intermediate points for standard ironing reasons. Lemma 3 highlights, however, that the bunching at the top and bottom is due not to failures of standard regularity conditions—which are satisfied by the uniform distribution—but precisely to the problem of optimally screening time inconsistency.

4.3 The Optimal Device for the Consistent Agent

This section characterizes the optimal screening device for type C. It shows that, to deter I from mimicking C, the provider may have to add unused options to an otherwise efficient C-device, and if these options are not deterring enough, she may also distort C's behavior.

By Lemma 2, when designing the C-device, the provider wants to maximize the surplus with C, but also has to worry about jeopardizing I's incentives for revealing

[19] For the proof, see Online Appendix A.

his time inconsistency—by (R), how much a dishonest $\mathtt{I}$ expects to lose, relative to $\mathtt{C}$'s payoff, must exceed $\mathtt{C}$'s rent. A natural benchmark of an efficient $\mathtt{C}$-device is the first-best device, defined (up to $k^{\mathtt{C}}$) by the allocation $\mathbf{a}^{\mathtt{C}*}$ with $\mathbf{a}^{\mathtt{C}*}(v) = \mathbf{a}^*(v)$ for $v \in [\underline{v}^{\mathtt{C}}, \overline{v}^{\mathtt{C}}]$ and $\mathbf{a}^{\mathtt{C}*}(v) = \mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})$ otherwise. By the next lemma, however, $\mathbf{a}^{\mathtt{C}*}$ can violate (R) for some $\mathbf{a}^{\mathtt{I}}$.

Lemma 4 *Let $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}*}$ be the allocations associated with $\mathbf{a}^*$. There is a family of distributions F such that $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}*}$ violate (R) and are therefore infeasible.*

To see the intuition for Lemma 4, recall that a dishonest $\mathtt{I}$ loses relative to $\mathtt{C}$'s payoff, to the extent that $\mathtt{I}$ and $\mathtt{C}$ behave differently under the $\mathtt{C}$-device. Given $\mathbf{a}^{\mathtt{C}*}$, this difference is small in states close to $\underline{s}$. Therefore, if these states are likely enough according to F, then ex ante $\mathtt{I}$ values the $\mathtt{C}$-device almost as much as $\mathtt{C}$. On the other hand, since $\mathbf{a}^{\mathtt{I}*}$ is flexible, $\mathtt{C}$'s rent must be positive and can be large enough to lure $\mathtt{I}$ to mimic $\mathtt{C}$. One can show, for example, that if $b(a) = \sqrt{a}$ and F is uniform, then $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}*}$ are infeasible.

Although Lemma 4 looks at the extreme case of $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}*}$, its conclusion holds more generally. By Proposition 3, as the provider cares more about $\mathtt{C}$'s rent ($r^{\mathtt{C}} \to +\infty$), she tends to reduce the $\mathtt{I}$-device to a single option; therefore, by continuity, $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ are always feasible for $r^{\mathtt{C}}$ large enough. On the other hand, as the provider cares less about $\mathtt{C}$'s rent ($r^{\mathtt{C}} \to 0$), she tends to design an $\mathtt{I}$-device similar to the first-best one. So, if $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}*}$ are infeasible, then by continuity, $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ are also infeasible for $r^{\mathtt{C}}$ small enough. In this case, the key question is how the provider changes $\mathbf{a}^{\mathtt{C}*}$ to avoid $\mathtt{I}$'s mimicking $\mathtt{C}$.

The first strategy she adopts is unconventional. In models in which the agent is always time consistent, the provider must distort the offer for one type—hence his choices—whenever that offer makes another type mimic. In the present model, instead, the provider may be able to avoid $\mathtt{I}$'s mimicking and to offer, at the same time, an efficient $\mathtt{C}$-device. To do that, she adds to the device options that $\mathtt{C}$ never uses, but that make a dishonest $\mathtt{I}$ behave more differently from $\mathtt{C}$. Intuitively, these options lower the degree of commitment that $\mathtt{I}$ already finds too low in an efficient $\mathtt{C}$-device, thus making a dishonest $\mathtt{I}$ lose more relative to $\mathtt{C}$'s payoff. Since these options are off path, their production cost is irrelevant and the provider can usefully employ them in many ways. The way that maximizes a dishonest $\mathtt{I}$'s loss is presented in the next proposition: To maximally deter $\mathtt{I}$'s mimicking $\mathtt{C}$, the provider needs to add to an efficient $\mathtt{C}$-device only one option, which a dishonest $\mathtt{I}$ would choose in sufficiently low states.

Proposition 4 (Usefulness of Unused Options) *An increasing* $\mathbf{a}^{\mathtt{C}}$ *sustains* $\mathbf{a}^*$ *with* $\mathtt{C}$ *and maximizes* $-R^{\mathtt{I}}(\hat{\mathbf{a}}^{\mathtt{C}})$ *if and only if* $\mathbf{a}^{\mathtt{C}}(v) = \underline{a}$ *for* $v < v_{\mathrm{u}}$ *and* $\mathbf{a}^{\mathtt{C}}(v) = \mathbf{a}^{\mathtt{C}*}(v)$ *for* $v \geq v^{\mathrm{u}}$*, where* $\underline{v}^{\mathtt{I}} < v_{\mathrm{u}} \leq v^{\mathrm{u}} \leq \underline{v}^{\mathtt{C}}$*.*

One unused option involving $\underline{a}$ is enough to maximally deter I's mimicking, because I's self-1 views $\underline{a}$ as the action that would most tempt his self-2, who always prefers lower actions ($t^{\mathtt{I}} < 1$). The payment $\underline{p}$ for $\underline{a}$ plays a key role too, for it controls both the probability (by pinning down v_{u}) and the regret that I's self-1 assigns to choosing the unused option. Intuitively, if $\underline{p}$ is low, then I expects to choose $\underline{a}$ already for v close to $\underline{v}^{\mathtt{C}}$, but regrets it only a little. If instead $\underline{p}$ is high, then I expects to choose $\underline{a}$ only for v close to $\underline{v}^{\mathtt{I}}$—hence with lower probability—but regrets it more. Clearly, depending on the distribution F, a high payment for $\underline{a}$ may deter more I's mimicking. This explains why the provider may actually *restrict*—by setting $v^{\mathrm{u}} < \underline{v}^{\mathtt{C}}$—the set of states in which I's self-2 has a valuation low enough to choose the unused option from the C-device.

Although by Proposition 4 unused options in the efficient C-device can make I less willing to mimic C, they must be tempting enough to keep I away. If they are, the provider can sustain the efficient outcome with C.

Corollary 3 *If* $\mathbf{a}^{\mathtt{C}*}$ *and* $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ *are infeasible, then there is* $d > 0$*—which depends on* $\mathbf{a}^{\mathtt{C}*}$ *and* $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$*—such that* $\mathbf{a}^{\mathtt{C}}$ *as in Proposition 4 and* $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ *are feasible if and only if* $b(\underline{a}) \leq b(\mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})) - d$*.*

The lowest feasible action $\underline{a}$ can be tempting enough, relative to the lowest efficient one $\mathbf{a}^*(\underline{s})$, for several reasons. For instance, suppose $b(a)$ captures the future consequences of some current action (like shopping with credit cards). The worst consequence $b(\underline{a})$ is then likely far worse than the efficient one $b(\mathbf{a}^*(\underline{s}))$, which takes into account the current utility $-a$ and the cost $c(a)$ (like default costs). More generally, no technological or legal constraint may prevent the provider from offering very unattractive actions.

Combining the previous insights, the next proposition summarizes how the screening C-device depends on C's probability γ and the weight π on profits. Recall that $r^{\mathtt{C}} = \pi\frac{\gamma}{1-\gamma} + \frac{\mu}{1-\gamma}$, where $\mu \geq 0$ is the Lagrange multiplier associated with condition (R).

Proposition 5 (Optimal C-device) *There are* r_1 *and* r_2*, where* $0 \leq r_1 \leq r_2 < +\infty$ *possibly with strict inequalities, such that the optimal* $\mathtt{C}$*-device sustains* $\mathbf{a}^*$ *with* $\mathtt{C}$ *if and only if* $\pi\frac{\gamma}{1-\gamma} \geq r_1$*, and must include unused options if and only if* $\pi\frac{\gamma}{1-\gamma} < r_2$*.*

Since r_1 can be strictly positive, the present model violates the 'no distortion at the top' property, common in standard screening models. In those models, the agent of the 'highest' type usually achieves an efficient outcome, as if types were observable; for example, in Mussa and Rosen (1978), the highest-valuation buyer always trades efficiently with the monopolist. Here, instead, although $\mathtt{C}$ is the 'high' type, the $\mathtt{C}$-device can be inefficient.

The $\mathtt{C}$-device is more likely to feature unused options and (possibly) be inefficient when the agent is less likely to be of type $\mathtt{C}$ (γ is lower), or the provider cares more about welfare (π is lower). Intuitively, when either γ or π are lower, the provider is willing to grant $\mathtt{C}$ a larger rent. But this rent makes $\mathtt{I}$ more willing to mimic $\mathtt{C}$, so the provider has to offset it by lowering the degree of commitment that $\mathtt{I}$ finds in the $\mathtt{C}$-device.

Corollary 4 *If the monopolist ($\pi = 1$) has to add unused options to the $\mathtt{C}$-device, so does the planner ($\pi < 1$). If the monopolist's solution violates the 'no distortion at the top' property, so does the planner's.*

Finally, when unused options alone cannot avoid $\mathtt{I}$'s mimicking $\mathtt{C}$—for example, because a represents consumption of a good that cannot be negative (healthy food), and self-1 deems consuming zero a minor temptation—then the provider has to distort $\mathbf{a}^{\mathtt{C}}$ on path. She does so to make $\mathtt{C}$ and $\mathtt{I}$ behave even more differently after choosing the $\mathtt{C}$-device—this is the only way to make a dishonest $\mathtt{I}$ lose even more relative to $\mathtt{C}$'s payoff and thus satisfy (R). By Lemma 2, the optimal $\mathbf{a}^{\mathtt{C}}$ maximizes $W^{\mathtt{C}}\left(\hat{\mathbf{a}}^{\mathtt{C}}\right) - r^{\mathtt{I}} R^{\mathtt{I}}\left(\hat{\mathbf{a}}^{\mathtt{C}}\right)$ with $r^{\mathtt{I}} > 0$, which can again be written as a virtual surplus to study the properties of $\mathbf{a}^{\mathtt{C}}$. The resulting $\mathtt{C}$-device, in general, can distort $\mathtt{C}$'s behavior both up and down relative to efficiency, in part also to make $\mathtt{I}$'s self-1 view the unused option with $\underline{a}$ as an even worse temptation.[20]

4.4 Discussion

When applied to concrete settings, the previous results can be interpreted as follows.[21]

Consider the savings example in Section 2. In contrast to the first-best devices, the screening devices sustain different behaviors. The $\mathtt{C}$-device should be able to

[20] Formal details are available upon request.

[21] Of course, this discussion is not meant to be an exact (let alone sole) explanation for the considered features of some real contracts, given that this paper intentionally omits important aspects of specific applications to focus on its main theoretical question.

make C save efficiently, as the lowest feasible savings most likely lead to very low utility at retirement (i.e., $b(\underline{a})$ is very low). The I-device, instead, curtails flexibility (i.e., liquidity) at both ends of the efficient savings range, which can be interpreted as restricting both withdrawals and deposits. Such restrictions are implemented by making the withdrawal penalties stiffer and the deposit rewards weaker for large enough amounts. This can create caps on withdrawals as well as on deposits. Lemma 1 highlights that, contrary to what one might think, this kind of restrictions may be introduced not to help inconsistent agents avoid depleting their savings, but to ward off consistent agents.

These implications for commitment policies involving savings differ substantively from related results in the literature. Amador et al. (AWA) (2006), Ambrus and Egorov (AE) (2013), and Bond and Sigurdsson (BS) (2013) give a justification for penalizing inconsistent agents when they tap their savings to pay for current overconsumption. In this paper, the first-best I-device similarly penalizes current overconsumption with withdrawal penalties, but it also promotes savings with deposit rewards (Lemma 1). In all four papers, penalties (and rewards) are justified by inconsistent agents' propensity to overconsume. But, in contrast to AWA, BS, and AE, this paper also gives a justification for inefficiently restricting inconsistent agents' ability not only to tap but also to deposit into their devices. These restrictions are justified by the need to dissuade consistent agents from choosing the devices for inconsistent agents. This need never arises in AWA, AE, and BS: One can easily see that, because those papers rule out payments across states, consistent agents would never want the commitment policies for inconsistent agents, and vice versa. That is, those policies create no adverse-selection problem with respect to agents' degree of inconsistency. Concretely, this implies that such problems do not arise with policies like Social Security and defined-benefit plans, which are better captured by models without payments. Instead, devices like defined-contribution plans—which have tax penalties and rewards varying with savings decisions and therefore are better captured by models with payments—are subject to the adverse-selection problem identified in this paper.

For real-life examples of savings devices that resemble—at least at a broad level—the differences between the screening C- and I-devices, consider the U.S. retirement market.[22] This market offers standard devices, called taxable accounts

[22] Other examples of savings devices that resemble the screening I-device include some Christmas-club accounts and individual-development accounts—a form of matched-savings accounts—offered, in the U.S., by some financial institutions (see also Ashraf, Gons, Karlan, and Yin (2003)).

(TAs), as well as special devices, called 'tax-shielded' accounts (TSAs)—like IRAs and 401(k) plans. Governed by different tax rules, TSAs and TAs differ in many ways. In particular, TSAs have all of the following features, while TAs don't. Consistent with Lemma 1, TSAs reward deposits and penalize withdrawals through taxation. However, consistent with Proposition 2 and Corollary 2, TSAs set dear tax penalties for deposits beyond certain amounts and also limit withdrawals—empirically these limits sometimes bind.[23] Moreover, for instance, IRA-backed loans are *de facto* prohibited, and 401(k)-backed loans are capped and subject to quick repayments. Finally, there is some evidence consistent with the principle that curtailing flexibility in the special devices curbs their appeal to less inconsistent agents. Amromin (2002, 2003) shows that a significant share (39%) of U.S. savers does not take full advantage of TSAs' tax benefits, and prefers to invest in TAs because of TSAs' liquidity constraints. These savers reveal that they value flexibility more than commitment, which may depend, among other things, on them being less time inconsistent.

Of course, many other reasons can explain each of these features. For example, deposits could be limited to prevent rich people from exploiting IRAs' tax benefits too much, thus avoiding large tax-revenue losses. But the possibility of observing people's income should, in principle, remove this issue. Indeed, one might wonder why deposit limits also apply to poor savers—this paper provides an answer. But even if other reasons led to the current regulation, this paper suggests that its broad features also appropriately screen among differently inconsistent savers. Moreover, this paper points out a single rationale that can account for all the mentioned features at once, as well as some, perhaps unexpected, consequences of modifying them. For example, relaxing the withdrawal (but not the deposit) penalties of TSAs can lure consistent people away from TAs, even though TAs have no such penalties to begin with.

Consider now the exercising example in Section 2. Although here $t^{\mathtt{I}} > 1$, the results in Sections 4.2-4.3 are qualitatively unchanged (see Section 5.4) and can be interpreted as follows. The screening $\mathtt{C}$-device features no monetary incentives to work out. Moreover, by allowing the agent to skip any number of workouts

[23] In 2007 (2008), about 59% (49%) of IRA-owners contributed at the limit (Holden et al. (2010b)), and roughly 11% of all 401(k) participants did so in 2004 (Munnell and Sundén (2006)). Except for few cases (like first-time home purchase for IRAs), any sum withdrawn before the age of $59\frac{1}{2}$ incurs tax penalties, which seem to actually limit access to these TSAs: Holden and Schrass (2008-2010a) report that the vast majority of IRA withdrawals are retirement related, and only about 5% occurs before the owner turns $59\frac{1}{2}$.

without penalties (intended as an unused option), the C-device may deter I from choosing it because, otherwise, he would end up exercising too little. On the other hand, the I-device curtails flexibility at both ends of the efficient range of workouts, by making the penalties for missed workouts stiffer and the rewards for attended workouts weaker as their number exceeds some level; again, this can create caps on maximum and minimum workouts. Finally, there are examples of devices that offer monetary incentives to work out, but also have restrictions that are evocative of the I-device (see, e.g., GymPact.com).

The curtailment of flexibility in the I-device should not be confused with an apparently similar result in O'Donoghue and Rabin (OR) (1999b). In OR, a firm designs contracts to incentivize a present-biased worker to complete a task at the most efficient time, which only the worker knows by observing the task cost over time. If the firm does not know the worker's propensity to procrastinate (which the worker learns only after signing a contract), then the best contract sets penalties that increase for longer delays, possibly involving deadlines. Although such a contract curtails the worker's flexibility, its rationale is based on learning, not on screening. As the worker continues to delay, the firm will learn that his propensity to procrastinate is high, which justifies stronger punishments; therefore, the firm wants to commit to a contract that will take that into account.

Finally, as noted, the result that optimal mechanisms may have to offer options that no type ever uses also appears in Esteban and Miyagawa (EM) (2006b). EM consider a classic monopolistic-screening model, except that buyers are prone to temptations as in Gul and Pesendorfer (2001). In such a model, the monopolist may be able to fully extract all buyers' surplus, by offering low-valuation buyers menus with unused options that ward off high-valuation buyers who view them as temptations.

EM's paper, however, differs substantively from this paper. First, in EM, the monopolist screens buyers' usual valuation, not their degree of inconsistency. Second, in EM, buyers value commitment but not flexibility, and consequently unused options are necessary to screen high-valuation buyers. In this paper, instead, since the agent also values flexibility, type I may find the efficient C-device unappealing even without unused options: To satisfy C's demand for flexibility, such a device must include many options and consequently can already create enough temptations from I's point of view. Finally, one can show that, in EM, maximizing welfare never requires adding unused options to any menu. In this paper, instead, the opposite is true: Maximizing welfare may require a C-device

with unused options, while maximizing profits may not (Proposition 5). This last difference explains why, as EM point out, their result on menus with unused options is not robust to competition, whereas the possibility of a C-device with unused options is (see Section 5.2).

5 Extensions

This section extends the model of Section 2 and its analysis in several natural directions. Section 5.1 allows for naive agents, in the sense that some agents may overestimate their self-control and thus undervalue commitment. In fact, some people do so (see, e.g., DellaVigna's (2009) survey and the references therein). Section 5.2 allows for competition among providers of commitment devices in period 1. The paper focuses on the case without competition, so as to isolate and thus better understand the screening problem created by the inability to observe people's time inconsistency. But, in many settings, commitment devices (like savings accounts) are provided by competing firms (like banks), whose incentives differ from those of a monopolist. Section 5.3 allows for more than two types of agents. The two-type model reveals important features of the problem of screening time inconsistency. But, in principle, richer heterogeneity across agents may generate other interesting results. Section 5.4 allows for inconsistent agents whose self-1's and self-2's preferences disagree in the opposite direction ($t > 1$) to that considered so far, so that self-2 tends to 'overconsume' a. As we saw in example 2, some interpretations of the model are consistent with $t > 1$. Section 5.5 allows different types to assign different values to the option of rejecting all the provider's devices. Intuitively, having no commitment device can be worse for more inconsistent types. This clearly changes the incentive constraints the provider faces. Finally, Section 5.6 relaxes the assumption that the agent's utility is transferable. Transferability is standard in many screening models and gives tractability, but it is not ideal for some applications of the present model.

Of course, these extensions add generality, nuances, and complications. Importantly, however, they do not change the main insights of the paper.[24]

[24] For reasons of space, the following discussion will be mainly informal and focused on intuitions.

5.1 Naïveté

One way to add naïveté to the present model is to assume that, with some probability, the agent of type $t < 1$ (or $t > 1$) believes in period 1 that his type is $\hat{t} \in (t, 1]$ (or $\hat{t} \in [1, t)$), but learns in period 2 that it is actually t (O'Donoghue and Rabin (2001)).

Naïveté affects the analysis as follows. On the one hand, it changes the provider's objective in the usual way: The provider now has to design the device for an agent who in period 1 believes his type to be $\hat{t}$, taking into account that (with some probability) his real type is t. So, depending on how much she cares about profits versus welfare, she will exploit or counteract the agent's naïveté. Of course, this implies that, even when the provider can observe $\hat{t}$ and t, she may offer a device that does not sustain $\mathbf{a}^*$. Intuitively, if an inconsistent agent believes himself to be consistent, he does not value the commitment that would allow his real self-2 to achieve $\mathbf{a}^*$, so providing such a commitment does not help to maximize profits. On the other hand, naïveté does not change the incentive constraints the provider has to satisfy: In period 2 only self-2's valuation v matters, and in period 1 only self-1's belief about t matters—not whether it is correct. In particular, Proposition 1 now says that, in period 1, an agent who *believes* himself to be less inconsistent values any flexible device more than an agent who *believes* himself to be more inconsistent. So, if the second gets a flexible device, then the first must get a positive information rent.

Therefore, the insights from Section 4 also apply to a setting with naïveté. Now the provider has to screen, in period 1, the agent's *perceived* degree of inconsistency $\hat{t}$. For each $\hat{t}$, she will face a trade-off between offering $\hat{t}$ the device that is optimal with observable types and extracting rents from an agent who believes himself to be less inconsistent than $\hat{t}$, while ensuring that an agent who believes himself to be more inconsistent than $\hat{t}$ does not mimic $\hat{t}$. Since, as noted, the incentive constraints are unchanged, the optimal way to solve this trade-off is also qualitatively unchanged. For instance, suppose in the two-type model $\mathtt{I}$ believes his type to be $\hat{t} \in (t^{\mathtt{I}}, t^{\mathtt{C}})$, and let $\mathbf{a}^{\mathtt{nI}*}$ be the optimal allocation for the naive $\mathtt{I}$ if all types are observable.[25] Then, to lower $\mathtt{C}$'s rent, the provider will curtail the flexibility of $\mathbf{a}^{\mathtt{nI}*}$ in the same way as she curtails that of $\mathbf{a}^{\mathtt{I}*}$ for the sophisticated $\mathtt{I}$. Finally, to better dissuade the naive $\mathtt{I}$ from mimicking $\mathtt{C}$, she may add unused options to the $\mathtt{C}$-device as shown in Section 4.3. Of course, design-

[25] If $\hat{t} = t^{\mathtt{C}}$, then in period 1 $\mathtt{I}$ and $\mathtt{C}$ believe that they are the same type. So the provider cannot screen them.

ing such options requires more caution with naïveté—especially when maximizing welfare—for a naive agent can incorrectly predict the likelihood of choosing them after committing to a device.

5.2 Competition

Competition in the provision of commitment devices raises natural questions. For example, can it lead firms to provide these devices efficiently, removing the distortions shown in Section 4? A complete answer is beyond the scope of this paper; therefore, the following only aims to give a rough intuition for why the answer may be no, even with perfect competition.

Consider a model that is identical to that in Section 2, with two exceptions: (1) there are many firms that, in period 1, can freely enter the market and offer devices as does the monopolist in the original setup; (2) there are many agents, each of type $\mathtt{I}$ or $\mathtt{C}$, who can freely interact with any firm. Ignoring technicalities, suppose a perfectly competitive equilibrium exists. First, one can see that the efficient outcome cannot arise with each type. Clearly, for this to happen, different types must choose different devices (Lemma 1); that is, the equilibrium must be separating. By free entry, each firm in the market must then break even on each device it offers, so j's expected payoff from the j-device equals the social surplus it creates with j: $U^j(\mathbf{a}^j, \mathbf{p}^j) = W^j(\mathbf{a}^j)$. Since the largest surplus $W^j(\mathbf{a}^{j*})$ is the same across types (by definition), it follows that $U^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}*}, \mathbf{p}^{\mathtt{C}}) = U^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}*}, \mathbf{p}^{\mathtt{I}})$. But by Proposition 1 $U^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*}, \mathbf{p}^{\mathtt{I}}) > U^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}*}, \mathbf{p}^{\mathtt{I}})$, so $\mathtt{C}$ agents would prefer the $\mathtt{I}$-devices, contradicting separation. We conclude that in any separating equilibrium

$$W^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}*}) \geq U^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}}, \mathbf{p}^{\mathtt{C}}) > U^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}, \mathbf{p}^{\mathtt{I}}) = W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}),$$

so $\mathbf{a}^{\mathtt{I}}$ is inefficient.

Second, one can see that the inefficiencies of the $\mathtt{I}$-devices should qualitatively match those in Section 4.2. For simplicity, assume that it is always possible to deter $\mathtt{I}$ agents from mimicking $\mathtt{C}$ agents by adding unused options to otherwise efficient $\mathtt{C}$-devices (Section 4.3). Then, in equilibrium, each $\mathtt{I}$-device should maximize efficiency with $\mathtt{I}$ (i.e., $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}})$), while ensuring that no $\mathtt{C}$ wants to mimic $\mathtt{I}$ (i.e., $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) + R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}) \leq W^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}*})$). Applying once more the methods used for Lemma 2, we get that $\mathbf{a}^{\mathtt{I}}$ must maximize, among all increasing functions, an objective of the form $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) - rR^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$ with $r > 0$, as in Section 4.2.

5.3 Many Degrees of Inconsistency

Consider a model that is identical to that in Section 2, except that now the agent can be one of $N > 2$ types with N finite. For simplicity, index the types from the lowest to the highest degree of inconsistency: $1 \geq t^1 > t^2 > \ldots > t^N > 0$. Finally, let j's probability be $\gamma^j > 0$.

Most of the analysis of the screening problem generalizes easily. A DM is now an array $\{\mathbf{A}, \mathbf{P}\} = (\mathbf{a}^j, \mathbf{p}^j)_{j=1}^N$ with $\mathbf{a}^j : [\underline{v}, \overline{v}] \to [\underline{a}, \overline{a}]$ and $\mathbf{p}^j : [\underline{v}, \overline{v}] \to \mathbb{R}$, where $[\underline{v}, \overline{v}] = [\underline{s}t^N, \overline{s}t^1]$. As in problem $\mathcal{P}$, each DM must satisfy constraints (IC_2^j) and (IR^j) for every j, and constraint IC_1^{jm}, given by $U^j(\mathbf{a}^j, \mathbf{p}^j) \geq U^j(\mathbf{a}^m, \mathbf{p}^m)$, for every j and m. The provider's objective changes only to the extent that she now takes expectations over N types. Finally, a generalization of Proposition 1 says that if (IC_2^j) holds, then $U^i(\mathbf{a}^j, \mathbf{p}^j) \geq U^m(\mathbf{a}^j, \mathbf{p}^j)$ for $i < m$, with equality if and only if $\mathbf{a}^j$ is constant over $(\underline{v}^m, \overline{v}^i)$ (recall that $\underline{v}^m = \underline{s}t^m$ and $\overline{v}^i = \overline{s}t^i$). Intuitively, outside this range neither i nor m get to choose $\mathbf{a}^j$ (v), so neither i's nor m's payoff depends on it.[26]

Screening more than two degrees of inconsistency, however, raises one additional complication, which this paper solves with a novel strategy. Not only can the constraints IC_1^{jm} bind in both directions between adjacent types—as in the two-type model. But they can also be violated between nonadjacent types, even though they hold between all adjacent ones. A similar issue appears in the literature on dynamic mechanism design with only time-consistent agents, which has developed a specific strategy to handle it (Courty and Li (2000); Pavan, Segal, and Toikka (2012)). This strategy involves restricting the primitives (potentially in an ad-hoc way), so that local incentive compatibility implies global incentive compatibility. Using this strategy to study a new screening problem seems unappealing; moreover, finding effective—let alone reasonable—restrictions is hard in the present model. Therefore, this paper adopts a different strategy that does not restrict the primitives and deals with the constraints IC_1^{jm} all at once; it uses Lagrangian methods and relies on having finitely many types.

Applying these methods reveals that, when designing each j-device, the provider trades off the surplus with j and the rents the device causes for (some of) the less inconsistent types (see Online Appendix A); this generalizes the trade-off highlighted in the two-type model. The provider also has to ensure that each j-device deters types who are more inconsistent than j from mimicking j. Based on Sec-

[26] It is easy to see this generalization from the proof of Proposition 1.

tion 4.3, the next proposition looks at the case in which unused options suffice to ensure this property.[27]

Proposition 6 *Suppose $b(\underline{a})$ is low so that unused options suffice to satisfy IC_1^{jm} for $j > m$. An optimal mechanism $\{\mathbf{A}, \mathbf{P}\}$ exists with $\mathbf{a}^j$ unique over $(\underline{v}^j, \overline{v}^j)$, for every j. Moreover, (i) the 1-device sustains $\mathbf{a}^*$; (ii) all devices sustain $\mathbf{a}^*$ if $\pi = 0$, otherwise at least the N-device sustains a distorted outcome; (iii) if the j-device is distorted, it curtails j's flexibility as in Proposition 2; and (iv) for $j < N$, the j-device may have to include unused options with $a < \mathbf{a}^j(\underline{v}^j)$.*

As long as the provider cares about profits ($\pi > 0$), she will curtail flexibility for the most inconsistent type in the same way as she did for $\mathtt{I}$ in the two-type model. This also happens for an intermediate type m, whenever at the optimum IC_1^{jm} binds for some j less inconsistent than m.

5.4 Other Directions of Time Inconsistency

To see what happens when time inconsistency induces self-2 to overconsume a, consider a two-type model with $t^{\mathtt{I}} > 1 = t^{\mathtt{C}}$. In period 1, $\mathtt{C}$ values any flexible device more than $\mathtt{I}$, for the same logic behind Proposition 1. So the provider has to trade-off the surplus with $\mathtt{I}$ and the rent to $\mathtt{C}$ when designing the $\mathtt{I}$-device, which calls for curtailing its flexibility for the same reason as in Section 4.2. The optimal way to curtail flexibility is also qualitatively the same. The only change is that now there is always no flexibility at the bottom (rather than at the top) and possibly no flexibility at the top (rather than at the bottom). Moreover, Lemma 4 and its implications continue to hold. So, if the provider has to deter $\mathtt{I}$ from mimicking $\mathtt{C}$, she again designs the $\mathtt{C}$-device as shown in Section 4.3. The only change here is that the unused option features $\overline{a}$ (rather than $\underline{a}$), because now $\mathtt{I}$'s self-1 deems $\overline{a}$ as the action that would most tempt his self-2, who always prefers higher actions ($t^{\mathtt{I}} > 1$).

Finally, to see what happens when time inconsistency induces some agents to underconsume a and others to overconsume it in period 2, consider a three-type model with $t^{\mathtt{I}_1} > 1 = t^{\mathtt{C}} > t^{\mathtt{I}_2}$. Again, in period 1, $\mathtt{C}$ values flexible devices more than both $\mathtt{I}_1$ and $\mathtt{I}_2$; so the provider has to grant $\mathtt{C}$ an information rent. However, she may be able to extract the whole surplus from both $\mathtt{I}_1$ and $\mathtt{I}_2$. Indeed, neither $\mathtt{I}_1$ nor $\mathtt{I}_2$ may value the device for any other type j more than j. Intuitively, $\mathtt{I}_1$

[27] For the proof, see Online Appendix A.

and $\mathtt{I}_2$ can behave differently after choosing the device for j, but from self-1's point of view, neither may be able to improve on j's choices (see Online Appendix B). To fully analyze the case with both $t < 1$ and $t > 1$, one can again use the methods in Section 5.3.

5.5 Outside Option with Type-Dependent Values

After rejecting all the provider's devices in period 1, the agent will make certain state-contingent choices in period 2, which can be described with the pair $(\mathbf{a}_0, \mathbf{p}_0)$ using the formalism of Section 4.1. For simplicity, consider again the two-type model. By Proposition 1, $U^{\mathtt{C}}(\mathbf{a}_0, \mathbf{p}_0) \geq U^{\mathtt{I}}(\mathbf{a}_0, \mathbf{p}_0)$ with equality if and only if $\mathbf{a}_0$ is constant over $(\underline{v}, \overline{v})$. So $\mathtt{C}$ and $\mathtt{I}$ value the outside option differently, unless they always end up making the same choice.

When $\mathtt{C}$ values the outside option more than $\mathtt{I}$, the analysis in Section 4 can be easily adjusted without changing its thrust. Recall that the constraints ($\text{IR}^{\mathtt{C}}$) and ($\text{IC}_1^{\mathtt{C}}$) set two lower bounds on $\mathtt{C}$'s payoff from the $\mathtt{C}$-device: one endogenous, given by $U^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}, \mathbf{p}^{\mathtt{I}}) = U^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}, \mathbf{p}^{\mathtt{I}}) + R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$, and one exogenous, given by $U^{\mathtt{C}}(\mathbf{a}_0, \mathbf{p}_0) = U^{\mathtt{I}}(\mathbf{a}_0, \mathbf{p}_0) + R^{\mathtt{C}}(\mathbf{a}_0)$. The question is which binds first. In Section 4, the endogenous bound always binds first, for ($\text{IR}^{\mathtt{I}}$) and ($\text{IC}_1^{\mathtt{C}}$) imply ($\text{IR}^{\mathtt{C}}$). Now this is no longer the case. But, intuitively, whenever the endogenous bound binds first, we are in a situation similar to Section 4; so the provider will distort the $\mathtt{I}$-device as shown in Section 4.2.[28] On the other hand, if the exogenous bound binds first, then obviously the provider has no reason to distort the $\mathtt{I}$-device. For example, the provider will never distort the $\mathtt{I}$-device, if the outside option already sustains the efficient outcome with $\mathtt{I}$—i.e., $\mathbf{a}_0 = \mathbf{a}^{\mathtt{I}*}$ over $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. In this case, the provider must grant $\mathtt{C}$ at least the surplus $R^{\mathtt{C}}(\mathbf{a}_0)$, which exceeds $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*})$, that $\mathtt{C}$ can get from $(\mathbf{a}_0, \mathbf{p}_0)$ relative to $U^{\mathtt{I}}(\mathbf{a}_0, \mathbf{p}_0)$. Finally, if ($\text{IC}_1^{\mathtt{I}}$) binds, then the provider will design the $\mathtt{C}$-device as shown in Section 4.3.[29]

5.6 Nontransferable Utility

The assumption of transferable utility between the provider and the agent is not ideal for some applications. In the case of savings devices, for example, it seems more natural to let self-1's and self-2's preferences be $b(y - a - p) + sb(a)$ and

[28] This is more likely to happen when the outside option involves little flexibility, so that $\mathbf{a}_0$ is almost constant and $R^{\mathtt{C}}(\mathbf{a}_0)$ is small. Recall that $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*}) > 0$.

[29] This argument can be extended to settings in which, in period 1, the agent has access to other devices if he rejects the provider's ones. In these settings, $(\mathbf{a}_0, \mathbf{p}_0)$ can be type dependent.

$b(y-a-p)+tsb(a)$, where the instantaneous utility $b(\cdot)$ is nonlinear. It is well known, however, that screening models without transferability can be complicated to analyze; this paper is no exception.

To some extent, characterizing the agent's incentives to use commitment devices may be straightforward even without transferability. First, in a model with types $\mathtt{I}$ and $\mathtt{C}$, the same revealed-preference logic behind Proposition 1 implies that $\mathtt{C}$ must enjoy an information rent; and the same logic behind Lemma 4 implies that $\mathtt{C}$'s rent can again jeopardize $\mathtt{I}$'s incentives to reveal his type. Second, we can still use truthful direct mechanisms to describe each device with a pair $(\mathbf{a}^j, \mathbf{p}^j)$ as in Section 4.1. More specifically, suppose in the previous example $b(\cdot)$ is strictly increasing with range $\mathbb{R}$. Then, we can let $\hat{\mathbf{p}}^j(\cdot) = b(y - \mathbf{a}^j(\cdot) - \mathbf{p}^j(\cdot))$ and work with $(\mathbf{a}^j, \hat{\mathbf{p}}^j)$ to characterize the incentive constraints as before.

On the other hand, characterizing the optimal screening devices is harder without transferability. As usual, this is because we lose the linear relationship between utils for the agent and profits for the provider. To see this, consider again the previous example with additive utility functions. Using the formalism $(\mathbf{a}^j, \hat{\mathbf{p}}^j)$, standard arguments imply that

$$\hat{\mathbf{p}}^j(v) = vb(\mathbf{a}^j(v)) + \textstyle\int_v^{\overline{v}} b(\mathbf{a}^j(x))dx + \hat{k}^j,$$

where $\hat{k}^j$ controls the level of j's expected payoff. For $\mathtt{C}$, this level depends on the rent caused by the $\mathtt{I}$-device, so $\hat{k}^{\mathtt{C}}$ depends on $\mathbf{a}^{\mathtt{I}}$—similarly, $\hat{k}^{\mathtt{I}}$ can depend on $\mathbf{a}^{\mathtt{C}}$. Now consider the profit from $(\mathbf{a}^{\mathtt{C}}(v), \hat{\mathbf{p}}^{\mathtt{C}}(v))$, which equals $y - \mathbf{a}^{\mathtt{C}}(v) - b^{-1}(\hat{\mathbf{p}}^{\mathtt{C}}(v)) - c(\mathbf{a}^{\mathtt{C}}(v))$. Since $b(\cdot)$ is nonlinear, the profit is not separable in $\mathbf{a}^{\mathtt{C}}$ and $\hat{k}^{\mathtt{C}}$ and therefore the trade-offs that define each $\mathbf{a}^{\mathtt{C}}(v)$ depend on the entire function $\mathbf{a}^{\mathtt{I}}$. In the end, this implies that—in contrast to Lemma 2—we cannot characterize each $\mathbf{a}^j$ by solving a distinct maximization that fully determines its properties.

However, recall that the logic behind the properties of the optimal devices in Sections 4.2 and 4.3 does not rely on transferability. Therefore, it seems reasonable to expect that the key insights of this paper should also hold in settings without transferability.

6 Conclusion

Commitment devices that rely on monetary incentives can offset people's time inconsistency by realigning their preferences at different dates; such devices can

thus avoid tensions between people's demand for commitment and for flexibility. When aware of their inconsistency, people are ready to pay more for devices that offset it; this gives profit-maximizing firms an incentive to offer such devices. But the combination of people's demand for flexibility and superior information on their degree of inconsistency creates an adverse-selection problem. Its profit-maximizing solution involves devices for more inconsistent people that curtail flexibility at both ends of the efficient choice range, and devices for less inconsistent people that may include unused options, and possibly distort their behavior; these properties can mark the welfare-maximizing solution too. By curtailing flexibility, a device for more inconsistent people makes a less inconsistent person less willing to mimic them, because it gives him fewer chances to make better decisions. By including unused options, a device for less inconsistent people makes a more inconsistent person less willing to mimic them, because it gives him more chances to make worse decisions.

The theory developed in this paper can be a basis to think about how public and private institutions (should) provide commitment devices when, ex ante, they cannot easily detect each person's degree of inconsistency (or self-control). For example, governments may want to offer savings accounts with tax incentives that help inconsistent people save enough for retirement; gyms may want to offer memberships with monetary incentives that help inconsistent people work out regularly.

Alternatively, one could also interpret the model in this paper as capturing 'nested-agency' situations like the following. A local state (corresponding to self-1 in the model) has to delegate some decision to a better informed firm (self-2). To do that, however, the local state must comply with delegation rules designed by a federal regulator (the provider). This paper could then offer insights on how the regulator should design these rules, when the local state knows better how the firm's goals differ from the state and federal goals.

One key assumption of this paper is that the provider can offer people commitment devices to begin with. But this ability may be limited if the provider herself lacks commitment power, or if people can contract with other parties in the future—people can open illiquid savings accounts at a bank today, and then get credit cards from the same or another bank. As in other dynamic-contracting settings, ex post, the provider may want to renegotiate her ex-ante offers. Moreover, since here people can be time inconsistent, they themselves may want to undo, ex post, the commitment they took on ex ante—either by renegotiating it or by

trading with new parties. Gottlieb (2008) and Zhang (2012) tackle some of these issues in settings in which everybody knows each person's degree of inconsistency. Investigating how renegotiation and ex-post contracting affects the provision of commitment devices, when people privately know their degree of inconsistency, is left for future work.

A Proofs

A.1 Proof of Lemma 1

Consider the provider's problem described in the main text. If $\pi > 0$, (IR) must bind; if $\pi = 0$, assume w.l.o.g. that (IR) holds with equality. Thus, the problem becomes

$$\max_{\mathbf{a}} \left\{ \int_{\underline{s}}^{\overline{s}} [u_1 (\mathbf{a}(s); s) - c(\mathbf{a}(s))] dF \right\} \quad \text{s.t. (IC)}.$$

The relaxed problem without (IC) has a unique solution (up to $\{\underline{s}, \overline{s}\}$): $\mathbf{a} \equiv \mathbf{a}^*$. It remains to show that there is $\mathbf{p}^*$ such that $(\mathbf{a}^*, \mathbf{p}^*)$ satisfies (IC). By standard arguments, for any $t > 0$ such a $\mathbf{p}_t^*$ exists if and only if $\mathbf{a}$ increases in s, a property satisfied by $\mathbf{a}^*$. Moreover, for every s,

$$\mathbf{p}_t^*(s) = u_2(\mathbf{a}^*(s); s, t) - \int_{\underline{s}}^{s} t b(\mathbf{a}^*(s)) - k,$$

where k is a scalar. By standard arguments, $\mathbf{a}^*$ is differentiable and

$$\frac{d\mathbf{p}_t^*(s)}{ds} = \frac{\partial u_2(\mathbf{a}^*(s); s, t)}{\partial a} \frac{d\mathbf{a}^*(s)}{ds},$$

which equals $c'(\mathbf{a}^*(s)) \frac{d\mathbf{a}^*(s)}{ds}$ if and only if $t = 1$, by the definition of $\mathbf{a}^*$ and Assumption 1. Finally, the relation between $\frac{d\mathbf{p}_t^*}{ds}$ and $\frac{d\mathbf{p}_1^*}{ds}$ is immediate from the definition of u_1 and u_2.

A.2 Proof of Proposition 1

By standard arguments, (IC_2^j) holds if and only if $\mathbf{a}^j$ is increasing and, for $v \in [\underline{v}, \overline{v}]$,

$$\mathbf{p}^j(v) = u_2(\mathbf{a}^j(v); v) + \int_v^{\overline{v}} b(\mathbf{a}^j(x)) dx - k^j \tag{1}$$

where $k^j = u_2(\mathbf{a}^j(\overline{v}); \overline{v}) - \mathbf{p}^j(\overline{v})$. Using (1), we get

$$U^i(\mathbf{a}^j, \mathbf{p}^j) = \int_{\underline{v}^i}^{\overline{v}^i} \left[u_1(\mathbf{a}^j(v); v/t^i) - u_2(\mathbf{a}^j(v); v) - \int_v^{\overline{v}} b(\mathbf{a}^j(x))dx \right] dF^i(v) + k^j, \tag{2}$$

where F^i is the distribution F induces on $[\underline{v}, \overline{v}]$, conditional on being type i. Changing variables, we get

$$U^i(\mathbf{a}^j, \mathbf{p}^j) = \int_{\underline{s}}^{\overline{s}} [sb(\mathbf{a}^j(t^i s)) - t^i s b(\mathbf{a}^j(t^i s)) - \int_{t^i s}^{\overline{v}} b(\mathbf{a}^j(x))dx] dF(s) + k^j. \tag{3}$$

For each s, define

$$\begin{aligned} \Delta(s|\,\mathbf{a}^j) &= s(1-t^{\mathtt{C}})b(\mathbf{a}^j(t^{\mathtt{C}}s)) - \int_{t^{\mathtt{C}}s}^{\overline{v}} b(\mathbf{a}^j(x))dx \\ &\quad -s(1-t^{\mathtt{I}})b(\mathbf{a}^j(t^{\mathtt{I}}s)) + \int_{t^{\mathtt{I}}s}^{\overline{v}} b(\mathbf{a}^j(x))dx \\ &= s(1-t^{\mathtt{C}})[b(\mathbf{a}^j(t^{\mathtt{C}}s)) - b(\mathbf{a}^j(t^{\mathtt{I}}s))] \\ &\quad + \int_{t^{\mathtt{I}}s}^{t^{\mathtt{C}}s} [b(\mathbf{a}^j(x)) - b(\mathbf{a}^j(t^{\mathtt{I}}s))]dx. \end{aligned} \tag{4}$$

Since $\mathbf{a}^j$ is increasing and $t^{\mathtt{I}} < t^{\mathtt{C}} \leq 1$, $\Delta(s|\,\mathbf{a}^j) \geq 0$. Also, if $\mathbf{a}^j(v) = a$ on $(\underline{v}, \overline{v})$, then $\Delta(s|\,\mathbf{a}^j) = 0$ on $(\underline{s}, \overline{s})$. Using (3) and (4), we have

$$U^{\mathtt{C}}(\mathbf{a}^j, \mathbf{p}^j) - U^{\mathtt{I}}(\mathbf{a}^j, \mathbf{p}^j) = \int_{\underline{s}}^{\overline{s}} \Delta(s|\,\mathbf{a}^j) dF \geq 0,$$

with equality if $\mathbf{a}^j(v) = a$ on $(\underline{v}, \overline{v})$. Now suppose $\mathbf{a}^j$ is not constant on $(\underline{v}, \overline{v})$. Since $\mathbf{a}^j$ is increasing, there is $\tilde{v} \in (\underline{v}, \overline{v})$ such that $v < \tilde{v} < v'$ implies $\mathbf{a}^j(v) < \mathbf{a}^j(v')$. Let $\tilde{s}_1 = \tilde{v}/t^{\mathtt{C}}$ and $\tilde{s}_2 = \tilde{v}/t^{\mathtt{I}}$, and consider interval $\mathcal{I} = (\tilde{s}_1, \tilde{s}_2) \cap [\underline{s}, \overline{s}] \neq \varnothing$. For $s \in \mathcal{I}$, $t^{\mathtt{I}}s < \tilde{v} < t^{\mathtt{C}}s$ and $\mathbf{a}^j(t^{\mathtt{I}}s) < \mathbf{a}^j(t^{\mathtt{C}}s)$. To prove that $U^{\mathtt{C}}(\mathbf{a}^j, \mathbf{p}^j) > U^{\mathtt{I}}(\mathbf{a}^j, \mathbf{p}^j)$, it is enough to show that

$$\int_{\mathcal{I}} \left[\int_{t^{\mathtt{I}}s}^{t^{\mathtt{C}}s} [b(\mathbf{a}^j(x)) - b(\mathbf{a}^j(t^{\mathtt{I}}s))]dx \right] dF > 0. \tag{5}$$

For $s \in \mathcal{I}$,

$$\int_{t^{\mathtt{I}}s}^{t^{\mathtt{C}}s} [b(\mathbf{a}^j(x)) - b(\mathbf{a}^j(t^{\mathtt{I}}s))]dx \geq \int_{\tilde{v}}^{t^{\mathtt{C}}s} [b(\mathbf{a}^j(x)) - b(\mathbf{a}^j(t^{\mathtt{I}}s))]dx > 0,$$

where the first inequality follows from $\mathbf{a}^j$ being increasing and the last from $\mathbf{a}^j(x) > \mathbf{a}^j(t^{\mathtt{I}}s)$ for $x \in (\tilde{v}, t^{\mathtt{C}}s)$. Since $\mathcal{I}$ has positive measure, (5) follows.

A.3 Proof of Lemma 2

Let $\mathcal{B} = \{\mathbf{b} : [\underline{v}, \overline{v}] \to [b(\underline{a}), b(\overline{a})] \mid \mathbf{b} \text{ increasing}\}$. If $\mathbf{a}$ is increasing, then $\mathbf{b}(v) = b(\mathbf{a}(v)) \in \mathcal{B}$; if $\mathbf{b} \in \mathcal{B}$, then $\mathbf{a}(v) = b^{-1}(\mathbf{b}(v))$ is increasing. Let $\widetilde{W}^i(\mathbf{b}) = W^i(b^{-1}(\mathbf{b}))$ and $\widetilde{R}^i(\mathbf{b}^j) = R^i(b^{-1}(\mathbf{b}^j))$. Then, $\mathcal{P}'$ is equivalent to

$$\mathcal{P}^{\mathtt{b}} = \begin{cases} \max \gamma \widetilde{W}^{\mathtt{C}}(\mathbf{b}^{\mathtt{C}}) + (1-\gamma)\left[\widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}) - \frac{\pi\gamma}{1-\gamma}\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}})\right] \\ \text{s.t. } \mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}} \in \mathcal{B} \text{ and } \widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) \leq 0. \end{cases}$$

The space $\mathcal{X} = \{(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}) \mid \mathbf{b}^i : [\underline{v}, \overline{v}] \to \mathbb{R}\}$ is linear and $\mathcal{Y} = \mathcal{B} \times \mathcal{B}$ is a convex subset of $\mathcal{X}$. The objective is concave, as b^{-1} and c are convex; $\widetilde{R}^{\mathtt{I}}(\cdot) + \widetilde{R}^{\mathtt{C}}(\cdot)$ is linear. Moreover, there is $(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}) \in \mathcal{Y}$ such that $\widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) + \widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) < 0$: e.g., $\mathbf{b}^{\mathtt{C}} = b(\mathbf{a}^{\mathtt{C}*})$ and $\mathbf{b}^{\mathtt{I}}$ constant. Let $\mu \geq 0$ and define

$$\begin{aligned} L(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu) &= \gamma\widetilde{W}^{\mathtt{C}}(\mathbf{b}^{\mathtt{C}}) + (1-\gamma)\left[\widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}) - \frac{\pi\gamma}{1-\gamma}\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}})\right] \\ &\quad -\mu[\widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) + \widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}})] \\ &= \gamma[\widetilde{W}^{\mathtt{C}}(\mathbf{b}^{\mathtt{C}}) - r^{\mathtt{I}}\widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}})] + (1-\gamma)\left[\widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}) - r^{\mathtt{C}}\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}})\right]. \quad (6) \end{aligned}$$

By Corollary 1, p. 219, and Theorem 2, p. 221, of Luenberger (1969), $(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}})$ solve $\mathcal{P}^{\mathtt{b}}$ if and only if there is $\mu \geq 0$ such that, for all $(\mathbf{b}_0^{\mathtt{C}}, \mathbf{b}_0^{\mathtt{I}}) \in \mathcal{Y}, \mu_0 \geq 0$, both $L(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu) \geq L(\mathbf{b}_0^{\mathtt{C}}, \mathbf{b}_0^{\mathtt{I}}, \mu)$ and $L(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu_0) \geq L(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu)$. Given $\mu \geq 0$, the first inequality holds if and only if $\mathbf{b}^{\mathtt{C}}$ and $\mathbf{b}^{\mathtt{I}}$ maximize, within $\mathcal{B}$, the first and second term in brackets of (6). The second inequality holds if $\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) \leq 0$ and $\mu[\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}})] = 0$. Finally, if $(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}})$ solves $\mathcal{P}^{\mathtt{b}}$, then $\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) \leq 0$. And if $\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) < 0$, then μ must be zero: otherwise, there is $\mu_0 \in [0, \mu)$ such that

$$L(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu_0) - L(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu) = (\mu - \mu_0)\left[\widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) + \widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}})\right] < 0.$$

Finally, if $(\mathbf{b}^{\mathtt{C}}, \mathbf{b}^{\mathtt{I}}, \mu)$ satisfies $\mathbf{b}^i \in \arg\max_{\mathbf{b} \in \mathcal{B}}\{\widetilde{W}^i(\mathbf{b}^i) - r^{-i}\widetilde{R}^{-i}(\mathbf{b}^i)\}$, $\mu \geq 0$, $\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) \leq 0$, and $\mu[\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}})] = 0$, then $\mathbf{a}^{\mathtt{C}} = b^{-1}(\mathbf{b}^{\mathtt{C}})$, $\mathbf{a}^{\mathtt{I}} = b^{-1}(\mathbf{b}^{\mathtt{I}})$, and μ satisfy the conditions in Lemma 2. Similarly, if $(\mathbf{a}^{\mathtt{C}}, \mathbf{a}^{\mathtt{I}}, \mu)$ satisfies the conditions in Lemma 2, then $\mathbf{b}^i = b(\mathbf{a}^i) \in \arg\max_{\mathbf{b} \in \mathcal{B}}\{\widetilde{W}^i(\mathbf{b}^i) - r^{-i}\widetilde{R}^{-i}(\mathbf{b}^i)\}$, $\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}}) \leq 0$, and $\mu[\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}) + \widetilde{R}^{\mathtt{I}}(\mathbf{b}^{\mathtt{C}})] = 0$.

A.4 Proof of Propositions 2 and 3

Step 0: Writing $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) - r^{\mathtt{C}} R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$ in Lemma 2 as an expected virtual surplus. To do so, note that

$$W^i(\mathbf{a}^i) = \int_{\underline{v}}^{\overline{v}} [u_1(\mathbf{a}^i(v); v/t^i) - c(\mathbf{a}^i(v))] dF^i. \tag{7}$$

Use (2) to express $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$. Then, changing order of integration and rearranging yields

$$R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}) = -\int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} b(\mathbf{a}^{\mathtt{I}}(v)) g^{\mathtt{C}}(v)\, dv - \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} b(\mathbf{a}^{\mathtt{I}}(v)) G^{\mathtt{I}}(v)\, dF^{\mathtt{I}}, \tag{8}$$

where $g^{\mathtt{C}}(v) : (\overline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}] \to \mathbb{R}$ and $G^{\mathtt{I}} : [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}] \to \mathbb{R}$ are given by

$$g^{\mathtt{C}}(v) = \tfrac{t^{\mathtt{C}}-1}{t^{\mathtt{C}}} v f^{\mathtt{C}}(v) - (1 - F^{\mathtt{C}}(v)) \quad \text{and} \quad G^{\mathtt{I}}(v) = q^{\mathtt{I}}(v) - \tfrac{f^{\mathtt{C}}(v)}{f^{\mathtt{I}}(v)} q^{\mathtt{C}}(v), \tag{9}$$

with $q^i(v) = v/t^i - v - F^i(v)/f^i(v)$. By (7) and (8), $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) - r^{\mathtt{C}} R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$ equals the expected virtual surplus

$$\begin{aligned} VS^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}; r^{\mathtt{C}}) \;=\; & \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \left[b(\mathbf{a}^{\mathtt{I}}(v)) w^{\mathtt{I}}(v; r^{\mathtt{C}}) - \mathbf{a}^{\mathtt{I}}(v) - c(\mathbf{a}^{\mathtt{I}}(v)) \right] dF^{\mathtt{I}} \\ & - r^{\mathtt{C}} \int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} b(\mathbf{a}^{\mathtt{I}}(v))(1 - F^{\mathtt{C}}(v)) dv, \end{aligned} \tag{10}$$

where, on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, $w^{\mathtt{I}}(v; r^{\mathtt{C}}) = v/t^{\mathtt{I}} + r^{\mathtt{C}} G^{\mathtt{I}}(v)$ is the *virtual valuation* of $b(\mathbf{a}^{\mathtt{I}}(v))$.

As in the proof of Lemma 2, it is convenient to work in terms of the functions $\mathbf{b}^{\mathtt{I}} \in \mathcal{B}$. The properties of the corresponding allocation follow by letting $\mathbf{a}^{\mathtt{I}} = b^{-1}(\mathbf{b}^{\mathtt{I}})$.

Part 1: Existence and Uniqueness.

Step 1: Constructing the generalized version of $VS^{\mathtt{I}}$ using Toikka's (2011) method on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. Since f is strictly positive, the inverse function $(F^{\mathtt{I}})^{-1} : [0,1] \to [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$ is well-defined, strictly increasing, and continuous. Fix $r^{\mathtt{C}}$ and define, for $x \in [0,1]$,

$$z(x; r^{\mathtt{C}}) = w^{\mathtt{I}}((F^{\mathtt{I}})^{-1}(x); r^{\mathtt{C}}) \quad \text{and} \quad Z(x; r^{\mathtt{C}}) = \int_0^x z(y; r^{\mathtt{C}}) dy.$$

Then, z is continuous in x, except possibly at $x^m = F^{\mathtt{I}}(v^m) > 0$, where $v^m = \min\{\overline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}\}$: If $t^{\mathtt{C}} < 1$ and $\underline{v}^{\mathtt{C}} < \overline{v}^{\mathtt{I}}$,

$$\lim_{v \downarrow \underline{v}^{\mathtt{C}}} w^{\mathtt{I}}(v; r^{\mathtt{C}}) = \lim_{v \uparrow \underline{v}^{\mathtt{C}}} w^{\mathtt{I}}(v; r^{\mathtt{C}}) - r^{\mathtt{C}} \frac{f^{\mathtt{C}}(\underline{v}^{\mathtt{C}})}{f^{\mathtt{I}}(\underline{v}^{\mathtt{C}})} \left[\frac{1 - t^{\mathtt{C}}}{t^{\mathtt{C}}} \underline{v}^{\mathtt{C}} \right]. \tag{11}$$

Let Ω be the convex hull of Z: Ω is the highest convex function such that $\Omega \leq Z$ (Rockafellar (1970), p.36). Define $\omega : [0,1] \to \mathbb{R}$ as $\omega(x; r^{\mathtt{C}}) = \Omega'(x; r^{\mathtt{C}})$, whenever $\Omega'(x; r^{\mathtt{C}})$ exists. W.l.o.g., extend $\omega(x; r^{\mathtt{C}})$ by right-continuity on $[0,1)$ and by left-continuity at 1.

Lemma 5 *The function ω is continuous in x and $r^{\mathtt{C}}$.*

Proof. (Continuity in x). Suppress $r^{\mathtt{C}}$. Continuity at 0 and 1 holds by construction. For $x \in (0,1) \setminus \{x^m\}$, z is continuous, so $Z'(x) = z(x)$. First, suppose $\Omega(x) < Z(x)$. By definition, $\omega(\cdot)$ is constant at $\omega(x)$ in a neighborhood of x; so ω is continuous at x. Second, suppose $\Omega(x) = Z(x)$. Since Ω is convex and $\Omega \leq Z$,

$$\Omega^+(x) = \lim_{y \downarrow x} \frac{\Omega(y) - \Omega(x)}{y - x} \leq \lim_{y \downarrow x} \frac{Z(y) - Z(x)}{y - x} = Z^+(x),$$

and similarly $\Omega^-(x) \geq Z^-(x)$. Since $\Omega^-(x) \leq \Omega^+(x)$ and Z is differentiable at x, $\Omega^-(x) = \Omega^+(x)$; so ω is continuous at x. Finally, consider x^m. If $v^m = \overline{v}^{\mathtt{I}}$, then $x^m = 1$ and we are done. For $x^m \in (0,1)$, ω is continuous if $\Omega(x^m) < Z(x^m)$ when z jumps at x^m. Recall that

$$\begin{aligned} z(x^m-) &= \lim_{x \uparrow x^m} z(x) = \lim_{v \uparrow v^m} w^{\mathtt{I}}(v; r^{\mathtt{C}}), \\ z(x^m+) &= \lim_{x \downarrow x^m} z(x) = \lim_{v \downarrow v^m} w^{\mathtt{I}}(v; r^{\mathtt{C}}). \end{aligned}$$

By (11), z can only jump down at x^m, so $z(x^m-) > z(x^m+)$. Also, $z(x^m+) = z(x^m)$. Suppose $\Omega(x^m) = Z(x^m)$. By the same steps as before, $\Omega^+(x^m) \leq Z^+(x^m) = z(x^m)$. By convexity, $\omega(x) \leq \Omega^-(x^m)$ for $x \leq x^m$. So, for x close to x^m from the left, we obtain the following contradiction:

$$\Omega(x) = \Omega(x^m) - \textstyle\int_x^{x^m} \omega(y)\, dy > Z(x^m) - \int_x^{x^m} z(y)\, dy = Z(x).$$

(Continuity in $r^{\mathtt{C}}$). Given x, $Z(x; r^{\mathtt{C}})$ is continuous in $r^{\mathtt{C}}$. So Ω is continuous if $x \in \{0,1\}$, since $\Omega(0; r^{\mathtt{C}}) = Z(0; r^{\mathtt{C}})$ and $\Omega(1; r^{\mathtt{C}}) = Z(1; r^{\mathtt{C}})$. Consider $x \in (0,1)$. For $r^{\mathtt{C}} \geq 0$, by definition,

$$\Omega(x; r^{\mathtt{C}}) = \min \alpha Z(x_1; r^{\mathtt{C}}) + (1 - \alpha) Z(x_2; r^{\mathtt{C}})$$

over all $\alpha, x_1, x_2 \in [0,1]$ such that $x = \alpha x_1 + (1 - \alpha) x_2$. By continuity of $Z(x; r^{\mathtt{C}})$ and the Maximum Theorem, $\Omega(x, \cdot)$ is continuous in $r^{\mathtt{C}}$ for every x. Moreover,

$\Omega(\cdot; r^{\mathtt{C}})$ is differentiable in x with derivative $\omega(\cdot; r^{\mathtt{C}})$. Fix $x \in (0,1)$ and any sequence $\{r_n^{\mathtt{C}}\}$ such that $\lim_{n\to\infty} r_n^{\mathtt{C}} = r^{\mathtt{C}}$. Since $\lim_{n\to\infty} \Omega(x; r_n^{\mathtt{C}}) = \Omega(x; r^{\mathtt{C}})$, Theorem 25.7, p. 248, of Rockafellar (1970) implies $\lim_{n\to\infty} \omega(x; r_n^{\mathtt{C}}) = \omega(x; r^{\mathtt{C}})$. ■

On $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, define the generalized virtual valuation

$$\overline{w}^{\mathtt{I}}(v; r^{\mathtt{C}}) = \omega(F^{\mathtt{I}}(v); r^{\mathtt{C}}),$$

which is increasing by construction and continuous by Lemma 5. Let $-\xi(\cdot) = b^{-1}(\cdot) + c(b^{-1}(\cdot))$ and replace $w^{\mathtt{I}}$ with $\overline{w}^{\mathtt{I}}$ and $\mathbf{a} = b^{-1}(\mathbf{b})$ in $VS^{\mathtt{I}}$ to get

$$\overline{VS}^{\mathtt{I}}(\mathbf{b}; r^{\mathtt{C}}) = \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \left[\mathbf{b}(v)\, \overline{w}^{\mathtt{I}}(v; r^{\mathtt{C}}) + \xi(\mathbf{b}(v))\right] dF^{\mathtt{I}} + r^{\mathtt{C}} \int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} \mathbf{b}(v)\, g^{\mathtt{C}}(v)\, dv.$$

Step 2: Deriving a candidate solution that maximizes $\overline{VS}^{\mathtt{I}}$. On $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, define $\varphi(y; v; r^{\mathtt{C}}) = y\overline{w}^{\mathtt{I}}(v; r^{\mathtt{C}}) + \xi(y)$ and let

$$\overline{\mathbf{b}}^{\mathtt{I}}(v; r^{\mathtt{C}}) = \arg \max_{y \in [b(\underline{a}), b(\overline{a})]} \varphi(y; v; r^{\mathtt{C}}); \tag{12}$$

let $\overline{\mathbf{b}}^{\mathtt{I}}(v; r^{\mathtt{C}}) = b(\underline{a})$ on $(\overline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$. Then, $\overline{\mathbf{b}}^{\mathtt{I}}$ is the unique pointwise maximizer of $\overline{VS}^{\mathtt{I}}$. Although $\overline{\mathbf{b}}^{\mathtt{I}}(r^{\mathtt{C}})$ is increasing on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, it may not be increasing on $[\underline{v}, \overline{v}]$. The next lemma characterizes any increasing maximizer of $\overline{VS}^{\mathtt{I}}$.

Lemma 6 *If* $\overline{VS}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}; r^{\mathtt{C}}) = \max_{\mathbf{b}\in\mathcal{B}} \overline{VS}^{\mathtt{I}}(\mathbf{b}; r^{\mathtt{C}})$, *then* $\mathbf{b}^{\mathtt{I}}$ *must satisfy* $\mathbf{b}^{\mathtt{I}}(v; r^{\mathtt{C}}) = \overline{\mathbf{b}}^{\mathtt{I}}(v; r^{\mathtt{C}})$ *if* $\underline{v}^{\mathtt{I}} < v < v^{\mathtt{b}}$, *and* $\mathbf{b}^{\mathtt{I}}(v; r^{\mathtt{C}}) = y^{\mathtt{b}}(r^{\mathtt{C}})$ *if* $v^{\mathtt{b}} \le v < \overline{v}^{\mathtt{C}}$, *where* $v^{\mathtt{b}} \in [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$ *and* $y^{\mathtt{b}}(r^{\mathtt{C}}) \le \overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathtt{b}}; r^{\mathtt{C}})$. *If* $v^{\mathtt{b}} > \underline{v}^{\mathtt{I}}$, *then* $y^{\mathtt{b}}(r^{\mathtt{C}}) = \overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathtt{b}}; r^{\mathtt{C}})$.

Proof. Drop $r^{\mathtt{C}}$ and suppose $\mathbf{b}^{\mathtt{I}} \in \mathcal{B}$ maximizes $\overline{VS}^{\mathtt{I}}$. First, $\mathbf{b}^{\mathtt{I}}(v) = \mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ on $(\overline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$. Otherwise, there is $v' \in (\overline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$ such that $\mathbf{b}^{\mathtt{I}}(v) > \mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ for $v > v'$. But then $\mathbf{b}^{\mathtt{I}}$ cannot be optimal in $\mathcal{B}$, as

$$\int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} \left[\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) - \mathbf{b}^{\mathtt{I}}(v)\right] g^{\mathtt{C}}(v)\, dv \ge \int_{v'}^{\overline{v}^{\mathtt{C}}} \left[\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) - \mathbf{b}^{\mathtt{I}}(v)\right] g^{\mathtt{C}}(v)\, dv > 0.$$

Consider $\mathbf{b}^{\mathtt{I}}(v)$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. Recall that $\varphi(y, v)$ in (12) is strictly concave in y and continuous in v. Since $\overline{\mathbf{b}}^{\mathtt{I}}(v)$ is continuous and increasing on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, only two cases can arise.

Case 1: $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) \le \overline{\mathbf{b}}^{\mathtt{I}}(\underline{v}^{\mathtt{I}})$. Then $\mathbf{b}^{\mathtt{I}}(v) = \mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. If not, there is $v' > \underline{v}^{\mathtt{I}}$ such that $\mathbf{b}^{\mathtt{I}}(v) < \mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) \le \overline{\mathbf{b}}^{\mathtt{I}}(v)$ for $v \le v'$. By strict concavity, for $v \in (\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, $\varphi(\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}), v) \ge \varphi(\mathbf{b}^{\mathtt{I}}(v), v)$, with strict inequality for $v \le v'$; so $\int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \varphi(\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}), v) dF^{\mathtt{I}} > \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \varphi(\mathbf{b}^{\mathtt{I}}(v), v) dF^{\mathtt{I}}$, contradicting the optimality of $\mathbf{b}^{\mathtt{I}}$.

Case 2: $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) = \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right) > \overline{\mathbf{b}}^{\mathtt{I}}(\underline{v}^{\mathtt{I}})$ for some $v^{\mathrm{b}} \in (\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. So $\mathbf{b}^{\mathtt{I}}(v) = \min\{\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right), \overline{\mathbf{b}}^{\mathtt{I}}(v)\}$ on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. Suppose not. First, consider $(v^{\mathrm{b}}, \overline{v}^{\mathtt{I}}]$ and suppose $\mathbf{b}^{\mathtt{I}}(v) < \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$ for some $v > v^{\mathrm{b}}$. Then, by the argument used in case 1, setting $\mathbf{b}^{\mathtt{I}}(v) = \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$ on $(v^{\mathrm{b}}, \overline{v}^{\mathtt{I}}]$ strictly improves on $\mathbf{b}^{\mathtt{I}}$: The resulting function is in $\mathcal{B}$ and $\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \varphi(\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right), v) dF^{\mathtt{I}} > \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \varphi\left(\mathbf{b}^{\mathtt{I}}(v), v\right) dF^{\mathtt{I}}$. Second, consider $(\underline{v}^{\mathtt{I}}, v^{\mathrm{b}}]$ and suppose $\mathbf{b}^{\mathtt{I}}(v') \neq \overline{\mathbf{b}}^{\mathtt{I}}(v')$ for some v'. If $\mathbf{b}^{\mathtt{I}}(v') > \overline{\mathbf{b}}^{\mathtt{I}}(v')$, then by continuity of $\overline{\mathbf{b}}^{\mathtt{I}}$ and monotonicity of $\mathbf{b}^{\mathtt{I}}$, there is a $v'' > v'$ such that $\mathbf{b}^{\mathtt{I}}(v) > \overline{\mathbf{b}}^{\mathtt{I}}(v)$ on (v', v''). Similarly, if $\mathbf{b}^{\mathtt{I}}(v') < \overline{\mathbf{b}}^{\mathtt{I}}(v')$, then there is $v''' < v'$ such that $\mathbf{b}^{\mathtt{I}}(v) < \overline{\mathbf{b}}^{\mathtt{I}}(v)$ on (v''', v'). In either case, since $\overline{\mathbf{b}}^{\mathtt{I}}$ is the unique maximizer of $\varphi(y, v)$, for $v \in (\underline{v}^{\mathtt{I}}, v^{\mathrm{b}}]$, $\varphi(\overline{\mathbf{b}}^{\mathtt{I}}(v), v) \geq \varphi(\mathbf{b}^{\mathtt{I}}(v), v)$, with strict inequality for $v \in (v''', v')$ or $v \in (v', v'')$; so $\int_{\underline{v}^{\mathtt{I}}}^{v^{\mathrm{b}}} \varphi(\overline{\mathbf{b}}^{\mathtt{I}}(v), v) dF^{\mathtt{I}} > \int_{\underline{v}^{\mathtt{I}}}^{v^{\mathrm{b}}} \varphi(\mathbf{b}^{\mathtt{I}}(v), v) dF^{\mathtt{I}}$, contradicting the optimality of $\mathbf{b}^{\mathtt{I}}$.

It remains to show that $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) > \overline{\mathbf{b}}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ is impossible. Suppose not. By the argument used in case 2, $\mathbf{b}^{\mathtt{I}}(v) = \overline{\mathbf{b}}^{\mathtt{I}}(v)$ on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}})$. Then, setting $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}}) > \overline{\mathbf{b}}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ cannot be optimal: Since $\mathbf{b}^{\mathtt{I}}(v) = \mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ on $(\overline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$, and $g^{\mathtt{C}}(v)$ is negative, reducing $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ to $\overline{\mathbf{b}}^{\mathtt{I}}(\overline{v}^{\mathtt{I}})$ satisfies monotonicity and strictly improves $\overline{VS}^{\mathtt{I}}$. ■

By Lemma 6, $\mathbf{b}^{\mathtt{I}}$ is continuous on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$. Lemma 6 doesn't pin down $\mathbf{b}^{\mathtt{I}}$ at $\underline{v}^{\mathtt{I}}$ and $\overline{v}^{\mathtt{C}}$, but it is w.l.o.g. to extend $\mathbf{b}^{\mathtt{I}}$ at $\underline{v}^{\mathtt{I}}$ and $\overline{v}^{\mathtt{C}}$ by continuity. The next lemma proves that a maximizer of $\overline{VS}^{\mathtt{I}}$ exists, and shows that it is unique on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$, and so on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$ w.l.o.g..

Lemma 7 *There is* $\mathbf{b}^{\mathtt{I}}$ *such that* $\overline{VS}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}; r^{\mathtt{C}}) = \max_{\mathbf{b} \in \mathcal{B}} \overline{VS}^{\mathtt{I}}(\mathbf{b}; r^{\mathtt{C}})$*; such a* $\mathbf{b}^{\mathtt{I}}$ *is unique.*

Proof. Drop $r^{\mathtt{C}}$. By Lemma 6, if a solution $\mathbf{b}^{\mathtt{I}}$ exists, then either (1) $\mathbf{b}^{\mathtt{I}}$ is constant at $y \leq \overline{\mathbf{b}}^{\mathtt{I}}(\underline{v}^{\mathtt{I}})$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$, or (2) $\mathbf{b}^{\mathtt{I}}$ is constant at $\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$ on $[v^{\mathrm{b}}, \overline{v}^{\mathtt{C}}]$, with $v^{\mathrm{b}} \leq \overline{v}^{\mathtt{I}}$, and equals $\overline{\mathbf{b}}^{\mathtt{I}}(v)$ for $v \leq v^{\mathrm{b}}$.

Case (1): In this case, $\overline{VS}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}) = VS^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}) = \widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}})$. The first equality follows because $\int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} w^{\mathtt{I}}(v) dF^{\mathtt{I}} = \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \overline{w}^{\mathtt{I}}(v) dF^{\mathtt{I}}$ since $\int_0^1 z(x) dx = \int_0^1 \omega(x) dx$; the second follows from Proposition 1. Moreover,

$$\widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}) = y \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} (v / t^{\mathtt{I}}) dF^{\mathtt{I}} + \xi(y). \tag{13}$$

By continuity and strictly concavity, there is a unique constant maximizer of $\overline{VS}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}})$. Call it $\mathbf{b}_1^{\mathtt{I}}$.

Case (2): Using $\varphi(y, v)$ in (12), $\overline{VS}^{\mathtt{I}}$ equals

$$\Upsilon\left(v^{\mathrm{b}}\right) = \int_{\underline{v}^{\mathtt{I}}}^{v^{\mathrm{b}}} \varphi(\overline{\mathbf{b}}^{\mathtt{I}}(v), v) dF^{\mathtt{I}} + \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \varphi(\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right), v) dF^{\mathtt{I}} + \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right) K, \tag{14}$$

where $K = r^{\mathtt{C}} \int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} g^{\mathtt{C}}(v)\, dv$. By continuity of $\overline{\mathbf{b}}^{\mathtt{I}}$, $\Upsilon\left(v^{\mathrm{b}}\right)$ is continuous. So there is $v^{\mathrm{b}} \in [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$ that fully identifies a maximizer for case (2). Since $\overline{\mathbf{b}}^{\mathtt{I}}$ can be locally flat, v^{b} need not be unique. However, there cannot be two optimal v_1^{b} and v_2^{b} such that $\overline{\mathbf{b}}^{\mathtt{I}}(v_1^{\mathrm{b}}) \neq \overline{\mathbf{b}}^{\mathtt{I}}(v_2^{\mathrm{b}})$. Suppose to the contrary that $v_1^{\mathrm{b}} < v_2^{\mathrm{b}}$ both maximize $\Upsilon\left(v^{\mathrm{b}}\right)$, and $\overline{\mathbf{b}}^{\mathtt{I}}(v_1^{\mathrm{b}}) < \overline{\mathbf{b}}^{\mathtt{I}}(v_2^{\mathrm{b}})$. W.l.o.g., let v_1^{b} be the largest v such that $\overline{\mathbf{b}}^{\mathtt{I}}(v) = \overline{\mathbf{b}}^{\mathtt{I}}(v_1^{\mathrm{b}})$. Let $\mathbf{b}_1$ and $\mathbf{b}_2$ be the functions identified by v_1^{b} and v_2^{b}, and for $\alpha \in (0,1)$, let $\widetilde{\mathbf{b}} = \alpha \mathbf{b}_1 + (1-\alpha)\mathbf{b}_2 \in \mathcal{B}$. On $(v_1^{\mathrm{b}}, \overline{v}^{\mathtt{C}}]$, $\mathbf{b}_2(v) \neq \mathbf{b}_1(v)$, whereas on $[\underline{v}^{\mathtt{I}}, v_1^{\mathrm{b}}]$, $\mathbf{b}_2(v) = \mathbf{b}_1(v) = \overline{\mathbf{b}}^{\mathtt{I}}(v)$. By strict concavity of $\varphi(y, v)$,

$$\int_{\underline{v}^{\mathtt{I}}}^{v_1^{\mathrm{b}}} \varphi(\widetilde{\mathbf{b}}(v), v) dF^{\mathtt{I}} + \int_{v_1^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \varphi(\widetilde{\mathbf{b}}(v), v) dF^{\mathtt{I}} + \widetilde{\mathbf{b}}(v) K > \alpha \Upsilon(v_1^{\mathrm{b}}) + (1-\alpha)\Upsilon(v_2^{\mathrm{b}}).$$

Note that $\widetilde{\mathbf{b}}$ is constant on $[v_2^{\mathrm{b}}, \overline{v}^{\mathtt{C}}]$ at some $\overline{\mathbf{b}}^{\mathtt{I}}(\tilde{v}^{\mathrm{b}})$, with $\tilde{v}^{\mathrm{b}} \in (v_1^{\mathrm{b}}, v_2^{\mathrm{b}})$. So $\mathbf{b}^{\mathtt{I}}(v)$ equals $\min\{\overline{\mathbf{b}}^{\mathtt{I}}(\tilde{v}^{\mathrm{b}}), \overline{\mathbf{b}}^{\mathtt{I}}(v)\}$ satisfies case (2) and, by the argument used in Lemma 6 (Case 2),

$$\Upsilon(\tilde{v}^{\mathrm{b}}) \geq \int_{\underline{v}^{\mathtt{I}}}^{v_1^{\mathrm{b}}} \varphi(\widetilde{\mathbf{b}}(v), v) dF^{\mathtt{I}} + \int_{v_1^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \varphi(\widetilde{\mathbf{b}}(v), v) dF^{\mathtt{I}} + \widetilde{\mathbf{b}}(v) K > \Upsilon(v_1^{\mathrm{b}}).$$

The claim follows. So any maximizer of $\Upsilon\left(v^{\mathrm{b}}\right)$ identifies a unique $\mathbf{b}^{\mathtt{I}}$ for case (2). Call it $\mathbf{b}_2^{\mathtt{I}}$.

By an argument similar to that for the uniqueness of $\mathbf{b}_2^{\mathtt{I}}$, $\overline{VS}^{\mathtt{I}}(\mathbf{b}_2^{\mathtt{I}}) = \overline{VS}^{\mathtt{I}}(\mathbf{b}_1^{\mathtt{I}})$ if and only if $\mathbf{b}_2^{\mathtt{I}} = \mathbf{b}_1^{\mathtt{I}}$ (on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$). So the overall maximizer of $\overline{VS}^{\mathtt{I}}$ is unique; it equals $\mathbf{b}_1^{\mathtt{I}}$ if $\overline{VS}^{\mathtt{I}}(\mathbf{b}_1^{\mathtt{I}}) \geq \overline{VS}^{\mathtt{I}}(\mathbf{b}_2^{\mathtt{I}})$, and $\mathbf{b}_2^{\mathtt{I}}$ otherwise. ■

Step 3: The unique maximizer of $\overline{VS}^{\mathtt{I}}$, denoted $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$, is also the unique maximizer of $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}))$. The argument modifies Toikka's (2011) proof of Theorem 3.7 and Corollary 3.9 to account for $(\overline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$.

Lemma 8 *The function $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ is the unique maximizer of $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}))$.*

Proof. Drop $r^{\mathtt{C}}$. Since $\mathbf{b} \in \mathcal{B}$, integrating by parts, we get

$$\begin{aligned}
\int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \mathbf{b}(v)\left[w^{\mathtt{I}}(v) - \overline{w}^{\mathtt{I}}(v)\right] dF^{\mathtt{I}} &= \mathbf{b}(v)\left[Z(F^{\mathtt{I}}(v)) - \Omega(F^{\mathtt{I}}(v))\right]\Big|_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \\
&\quad - \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \left[Z(F^{\mathtt{I}}(v)) - \Omega(F^{\mathtt{I}}(v))\right] d\mathbf{b}(v) \\
&= \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \left[\Omega(F^{\mathtt{I}}(v)) - Z(F^{\mathtt{I}}(v))\right] d\mathbf{b}(v) \leq 0.
\end{aligned}$$

The last equality follows from $Z(0) = \Omega(0)$ and $Z(1) = \Omega(1)$; the inequality follows from $\mathbf{b} \in \mathcal{B}$ and $\Omega \leq Z$. Rewriting $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}))$, we get

$$\sup_{\mathbf{b} \in \mathcal{B}} VS^{\mathtt{I}}(b^{-1}(\mathbf{b})) = \sup_{\mathbf{b} \in \mathcal{B}} \{\overline{VS}^{\mathtt{I}}(\mathbf{b}) + \int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} \left[\Omega(F^{\mathtt{I}}(v)) - Z(F^{\mathtt{I}}(v))\right] d\mathbf{b}(v)\}.$$

Since $\mathbf{b}^{\mathtt{I}} \in \mathcal{B}$ and achieves the supremum of $\overline{VS}^{\mathtt{I}}(\mathbf{b})$, we have to show that

$$\int_{\underline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{I}}} [\Omega(F^{\mathtt{I}}(v)) - Z(F^{\mathtt{I}}(v))] d\mathbf{b}^{\mathtt{I}}(v) = 0. \tag{15}$$

If $\mathbf{b}^{\mathtt{I}}$ is constant on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$, then $d\mathbf{b}^{\mathtt{I}} \equiv 0$ and we are done. Otherwise, consider the pointwise solution $\overline{\mathbf{b}}^{\mathtt{I}}$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$ as defined in (12), and a v such that $\Omega(F^{\mathtt{I}}(v)) < Z(F^{\mathtt{I}}(v))$. For some open interval N around v, $\overline{w}^{\mathtt{I}}(\cdot) = \omega(F^{\mathtt{I}}(v))$, and $\overline{\mathbf{b}}^{\mathtt{I}}$ is constant on N. So, $d\overline{\mathbf{b}}^{\mathtt{I}}(\cdot)$ assigns zero measure to any such N, and satisfies (15). For any such N, $d\mathbf{b}^{\mathtt{I}}$ does the same. Consider $[v^{\mathtt{b}}, \overline{v}^{\mathtt{C}}]$, on which $\mathbf{b}^{\mathtt{I}}$ is constant. If $N \subset [v^{\mathtt{b}}, \overline{v}^{\mathtt{C}}]$, the claim is immediate. The same holds if $N \cap [v^{\mathtt{b}}, \overline{v}^{\mathtt{C}}] = \varnothing$, because then $\mathbf{b}^{\mathtt{I}}(v) = \overline{\mathbf{b}}^{\mathtt{I}}(v)$ for $v \in N$. Finally, if both $N \cap [v^{\mathtt{b}}, \overline{v}^{\mathtt{C}}] \neq \varnothing$ and $N \cap [\underline{v}^{\mathtt{I}}, v^{\mathtt{b}}) \neq \varnothing$ (so $v^{\mathtt{b}} > \underline{v}^{\mathtt{I}}$), then $\mathbf{b}^{\mathtt{I}}$ is constant on $[v^{\mathtt{b}}, \overline{v}^{\mathtt{C}}] \cup N$, which implies the claim. So (15) holds also for a nonconstant $\mathbf{b}^{\mathtt{I}}$.

By Lemma 7, if $\widetilde{\mathbf{b}} \in \mathcal{B}$ differs from $\mathbf{b}^{\mathtt{I}}$ on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$, then $\overline{VS}^{\mathtt{I}}(\widetilde{\mathbf{b}}) < \overline{VS}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}})$. Uniqueness follows on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$; extending it to $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$ is w.l.o.g.. ■

Part 2: Continuity and Limit Behavior of $\mathbf{b}^{\mathtt{I}}$

Continuity in v follows from Part 1; consider continuity in $r^{\mathtt{C}}$. By the definition of $\overline{\mathbf{b}}^{\mathtt{I}}$ in (12) and the Maximum Theorem, $\overline{\mathbf{b}}^{\mathtt{I}}(v, \cdot)$ is continuous in $r^{\mathtt{C}}$ for $v \in [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. Now consider $\Upsilon(v^{\mathtt{b}}; r^{\mathtt{C}})$ in (14). By pointwise continuity of $\overline{w}^{\mathtt{I}}(v; r^{\mathtt{C}})$ and $\overline{\mathbf{b}}^{\mathtt{I}}(v; r^{\mathtt{C}})$, $\Upsilon(v^{\mathtt{b}}; r^{\mathtt{C}})$ is continuous in $r^{\mathtt{C}}$ and so $V^{\mathtt{b}}(r^{\mathtt{C}}) = \arg\max_{v \in [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]} \Upsilon(v; r^{\mathtt{C}})$ is u.h.c.. For $v, v' \in V^{\mathtt{b}}(r^{\mathtt{C}})$, $\overline{\mathbf{b}}^{\mathtt{I}}(v; r^{\mathtt{C}}) = \overline{\mathbf{b}}^{\mathtt{I}}(v'; r^{\mathtt{C}})$. Take any sequence $\{r_n^{\mathtt{C}}\}$ with $\lim_{n\to\infty} r_n^{\mathtt{C}} = r^{\mathtt{C}}$. Then, $\lim_{n\to\infty} v^{\mathtt{b}}(r_n^{\mathtt{C}}) = v^{\mathtt{b}} \in V^{\mathtt{b}}(r^{\mathtt{C}})$. The candidate $\mathbf{b}_2^{\mathtt{I}}(r^{\mathtt{C}})$ that maximizes $\Upsilon(v^{\mathtt{b}}; r^{\mathtt{C}})$ is such that $\mathbf{b}_2^{\mathtt{I}}(v; r^{\mathtt{C}})$ equals $\min\{\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathtt{b}}(r^{\mathtt{C}}); r^{\mathtt{C}}), \overline{\mathbf{b}}^{\mathtt{I}}(v; r^{\mathtt{C}})\}$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, and $\mathbf{b}_2^{\mathtt{I}}(v; r^{\mathtt{C}}) = \overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathtt{b}}(r^{\mathtt{C}}); r^{\mathtt{C}})$ for $v > \overline{v}^{\mathtt{I}}$. So, by continuity of $\overline{\mathbf{b}}^{\mathtt{I}}$, $\lim_{n\to\infty} \mathbf{b}_2^{\mathtt{I}}(v; r_n^{\mathtt{C}}) = \mathbf{b}_2^{\mathtt{I}}(v; r^{\mathtt{C}})$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$. Finally, the constant function $\mathbf{b}_1^{\mathtt{I}}$ in the proof of Lemma 7, as well as (13), is independent of $r^{\mathtt{C}}$. It remains to show that $\mathbf{b}^{\mathtt{I}}(r_n^{\mathtt{C}})$ converges pointwise to $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$. First, if $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_1^{\mathtt{I}})) > VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_2^{\mathtt{I}}(r^{\mathtt{C}}))) = \Upsilon(v^{\mathtt{b}}(r^{\mathtt{C}}); r^{\mathtt{C}})$, then by continuity of Υ, there is N such that $n \geq N$ implies $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_1^{\mathtt{I}})) > VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_2^{\mathtt{I}}(r_n^{\mathtt{C}})))$. So, for $n \geq N$, $\mathbf{b}^{\mathtt{I}}(v; r_n^{\mathtt{C}}) = \mathbf{b}_1^{\mathtt{I}}$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$. Second, if $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_1^{\mathtt{I}})) < VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_2^{\mathtt{I}}(r^{\mathtt{C}})))$, then again for n large $\mathbf{b}^{\mathtt{I}}(r_n^{\mathtt{C}}) = \mathbf{b}_2^{\mathtt{I}}(r_n^{\mathtt{C}})$, which converges pointwise to $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$. Finally, if $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_1^{\mathtt{I}})) = VS^{\mathtt{I}}(b^{-1}(\mathbf{b}_2^{\mathtt{I}}(r^{\mathtt{C}})))$, then $\mathbf{b}_1^{\mathtt{I}} \equiv \mathbf{b}_2^{\mathtt{I}}(r^{\mathtt{C}})$. So $|\mathbf{b}_1^{\mathtt{I}} - \mathbf{b}^{\mathtt{I}}(v; r_n^{\mathtt{C}})| \leq \max\{0, |\mathbf{b}_1^{\mathtt{I}} - \mathbf{b}_2^{\mathtt{I}}(v; r_n^{\mathtt{C}})|\} \to 0$ as $n \to \infty$.

To prove that $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}}) \to \mathbf{b}^{\mathtt{I}*} = b(\mathbf{a}^{\mathtt{I}*})$ pointwise as $r^{\mathtt{C}} \to 0$, note that $VS^{\mathtt{I}}(b^{-1}(\mathbf{b}), 0) = \widetilde{W}^{\mathtt{I}}(\mathbf{b})$. So $\mathbf{b}^{\mathtt{I}}(v, 0) = \mathbf{b}^{\mathtt{I}*}(v)$ on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}})$, which can be extended to $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}}]$ by letting $\mathbf{b}^{\mathtt{I}}(\underline{v}^{\mathtt{I}}, 0) = \mathbf{b}^{\mathtt{I}*}(\underline{v}^{\mathtt{I}})$ and $\mathbf{b}^{\mathtt{I}}(v, 0) = \mathbf{b}^{\mathtt{I}*}(\overline{v}^{\mathtt{I}})$ for $v \geq \underline{v}^{\mathtt{I}}$. To prove $\max_{[\underline{v}, \overline{v}]} \left|\mathbf{b}^{\mathtt{I}}(v; r^{\mathtt{C}}) - b(a^{\mathtt{nf}})\right| \to 0$ as $r^{\mathtt{C}} \to +\infty$, first recall that $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ maximizes

$VS^{\mathtt{I}}(b^{-1}(\mathbf{b}); r^{\mathtt{C}}) = \widetilde{W}^{\mathtt{I}}(\mathbf{b}) - r^{\mathtt{C}}\widetilde{R}^{\mathtt{C}}(\mathbf{b})$ and that, by Proposition 1, $\widetilde{R}^{\mathtt{C}}(\mathbf{b}) > 0$ for any $\mathbf{b} \in \mathcal{B}$ that is not constant on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$. Clearly, $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ cannot converge to a constant function with value $y_0 \neq b(a^{\mathrm{nf}})$, for $b(a^{\mathrm{nf}})$ is the unique maximizer of (13). Now, suppose $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ converges pointwise to a function, denoted $\mathbf{b}^{\mathtt{I}}_{\infty}$, that is not constant on $(\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{C}})$. Then, there is $\hat{r}^{\mathtt{C}}$ such that, for $r^{\mathtt{C}} > \hat{r}^{\mathtt{C}}$, $b(a^{\mathrm{nf}})$ strictly dominates $\mathbf{b}^{\mathtt{I}}_{\infty}$: Since $\widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}_{\infty})$ is bounded and $\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}_{\infty}) > 0$, $\widetilde{W}^{\mathtt{I}}(\mathbf{b}^{\mathtt{I}}_{\infty}) - \hat{r}^{\mathtt{C}}\widetilde{R}^{\mathtt{C}}(\mathbf{b}^{\mathtt{I}}_{\infty}) \leq \widetilde{W}^{\mathtt{I}}(b(a^{\mathrm{nf}}))$ for some $\hat{r}^{\mathtt{C}} \geq 0$. Finally, consider the unique extension of $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ by continuity. By monotonicity,

$$\max_{[\underline{v},\overline{v}]} \left|\mathbf{b}^{\mathtt{I}}(v; r^{\mathtt{C}}) - b(a^{\mathrm{nf}})\right| = \max\{\left|\mathbf{b}^{\mathtt{I}}(\underline{v}; r^{\mathtt{C}}) - b(a^{\mathrm{nf}})\right|, \left|\mathbf{b}^{\mathtt{I}}(\overline{v}; r^{\mathtt{C}}) - b(a^{\mathrm{nf}})\right|\}.$$

Part 3: Properties (a)-(c) of $\mathbf{b}^{\mathtt{I}}$

Property (b): Drop $r^{\mathtt{C}}$. Recall that (1) $\mathbf{b}^{\mathtt{I}}$ satisfies Lemma 6, (2) on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, $\overline{\mathbf{b}}^{\mathtt{I}}$ is defined by (12) and is continuous. Suppose $v^{\mathrm{b}} > \underline{v}^{\mathtt{I}}$. For $v < v^{\mathrm{b}}$, $\mathbf{b}^{\mathtt{I}}(v) = \overline{\mathbf{b}}^{\mathtt{I}}(v) < \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right) = \mathbf{b}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$ by Lemma 6, and $\Upsilon\left(v^{\mathrm{b}}\right) \geq \Upsilon(v)$ by construction (see (14)). So

$$\begin{aligned}
\frac{\Upsilon\left(v^{\mathrm{b}}\right) - \Upsilon(v)}{\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}}) - \overline{\mathbf{b}}^{\mathtt{I}}(v)} &= K + \int_v^{\overline{v}^{\mathtt{I}}} \overline{w}^{\mathtt{I}}(y)\, dF^{\mathtt{I}} - \int_v^{v^{\mathrm{b}}} \overline{w}^{\mathtt{I}}(y)\, dF^{\mathtt{I}} \\
&\quad + \frac{\xi(\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}})) - \xi(\overline{\mathbf{b}}^{\mathtt{I}}(v))}{\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}}) - \overline{\mathbf{b}}^{\mathtt{I}}(v)}(1 - F^{\mathtt{I}}(v)) \\
&\quad - \int_v^{v^{\mathrm{b}}} \frac{\xi(\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}})) - \xi(\overline{\mathbf{b}}^{\mathtt{I}}(y))}{\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}}) - \overline{\mathbf{b}}^{\mathtt{I}}(v)} dF^{\mathtt{I}} \geq 0.
\end{aligned}$$

Since $\overline{\mathbf{b}}^{\mathtt{I}}$ and $\overline{w}^{\mathtt{I}}$ are increasing,

$$\left|\int_v^{v^{\mathrm{b}}} \overline{w}^{\mathtt{I}}(y)\, dF^{\mathtt{I}}\right| \leq \max\{|\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}})|, |\overline{w}^{\mathtt{I}}(v^{\mathrm{b}})|\}(F^{\mathtt{I}}(v^{\mathrm{b}}) - F^{\mathtt{I}}(v)).$$

Since $\overline{\mathbf{b}}^{\mathtt{I}}$ is continuous, using the Mean Value Theorem, we get that for v close to v^{b}

$$\left|\int_v^{v^{\mathrm{b}}} \frac{\xi(\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)) - \xi(\overline{\mathbf{b}}^{\mathtt{I}}(y))}{\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}}) - \overline{\mathbf{b}}^{\mathtt{I}}(v)} dF^{\mathtt{I}}\right| \leq \max\{|\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}})|, |\overline{w}^{\mathtt{I}}(v^{\mathrm{b}})|\}(F^{\mathtt{I}}(v^{\mathrm{b}}) - F^{\mathtt{I}}(v)).$$

Therefore

$$\lim_{v \uparrow v^{\mathrm{b}}} \frac{\Upsilon\left(v^{\mathrm{b}}\right) - \Upsilon(v)}{\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}}) - \overline{\mathbf{b}}^{\mathtt{I}}(v)} = \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \overline{w}^{\mathtt{I}}(y)\, dF^{\mathtt{I}} + K + \xi'(\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right))(1 - F^{\mathtt{I}}\left(v^{\mathrm{b}}\right)) \geq 0. \quad (16)$$

It follows that $v^{\mathrm{b}} < \overline{v}^{\mathtt{I}}$ because $K < 0$ and $\xi'(\cdot) < 0$.

I claim that there is $v \in (v^{\mathrm{b}}, \overline{v}^{\mathtt{I}}]$ such that $\overline{\mathbf{b}}^{\mathtt{I}}(v) > \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$. Suppose not. If $\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$ is interior, $\overline{w}^{\mathtt{I}}(v) = -\xi'(\overline{\mathbf{b}}^{\mathtt{I}}(v))$ for $v \geq v^{\mathrm{b}}$, and (16) is violated. If $\overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right) = b(\overline{a})$ and the set $V_{\overline{a}} = \{v \in [\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}] \mid \overline{w}^{\mathtt{I}}(v) > -\xi'(b(\overline{a}))\}$ is nonempty—if $V_{\overline{a}} = \varnothing$, we are back to the previous case—then (16) is violated again. To see this, note that since v^{b} is the smallest v for which $\overline{w}^{\mathtt{I}}(v) = -\xi'(b(\overline{a}))$, $Z(F^{\mathtt{I}}\left(v^{\mathrm{b}}\right)) = \Omega(F^{\mathtt{I}}\left(v^{\mathrm{b}}\right))$, which implies $\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} w^{\mathtt{I}}(y)\, dF^{\mathtt{I}} = \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} \overline{w}^{\mathtt{I}}(y)\, dF^{\mathtt{I}}$; so

$$\begin{aligned}\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} [\overline{w}^{\mathtt{I}}(y) + \xi'(b(\overline{a}))] dF^{\mathtt{I}} + K &= \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} (y/t^{\mathtt{I}} + \xi'(b(\overline{a}))) dF^{\mathtt{I}} \\ &\quad + r^{\mathtt{C}} \left[\int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} g^{\mathtt{C}}(y)\, dy + \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} G^{\mathtt{I}}(y)\, dF^{\mathtt{I}} \right].\end{aligned}$$

By Assumption 1 and concavity of ξ, $\overline{v}^{\mathtt{I}}/t^{\mathtt{I}} + \xi'(b(\overline{a})) < 0$, so the first integral is negative. The term in brackets is too, contradicting 16. To see this, integrate by parts:

$$\int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} g^{\mathtt{C}}(v)\, dv = -\int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}} \left(v/t^{\mathtt{C}}\right) dF^{\mathtt{C}} + \overline{v}^{\mathtt{I}}(1 - F^{\mathtt{C}}(\overline{v}^{\mathtt{I}})),$$

$$\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} q^{i}(v)\, dF^{i} = \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} (v/t^{i} - v^{\mathrm{b}}) dF^{i} - (\overline{v}^{\mathtt{I}} - v^{\mathrm{b}}) F^{i}(\overline{v}^{\mathtt{I}}).$$

Thus, the term in brackets equals

$$\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} (v/t^{\mathtt{I}} - v^{\mathrm{b}}) dF^{\mathtt{I}} - \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{C}}} (v/t^{\mathtt{C}} - v^{\mathrm{b}}) dF^{\mathtt{C}} = -\int_{v^{\mathrm{b}}/t^{\mathtt{C}}}^{v^{\mathrm{b}}/t^{\mathtt{I}}} (s - v^{\mathrm{b}}) dF < 0, \tag{17}$$

where the equality is a change of variables, and the inequality follows from $\overline{v}^{\mathtt{I}} > v^{\mathrm{b}} > \underline{v}^{\mathtt{I}} > 0$ and $0 < t^{\mathtt{I}} < t^{\mathtt{C}} \leq 1$.

Define $v^1 = \max\{v \mid \overline{\mathbf{b}}^{\mathtt{I}}(v) = \overline{\mathbf{b}}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)\} < \overline{v}^{\mathtt{I}}$. For $v > v^1$, $\overline{\mathbf{b}}^{\mathtt{I}}(v) > \overline{\mathbf{b}}^{\mathtt{I}}(v^1)$ and $\Upsilon(v) \leq \Upsilon(v^1) = \Upsilon\left(v^{\mathrm{b}}\right)$. By the same steps that give (16),

$$\lim_{v \downarrow v^1} \frac{\Upsilon(v^1) - \Upsilon(v)}{\overline{\mathbf{b}}^{\mathtt{I}}(v^1) - \overline{\mathbf{b}}^{\mathtt{I}}(v)} = \int_{v^1}^{\overline{v}^{\mathtt{I}}} \overline{w}^{\mathtt{I}}(y)\, dF^{\mathtt{I}} + K + \xi'(\overline{\mathbf{b}}^{\mathtt{I}}(v^1))(1 - F^{\mathtt{I}}(v^1)) \leq 0.$$

As $\overline{\mathbf{b}}^{\mathtt{I}}(y)$ is interior and constant on $[v^{\mathrm{b}}, v^1]$, $\xi'(\overline{\mathbf{b}}^{\mathtt{I}}(y)) = -\overline{w}^{\mathtt{I}}(y) = -\overline{w}^{\mathtt{I}}(v^{\mathrm{b}}) = \xi'(\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}}))$, and

$$0 \geq \int_{v^1}^{\overline{v}^{\mathtt{I}}} [\overline{w}^{\mathtt{I}}(y) - \overline{w}^{\mathtt{I}}(v^1)] dF^{\mathtt{I}} + K = \int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}} [\overline{w}^{\mathtt{I}}(y) - \overline{w}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)] dF^{\mathtt{I}} + K \geq 0.$$

Finally, since $Z(F^{\mathtt{I}}\left(v^{\mathrm{b}}\right)) = \Omega(F^{\mathtt{I}}\left(v^{\mathrm{b}}\right))$, the argument used in Lemma 5 yields

$w^{\mathtt{I}}\left(v^{\mathrm{b}}\right)=\overline{w}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)$. So

$$\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}}[\overline{w}^{\mathtt{I}}\left(y\right)-\overline{w}^{\mathtt{I}}\left(v^{\mathrm{b}}\right)]dF^{\mathtt{I}}+K=\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}}[w^{\mathtt{I}}\left(y\right)-w^{\mathtt{I}}\left(v^{\mathrm{b}}\right)]dF^{\mathtt{I}}+K,$$

which gives, by rearranging $w^{\mathtt{I}}\left(y\right)$,

$$\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}}[w^{\mathtt{I}}(v^{\mathrm{b}};r^{\mathtt{C}})-v/t^{\mathtt{I}}]dF^{\mathtt{I}}\left(v\right)=r^{\mathtt{C}}\left[\int_{v^{\mathrm{b}}}^{\overline{v}^{\mathtt{I}}}G^{\mathtt{I}}\left(v\right)dF^{\mathtt{I}}-\int_{\overline{v}^{\mathtt{I}}}^{\overline{v}^{\mathtt{C}}}g^{\mathtt{C}}\left(v\right)dv\right]. \tag{18}$$

Property (a): Since $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ and $\mathbf{b}^{\mathtt{I}*}=b(\mathbf{a}^{\mathtt{I}*})$ are continuous and increasing, it is enough to prove that $\mathbf{b}^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})>\mathbf{b}^{\mathtt{I}*}(\underline{v}^{\mathtt{I}})$ and $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}};r^{\mathtt{C}})<\mathbf{b}^{\mathtt{I}*}(\overline{v}^{\mathtt{I}})$.

Case 1: $\mathbf{b}^{\mathtt{I}}$ not constant. This implies that $v^{\mathrm{b}}>\underline{v}^{\mathtt{I}}$, and by Lemma 6 $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}};r^{\mathtt{C}})=\mathbf{b}^{\mathtt{I}}(v^{\mathrm{b}};r^{\mathtt{C}})=\overline{\mathbf{b}}^{\mathtt{I}}(v^{\mathrm{b}};r^{\mathtt{C}})$. To show that $\mathbf{b}^{\mathtt{I}}(\overline{v}^{\mathtt{I}};r^{\mathtt{C}})<\mathbf{b}^{\mathtt{I}*}(\overline{v}^{\mathtt{I}})$ for $r^{\mathtt{C}}>0$, it is enough to prove that $\overline{w}^{\mathtt{I}}(v^{\mathrm{b}};r^{\mathtt{C}})<\overline{v}^{\mathtt{I}}/t^{\mathtt{I}}$. This inequality follows from (17) and (18), because $\overline{w}^{\mathtt{I}}(v^{\mathrm{b}};r^{\mathtt{C}})=w^{\mathtt{I}}(v^{\mathrm{b}};r^{\mathtt{C}})$. To show that $\mathbf{b}^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})>\mathbf{b}^{\mathtt{I}*}(\underline{v}^{\mathtt{I}})$, given $r^{\mathtt{C}}>0$, let $v_{\mathrm{b}}=\max\{v\mid\overline{w}^{\mathtt{I}}(v;r^{\mathtt{C}})=\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})\}$.

Lemma 9 *$\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})\leq w^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})$. If the inequality is strict, then $v_{\mathrm{b}}>\underline{v}^{\mathtt{I}}$.*

Proof. Drop $r^{\mathtt{C}}$ and recall that $\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}})=\omega\left(0\right)$ and $w^{\mathtt{I}}(\underline{v}^{\mathtt{I}})=z\left(0\right)$. If $\omega\left(0\right)>z\left(0\right)$, the argument used in Lemma 10 leads to a contradiction. Suppose $\omega\left(0\right)<z\left(0\right)$ and let $\hat{x}=\sup\{x\mid\forall x'<x,\omega(x')<z(x')\}$; by continuity, $\hat{x}>0$. Then, for $0<x<\hat{x}$,

$$Z\left(x\right)=Z\left(0\right)+\int_{0}^{x}z\left(y\right)dy>\Omega\left(0\right)+\int_{0}^{x}\omega\left(y\right)dy=\Omega\left(x\right).$$

It follows that $v_{\mathrm{b}}\geq\left(F^{\mathtt{I}}\right)^{-1}\left(\hat{x}\right)>\underline{v}^{\mathtt{I}}$. ■

So, if $\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})=w^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})$, then it equals $(\underline{v}^{\mathtt{I}}/t^{\mathtt{I}})(1+r^{\mathtt{C}}(1-t^{\mathtt{I}}))+r^{\mathtt{C}}(f^{\mathtt{I}}(\underline{v}^{\mathtt{I}}))^{-1}>\underline{v}^{\mathtt{I}}/t^{\mathtt{I}}$. If instead $\overline{w}^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})<w^{\mathtt{I}}(\underline{v}^{\mathtt{I}};r^{\mathtt{C}})$, then it is constant on $[\underline{v}^{\mathtt{I}},v_{\mathrm{b}}]$ at $\overline{w}^{\mathtt{I}}(v_{\mathrm{b}};r^{\mathtt{C}})$. Since $Z(F^{\mathtt{I}}(v_{\mathrm{b}});r^{\mathtt{C}})=\Omega(F^{\mathtt{I}}(v_{\mathrm{b}});r^{\mathtt{C}})$, v_{b} must satisfy

$$\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathrm{b}}}[w^{\mathtt{I}}(y;r^{\mathtt{C}})-\overline{w}^{\mathtt{I}}(v_{\mathrm{b}};r^{\mathtt{C}})]dF^{\mathtt{I}}=0 \tag{19}$$

or equivalently,

$$\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathrm{b}}}(y/t^{\mathtt{I}}-\overline{w}^{\mathtt{I}}(v_{\mathrm{b}};r^{\mathtt{C}}))dF^{\mathtt{I}}=-r^{\mathtt{C}}\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathrm{b}}}G^{\mathtt{I}}\left(y\right)dF^{\mathtt{I}}. \tag{20}$$

Integrating by parts,

$$\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathrm{b}}}G^{\mathtt{I}}\left(y\right)dF^{\mathtt{I}}=\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathrm{b}}}(y/t^{\mathtt{I}}-v_{\mathrm{b}})dF^{\mathtt{I}}-\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathrm{b}}}(y/t^{\mathtt{C}}-v_{\mathrm{b}})dF^{\mathtt{C}}=\int_{v_{\mathrm{b}}/t^{\mathtt{C}}}^{v_{\mathrm{b}}/t^{\mathtt{I}}}(s-v_{\mathrm{b}})dF>0,$$

where the last equality follows from a change of variables and the inequality from $v_{\mathrm{b}} > \underline{v}^{\mathtt{I}} > 0$ and $0 < t^{\mathtt{I}} < t^{\mathtt{C}} \leq 1$. So, by (20), $\overline{w}^{\mathtt{I}}(v_{\mathrm{b}}; r^{\mathtt{C}}) > \underline{v}^{\mathtt{I}}/t^{\mathtt{I}}$. In either case, $\mathbf{b}^{\mathtt{I}}(\underline{v}^{\mathtt{I}}; r^{\mathtt{C}})$ must be interior and strictly greater than $\mathbf{b}^{\mathtt{I}*}(\underline{v}^{\mathtt{I}})$.

Case 2: $\mathbf{b}^{\mathtt{I}}$ constant. From the proof of Lemma 7, $\mathbf{b}^{\mathtt{I}}(v; r^{\mathtt{C}})$ equals $b(a^{\mathrm{nf}})$ on $[\underline{v}, \overline{v}]$. Since $\underline{v}^{\mathtt{I}}/t^{\mathtt{I}} < \mathbb{E}(s) < \overline{v}^{\mathtt{I}}/t^{\mathtt{I}}$, Assumption 1 implies $\mathbf{b}^{\mathtt{I}*}(\underline{v}^{\mathtt{I}}) < b(a^{\mathrm{nf}}) < \mathbf{b}^{\mathtt{I}*}(\overline{v}^{\mathtt{I}})$.

Property (c): Let $\underline{v}^{\mathtt{I}} < v' \leq v^m$ (recall $v^m = \min\{\overline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}\}$) and consider

$$w^{\mathtt{I}}(v'; r^{\mathtt{C}}) - w^{\mathtt{I}}(\underline{v}^{\mathtt{I}}; r^{\mathtt{C}}) = \frac{v' - \underline{v}^{\mathtt{I}}}{t^{\mathtt{I}}}(1 + r^{\mathtt{C}}(1 - t^{\mathtt{I}})) - r^{\mathtt{C}}\left[\frac{F^{\mathtt{I}}(v')}{f^{\mathtt{I}}(v')} - \frac{F^{\mathtt{I}}(\underline{v}^{\mathtt{I}})}{f^{\mathtt{I}}(\underline{v}^{\mathtt{I}})}\right]. \quad (21)$$

The first part of (21) is positive since $0 < t^{\mathtt{I}} < 1$, but the second part can be negative. So $w(\cdot; r^{\mathtt{C}})$ can be decreasing in a neighborhood of $\underline{v}^{\mathtt{I}}$. If so, $\overline{w}^{\mathtt{I}}(v; r^{\mathtt{C}})$ and $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ are constant on $[\underline{v}^{\mathtt{I}}, v_{\mathrm{b}}] \neq \varnothing$.

Finally, note that $\frac{F^{\mathtt{I}}(v')/f^{\mathtt{I}}(v') - F^{\mathtt{I}}(v)/f^{\mathtt{I}}(v)}{v' - v} = \frac{F(v'/t^{\mathtt{I}})/f(v'/t^{\mathtt{I}}) - F\left(v/t^{\mathtt{I}}\right)/f\left(v/t^{\mathtt{I}}\right)}{v'/t^{\mathtt{I}} - v/t^{\mathtt{I}}}$. So, the condition in part (c) of Proposition 2 implies that, for $v' > v$ in $[\underline{v}^{\mathtt{I}}, \min\{t^{\mathtt{I}} s^{\dagger}, v^m\}]$,

$$\begin{aligned}\frac{w^{\mathtt{I}}(v'; r^{\mathtt{C}}) - w^{\mathtt{I}}(v; r^{\mathtt{C}})}{r^{\mathtt{C}}(v' - v)} &= \frac{1}{r^{\mathtt{C}} t^{\mathtt{I}}} + \frac{1 - t^{\mathtt{I}}}{t^{\mathtt{I}}} \\ &\quad + \frac{F(v'/t^{\mathtt{I}})/f(v'/t^{\mathtt{I}}) - F\left(v/t^{\mathtt{I}}\right)/f\left(v/t^{\mathtt{I}}\right)}{v'/t^{\mathtt{I}} - v/t^{\mathtt{I}}} \leq 0.\end{aligned}$$

So $\mathbf{b}^{\mathtt{I}}(r^{\mathtt{C}})$ must be constant on $[\underline{v}^{\mathtt{I}}, v_{\mathrm{b}}] \neq \varnothing$.

A.5 Proof of Corollary 2

Being increasing, $\mathbf{a}^{\mathtt{I}}$ is differentiable a. e. on $[\underline{v}, \overline{v}]$. If $\frac{d\mathbf{a}^{\mathtt{I}}}{dv} > 0$ at v, then using (E)

$$\frac{d\mathbf{p}^{\mathtt{I}}/dv}{d\mathbf{a}^{\mathtt{I}}/dv} = vb'(\mathbf{a}^{\mathtt{I}}(v)) - 1 \quad \text{and} \quad \frac{d\mathbf{p}^{\mathtt{I}*}/dv}{d\mathbf{a}^{\mathtt{I}*}/dv} = vb'(\mathbf{a}^{\mathtt{I}*}(v)) - 1.$$

The result follows from $b'' < 0$ and point (a) in Proposition 2.

A.6 Proof of Lemma 4

Using (4),

$$-R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}*}) = \int_{\underline{s}}^{\frac{t^{\mathtt{C}}}{t^{\mathtt{I}}}} \Delta(s|\,\mathbf{a}^{\mathtt{C}*}) dF + \int_{\underline{s}\frac{t^{\mathtt{C}}}{t^{\mathtt{I}}}}^{\overline{s}} \Delta(s|\,\mathbf{a}^{\mathtt{C}*}) dF.$$

For $s \leq \underline{s}\frac{t^{\mathtt{C}}}{t^{\mathtt{I}}}$, since $\mathbf{a}^{\mathtt{C}*}(st^{\mathtt{I}}) = \mathbf{a}^{\mathtt{C}*}(\underline{s}t^{\mathtt{C}})$,

$$\begin{aligned}\Delta(s|\,\mathbf{a}^{\mathtt{C}*}) &= \underline{s}t^{\mathtt{C}}b(\mathbf{a}^{\mathtt{C}*}(\underline{s}t^{\mathtt{C}})) - st^{\mathtt{C}}b(\mathbf{a}^{\mathtt{C}*}(st^{\mathtt{C}})) + \textstyle\int_{\underline{s}t^{\mathtt{C}}}^{st^{\mathtt{C}}} b(\mathbf{a}^{\mathtt{C}*}(y))dy \\ &\quad + s[b(\mathbf{a}^{\mathtt{C}*}(st^{\mathtt{C}})) - b(\mathbf{a}^{\mathtt{C}*}(\underline{s}t^{\mathtt{C}}))].\end{aligned}$$

Since $\mathbf{a}^{\mathtt{C}*}$ is continuous, $\Delta(s|\,\mathbf{a}^{\mathtt{C}*}) \to 0$ as $s \to \underline{s}$. Now consider $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*})$. Since $st^{\mathtt{I}} < st^{\mathtt{C}}$, $\Delta(s|\,\mathbf{a}^{\mathtt{I}*}) > 0$ for $s < \overline{s}$. Let $s_0 = \frac{1}{2}(\overline{s} + \underline{s})$. By continuity, $\min_{[\underline{s},s_0]} \Delta(s|\,\mathbf{a}^{\mathtt{I}*}) = \kappa > 0$. Choose $s_{\kappa/2} > \underline{s}$ so that $\Delta(s|\,\mathbf{a}^{\mathtt{C}*}) \leq \kappa/2$ for $s \in [\underline{s}, s_{\kappa/2}]$. Finally, let $s_1 = \min\{s_0, s_{\kappa/2}\}$. Then,

$$R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*}) \geq \textstyle\int_{\underline{s}}^{s_1} \Delta(s|\,\mathbf{a}^{\mathtt{I}*})dF \geq \kappa F(s_1),$$

$$-R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}*}) \leq \sup_{s>s_1} \Delta(s|\,\mathbf{a}^{\mathtt{C}*})(1 - F(s_1)) + \frac{\kappa}{2}F(s_1).$$

So $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*}) > -R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}*})$ if $F(s_1)/(1 - F(s_1)) > \frac{2}{\kappa}\sup_{s>s_1} \Delta(s|\,\mathbf{a}^{\mathtt{C}*})$.

A.7 Proof of Proposition 4

Use (2) to express $R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}})$. Then, changing order of integration and rearranging yields

$$R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}}) = -\textstyle\int_{\underline{v}^{\mathtt{I}}}^{\underline{v}^{\mathtt{C}}} b(\mathbf{a}^{\mathtt{C}}(v))g^{\mathtt{I}}(v)\,dv + \int_{\underline{v}^{\mathtt{C}}}^{\overline{v}^{\mathtt{C}}} b(\mathbf{a}^{\mathtt{C}}(v))G^{\mathtt{C}}(v)\,dF^{\mathtt{C}}, \tag{22}$$

where $g^{\mathtt{I}} : [\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}] \to \mathbb{R}$ is given by

$$g^{\mathtt{I}}(v) = \frac{t^{\mathtt{I}} - 1}{t^{\mathtt{I}}} v f^{\mathtt{I}}(v) + F^{\mathtt{I}}(v),$$

and $G^{\mathtt{C}} : [\underline{v}^{\mathtt{C}}, \overline{v}^{\mathtt{C}}] \to \mathbb{R}$ is given by

$$G^{\mathtt{C}}(v) = \frac{t^{\mathtt{C}} - 1}{t^{\mathtt{C}}} v - \frac{1 - F^{\mathtt{C}}(v)}{f^{\mathtt{C}}(v)} - \frac{f^{\mathtt{I}}(v)}{f^{\mathtt{C}}(v)}\left[\frac{t^{\mathtt{I}} - 1}{t^{\mathtt{I}}} v - \frac{1 - F^{\mathtt{I}}(v)}{f^{\mathtt{I}}(v)}\right].$$

Maximizing $-R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}})$ with an increasing $\mathbf{a}^{\mathtt{C}}$ that equals $\mathbf{a}^{\mathtt{C}*}$ on $[\underline{v}^{\mathtt{C}}, \overline{v}^{\mathtt{C}}]$ is equivalent to maximizing $\int_{\overline{v}^{\mathtt{C}}}^{\overline{v}^{\mathtt{I}}} b(\mathbf{a}(v))g^{\mathtt{I}}(v)\,dv$ with an increasing $\mathbf{a} : [\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}) \to [\underline{a}, \mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})]$. Although $g^{\mathtt{I}}(\underline{v}^{\mathtt{I}}) < 0$, $g^{\mathtt{I}}(v)$ may be strictly positive or decreasing. So, let $\widetilde{F}$ be the uniform distribution on $[\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}]$ and $\widetilde{F}^{-1} : [0,1] \to [\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}]$ be its inverse function. For $x \in [0,1]$, define $z(x) = g^{\mathtt{I}}(\widetilde{F}^{-1}(x))$ and $Z(x) = \int_0^x z(y)\,dy$. Let Ω be the convex hull of Z, and define $\omega : [0,1] \to \mathbb{R}$ as $\omega(x) = \Omega'(x)$, whenever $\Omega'(x)$ exists. W.l.o.g., extend $\omega(x)$ by right-continuity on $[0,1)$ and by left-continuity at 1. On $[\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}]$, let $\overline{g}^{\mathtt{I}}(v) = \omega(\widetilde{F}(v))$, which is increasing. Recall that

$v^m = \min\{\overline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}\} > \underline{v}^{\mathtt{I}}$, and let $x^m = \widetilde{F}(v^m) > 0$. Since $g^{\mathtt{I}}$ is continuous on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$, by the argument in Lemma 5, ω is continuous on $[0, x^m]$, so $\overline{g}^{\mathtt{I}}$ is continuous on $[\underline{v}^{\mathtt{I}}, v^m]$.

Lemma 10 $\overline{g}^{\mathtt{I}}(\underline{v}^{\mathtt{I}}) \leq g^{\mathtt{I}}(\underline{v}^{\mathtt{I}})$.

Proof. Otherwise, $\omega(0) > z(0)$. Since z is continuous on $[0, x^m]$ and ω is increasing, there is $x > 0$ such that $\omega(y) > z(y)$ for $y \leq x$. Since $Z(0) = \Omega(0)$, we get the contradiction

$$Z(x) = Z(0) + \int_0^x z(y)\, dy < \Omega(0) + \int_0^x \omega(y)\, dy = \Omega(x).$$

■

So $v_{\mathtt{u}} = \sup\{v \in [\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}] \mid \overline{g}^{\mathtt{I}}(v) < 0\} > \underline{v}^{\mathtt{I}}$. Similarly, define $v^{\mathtt{u}} = \inf\{v \in [\underline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}] \mid \overline{g}^{\mathtt{I}}(v) > 0\}$, if the set is nonempty, otherwise $v^{\mathtt{u}} = \underline{v}^{\mathtt{C}}$. By Theorem 3.7 of Toikka (2011), $\mathbf{a}^{\mathtt{C}}$ must satisfy $\mathbf{a}^{\mathtt{C}}(v) = \underline{a}$ for $v \in (\underline{v}^{\mathtt{I}}, v_{\mathtt{u}})$ and $\mathbf{a}^{\mathtt{C}}(v) = \mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})$ for $v \in (v^{\mathtt{u}}, \underline{v}^{\mathtt{C}})$, if any. Letting $\mathbf{a}^{\mathtt{C}}(v^{\mathtt{u}}) = \mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})$ is w.l.o.g.. For completeness, on $[v_{\mathtt{u}}, v^{\mathtt{u}})$, $\mathbf{a}^{\mathtt{C}}$ can be any increasing function, so long as it satisfies the necessary *pooling property* described by Toikka (see Definition 3.5). By Corollary 3.8 of Toikka (2011), it is w.l.o.g. to set $\mathbf{a}^{\mathtt{C}}(v) = \mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})$ on $[v_{\mathtt{u}}, v^{\mathtt{u}})$.

A.8 Proof of Corollary 3

Fix $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ and recall that it minimizes $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$ among all increasing $\mathbf{a}^{\mathtt{I}}$ that equal $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ on $[\underline{v}^{\mathtt{I}}, \overline{v}^{\mathtt{I}}]$. Using (22) and $\mathbf{a}^{\mathtt{C}}$ from Proposition 4, (R) becomes

$$\begin{aligned}[b(\underline{a}) - b(\mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}}))]\textstyle\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathtt{u}}} g^{\mathtt{I}}(v)\, dv \quad \geq \quad & R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})) + \textstyle\int_{\underline{v}^{\mathtt{C}}}^{\overline{v}^{\mathtt{C}}} b(\mathbf{a}^{\mathtt{C}*}(v)) G^{\mathtt{C}}(v)\, dF^{\mathtt{C}} \\ & - b(\mathbf{a}^{\mathtt{C}*}(\underline{v}^{\mathtt{C}})) \textstyle\int_{\underline{v}^{\mathtt{I}}}^{\underline{v}^{\mathtt{C}}} g^{\mathtt{I}}(v)\, dv.\end{aligned}$$

Since $\mathbf{a}^{\mathtt{C}*}$ and $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ are infeasible, the right-hand side is positive. $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$ has been minimized. The result follows, since $\int_{\underline{v}^{\mathtt{I}}}^{v_{\mathtt{u}}} g^{\mathtt{I}}(v)\, dv < 0$.

A.9 Proof of Proposition 5

First, the next lemma shows that $\mathbf{a}^{\mathtt{C}}$ sustains $\mathbf{a}^*$ with $\mathtt{C}$ if and only if (R) does not bind. Recall that $r^{\mathtt{I}} = \mu/\gamma$, where μ is the Lagrange multiplier associated with (R).

Lemma 11 *There is* $\mathbf{a}^{\mathtt{C}}$ *increasing such that* $\mathbf{a}^{\mathtt{C}}(v) = \mathbf{a}^{\mathtt{C}*}(v)$ *on* $[\underline{v}^{\mathtt{C}}, \overline{v}^{\mathtt{C}}]$ *and* $\mathbf{a}^{\mathtt{C}}$ *maximizes* $W^{\mathtt{C}}\left(\hat{\mathbf{a}}^{\mathtt{C}}\right) - r^{\mathtt{I}} R^{\mathtt{I}}\left(\hat{\mathbf{a}}^{\mathtt{C}}\right)$ *if and only if* $r^{\mathtt{I}} = 0$.

Proof. The proof uses functions $\mathbf{b} \in \mathcal{B}$ (see the proof of Lemma 2). Suppose $r^{\mathtt{I}} > 0$. Using $\widetilde{R}^{\mathtt{I}}(\mathbf{b}) = R^{\mathtt{I}}(b^{-1}(\mathbf{b}))$ in (22), write $\widetilde{W}^{\mathtt{C}}(\mathbf{b}) - r^{\mathtt{I}} \widetilde{R}^{\mathtt{I}}(\mathbf{b})$ as

$$VS^{\mathtt{C}}(b^{-1}(\mathbf{b}), r^{\mathtt{I}}) = \int_{\underline{v}^{\mathtt{C}}}^{\overline{v}^{\mathtt{C}}} [\mathbf{b}(v)\, w^{\mathtt{C}}(v, r^{\mathtt{I}}) + \xi(\mathbf{b}(v))] dF^{\mathtt{C}} + r^{\mathtt{I}} \int_{\underline{v}^{\mathtt{I}}}^{\underline{v}^{\mathtt{C}}} \mathbf{b}(v)\, g^{\mathtt{I}}(v)\, dv,$$

where $w^{\mathtt{C}}(v, r^{\mathtt{I}}) = v/t^{\mathtt{C}} - r^{\mathtt{I}} G^{\mathtt{C}}(v)$. Let $\mathbf{b}_{\mathtt{u}}^{\mathtt{C}} = b(\mathbf{a}^{\mathtt{C}})$ where $\mathbf{a}^{\mathtt{C}}$ is as in Proposition 4 and let $\mathcal{B}^*$ be the set of $\mathbf{b}^{\mathtt{C}} \in \mathcal{B}$ that equal $\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}$ on $[\underline{v}^{\mathtt{C}}, \overline{v}^{\mathtt{C}}]$. By construction, $VS^{\mathtt{C}}(b^{-1}(\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}), r^{\mathtt{I}}) = \max_{\mathbf{b} \in \mathcal{B}^*} VS^{\mathtt{C}}(b^{-1}(\mathbf{b}), r^{\mathtt{I}})$. But there is $\hat{\mathbf{b}}^{\mathtt{C}} \in \mathcal{B} \backslash \mathcal{B}^*$ such that $VS^{\mathtt{C}}(b^{-1}(\hat{\mathbf{b}}^{\mathtt{C}}), r^{\mathtt{I}}) > VS^{\mathtt{C}}(b^{-1}(\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}), r^{\mathtt{I}})$. Focus on $[v_m, \overline{v}^{\mathtt{C}}]$ with $v_m = \max\{\overline{v}^{\mathtt{I}}, \underline{v}^{\mathtt{C}}\}$, and let $\hat{w}^{\mathtt{C}}$ be the generalized version of $w^{\mathtt{C}}$ on this interval, obtained with the method used in the proof of Proposition 2. Since $w^{\mathtt{C}}$ is continuous on $[v_m, \overline{v}^{\mathtt{C}}]$, so is $\hat{w}^{\mathtt{C}}$ (Lemma 5). Since $r^{\mathtt{I}} > 0$, $G^{\mathtt{C}}$ implies $w^{\mathtt{C}}(v, r^{\mathtt{I}}) > v/t^{\mathtt{C}}$ for $v \in [v_m, \overline{v}^{\mathtt{C}})$. I claim that $\hat{w}^{\mathtt{C}}(v_m, r^{\mathtt{I}}) > v_m/t^{\mathtt{C}}$. By the argument in Lemma 9, $\hat{w}^{\mathtt{C}}(v_m, r^{\mathtt{I}}) \leq w^{\mathtt{C}}(v_m, r^{\mathtt{I}})$. If $\hat{w}^{\mathtt{C}}(v_m, r^{\mathtt{I}}) = w^{\mathtt{C}}(v_m, r^{\mathtt{I}})$, the claim follows. If $\hat{w}^{\mathtt{C}}(v_m, r^{\mathtt{I}}) < w^{\mathtt{C}}(v_m, r^{\mathtt{I}})$, then there is $v_0 > v_m$ such that $\hat{w}^{\mathtt{C}}(v, r^{\mathtt{I}}) = w^{\mathtt{C}}(v_0, r^{\mathtt{I}})$ on $[v_m, v_0]$; so, $\hat{w}^{\mathtt{C}}(v_m, r^{\mathtt{I}}) = w^{\mathtt{C}}(v_0, r^{\mathtt{I}}) \geq v_0/t^{\mathtt{C}} > v_m/t^{\mathtt{C}}$. Since $\hat{w}^{\mathtt{C}}$ is continuous and increasing, in either case there is $v_1 > v_m$ such that $\hat{w}^{\mathtt{C}}(v, r^{\mathtt{I}}) > v/t^{\mathtt{C}}$ on $[v_m, v_1]$. Construct $\hat{\mathbf{b}}^{\mathtt{C}}$ by letting $\hat{\mathbf{b}}^{\mathtt{C}}(v) = \arg\max_{y \in [b(\underline{a}), b(\overline{a})]} y \hat{w}^{\mathtt{C}}(v, r^{\mathtt{I}}) + \xi(y)$ if $v \in [v_m, \overline{v}^{\mathtt{C}}]$, and $\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}(v)$ if $v \in [\underline{v}^{\mathtt{I}}, v_m)$. Then, $\hat{\mathbf{b}}^{\mathtt{C}} \in \mathcal{B}$, but $\hat{\mathbf{b}}^{\mathtt{C}}(v) > \mathbf{b}_{\mathtt{u}}^{\mathtt{C}}(v)$ on $[v_m, v_1]$; so $\hat{\mathbf{b}}^{\mathtt{C}} \notin \mathcal{B}^*$. Finally, the difference between $VS^{\mathtt{C}}(b^{-1}(\hat{\mathbf{b}}^{\mathtt{C}}), r^{\mathtt{I}})$ and $VS^{\mathtt{C}}(b^{-1}(\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}), r^{\mathtt{I}})$ is

$$\int_{v_m}^{\overline{v}^{\mathtt{C}}} [\hat{\mathbf{b}}^{\mathtt{C}}(v)\, w^{\mathtt{C}}(v, r^{\mathtt{I}}) + \xi(\hat{\mathbf{b}}^{\mathtt{C}}(v))] dF^{\mathtt{C}} - \int_{v_m}^{\overline{v}^{\mathtt{C}}} [\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}(v)\, w^{\mathtt{C}}(v, r^{\mathtt{I}}) + \xi(\mathbf{b}_{\mathtt{u}}^{\mathtt{C}}(v))] dF^{\mathtt{C}},$$

which is strictly positive. So the increasing maximizer of $W^{\mathtt{C}}(\mathbf{a}^{\mathtt{C}}) - r^{\mathtt{I}} R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}})$ must differ from $\mathbf{a}^{\mathtt{C}*}(v)$ on a nonempty interval of $[\underline{v}^{\mathtt{C}}, \overline{v}^{\mathtt{C}}]$. ■

Recall that $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ maximizes $W^{\mathtt{I}}(\mathbf{a}^{\mathtt{I}}) - r^{\mathtt{C}} R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}})$ among all increasing $\mathbf{a}^{\mathtt{I}}$. By revealed optimality, $\hat{r}^{\mathtt{C}} > r^{\mathtt{C}}$ implies $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(\hat{r}^{\mathtt{C}})) \leq R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}}))$. By uniqueness of $\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})$ (Proposition 2), $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(\hat{r}^{\mathtt{C}})) = R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}}))$ if $\hat{r}^{\mathtt{C}} = r^{\mathtt{C}}$. So, $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}}))$ is a decreasing function of $r^{\mathtt{C}}$, $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})) \leq R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}*})$, and $\lim_{r^{\mathtt{C}} \to +\infty} R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})) = 0$ by Proposition 2 and 3. By Proposition 1, $R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}*}) < 0$. So, define $r_2 = \min\{r^{\mathtt{C}} \geq 0 \mid R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})) + R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}*}) \leq 0\}$. Now, let $\mathbf{a}^{\mathtt{C}}$ be as in Proposition 4, so $R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}}) < R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}*}) < 0$. Define $r_1 = \min\{r^{\mathtt{C}} \geq 0 \mid R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}})) + R^{\mathtt{I}}(\mathbf{a}^{\mathtt{C}}) \leq 0\}$. Lemma 4 and continuity of $R^{\mathtt{C}}(\mathbf{a}^{\mathtt{I}}(r^{\mathtt{C}}))$ imply that $r_1 < r_2$ in some environments and, together with Corollary 3, that $r_1 > 0$ in some environments. The result follows from

Lemma 11.

References

[1] Akerlof, G. A. (1991): "Procrastination and Obedience", *American Economic Review (Papers and Proceedings)*, 81, 2, 1-19.

[2] Amador, M., Werning, I. and Angeletos, G.M. (2006): "Commitment vs. Flexibility", *Econometrica*, 74, 2, 365-396.

[3] Ambrus, A., and G. Egorov (2013): "A Comment on Commitment vs. Flexibility," *Econometrica*, Forthcoming.

[4] Amromin, G. (2002): "Portfolio Allocation Choices in Taxable and Tax-Deferred Accounts: An Empirical Analysis of Tax Efficiency", Mimeo.

[5] Amromin, G. (2003): "Household Portfolio Choices in Taxable and Tax-Deferred Accounts: Another Puzzle?", *European Finance Review*, 7, 547-582.

[6] Ashraf, N., N. Gons, D. S. Karlan and W. Yin, (2003): "A Review of Commitment Savings Products in Developing Countries", ERD Working Paper No. 45.

[7] Ashraf, N., D. S. Karlan and W. Yin, (2006): "Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines", *Quarterly Journal of Economics*, 121, 2, 635-672.

[8] Battaglini, M. (2005): "Long-Term Contracting with Markovian Consumers," *American Economic Review*, 95, 3, 637-658.

[9] Bond, P. and Sigurdsson, G. (2013): "Commitment Contracts", Mimeo, University of Minnesota.

[10] Bryan, G., Karlan, D., and Nelson, S. (2010): "Commitment Devices", *Annual Review of Economics*, 2, 1, 671-698.

[11] Courty, P. and Li, H. (2000): "Sequential Screening", *Review of Economic Studies*, 67, 4, 697–717.

[12] DellaVigna, S. (2009): "Psychology and economics: Evidence from the field", *Journal of Economic Literature,* 47,2, 315-372.

[13] DellaVigna, S. and Malmendier, U., (2004): "Contract Design And Self-Control: Theory And Evidence", *Quarterly Journal of Economics,* 119, 2, 353-402.

[14] DellaVigna, S. and Malmendier, U., (2006): "Paying not to go to the gym", *American Economic Review*, 96, 3, 694-719.

[15] Eliaz, K. and Spiegler, R., (2006): "Contracting with diversely naive agents", *Review of Economic Studies*, 73, 3, 689-714.

[16] Eliaz, K. and Spiegler, R., (2008): "Consumer optimism and price discrimination", *Theoretical Economics*, 3, 4, 459–497.

[17] Esteban, S. and Miyagawa, E., (2006a): "Temptation, self-control, and competitive nonlinear pricing", *Economics Letters*, 90, 3, 348-355.

[18] Esteban, S. and Miyagawa, E., (2006b): "Optimal menu of menus with self-control preferences", Mimeo, Universitat Autònoma de Barcelona.

[19] Esteban, S., Miyagawa, E. and Shum, M. (2007): "Nonlinear pricing with self-control preferences", *Journal of Economic Theory*, 135, 1, 306-338.

[20] Gottlieb, D. (2008): "Competition over Time-Inconsistent Consumers," *Journal of Public Economic Theory*, 10,4, 673-684.

[21] Grubb, M. (2009): "Selling to Overconfident Consumers," *American Economic Review*, 99, 5, 1770-1870.

[22] Gul, F, and W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica*, 69, 1403-1435.

[23] Heidhues, P. and Koszegi, B. (2010): "Exploiting naivete about self-control in the credit market", *American Economic Review,* 2279-2303.

[24] Holden, S., Ireland, K., Leonard-Chambers, V., and Bodgan, M. (2005): "The Individual Retirement Account at Age 30: A Retrospective", Investment Company Institute, 11, 1.

[25] Holden, S., Schrass, D. (2008): "The Role of IRAs in U.S. Households' Saving for Retirement, 2008", Investment Company Institute, 18, 1.

[26] Holden, S., Schrass, D. (2009): "The Role of IRAs in U.S. Households' Saving for Retirement, 2009", Investment Company Institute, 19, 1.

[27] Holden, S., Schrass, D. (2010a): "The Role of IRAs in U.S. Households' Saving for Retirement, 2010", Investment Company Institute, 19, 8.

[28] Holden, S., Sabelhaus, J., and Bass, S. (2010b): "The IRA Investor Profile. Traditional IRA Investors' Contribution Activity, 2007 and 2008", Investment Company Institute.

[29] Jianye, Yan (2012): "Contracting with a Quasi-hyperbolic Agent under Incomplete Information: A Second-best Investment Good Pricing", *Journal of Economics*, 119, 353-402.

[30] Kreps, D. (1979): "A Representation Theorem for Preference for Flexibility", *Econometrica*, 47, 565-577.

[31] Laibson, D. (1997): "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics,* 112, 2, 443-478.

[32] Laibson, D. (1998): "Life-Cycle Consumption and Hyperbolic Discount Functions", *European Economic Review*, 42, 861-871.

[33] Luenberger, D. G. (1969): "Optimization by Vector Space Methods". New York: Wiley.

[34] Munnell, A. H., Sundén, A. E. (2006): "401(k) Plans Are Still Coming Up Short", Brookings Institute Press.

[35] Mussa, M., and Rosen, S. (1978): "Monopoly and Product Quality", *Journal of Economic Theory*, 18, 2, 301-317.

[36] Myerson, R. B. (1981): "Optimal Auction Design", *Mathematics of Operation Research*, 6, 1, 58-73.

[37] Myerson, R. B. (1986): "Multi Stage Games with Communication", *Econometrica*, 54, 2, 323-358.

[38] O'Donoghue, T. and Rabin, M. (1999a): "Doing It Now or Later", *American Economic Review*, 89, 1, 103-124.

[39] O'Donoghue, T. and Rabin, M. (1999b): "Incentives for Procrastinators", *Quarterly Journal of Economics*, 114, 3, 769-816.

[40] O'Donoghue, T. and Rabin, M. (2001): "Choice and Procrastination", *Quarterly Journal of Economics*, 116, 1, 121-160.

[41] O'Donoghue, T. and Rabin, M. (2003): "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes", *American Economic Review*, 93, 2, 186-191.

[42] O'Donoghue, T. and Rabin, M. (2007): "Incentives and Self Control", in Richard Blundell, Whitney Newey, and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications* (Ninth World Congress), Cambridge University Press, August 2007.

[43] Pavan, A. (2007): "Long-Term Contracting in a Changing World", Mimeo. Northwestern University.

[44] Pavan, A., Segal, I., and J. Toikka (2012): "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization, and Information Disclosure." Mimeo, Northwestern University.

[45] Phelps, E. and R. Pollack (1968): "On Second Best National Savings and Game-Equilibrium Growth," *Review of Economic Studies*, 35, 185-199.

[46] Rockafellar, R. T. (1970): "Convex Analysis", Princeton University Press.

[47] Spiegler, R. (2011): "Bounded Rationality and Industrial Organization," Oxford University Press.

[48] Strotz, R. H. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23, 165-180.

[49] Toikka, J. (2011): "Ironing Without Control", *Journal of Economic Theory*, 146, 6, 2510-2526.

[50] Zhang, W. (2012): "Endogenous Preference and Dynamic Contract Design," *The B.E. Journal of Theoretical Economics*, 12, 1.