Learning Self-Control*

S. Nageeb Ali†
University of California, San Diego

February 2011

Abstract

This paper examines how a decisionmaker who is only partially aware of his temptations learns about them over time. In facing temptations, individuals use their experiences to forecast future self-control problems and choose the appropriate level of commitment. I demonstrate that rational learning can be perpetually partial and need not result in full sophistication. The main result of this paper characterizes necessary and sufficient conditions for learning to converge to full sophistication. I apply this result to a consumption-savings environment in which a decisionmaker is tempted by present-bias and establish a learning-theoretic justification for assuming sophistication in this setting.

Keywords: self-control, partial awareness, Bayesian learning, sophistication, experimentation, dual-selves, self-confirming equilibrium

---

*This paper is a revised version of Chapter 3 of my dissertation. I am grateful to my advisers, Susan Athey and Doug Bernheim, for their support and encouragement, and to Drew Fudenberg and Paul Niehaus for numerous suggestions that greatly improved the paper. I thank Ricardo Alonso, Manuel Amador, Dan Benjamin, Aislinn Bohren, Juan Carrillo, Chris Chambers, Vince Crawford, Stefano DellaVigna, Ben Ho, Shachar Kariv, Navin Kartik, Botond Köszegi, Troy Kravitz, Jon Levin, Charles Lin, Ulrike Malmendier, Andres Santos, Josh Schwartzstein, Shamim Sinnar, Joel Sobel, Joel Watson, Tom Wiseman, the Editor (Robert Barro), and three referees for helpful comments. I acknowledge financial support from the UCSD Academic Senate, the UCSD Hellman Fund, the Stanford Institute for Economic Policy Research, and the Institute for Humane Studies.

†Email: snali@ucsd.edu; Web: http://econ.ucsd.edu/∼snali; Address: 9500 Gilman Drive, Department of Economics, La Jolla, CA 92093-0508.

Contents

1 Introduction **1**

2 A Simple Example **6**

3 When is Learning Adequate? **9**

3.1 Adequate Vs. Partial Learning . . . 13

3.2 Full Commitment Distinguishability . . . 14

3.3 Main Results . . . 16

3.4 Remarks . . . 18

4 Applications **21**

4.1 Consumption-Savings . . . 21

4.2 Costly Self-Control . . . 26

5 Discussion **27**

References **30**

A Online Appendix **35**

A.1 Section 2 . . . 35

A.2 Section 3 . . . 37

A.3 Section 4.1 . . . 46

A.4 Section 4.2 . . . 50

A.5 Section 5 . . . 52

"An individual who finds himself continuously repudiating his past plans may learn to distrust his future behavior, and may do something about it."
— Strotz (1955)

1 Introduction

Questions of temptation and self-control are at the forefront of psychology and economics. When studying self-control, a modeler must decide just how much individuals can be assumed to know about their temptations since different assumptions of self-awareness translate into different behavioral predictions. The assumption most commonly made, referred to as *sophistication*, is that the decisionmaker perfectly anticipates future self-control problems. Yet, sophistication is often seen as an inaccurate model of a decisionmaker's awareness, especially in the short- and medium-run when the decisionmaker may lack experience. An alternative assumption is that of *naivete*, whereby a decisionmaker anticipates having perfect self-control in the future. Bridging the gap between these two extreme assumptions, O'Donoghue and Rabin (2001) propose a framework of partial naivete in which a decisionmaker with (β, δ) preferences assigns probability 1 to a particular level of present bias $(\hat{\beta})$ less than his true present-bias $(\hat{\beta} > \beta)$.

With few exceptions, most models of imperfect self-control make one of the above assumptions. That derived behavior is sensitive to the specification of beliefs has been widely recognized, including by those that first modeled self-control (Strotz, 1955; Pollak, 1968). Nevertheless, the practice has been to treat a decisionmaker's belief about his self-control as exogenous and model choice given that exogenous belief. While this practice makes analysis tractable, it raises two conceptual issues. First, in models of partial sophistication, a decisionmaker's beliefs may be incompatible with what is observed over time, and so ancillary assumptions are necessary for analysis.[1] Second, given the many different awareness assumptions that could be made, it has been difficult to assess which assumption is appropriate for a particular environment and why. Fudenberg (2006), in his recent discussion of behavioral economics, voices this concern.

> "I think that behavioral economics would be well served by concerted attempts to provide learning-theoretic (or any other foundations) for its equilibrium concepts. At the least, this process might provide a better understanding of when the currently used concepts apply...."

[1] For example, a partially naive decisionmaker may see behavior that contradicts his view of the world in each and every period. The literature typically addresses this issue by assuming that his beliefs are not revised, which implicitly assumes that he receives no feedback about payoffs or the past history.

This paper proposes a simple framework to address these concerns. In this approach, beliefs and choices are derived endogenously and jointly evolve based on the decisionmaker's experience. Endogenizing beliefs in this way allows one to pose and answer the question of whether and when sophistication closely approximates the decisionmaker's self-awareness once he has had many opportunities to learn. Although one may be tempted to conclude that Bayesian learning should always engender sophistication, I demonstrate that tradeoffs between commitment and flexibility inherent in self-control environments may actually impede learning. The main result of this paper is to offer a necessary and sufficient condition—across self-control environments—under which learning is not impeded and inexorably leads to sophistication; this condition can be checked in applications to assess the appropriateness of assuming sophistication.

As the main application of this approach, I analyze a standard consumption-savings environment in which a decisionmaker is tempted by immediate consumption and purchases illiquid assets to commit towards future consumption. This setting has been the focus of many papers in the quasi-hyperbolic literature (e.g. Laibson, 1997; Barro, 1999), almost all of which assume sophistication. I demonstrate that this canonical setting satisfies the condition for adequate learning: thus, a decisionmaker who learns about his tendency to overconsume from his past choices eventually chooses commitment as if he could perfectly forecast his temptation to overconsume. This result therefore offers a learning-theoretic foundation for sophistication in consumption and savings decisions.

The general framework that I develop builds on the Planner-Doer approach to self-control: decisions are made by a single long-run Planner with dynamically consistent preferences and a myopic Doer. Prior research using such models, particularly Fudenberg and Levine (2006), has demonstrated that such models offer analytically simple and tractable frameworks to understand self-control in a variety of settings,[2] and this paper illustrates how this approach is useful for studying questions of "self-awareness." The Planner represents the forward-looking rational individual that chooses how much to commit by investing in illiquid assets, signing contracts, making promises that are costly to betray, etc. In contrast, the Doer represents the instinctive response of the individual that makes the daily choices, and is presented with stimuli, situations, and choices; the Doer is best thought of as a short-run player or behavioral type. The Planner is uncertain about the extent to which the Doer resists temptation and uses the Doer's behavior to learn over time.

[2]Thaler and Shefrin (1981) were the first to propose a dual-self approach to imperfect self-control. The recent literature also includes Benabou and Pycia (2002), Benhabib and Bisin (2005), Bernheim and Rangel (2004), Brocas and Carrillo (2008), Chatterjee and Krishna (2009), Dekel and Lipman (2010), Loewenstein and O'Donoghue (2007).

To fix ideas, I describe the example studied in Section 2: suppose that an individual faces a consumption choice between fish and steak in each period. In the beginning of the period, the Planner makes a reservation either at a restaurant that serves only fish or one that serves both fish and steak. The payoff from eating fish is constant over time and across the two restaurants, but that from eating steak depends on an i.i.d. taste shock that is realized after the Planner selects the menu. Eating fish is *ex ante* optimal given its long-term health benefits, but there are contingencies in which steak may be preferred. As in Kreps (1979), the Planner strictly prefers the flexibility associated with the larger menu were self-control not an issue; however, once at the restaurant, the Doer chooses what to eat (if there is a choice to make). The Doer's preferences are partially aligned with that of the Planner, and partially reflect a temptation to eat steak in contingencies in which the Planner prefers fish. The Planner is uncertain about the strength of this temptation and would benefit from knowing it since that helps him plan. Were the Planner to know that in the restaurant, temptation is resisted, he may strictly prefer the larger menu for the sake of flexibility. In contrast, if he believes that the Doer would too easily succumb to the temptation to eat steak, the Planner may find it *ex ante* optimal to select the smaller menu, which offers commitment value.

Learning about one's self-control is costly experimentation (and, in this case, a two-armed bandit): the Planner observes the Doer's self-control *only* when the Planner chooses the larger menu and exposes himself to temptation. Accordingly, once the Planner becomes sufficiently pessimistic about the Doer's type, he chooses the smaller menu since the value of flexibility and learning no longer outweigh the expected cost of the Doer succumbing to temptation. Notably, the Planner may make this decision with positive probability even when the Doer is a "good type" who resists temptation. Despite the infinite possibilities to learn, the Planner may decide it not worthwhile to do so, and make forever inferior commitment choices.

The direction of the skewness in these beliefs merits discussion. The Planner cannot perpetually overestimate the Doer's self-control and undercommit relative to how he would choose in the full information benchmark. To see why, notice that any belief that rationalizes flexibility will allow the Planner to continue to passively learn and therefore update his beliefs. Thus, if the Doer is unable to resist temptation, the Planner almost-surely learns this over time, and eventually makes the same commitment choice that he would in the full information benchmark.

On the other hand, the Planner may perpetually underestimate the Doer's self-control even if he began with an overoptimistic prior: once the Planner has observed the Doer choose steak too often, flexibility appears costly, and so a Planner will choose to actively learn as long as the future value of learning outweighs its cost. If the Planner is not perfectly patient, he is willing to undertake a finite number of trials after which he may commit to the singleton menu and never revise his belief thereafter. Thus, partial awareness and learning can endogenously lead to perpetual *overregulation* whereby individuals choose rigid commitments or lifestyles that they would not were they more self-aware. This possibility resonates with the perspective that obsessions, compulsions, and rigidity emerge from erroneous beliefs that one lacks control otherwise (Baumeister et al., 1994).[3]

A feature of the setting described above is that the Planner chooses between full flexibility or full commitment, but has no ability to partially commit. In contrast, many settings feature partial commitments that retain some flexibility for the decisionmaker while offering a measure of commitment. For example, in a savings environment, an individual can purchase illiquid assets to ensure some minimal savings without relinquishing flexibility altogether. Contractual mechanisms, such as recent innovations like *stickK.com*,[4] can penalize certain choices and provide incentives to counteract temptation while retaining some flexibility. Social mechanisms—through promises, shame, and peer groups—also provide partial commitment insofar as they change the benefits and costs of particular actions, but are not completely binding. Apart from external commitments, a decisionmaker may also rely on internal commitments, such as *costly self-control* (Gul and Pesendorfer, 2001; Fudenberg and Levine, 2006), which influence the choice from a menu without relinquishing flexibility altogether. To capture possibilities for partial commitment, I allow the Planner to affect the Doer's choice through *menus* and *nudges* that restrict and influence the Doer's choice. The framework here demonstrates that these partial commitments play an important role in the decisionmaker's long-run behavior and beliefs.

The main result of this paper identifies when partial commitments induce efficient learning regardless of the Planner's patience. I identify a condition, *Full Commitment Distinguishability* (FCD), that is sufficient for learning to engender sophistication, and may also be necessary. FCD is a condition on how fully informed Planners behave, and relates to the richness of the set of partial commitments. The necessity and sufficiency of FCD offers a simple criterion to understand whether learning leads to sophistication in applications: instead of solving a more intricate model in which the Planner is uncertain and updates beliefs over time, a modeler needs to solve only the full information model

[3]A complementary channel for over-regulation is explored by Benabou and Tirole (2004), who study a self-signaling mechanism in which temptation is mitigated by the adverse reputation effect it induces in future incarnations.

[4]*StickK.com* is designed to "...help people achieve their goals and objectives by enabling them to form Commitment Contracts." (http://www.stickk.com/about.php).

in which the Planner knows the Doer's type and check whether that satisfies FCD. Using this characterization, I study two applications that have been examined in prior work assuming sophistication: consumption-savings with illiquid assets and a costly self-control environment. Within these settings, I illustrate how to check the validity of FCD and derive its implications for learning.

The results herein underscore the insights that emerge from deriving the decision-maker's perception of his temptations from his environment. In contrast to prior research that divorces a decisionmaker's self-awareness from the fundamentals of the setting, whether incorrect beliefs persist is determined endogenously within this framework. Based on these results, one may expect that in those settings in which an individual has access to a wide range of partial commitments, he may know more about his temptations than in settings in which partial commitments are ineffectual or lacking.[5] From a normative perspective, the role that partial commitments play in learning has new implications for the design of commitment, and suggests when interventions enhance learning. Endogenizing beliefs also makes it possible to connect a decisionmaker's perceptions of his temptation to other aspects of his preference; for example, in Section 3.4.1, I highlight how patience fosters learning and sophistication.

Some readers may be troubled by the disappearance of overoptimism at the limits of Bayesian learning. While I use the approach developed here to study asymptotic behavior, the framework is sufficiently flexible and tractable to model a decisionmaker's initial response and these initial responses may reflect initial optimism. Indeed, there are numerous reasons to believe that decisionmakers may begin with optimistic priors and learn slowly,[6] and so findings in the field are consistent with the short-run behavior predicted by this theory. Moreover, when a decisionmaker makes choices in an environment in which he has had little prior experience, he has to simultaneously learn about how tempted he is by different choices in that setting and the payoffs of different actions. I show by example that such multidimensional learning can be impeded by challenges of identification that slow down the rate of learning. Thus, there are natural reasons to expect decisionmakers to appear to undercommit in the short-run as they simultaneously learn about self-control and the benefits and costs of actions.

Section 2 presents a simple example of the impediment to learning introduced by

[5] Just as the nature and severity of an individual's temptations varies across decision problems, it is likely that his perception and awareness of his temptations also differs across these settings. Thus, someone could be sophisticated about his tendency to procrastinate, and yet, still have erroneous beliefs about the extent to which he may become addicted to particular substances.

[6] Optimism could emanate from a number of different sources, if a positive belief about one's attributes has some intrinsic value (Brunnermeier and Parker, 2005; Kőszegi, 2006) or induces motivation (Carrillo and Mariotti, 2000; Benabou and Tirole, 2002; Compte and Postlewaite, 2004).

imperfect self-control. Section 3 studies a general framework with partial commitments, in which experimentation takes a rich form. That section contains the main results of this paper, and also describes the implications of patience, offers some suggestions on how a modeler may make partial inferences about a decisionmaker's awareness, and describes the connection of the results here with the steady-state solution concept of self-confirming equilibrium. Section 4 applies the framework to savings behavior and costly self-control. Section 5 discusses the results of this paper in light of the related literature and illustrates the identification challenged induced by multidimensional learning. The proofs for all results are collected in an Online Appendix.

2 A Simple Example

I begin with a simple example to illustrate the mechanism for incomplete learning in the most transparent way. Consider an infinitely-lived individual who chooses between undertaking an activity ($a_t = 1$) or not ($a_t = -1$) in period $t = 0, 1, 2, ...$ Undertaking the activity in period t generates a deterministic reward $b \in \left(\frac{1}{2}, 1\right)$ but involves a stochastic cost s_t uniformly drawn from $[0, 1]$. This activity can represent, for example, the choice to consume fish from the example described in the introduction or to exercise, in which case b captures the discounted long-run gain from the activity.

The decision in each period is made through the conjunction of two systems, the *Planner* and the *Doer*, who act sequentially in each period. The Planner is a long-run agent with an exponential discount factor $\delta \in (0, 1)$, and in period t, he obtains a payoff of $(b - s_t)$ if $a_t = 1$ and 0 otherwise. In the 1$^{\text{st}}$ sub-period of each period t, the Planner chooses a menu prior to the realization of s_t, after which the Doer selects an alternative from the menu. The Planner chooses between full commitment to undertaking the activity by selecting the singleton menu $\{1\}$, full commitment to not undertaking the activity by selecting the singleton menu $\{-1\}$, and *flexibility* by selecting the menu $\{-1, 1\}$ thereby permitting the Doer to choose either action. When the Planner chooses to commit to an action, that action is undertaken regardless of the Doer's type or the realized cost.

The Doer in each period is not a strategic actor but can be thought of as a behavioral type or short-run player. Nature selects the type, θ, of the Doer from $\left\{\underline{\theta}, \bar{\theta}\right\}$ where $0 \leq \underline{\theta} < \bar{\theta} \leq 1$, and this type persists through time. When the Planner is flexible, the Doer of type θ observes s_t and chooses $a_t = 1$ if and only if θb exceeds s_t. Thus, the Doer is tempted towards inactivity, and undertakes the activity in fewer contingencies than the Planner would wish to do so. This temptation may emerge from a manifestation of a present-bias when costs and rewards are temporally separated. The perspective here is that the Doer being myopic (representing the individual's *instincts*) does not account for how its actions influence the Planner's future commitment choices.

The Planner faces a tradeoff between commitment and flexibility: while he would like to exploit the Doer's informational advantage, flexibility allows temptations to guide that choice. Since the payoff from flexibility depends on the Doer's susceptibility to temptation, it is valuable for the Planner to try to learn about θ so as to optimally choose commitment in the future. I assume that the Planner would prefer to commit to undertaking the activity if he were perfectly confident that $\theta = \underline{\theta}$ and would prefer to remain flexible if he were confident that $\theta = \bar{\theta}$.[7]

He begins with a prior $\mu_0 \equiv \Pr\left(\theta = \bar{\theta}\right)$ that ascribes positive probability to both types. For purposes of simplicity, suppose that all that he observes over time is behavior and not past realizations of costs (I consider more general informational structures in Section 3). Since full commitments override the Doer, the Planner learns about θ only when he chooses to be flexible. I use h^t to summarize the relevant history for the Planner when acting at time t, and μ_t to denote the Planner's posterior belief that the Doer's type is $\bar{\theta}$. Given a prior belief μ, the Planner updates his beliefs to μ^+ and μ^- when he is flexible and observes the Doer choose $a_t = 1$ and $a_t = -1$ respectively. The dynamic decision problem of the Planner is described by the value function

$$V(\mu) = \max \left\{ \begin{array}{l} b - E[s] + \delta V(\mu), \\ \sum_{\theta \in \{\underline{\theta}, \bar{\theta}\}} \Pr^{\mu}(\theta) \left(\begin{array}{c} \theta b \left(b - E[s | s \leq \theta b] + \delta V(\mu^+)\right) \\ + (1 - \theta b) \delta V(\mu^-) \end{array} \right) \end{array} \right\}, \tag{1}$$

where the 1^{st} term is the value of committing to undertake the action, and the 2^{nd} term is the value of flexibility. The Planner never commits to $a = -1$ since *ex ante*, this is dominated by committing to $a = 1$. Standard arguments ensure that the value function exists, is unique, continuous, and non-decreasing in μ, and thus, the optimal decision takes the form of a simple threshold rule.

Proposition 1. *There exists $\mu^* \in (0, 1)$ such that for all $\mu < \mu^*$, the Planner's optimal choice is to commit to the activity and for $\mu \geq \mu^*$, the Planner's optimal choice is flexibility.*

When the Planner is more optimistic about the Doer's ability to resist temptation or values the option to learn, he is more willing to remain flexible. Because of the option-

[7] The relevant condition is $b \in \left(\frac{1}{2-\underline{\theta}}, \frac{1}{2-\bar{\theta}}\right)$; when this does not hold, the Planner's optimal commitment choice is independent of his belief about the Doer's type.

value associated with flexibility and learning, a forward-looking Planner is willing to choose flexibility even if he expects that committing yields greater short-run payoffs. For less optimistic beliefs, when $\mu < \mu^*$, the benefits of flexibility and learning do not offset the expected costs of temptation and therefore the Planner chooses to commit. Since the choice of commitment shuts down the channel for learning, once he chooses to commit in one period, he finds it optimal to commit in every subsequent period. Because the choice to commit is endogenous, eventual beliefs are endogenous, and therefore evolve differently for each type of the Doer (assuming that the prior μ_0 exceeds μ^*).

Consider the case in which the Doer is of type $\underline{\theta}$. The Planner's beliefs will fluctuate so long as he remains flexible and his beliefs remain bounded away from putting probability 1 on $\bar{\theta}$. The Martingale Convergence Theorem ensures that the Planner's beliefs eventually settle but do not converge to the completely incorrect belief of $\mu = 1$. Therefore, the Planner eventually commits, making the same choice that he would in the full information benchmark; initial partial awareness does not forge any long-lasting differences from the standard benchmark.

Now suppose that the Doer is of type $\bar{\theta}$. The Martingale Convergence Theorem again implies eventual convergence of the Planner's beliefs, which obtains in two distinct ways: either the Planner's beliefs converge to the truth ($\mu_t \to 1$) or the Planner's belief μ_t falls below μ^*, leading him to commit. Both events occur with positive probability. Recall that in the full information environment, a Planner who knew that his self-control problem corresponds to $\bar{\theta}$ would never choose to commit. Relative to this benchmark, partial awareness and learning introduce a possibility for long-run inefficiency and overcommitment. The preceding ideas are summarized below.

Theorem 1. *For any type $\theta \in \left\{\underline{\theta}, \bar{\theta}\right\}$, a Planner eventually chooses to commit with strictly positive probability.*

1. *If the Doer is of type $\underline{\theta}$, then almost-surely, the Planner eventually chooses to commit.*
2. *If the Doer is of type $\bar{\theta}$, then the Planner either chooses to commit or he learns the Doer's type. Both events occur with strictly positive probability if $\mu_0 \geq \mu^*$.*

This result illustrates how beliefs that induce excessive flexibility eventually dissipate: if the Doer is of type $\underline{\theta}$, any belief of the Planner that rationalizes flexibility is refined over time as the Planner infers the Doer's type from its choices. Almost-surely, this information leads the Planner to conclude that the Doer's type does not warrant flexibility, and therefore, the Planner eventually chooses to commit.[8]

[8] While such optimism disappears eventually, it may emerge in the short-run especially if individuals begin with priors that overestimate self-control.

What is the source of the friction that leads to perpetual overcommitment? The principal challenge that the Planner faces is that when he optimally chooses to commit (believing that he faces a Doer of type $\underline{\theta}$), his commitment choice leaves no opportunity for him to retain flexibility in case the Doer proves to be type $\bar{\theta}$. Thus, the Planner faces an experimentation challenge that traps learning with positive probability. Augmenting this environment with partial commitments offer an escape from this experimentation trap.

Consider the inclusion of a simple bond commitment contract. For simplicity, suppose that the individual has risk-neutral and additive preferences, and consider a one-period bond contract in which the Planner pays a lump-sum amount L and a prize x is returned to the decisionmaker if the Doer chooses $a = 1$. There are various ways to model how such a contract affects the Doer; for simplicity, suppose that a Doer of type θ chooses $a = 1$ if $\theta b + x$ exceeds the cost s. In this case, a fully informed Planner regardless of the Doer's type implements the first-best by setting x to equal $(1 - \theta)\, b$, and paying a prior lump-sum amount of $(1 - \theta)\, b^2$. This bond contract leads to zero expected transfers and induces the Doer of type θ to choose $a = 1$ whenever the Planner would like to do so.

Not only does such a bond contract allow the attainment of first-best, but it also helps a partially aware Planner learn efficiently. Regardless of the Planner' beliefs, it is now never optimal to restrict the Doer's choice set to a singleton: even when the Planner becomes quite confident that the Doer's type is $\underline{\theta}$, he is better off, even in the short-run, by partially committing through a bond contract. Over time, as the Planner uses bond contracts repeatedly, he can observe the Doer's behavior and therefore continue learning; almost-surely, his beliefs converge to the truth.

The possibility for partial commitments to dynamically improve learning has implications for commitment design. Recent years have seen the advent and study of commitment contracts that incentivize good behavior and mitigate self-control problems. The following section analyzes generally when partial commitments facilitate sophistication.

3 When is Learning Adequate?

The framework builds on the above example in several ways. The action-space can be binary or a continuum. The commitment technology can be one of two types: first, as in the prior section, the Planner can set incentives to *nudge* the Doer to take a particular action, and second, he can choose *menus* that restrict the Doer's choice. Both commitment technologies capture the essence of partial commitment, albeit in different forms, and the generality of this framework is useful in applying these results to distinct, but commonly studied, applications. The Doer's type is also drawn from a continuum. Finally, the Planner observes signals about past taste shocks.

As before, time is discrete and unbounded, with choices being made in period $t = 0, 1, 2, \ldots$. The Doer makes choices from a fixed, time, and history-invariant action set A: this is assumed to be a compact subset of $\Re$. In each period, a state s is realized that determines the Planner's preferences for that period: s is drawn from $\mathcal{S} \equiv [\underline{s}, \bar{s}]$ and is distributed i.i.d. with measure ν that has cdf $F(\cdot)$ and a continuous and strictly positive density f. The Planner's state-dependent preferences over actions, $u(a, s)$, are bounded, strictly quasi-concave in a for each state s, jointly continuous, and satisfy the strict single-crossing condition in (a, s). Accordingly, for each state s, the Planner has a most preferred action $a_P(s)$ which is non-decreasing in s. Let $\hat{a}$ denote the action that the Planner would choose if the Planner had to prescribe a single action for every contingency; for simplicity, I assume that $\hat{a}$ is unique, and I denote by $\hat{\pi}$ the expected payoff of this action.[9]

The Planner cannot directly choose actions but can choose commitment. The Planner has two forms of commitment: *menus* and *nudges*. A menu, M, is a non-empty closed subset of A, which restricts the Doer's choice of action. Let $\mathcal{F}$ be the collection of all non-empty closed subsets of A; this is the collection of all *feasible* menus. The Planner may or may not have access to all feasible menus: the set of menus from which the Planner can make a choice is $\mathcal{M}$, a closed subset of $\mathcal{F}$ that contains full flexibility ($M = A$) and the optimal full commitment ($M = \{\hat{a}\}$).[10] In addition to restricting choices through menus, the Planner can *nudge* the Doer into taking particular actions through incentives: the Planner can sign commitment contracts, make promises to peers that are costly to betray, or exert self-control. A nudge, η, directly affects the Doer's choices between actions in a state-dependent manner that is specified below. The set of nudges is $\mathcal{N}$, a closed subset of $[0, 1]$, which includes the case of no nudge ($\eta = 0$). I let $\mathcal{C} = \mathcal{M} \times \mathcal{N}$ denote the full set of commitments available, and $c = (M, \eta)$ denote a generic element.

As before, the Doer is a myopic agent that does not account for how its actions affect the Planner's choices. The Doer's behavior is governed by its type θ drawn from $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$. The Doer observes the state s and nudge η, and chooses the action from the menu M that maximizes the payoff function $W(a, s, \theta, \eta)$; when indifferent between two actions, the Doer chooses the action that the Planner prefers. The function W is continuous and satisfies the strict single-crossing property in (a, s), (a, θ), and (a, η), and

[9] In other words, $\hat{a}$ is the unique solution to $\max_{a \in A} \int_S u(a, s)\, d\nu$, and $\hat{\pi}$ is the associated expected payoff.

[10] Distance is defined by the Hausdorff metric. By Theorem 3.85 of Aliprantis and Border (2006), $\mathcal{F}$ is compact, and $\mathcal{M}$, being a closed subset of a compact set, inherits compactness.

for each (s, θ, η), W is strictly quasi-concave in a. The action selected by the Doer facing a commitment c is denoted by $a_D(s, \theta, c)$.

These assumptions ensure that the Doer's preferred action is increasing in the state s for a fixed type θ and nudge η. Thus, both the Doer and Planner prefer higher actions in higher states. For a fixed state, the Doer's preferred action is increasing in its type θ and in the nudge η chosen by the Planner. I assume that in the absence of all commitment, the Doer prefers to choose a lower action than the Planner.

Assumption 1. *The Doer is tempted to undertake lower actions than the Planner: $u(a, s)$ dominates $W(a, s, \theta, 0)$ by the single-crossing property for every type θ of the Doer.*[11]

Assumption 1 implies that the Planner knows the direction of temptation but is uncertain about its strength, which varies with the type of the Doer. Such an assumption is consistent with most applications of this model, e.g., in which temptation is driven by some form of present-bias, or identification of a tempting object. A special case of this framework, also consistent with most applications, accords the highest type, $\bar{\theta}$, with the Planner's preferences over actions, and thus, no temptation.

I denote a full information environment—in which the Planner knows the Doer's type—by $\Gamma = (\mathcal{S}, \nu, A, u, W, \Theta; \mathcal{C})$, a constellation of parameters that satisfy the above conditions. The Planner begins with a prior μ_0; throughout the paper, I restrict attention to priors μ_0 with a continuous and strictly positive density on Θ.

After each period t, the Planner observes a_t and obtains some information about s_t. The setting of Section 2 considered an extreme case in which the Planner learns nothing about the state; one could equally consider the other extreme in which the Planner observes the state perfectly. Spanning these two extremes is an *imperfect attribution* environment in which the Planner observes the state with noise. Generally, suppose that the Planner obtains a signal $\sigma_t(a_t, s_t)$ that comes from the space $\mathcal{S}$, and is distributed according to $\tilde{F}(\sigma|s, a)$.

For $t > 0$, h^t denotes the relevant history for the Planner in period t, which comprises past commitments, actions, and signals. The Planner's posterior belief about the Doer's type conditional on history h^t is denoted by μ_t.

I analyze two special cases of this model relevant for the applications herein.

Model 1—Actions are chosen from a binary set $A = \{-1, 1\}$ in which the *ex ante* optimal choice, $\hat{a}$, is 1. The Planner's payoff satisfies $u(-1, s) = 0$ for all s, and $u(1, s) =$

[11] In other words, for every type θ, state s, and pair of actions a and $a' > a$, $W(a', s, \theta, 0) - W(a, s, \theta, 0) \geq (>)\, 0$ implies $u(a', s) - u(a, s) \geq (>)\, 0$.

$U(s)$, where U is smooth, strictly increasing in s, and satisfies $U(\underline{s}) < 0 < U(\bar{s})$. The Planner can choose from all possible menus—$\mathcal{M} = \{\{-1,1\}, \{-1\}, \{1\}\}$—and can choose nudges from $\mathcal{N}$.

Model 2—Actions are chosen from a continuum $A = [\underline{a}, \bar{a}]$, and the Planner's and Doer's state-contingent payoffs satisfy the assumptions of the general model and are smooth. The Planner does not have access to nudges ($\mathcal{N} = \{0\}$), but has access to a closed set of menus, $\mathcal{M}$. For convenience, suppose that $\underline{a} = a_D(\underline{s}, \underline{\theta}, A, 0)$, and $\bar{a} = a_P(\bar{s})$. Two special cases of interest are those in which $\mathcal{M}$ is the set of all feasible interval menus, which I refer to as Model $2a$, and in which $\mathcal{M}$ comprises all feasible menus including non-intervals, which I refer to as Model $2b$.[12]

Model 1 generalizes the example in Section 2 to a continuum of types.[13] In the binary choice environment, menus without nudges offer the Planner a choice between commitment to a singleton menu or full flexibility. Nudges introduce partial commitment by allowing the Planner to influence how the Doer chooses actions in each state. Apart from its usefulness in understanding the impact of incentive-based commitment mechanisms, this model provides insight into how a decisionmaker learns to exert internal self-control and willpower, as studied in Section 4.2

Model 2 offers a continuous action space in which the Planner commits by restricting the Doer's choices and has no other way to influence the Doer. This framework is particularly applicable to study savings environments, as I demonstrate in Section 4.1, in which the Planner can purchase illiquid assets to commit the Doer to some minimal savings but still retain some flexibility for the Doer to save more than that minimum.

Since the applications of interest in the literature satisfy either Model 1 or Model 2, I characterize learning only for these two special cases. In both cases, the Planner faces an exercise of optimal experimentation similar to but richer than a standard multi-armed bandit. Naturally, my work builds on insights derived from the experimentation literature, particularly Easley and Kiefer (1988), and Aghion et al. (1991).

[12]Specifically, Model $2a$ corresponds to $\mathcal{M} = \mathcal{F}_I = \{[\alpha, \beta] : \alpha, \beta \in A \text{ and } \beta \geq \alpha\}$, and Model $2b$ corresponds to $\mathcal{M} = \mathcal{F}$.

[13]Section 2, restricted to two types, corresponds to $S = [-1, 0]$, $u(-1, s) = 0$ for all s, and $u(1, s) = b + s$, where $b \in \left(\frac{1}{2}, 1\right)$. The Doer's payoff is $W(a, s, \theta, \eta) = \mathbf{1}_{\{a=1\}}(\theta b + \eta + s)$ and thus, the ideal action $a_D^*(s, \theta, \eta)$ takes on a value of 1 if $0 \leq \theta b + \eta + s$ and -1 otherwise. In the absence of the bond contract, the set of nudges is $\mathcal{N} = \{0\}$, but with the bond contract, the set of nudges is expanded to $\mathcal{N}' = [0, (1 - \underline{\theta}) b]$.

3.1 Adequate Vs. Partial Learning

The central issue analyzed is whether the Planner eventually attains the same payoffs that he would in the full information benchmark in which the Doer's type is known. Answering this question sheds light on when an assumption of sophistication can be justified through long-run learning.

The full information payoff is defined as follows: for each type θ, when the Planner selects commitment c, his expected static payoff is $\pi(\theta, c) = \int_S u(a_D(s, \theta, c), s)\, d\nu$. Given a Doer of type θ, denote the Planner's *full information payoff* by $\pi^*(\theta) = \max_{c \in \mathcal{C}} \pi(\theta, c)$, and let $C^*(\theta)$ denote the set of commitments that attain this payoff. This payoff corresponds to the Planner's optimal balance between flexibility and commitment when the Planner knows the Doer's type; its existence is guaranteed by continuity and compactness assumptions.

When the Planner is uncertain about θ and has beliefs μ, he optimizes with respect to his beliefs. Denote the payoff in a single period from a commitment choice c by $m(\mu, c) = \int_\theta \pi(\theta, c)\, d\mu$; however, when the Planner is even slightly patient, he also values learning. Given a commitment choice c and prior μ, the Planner may revise his beliefs when observing the Doer's chosen action and the realized signal; $Q(\mu, c)$ denotes the probability distribution over posteriors induced by (μ, c), and $P(\Theta)$ denotes the set of Borel probability measures on Θ. The Planner solves the Bellman Equation,

$$V(\mu; \delta) = \max_{c \in \mathcal{C}} \left\{ (1-\delta)\, m(\mu, c) + \delta \int_{P(\Theta)} V(\tilde{\mu}; \delta)\, dQ(\mu, c) \right\}. \tag{2}$$

The Planner's beliefs, commitment choices, and therefore, value function evolve stochastically; the object of interest is their long-run distribution conditional on the true θ. We can partition the set of types based on whether learning is necessarily efficient, yielding the full information payoffs in the long-run.

Definition 1. *Learning is **adequate** for a type θ if the Planner's payoffs eventually converges to the full information benchmark with probability* 1*:*

$$\Pr\left(\lim_{t \to \infty} V(\mu_t; \delta) = \pi^*(\theta) \,|\theta \right) = 1,$$

*and otherwise, learning is **partial** for type θ. Learning is **globally adequate** for a prior μ_0 and discount factor δ if the set of types for which learning is partial has μ_0-measure* 0*, and otherwise, learning is **inadequate** for that prior and discount factor.*

Adequate learning is weaker than *complete learning* insofar as the eventual belief need not identify the type but the Planner nevertheless obtains the same payoffs as in the full information benchmark. Globally adequate learning requires more than the mere existence of strategies that ensure learning; indeed, it must be that at least one such strategy is optimal for the Planner.

3.2 Full Commitment Distinguishability

The possibility for globally adequate learning relies critically on the commitment set, and one can assess if it is conducive to learning by examining how a fully informed Planner behaves. Before deriving the precise condition, I offer a heuristic explanation of its connection to globally adequate learning. Whenever the Planner remains partially flexible so that the Doer can choose different actions in different states, the Planner eventually learns the Doer's type from the empirical frequency of its action choices. It is only when the Planner chooses to constrain the Doer to a single action in all states that learning is possibly impeded. Since it may be optimal to fully commit some types of the Doer to choose $\hat{a}$, the Planner can perpetually *distinguish* these types from others if there is a way to fully commit those types while retaining flexibility for others. The presence of such expedient commitment is formalized as the property of *Full Commitment Distinguishability*.

Formally, let $\hat{C}(\theta) = \{c \in \mathcal{C} : a_D(s, \theta, c) = \hat{a} \text{ for almost all } s\}$ denote the set of commitments that induce a Doer of type θ to choose the action $\hat{a}$ with probability 1; for each type, this set is non-empty because $\mathcal{C}$ includes the singleton menu $\{\hat{a}\}$. The set of types for which a fully informed Planner finds it optimal to fully commit to the action $\hat{a}$—and relinquish flexibility altogether—is denoted by $\hat{\Theta} = \{\theta \in \Theta : \pi^*(\theta) = \hat{\pi}\}$.

Definition 2. *An environment* Γ *satisfies* ***Full Commitment Distinguishability*** *(FCD) if for almost every* $\hat{\theta}$ *in* $\hat{\Theta}$ *and every* θ *not in* $\hat{\Theta}$*,* $\hat{C}(\hat{\theta}) \nsubseteq \hat{C}(\theta)$*.*

Under FCD, the Planner can carefully select commitment so as to fully commit almost every "bad type" in $\hat{\Theta}$ while retaining partial flexibility for some "good type" outside $\hat{\Theta}$. A sufficient but unnecessary condition for FCD is that a fully informed Planner never chooses full commitment. This strong form of FCD manifested in Section 2 when the bond commitment contract was introduced. An environment fails FCD if the only commitments that fully commit a strictly positive measure of types in $\hat{\Theta}$ also fully commit some types not in $\hat{\Theta}$. Since the Planner's preference for commitment is determined by how the Doer maps taste shocks into actions, the particular representation of the Doer's preferences do not affect whether an environment satisfies FCD.

To clarify what is being assumed in FCD, I detail when the condition holds in each of these two models below. These details are not entirely necessary to understand the connections drawn between FCD and globally adequate learning, and so some readers may wish to skip to Section 3.3.

3.2.1 When Does FCD Hold?

Consider the binary action environment of Model 1: because W satisfies the single-crossing condition in (a, θ), any commitment that induces the Doer of type θ to always select $a = 1$ also induces all higher types to choose that higher action in every state. Thus, FCD can be satisfied only if whenever a type θ is in $\hat{\Theta}$, all higher types are also in $\hat{\Theta}$.

Observation 1. *In Model* 1*, an environment* Γ *satisfies FCD if and only if for almost all* θ *in* $\hat{\Theta}$*, all higher types* θ' *are also in* $\hat{\Theta}$*.*

The following examples highlight aspects of FCD in Model 1: let the Planner's payoff from $a = 1$ in period t be $b + s_t$ in which $b = \frac{3}{4}$ and s_t is drawn uniformly from $S = [-1, 0]$, and recall that the Planner's payoff from $a = 0$ is normalized to 0. The Doer's type is drawn from $\Theta = \left[0, \frac{1}{2}\right]$ and the Doer chooses $a = 1$ if and only if $\theta b + s + \eta$ is positive. The Planner selects nudges η from $\mathcal{N}$, and depending on $\mathcal{N}$, FCD is satisfied or violated. In each of these examples, the Planner has access to all possible menus but does at least as well by choosing the menu $\{-1, 1\}$ and using the nudges specified below.

Example 1 (FCD with no commitment)**.** *Suppose* $\mathcal{N} \supseteq \left[\frac{3}{8}, \frac{3}{4}\right]$*: then a fully informed Planner sets* $\eta^*(\theta) = (1 - \theta)\, b$ *to align every type's preference perfectly with the Planner. Since* $\hat{\Theta} = \emptyset$*, FCD is trivially satisfied.*

Example 2 (FCD with commitment)**.** *Suppose that* $\mathcal{N} = \{0\} \cup \left[\frac{3}{4}, 1\right]$*: then a fully informed Planner uses a nudge of* $\frac{3}{4}$ *for each type. Such a commitment choice offers flexibility to all types below* $\frac{1}{3}$ *and fully commits all higher types. Thus,* $\hat{\Theta} = \left[\frac{1}{3}, \frac{1}{2}\right]$ *and satisfies FCD.*

Example 3 (FCD not preserved under expansion)**.** *Suppose that the nudge* $\left\{\frac{3}{16}\right\}$ *is added to* $\mathcal{N}$ *above: then a fully informed Planner uses a nudge of* $\frac{3}{16}$ *for types in* $\left(\frac{5}{12}, \frac{1}{2}\right]$*, thereby offering flexibility to those types, and a nudge of* $\frac{3}{4}$ *for any other type. FCD is violated because* $\hat{\Theta} = \left[\frac{1}{3}, \frac{5}{12}\right]$*.*

Examples 2 and 3 highlight a distinction between FCD and standard notions of richness: FCD in Model 1 is not necessarily preserved when $\mathcal{N}$ is expanded.

In Model 2, it is difficult to offer a simple criterion for FCD generally. One can simplify having to check every potential pair of types $\hat{\theta}$ in $\hat{\Theta}$ and θ not in $\hat{\Theta}$: FCD holds if and only if it is possible to fully commit the type with the least self-control ($\underline{\theta}$) while retaining flexibility for any type not in $\hat{\Theta}$. FCD simplifies further in the special cases of Models 2*a*-*b* in which $\mathcal{M}$ comprises all feasible interval menus or all feasible menus respectively. In these cases, a Planner has no reason to fully commit a Doer that would freely pick actions higher than $\hat{a}$ in some states: by Assumption 1, the Doer chooses actions higher than $\hat{a}$ only in those states that the Planner prefers higher actions to $\hat{a}$. Thus, a Planner is better off by offering the menu $[\hat{a}, \bar{a}]$ rather than committing the Doer to choose $\hat{a}$ in every state. So the set of types for which full commitment is optimal is a subset of those that choose actions less than $\hat{a}$ in every contingency. In this context, FCD is satisfied if and only if these two sets of types are identical.

Proposition 2. *An environment Γ satisfies FCD in Model 2 if and only if $\hat{\Theta}$ has zero measure or for every θ not in $\hat{\Theta}$, it is the case that $\hat{C}(\underline{\theta}) \nsubseteq \hat{C}(\theta)$. In Models 2a-b, FCD is satisfied if and only if the set of types for which full commitment is optimal, $\hat{\Theta}$, coincides with the set of all types whose preferred action is below $\hat{a}$ in every state.*

3.3 Main Results

I describe the connection between FCD and globally adequate learning: FCD is sufficient for globally adequate learning and may also be necessary in particular cases. When FCD is satisfied, these results offer Bayesian learning foundations for sophistication. Simultaneously, the results indicate that individuals may not be fully sophisticated in environments in which FCD fails, and that interventions may help even experienced individuals develop a better understanding of their temptations.

I begin by demonstrating sufficiency.

Theorem 2. *If an environment Γ satisfies FCD, then for all priors μ_0 and discount factors δ, learning is globally adequate in Models 1 and 2.*

When FCD is satisfied, even a slightly patient Planner eventually chooses commitment as if he were fully sophisticated about his imperfect self-control. Moreover, under FCD, globally adequate learning may obtain even if the Planner is not sufficiently forward-looking to actively learn and experiment: a *passive learner* who myopically best-responds to his beliefs eventually behaves as if he perfectly understood his temptation. As such, the introduction of partial commitments that induce FCD can have strong implications for individuals regardless of their patience or prior.

The essence of the argument is that if FCD is satisfied, learning has no short-term costs but offers short-term and long-term benefits; in other words, the commitment structure facilitates cheap experimentation. Under FCD, the Planner can select commitments that fully commit those types in $\hat{\Theta}$—for which full commitment is optimal—while retaining partial flexibility to some types not in $\hat{\Theta}$. By choosing commitment in this way, the Planner can learn about the Doer based on the empirical frequency of its choices. Should the Doer choose different actions over time, that frequency identifies its type. On the other hand, if the Doer chooses the same action in every period, the Planner learns that the Doer's type is one that has no flexibility at that commitment choice. Based on this updated belief, the Planner can change his commitment choice so as to still guarantee that types in $\hat{\Theta}$ fully commit while offering partial flexibility to types not in $\hat{\Theta}$ that remain in the support of his posterior. By iterating this procedure, the Planner eventually makes choices that converge to those in the full information benchmark.

Notice that such an argument applies independently of the accuracy with which the Planner observes past states. The signal $\sigma_t(a_t, s_t)$ could be a noisy signal of the state s_t or uninformative altogether and globally adequate learning is nevertheless ensured. Insofar as there are substantive reasons to consider noisy attribution structures, and prior work has focused on such cases,[14] it is noteworthy that such noise does not prevent long-run learning in the presence of FCD, although it likely interferes with the speed of learning.

In the absence of FCD, challenges of costly experimentation emerge. When the cost of experimentation outweighs the benefits of learning, the Planner may choose to sacrifice learning for the sake of short-term payoffs if he is not perfectly patient. Since sets of menus can assume many different forms in Model 2, it is challenging to demonstrate the necessity of FCD generally. Accordingly, I demonstrate the necessity of FCD in the special cases of Model 2*a*-*b*, in which $\mathcal{M}$ comprises all feasible interval menus or all feasible menus respectively.

Theorem 3. *Suppose that an environment Γ fails FCD. Then for every discount factor $\delta < 1$, learning is inadequate for some open set of priors in Models 1 and 2a-b.*

The inadequacy of learning highlights that when FCD fails, a modeler cannot preclude the possibility of partial learning. When learning fails, it does so in a systematic way: the Planner fully commits a Doer type that he would have offered flexibility had he

[14] Noisy attribution has featured in, for example Benabou and Tirole (2004), where the assumption is that a decisionmaker learns about his self-control using a "revealed preference" approach. One reason to study noise in attribution is that if past circumstances and taste shocks are representations of subjective states and emotions, it may be difficult for an individual with imperfect self-control to obtain perfect information about past circumstances. Various studies highlight the difficulty that individuals face in separating situational factors from aspects of one's innate character (Baumeister et al., 1994; Kahneman et al., 1997; Ameriks et al., 2003). Moreover, if the Doer observe a noisy signal of s_t and the Planner observes s_t but not the signal of the Doer, the imperfect monitoring model that emerges is equivalent to a noisy attribution framework.

been fully informed. The likelihood with which learning fails necessarily depends on the discount factor, the Planner's prior, and the informativeness of the feedback that the Planner receives; I show in Theorem 4 below that such inefficiencies vanish as the Planner approaches perfect patience.

3.4 Remarks

3.4.1 The Role of Patience

Posing self-awareness as an issue of experimentation permits the study of the connection between the decisionmaker's patience and his sophistication. As the Planner becomes more patient, he increasingly values sacrificing current payoffs to gain information that helps improve commitment choices in the future. This logic suggests that a more patient Planner is less likely to fully commit when the Doer's type warrants flexibility, and by the same token, shall experiment more even if the Doer's type warrants full commitment. I offer a limit result in this vein.

Theorem 4. *As the Planner approaches perfect patience, his payoffs and choices approach the full information benchmark with probability* 1 *regardless of whether FCD is satisfied: for every prior* μ_0, $\lim_{\delta \to 1} V(\mu_0; \delta) = \int_{\Theta} \pi^*(\theta)\, d\mu_0$.

Theorem 4 follows from the observation that if a fully informed Planner is forced to select from a finite number of commitments $\mathcal{C}_{\mathcal{N}}$ in $\mathcal{C}$, his payoff loss from this restriction can be made arbitrarily small regardless of the Doer's type. Analogous to results derived elsewhere in the experimentation framework,[15] the Planner can then approach the full information payoffs (restricted to a finite set of commitment choices) by conducting a large number of experiments and then forever selecting what appears to be the best commitment thereafter.

This result indicates the virtue of patience when it comes to learning; a patient Planner tolerates many lapses of self-control so as to optimally commit in the future. Generally, it is ambiguous as to whether the increased experimentation takes the form of flexibility for a greater number of periods or greater flexibility per period. However, for the simplest commitment setting studied in Section 2, this result would suggest that the patient Planner will choose to remain fully flexible for many periods even when he sees the Doer choosing $a = 1$ with low probability. Accordingly, it would be difficult to distinguish between naive and patient Planners merely on the basis of short-run commitment choices.

[15] For example, see Aghion et al. (1991) and Fudenberg and Levine (1993b).

3.4.2 Identifying Awareness

The necessary and sufficient condition for adequate learning, FCD, is defined in terms of how a fully sophisticated Planner copes with the full range of possible temptations. As in the applications in Section 4, it is relatively straightforward to check this condition in full information models that we typically use; in many such contexts, specifying the full information model of temptation and commitment implicitly assumes FCD or its absence.

When a modeler or policy maker does not have a particular full information model, checking for FCD may be challenging since how a decisionmaker copes with different levels of temptation is not directly observed. Nevertheless, within some settings, an individual's long-run preference for commitment can help identify his eventual awareness.

To illustrate, if a decisionmaker strictly prefers some non-singleton menu c to every singleton menu, then the decisionmaker is partially committing and therefore must have adequately learned. In the opposite case, if the decisionmaker weakly prefers some singleton menu to every other menu, then he is fully committing and so he may or may not be sophisticated. In this case, the policy maker can learn about the decisionmaker's awareness by *exposing* him to different commitments and observing how he responds. If temporary subsidies to commitments not in his most preferred set have long lasting implications on the decisionmaker's commitment choices, then learning was inadequate.[16]

3.4.3 Relationship to Self-Confirming Equilibrium

The dynamic process of learning that I have studied bears a resemblance to steady-state solution concepts motivated by learning, in particular to the concept of *self-confirming equilibrium* (SCE), developed by Fudenberg and Levine (1993a). One way to view these results is as foundations for studying SCE (or a variant thereof) in dynamic self-control environments. In this section, I compare the limiting behavior and beliefs of the dynamic model with those that emerge from the steady-state concept.

In an SCE, players do not deviate given their beliefs about how others behave, and each player's belief is consistent with experiences on the path of play; unlike a Nash equilibrium, these beliefs may err on how others *would* behave off the equilibrium path. Adapting this concept to the framework here requires additional restrictions: the Planner's beliefs both on and off the path of play must be compatible with his knowledge

[16] Consistent with this possibility, temporary subsidies to drug rehabilitation are often believed to have effects after the subsidy has ended (Kaper et al., 2005). Given the experimentation challenge that individuals face in finding the appropriate commitment devices and therapy for addiction, it is natural that learning is inadequate.

of the Doer's type space, i.e., how different levels of temptation lead to different distributions of behavior. Encoding these restrictions yields an adaptation of *rationalizable self-confirming* equilibria (RSCE) to Bayesian games: an RSCE refines the set of SCE by imposing the further restriction that any player's belief about others' actions, on or off the path of play, must be rationalized by her knowledge of their payoffs.[17]

Thus, the appropriate analogue to the framework here is a Bayesian game in which: (a) Nature chooses the Doer's type, which is fixed once and for all, (b) the Planner and Doer have payoff functions specified above, (c) the Planner knows Nature's distribution of states s and how the Doer chooses actions as a function of its type, and (d) the Planner observes only the action chosen by the Doer and the signal $\sigma(a, s)$ at the end of each round. An RSCE of this game consists of a commitment choice c rationalized by some consistent belief μ about the Doer's type, and θ, the actual type of the Doer. The Planner's belief μ is *consistent* with the path of play if the distribution over actions and signals generated by the commitment choice c and the Doer's type θ coincide with what the Planner expects given belief μ.[18] The following describes the RSCE of this setting.

Observation 2. *The set of RSCE in which the Doer's type is θ comprises:*

1. *Complete learning: a commitment c in $C^*(\theta)$, rationalized by the belief that ascribes probability 1 to θ.*
2. *Full commitment: a commitment c rationalized by some belief μ, such that c induces type θ and all types in the support of μ to fully commit, i.e., c is in $\hat{C}(\theta) \cap \left(\cap_{\theta' \in Supp(\mu)} \hat{C}(\theta')\right)$.*

Since the Planner's beliefs must be consistent with observations from repeated play, if the Planner retains some flexibility for the Doer, his steady-state beliefs must ascribe probability 1 to the true type of the Doer (because different Doer types respond differently to flexibility). The Planner may also hold a belief that rationalizes full commitment, which precludes the possibility for further learning. In contrast to the dynamic setting, full commitment may be chosen even if FCD is satisfied and the Doer's type does not warrant it. A simple illustration of this RSCE is one in which the Planner assigns probability 1 to $\hat{\Theta}$ and chooses the singleton menu $\{\hat{a}\}$. Regardless of the Doer's actual type, these potentially incorrect beliefs are consistent with experiences on the path of play and rationalize the commitment choice. Because Theorem 2 demonstrates that such outcomes cannot emerge in the dynamic framework unless the Doer's type is indeed

[17] RSCE is developed by Dekel et al. (1999). Dekel et al. (2004) adapt SCE to Bayesian games, including those in which a player's type is fixed once and for all; however, they do not study RSCE in this setting. Fudenberg and Levine (2009) offer a survey of this literature.

[18] Formally, when the Planner chooses commitment c, and the Doer's type is θ, consistency of μ demands that if θ' is in the support of μ, then $\nu(s : a_D(s, \theta, c) \neq a_D(s, \theta', c)) = 0$.

in $\hat{\Theta}$, this inconsistency suggests that the limits of learning in the dynamic framework correspond to a particular selection of RSCE.

The source of this tension is that an RSCE permits the Planner to ascribe zero probability to the truth in contrast to the dynamic framework in which the Planner begins with a full support prior, and almost-surely includes the truth within the support of his limit beliefs. Selecting those RSCE that contain a "grain of truth" is sufficient to reconcile the two approaches. Define an RSCE to be *absolutely continuous (AC-RSCE)* if the commitment choice c is rationalized by at least one belief that is both consistent and includes in its support the Doer's actual type θ. Although the full implications of this selection are better left for future work, it captures the implications of FCD in a steady-state setting.

Theorem 5. *If an environment Γ satisfies FCD in Models 1 and 2, then in every AC-RSCE in which the Doer's type is θ, the Planner's commitment choice is from $C^*(\theta)$ yielding the full information payoffs, regardless of θ. If an environment Γ fails FCD in Models 1 and 2a-b, then there exist AC-RSCE for a strictly positive measure of θ in which the Planner fails to attain full information payoffs.*

4 Applications

4.1 Consumption-Savings

The canonical application of self-control models is to consumption and savings decisions. Evidence suggests that individuals are tempted by immediate consumption and a recent literature has studied how access to an illiquid asset offers a powerful instrument for commitment to such individuals.[19] Holding wealth in illiquid form effectively sets a minimal level of savings and thereby protects some wealth from consumption splurges. At the same time, excessive illiquidity impedes a decisionmaker from responding to uninsurable risk, and so a delicate balance between commitment and flexibility must be maintained. The burgeoning literature on these issues has largely focused on decisionmakers who are fully sophisticated about their imperfect self-control. While this is the natural starting point, it raises the question as to whether experience would lead a decisionmaker who is initially uncertain about his self-control to eventually act as if he were fully sophisticated.

[19] The role of illiquid assets as commitment was discussed in some of the earliest work on self-control—e.g., Strotz (1955) and Thaler and Shefrin (1981)—and studied in a number of subsequent papers, including Laibson (1997), Barro (1999), Gul and Pesendorfer (2005), Fudenberg and Levine (2006), and Amador et al. (2006). Harris and Laibson (2003) offer a survey and discussion.

This section answers the question in the affirmative in this canonical setting: when a decisionmaker is uncertain about his tendency to overconsume and has access to an illiquid asset, he eventually makes commitment choices as if he were fully informed. I demonstrate this positive learning result first in the setting studied in the prior literature—in which commitment corresponds to the purchase of illiquid assets—and then show that it also extends to a setting in which singleton consumption sets are also available.

The simple savings environment that I study is a decisionmaker who faces a tradeoff between commitment and flexibility in each period. His expected payoff from a consumption stream $\{c_t\}_{t=0}^{\infty}$ is $E\left[\sum_{t=0}^{\infty} \delta^t u_t U\left(c_t\right)\right]$, where U is a CRRA utility function with a coefficient of relative risk aversion $\sigma > 0$, and u_t is a taste shock. The taste shock in each period affects the marginal utility of current consumption and follows an i.i.d. process represented by a continuous density on $[\underline{u}, \bar{u}]$ with $\underline{u} > 0$, and normalized so that $E\left[u\right] = 1$. The Planner begins with an initial wealth y_0, and subsequent wealth is generated by the returns from savings in the prior period at gross interest rate R. To ensure that the transversality condition is satisfied, I assume that $\delta R^{1-\sigma} < 1$. The restriction to CRRA utility ensures that in the absence of a self-control problem, a decisionmaker would have a consumption *rate* that depends only on the taste shock: in period t, if he has wealth y_t and faces a taste shock u_t, he consumes $c_P\left(u_t\right) y_t$.[20]

In this dual self framework, consumption choices are made as follows: each period is divided into two sub-periods. Wealth at the beginning of the period is denoted by y_t which can be invested by the Planner in illiquid assets or kept in a liquid form in the first sub-period. The Planner effectively selects a minimal savings rate, s_t, by investing an amount $s_t y_t$ in illiquid assets. In the second sub-period, the taste shock is realized, then the Doer selects consumption c_t from the feasible set—which is $[0, (1-s_t) y_t]$—and invests the remainder in the asset. For expository convenience, assume that returns from saving occur at the same rate regardless of whether the choice was made by the Planner or Doer.

As before, the Planner has the standard long-run preferences whereas the Doer is a short-run behavioral type that is tempted by immediate consumption. Formally, a Doer of type θ will aim to consume $c\left(u,\theta\right) y$ when total wealth is y and selects the feasible consumption closest to this ideal. Doer types θ are in $[\underline{\theta}, 1]$ drawn according to μ_0. I assume that $c\left(u,\theta\right)$ is strictly positive, smooth in (u,θ), strictly decreasing in θ, and strictly increasing in u; morever, $c\left(u,1\right)$ coincides with the Planner's ideal, $c_P\left(u\right)$. Accordingly, for every taste-shock, every Doer type has a constant saving rate that is

[20] A guess and verify approach to solving the stochastic Euler equation yields that $c_P\left(u\right) = \frac{u^{1/\sigma} C}{u^{1/\sigma} C + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}}$, where C solves $E\left[\left(C u^{1/\sigma} + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}\right)^{\sigma}\right] = 1$.

(weakly) less than the Planner's optimal saving rate and strictly increasing in the Doer's type. One special case is a linear specification $c(u,\theta) = \left(\frac{1}{\theta}\right) c_P(u)$; another special case inheres to standard formulations of present-bias in which $c(u,\theta)$ corresponds to how the Planner would have chosen had his discount factor been $\theta\delta$.[21]

Since my goal is to establish the inevitability of globally adequate learning, I make the most stark of informational assumptions: the Planner observes only past consumption and not past taste shocks. Since learning emerges here, it continues do so in richer observational environments.

The setting here is similar to Model 2a of Section 3 in which partial commitment takes the form of interval menus. By investing in illiquid assets, the Planner specifies a consumption cap (or a savings floor), and for every taste shock, the Planner would wish to consume less than the Doer. The contrast to Section 3 emerges in the introduction of a state variable other than beliefs, namely the Planner's wealth, although the restriction to CRRA utility affords tractability. Although the set of feasible consumption choices changes over time, the preferred savings and consumption rates for the Planner and Doer are time and history invariant. To apply the methods developed earlier, one can formulate actions as savings rates and states as $(u)^{-1}$; with such transformations, both the Planner and Doer prefer higher savings rates in higher states, the Doer's preferred savings rate in each state is increasing in its type, and the Planner prefers a higher savings rate than the Doer in every state.

Directly solving the model in which the Planner is uncertain about the Doer's temptation is challenging because the specification of any menu induces different separating and bunching regions for different types of the Doer. Accordingly, the selection of a menu affects the Planner's learning in intricate ways. The tools developed in Section 3 sidestep these complications altogether: I demonstrate that the simpler full information model (in which the Planner knows the Doer's type) satisfies FCD, and can then apply Theorem 2. Verifying FCD takes the following steps.

First, suppose that the Planner could not offer any flexibility and were forced to commit to a particular consumption path irrespective of taste shocks; this selection corresponds to finding the analogue of the optimal full commitment $\hat{a}$ from Section 3. The optimal binding consumption path in this case is one in which the Planner ensures that the fraction $\hat{s} = \delta^{1/\sigma} R^{(1-\sigma)/\sigma}$ is saved in each period, and consumes the remainder.

Second, I solve for how the Planner resolves the tradeoff between commitment and flexibility when he knows that the Doer is of type θ. In the Online Appendix, I demon-

[21]Specifically, $c(u,\theta) = \frac{u^{1/\sigma} C(\theta)}{u^{1/\sigma} C(\theta) + (\theta\delta)^{1/\sigma} R^{(1-\sigma)/\sigma}}$, where $C(\theta)$ solves $E\left[\left(C u^{1/\sigma} + (\theta\delta)^{1/\sigma} R^{(1-\sigma)/\sigma}\right)^{\sigma}\right] = 1$.

strate that he purchases the same fraction of illiquid assets, $s^*(\theta)$, in each period so as to induce some commitment. Accordingly, the Doer can never consume more than $1 - s^*(\theta)$ of the current wealth, but can choose to consume less when consumption is less valuable. Without temptation towards present consumption ($\theta = 1$), the Planner chooses to keep all wealth liquid and affords the Doer maximal flexibility; otherwise, the Planner may use a strictly positive level of illiquid assets as commitment. If it so happens that for all taste shocks, the Doer wish to consume more than the amount of wealth that is kept in a liquid form, the Doer is fully committed to a particular consumption path.

Using the above characterizations, one can then partition Θ into $\hat{\Theta}$, the set of types that a fully informed Planner chooses to leave with no flexibility, and its complement, to wit, types that the Planner offers (partial) flexibility. In the Online Appendix, I demonstrate that $\hat{\Theta}$ comprises all types that consume at least $(1 - \hat{s})\, y_t$ for every realization of u_t; to effectively fully commit such types, the Planner only keeps that level of wealth liquid, and holds illiquid assets of $\hat{s} y_t$ in each period. In contrast, if the Doer consumes less than $(1 - \hat{s})\, y_t$ for some taste shocks, the Planner benefits from maintaining some flexibility. When facing this milder temptation to overconsume, the Planner purchases a smaller holding of illiquid assets and thereby retains greater flexibility for the Doer. Thus, the types that are fully committed, $\hat{\Theta}$, corresponds to the set $\{\theta \in \Theta : c\,(\underline{u}, \theta) \geq 1 - \hat{s}\}$.

Figure I illustrates these consumption caps. If a Planner faces a Doer of type θ, he sets the consumption cap so as to permit the Doer to select its most preferred consumption for all taste shocks $u < \hat{u}$, and constrains the Doer to consume $(1 - s\,(\theta))$ otherwise. In contrast, when facing a Doer of type $\hat{\theta}$, who has less self-control, the Planner's optimally chosen consumption cap binds the Doer for every taste shock, inducing it to select $1 - \hat{s}$ for every taste shock; this latter type is in $\hat{\Theta}$.

It is straightforward to see how FCD is satisfied in this setting: when the Planner purchases $\hat{s} y_t$ of illiquid assets in period t, he simultaneously fully commits every type in $\hat{\Theta}$ while retaining flexibility for types not in $\hat{\Theta}$. As shown in Figure I, when choosing from this menu, a Doer of type θ' consumes the remaining wealth with probability 1, whereas the type θ chooses below the consumption cap with strictly positive probability. Thus, even when selecting this commitment, the Planner can distinguish types in $\hat{\Theta}$ (for which such full-commitment is optimal) from types not in $\hat{\Theta}$ ensuring globally adequate learning.

Theorem 6. *Learning is globally adequate in this consumption-savings framework for all priors and discount factors.*

This positive learning result demonstrates that if a decisionmaker begins with uncertainty about his self-control, eventually, he purchases the same fraction of illiquid assets

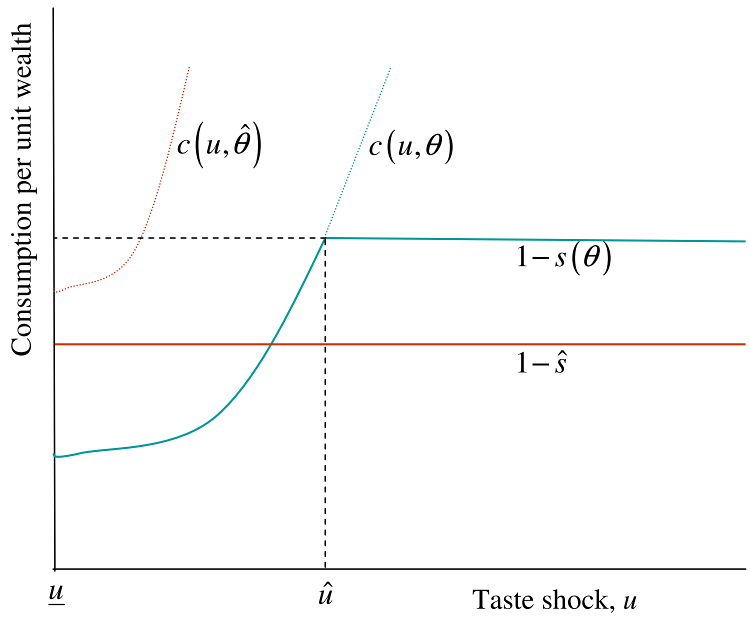

Figure I: Consumption Caps

Dotted curves indicate the Doer's ideal consumption for each taste-shock, and solid lines indicate the Doer's actual consumption when the Planner selects commitment optimally.

that he would if he could forecast his temptations perfectly. Thus, assuming sophistication in this important setting can be justified by Bayesian learning. Of course, while learning, the Planner may make (*ex post*) inefficient choices that adversely affect his future income distribution relative to the full information benchmark.

The commitment mechanism studied so far in this application is empirically relevant to savings decisions, and has been the focus of prior work. However, an inconsistency with the model of Section 3 is that this mechanism excludes singleton menus, which fully commit *every* type of the Doer. Establishing that the above results extend to a setting that includes both illiquid assets and singleton menus requires additional analysis; in the Online Appendix, I demonstrate that learning is globally adequate in this richer setting if the coefficient of relative risk aversion, σ, is at least 1.

The full information model to which behavior eventually converges shares many predictions of sophistication in the quasi-hyperbolic savings model but is analytically distinct. It has a unique solution and a constant commitment rate with CRRA utility (similar to Fudenberg and Levine 2006). In contrast, the quasi-hyperbolic savings model in discrete-time has a continuum of sophisticated solutions, even after assuming Markovian behavior (Krusell and Smith, 2003). Thus, equilibrium selection remains a challenge in the quasi-hyperbolic model thereby making it difficult to study learning in that model but does not pose an issue here.

4.2 Costly Self-Control

Fudenberg and Levine (2006) offer a simple but powerful Planner-Doer framework in which the Doer's preferences are completely characterized by the present and the Planner exerts costly self-control so as to induce the Doer to take actions more consistent with the Planner's preferences. Here, I analyze how a Planner who is uncertain about the cost or efficacy of self-control learns over time how much self-control to use, building on the binary action environment studied in Model 1 of Section 3.

Consider a decisionmaker who faces an infinite sequence of tasks. Completing a task in any period ($a_t = 1$) leads to a *future* benefit with a present value of b, which is not valued by the Doer. Not completing the task ($a_t = -1$) has the benefit of immediate leisure s_t, where s_t is uniform drawn from $[0, 1]$. I assume that b is in $\left(\frac{1}{2}, 1\right)$, generating a tradeoff between commitment and flexibility. Prior to the realization of s_t, the Planner chooses a menu from $\{-1, 1\}$. After the realization of s_t, the Planner can exert costly self-control to nudge the Doer to complete the task. If the Planner exerts self-control of magnitude $\Gamma \geq 0$, self-control is *successful* for a Doer of type θ if Γ is at least $\dfrac{s_t}{\theta}$; otherwise, it *fails*. The Planner's payoff from completing the task at date t is $b - \Gamma$ and the payoff from not completing the task is $s_t - \Gamma$. I assume that θ is drawn from an interval $\left[\underline{\theta}, \bar{\theta}\right]$ in which $\underline{\theta} > 0$; lower types require the Planner exert greater self-control to complete the task, whereas higher types require less self-control.

The full information model, in which the Planner knows θ, is similar to that considered by Fudenberg and Levine (2006). If θ is sufficiently high, then the Planner chooses the menu $\{-1, 1\}$ and has an associated cutoff s^* such that he uses costly self-control to complete the task in period t if and only if s_t is less than s^*. Such a choice has the benefit of flexibility, since the Planner can choose whether to exert costly self-control depending on the value of leisure. In contrast, when the cost of self-control is sufficiently high, the Planner may prefer to commit through the singleton menu. While such a choice relinquishes flexibility, it obviates costly self-control, since the Doer no longer needs to be nudged to complete the task. The Planner prefers flexibility to full commitment if θ^* is at least $\frac{(1-b)^2}{2b-1}$.

When the Planner is uncertain about the cost of self-control, he may use too much or too little self-control to influence the Doer. Employing too little self-control leads the agent to not complete the task while the costs of self-control are still borne. In contrast, too much self-control guarantees that the task is completed but at unnecessary expense.

Over time, the Planner learns about the efficacy of self-control based on experience. As before, it is possible to assess eventual beliefs and choices by studying the full information model rather than the more intricate model that embeds uncertainty about the Doer's type.

Before applying the tools developed for Model 1 in Section 3, let me highlight two differences between costly self-control and nudges. In Section 3, a nudge is costly if it induces the Doer to choose $a = 1$ too often, but otherwise is intrinsically costless. In contrast, the self-control studied in this section is intrinsically costly. Although this distinction is of conceptual importance, it does not create any difficulty in applying the earlier results. The challenge arises from the distinction in timing. In the earlier analyses, the Planner's choice of nudges and menus is made before the resolution of uncertainty; as such, if the Planner partially commits, his commitment choice induces a non-trivial distribution of behavior that varies by type. In contrast, in the setting here, the Planner's choice to exert costly self-control is made after uncertainty about the state is resolved. Thus, if the Planner chooses to exert costly self-control of magnitude Γ whenever the value of leisure is less than some level s, and this is successful for types θ and θ' of the Doer, the Planner cannot distinguish these types by behavior even though he is partially committing for both. To generate distinguishability with partial commitment, I assume that if the Planner remains flexible and exerts costly self-control, he observes some signal σ with support in $\Re$ and whose distribution is increasing in $\Gamma - \frac{s}{\theta}$ (in terms of First Order Stochastic Dominance). The interpretation of this signal is that the Planner obtains partial information about the impact of self-control beyond the action choice. With this modification, results analogous to Theorems 2 and 3 hold.

Theorem 7. *If $\underline{\theta} > \theta^*$, then a fully informed Planner always chooses a flexible menu and exercises self-control. Thus, no type is fully committed, and so FCD is satisfied. Learning is globally adequate for all priors and discount factors. In contrast, if $\underline{\theta} < \theta^*$, then the Planner chooses the menu $\{1\}$ for sufficiently low types. FCD is violated and for each $\delta > 0$, learning may be inadequate for an open set of priors.*

5 Discussion

This paper offers a simple and tractable non-equilibrium approach to study the evolution of beliefs, self-control, and commitment choices as a decisionmaker learns from experience. I derive a condition that reveals whether such learning inexorably leads to sophistication. When FCD is satisfied, even a passive learner eventually makes choices as if he were in the full information benchmark. As such, learning-based foundations for sophistication in these settings, such as a savings environment, can help address the criticism that it involves an unrealistic level of rationality.[22] In contrast, when FCD fails, then an impatient decisionmaker may fail to learn adequately, and make commitment choices that are inefficient relative to the full information benchmark. In these cases, policy interventions that subsidize experimentation may facilitate greater awareness.

In this paper, I have adopted a dual self approach to understand learning. By treating a Doer as non-strategic, dual self models are analytically simple and generally have unique solutions when a decisionmaker is sophisticated. From the perspective of learning, this simplicity permits the investigation of whether the *unique* solution under partial awareness converges to the *unique* solution under sophistication. In contrast, the quasi-hyperbolic approach has multiple solutions even after imposing sophistication and a Markovian restriction on behavior. This challenge of multiplicity is exacerbated when beliefs are endogenized because in the resulting intrapersonal game, uncertainty about self-control induces a motive for self-signaling. Multiplicity thus impedes the study of learning in the quasi-hyperbolic approach, especially in settings such as consumption-savings in which the entire set of sophisticated solutions has not been characterized. Since the dual self framework captures issues of temptation and commitment with greater parsimony, this approach facilitates a more direct study of learning.

When learning fails at the limit, I find that it occurs in a particular direction, namely that individuals underestimate their self-control. It is important to understand the extent to which these results can be reconciled with findings in the field. Some of the evidence on self-awareness comes from individuals' contractual choices and their subsequent decisionmaking. DellaVigna and Malmendier (2006) demonstrate that many individuals choose monthly gym memberships but given their usage of the gym, would have paid less per visit had they selected the pay-per-visit option. Similarly, Shui and Ausubel (2005) find that individuals accept introductory credit card offers with lower interest rates for a shorter duration rather than a higher interest rate with a longer duration though they may be better off with the latter choice. In both cases, behavior is consistent with individuals being sophisticated about their self-control—valuing the commitment offered by a monthly membership or a shorter duration period—as well as them overestimating their self-control. In a model with overlapping generations, it is quite likely that both

[22] Rubinstein (2005) offers such a criticism:

> "Sophistication is unrealistic since it suffers from the problems of subgame perfection. An agent is super-rational in the sense that he perfectly anticipates his future selves and arrives at equilibrium between them. Present-bias is a realistic phenomenon, but the combination of the β, δ preferences with naivete or sophistication assumptions makes the model even more unrealistic than time consistency models."

effects are present insofar as less experienced customers who have not had opportunities to learn may overestimate self-control. DellaVigna and Malmendier (2006) argue that delays in canceling contracts are more consistent with partial naivete. However, even when a decisionmaker does not overestimate self-control, he has an incentive to delay canceling since there is an "option value" from learning, particularly in this sample that comprises first-time customers. This option value increases with the decisionmaker's patience making it difficult to distinguish patience from naivete. Moreover, it would be natural to assume that many such individuals may be learning both about their self-control and the costs and benefits of going to a health club (as also argued by Fudenberg and Levine 2006).

Indeed, this point raises the tension of assuming that a decisionmaker is partially sophisticated about his self-control but somehow fully sophisticated about the payoffs from his choices. Although this paper, like the prior literature, abstracts from this issue, it is surely realistic that individuals often have to learn both the payoffs of choices and whether their decisions are vulnerable to temptation. In settings in which feedback about payoffs are delayed, we should expect learning to be slow. Although I defer a fuller treatment of this issue, I illustrate its potential implications through the following example.

Consider the two-type model from Section 2, and suppose for simplicity that $\delta = 0$. Instead of the distribution of costs being known by the Planner, suppose that it depends on the realization of a random variable ξ; in particular, let s_t be distributed uniformly with support $[\xi, 1 + \xi]$, and suppose that ξ itself is 0 with probability $1 - \lambda_0$ and $\left(\bar{\theta} - \underline{\theta}\right) b$ with probability λ_0. This particular support is assumed for simplifying reasons and is unnecessary for the analysis.[23] To understand the implications of delayed feedback about payoffs transparently, consider a limit model in which information about past costs does not emerge.[24] In this setting, a complete breakdown of learning occurs because a Planner who observes the Doer choose $a = 1$ with frequency $\underline{\theta} b$ cannot distinguish his lack of self-control from a high average cost of taking the action. Prior beliefs then determine his choice to remain flexible or commit (in the long-run limit), and under certain conditions, such a Planner may perpetually undercommit relative to a full information benchmark.

Proposition 3. *If μ_0 and λ_0 are sufficiently high, and $(\underline{\theta}, 0)$ is the realized state, then a partially aware Planner always chooses flexibility. If, however, a sophisticated Planner knew that $\xi = 0$ or that $\theta = \underline{\theta}$, then almost-surely, the Planner would choose to commit.*

[23] It suffices if ξ is drawn from a continuum with values in its support that are $\left(\bar{\theta} - \underline{\theta}\right) b$ apart.

[24] This stark setting is similar to cases studied by Dekel et al. (2004) in which payoffs are not observed, and to the identification challenges that arise in Acemoglu et al. (2009).

The assumption that a Planner never observes any ex post information about payoffs is critical and instructive: when the Planner obtains informative signals about payoffs, this sequence of signals identify ξ in the long-run limit, eliminating the possibility for undercommitment. As such, if feedback is delayed but eventually forthcoming, one should expect undercommitment to persist for some length of time but to not be perpetual. More generally, if the benefits and costs of an action are uncertain and take time to be realized, and a decisionmaker has limited experience with the setting, learning should not be expected to be fast.

Finally, it is worth highlighting that the nature of long-run beliefs does not imply that a decisionmaker should never undercommit. Commitment often has costs apart from monetary or its loss of flexibility; in many contexts, such as addiction, the choice to commit is observable, and so a decisionmaker who recognizes his imperfect self-control may nevertheless choose not to commit so as to hide it from others. Moreover, to the extent that commitment requires self-control (Noor, 2007), even a fully sophisticated decisionmaker may not commit. Allowing for these considerations can explain why individuals may not commit to the extent that our models (which omit these features) predict, even when individuals learn about their temptations over time.

In spite of all the potential reasons for people not to commit, many recent studies highlight the demand for commitment that individuals exhibit.[25] Such demand for commitment is consistent with both sophistication and the failure of learning that can emerge in tradeoffs between commitment and flexibility.

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2009): "Fragility of Asymptotic Agreement under Bayesian Learning," Mimeo, MIT.

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991): "Optimal Learning by Experimentation," *Review of Economic Studies*, 58, 621–654.

Aliprantis, C. and K. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, New York, NY: Springer Verlag.

Amador, M., I. Werning, and G.-M. Angeletos (2006): "Commitment Vs. Flexibility," *Econometrica*, 74, 365–396.

[25]For instance, see Angeletos et al. (2001), Ariely and Wertenbroch (2002), Ashraf et al. (2006), Giné et al. (2009), Thaler and Benartzi (2004), Wertenbroch (1998); DellaVigna (2009) offers a survey.

Ameriks, J., A. Caplin, and J. Leahy (2003): "Wealth Accumulation and the Propensity to Plan," *Quarterly Journal of Economics*, 118, 1007–1047.

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001): "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15, 47–68.

Ariely, D. and K. Wertenbroch (2002): "Procrastination, Deadlines and Performance: Self-Control by Precommitment," *Psychological Science*, 13, 219–224.

Ashraf, N., D. Karlan, and W. Yin (2006): "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 121, 635–672.

Barro, R. (1999): "Ramsey Meets Laibson in The Neoclassical Growth Model*," *Quarterly Journal of Economics*, 114, 1125–1152.

Baumeister, R. F., T. F. Heatherton, and D. M. Tice (1994): *Losing Control: How and Why People Fail at Self-regulation*, San Diego, CA: Academic Press.

Benabou, R. and M. Pycia (2002): "Dynamic Inconsistency and Self-Control: A Planner-Doer Interpretation," *Economics Letters*, 77, 419–424.

Benabou, R. and J. Tirole (2002): "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117, 871–915.

——— (2004): "Willpower and Personal Rules," *Journal of Political Economy*, 112, 848–886.

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

Bernheim, B. D. and A. Rangel (2004): "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 94, 1558–1590.

Billingsley, P. (1995): *Probability and Measure*, New York, NY: Wiley, 3rd ed.

Brocas, I. and J. D. Carrillo (2008): "The Brain as a Hierarchical Organization," *American Economic Review*, 98, 1312–1346.

Brunnermeier, M. K. and J. A. Parker (2005): "Optimal Expectations," *The American Economic Review*, 95, 1092–1118.

Carrillo, J. D. and T. Mariotti (2000): "Strategic Ignorance as a Self-Disciplining Device," *The Review of Economic Studies*, 67, 529–544.

Chatterjee, K. and R. Krishna (2009): "A "Dual Self Representation for Stochastic Temptation," *American Economic Journal: Microeconomics*, 1, 148–167.

Compte, O. and A. Postlewaite (2004): "Confidence-Enhanced Performance," *American Economic Review*, 94, 1536–1557.

Dekel, E., D. Fudenberg, and D. K. Levine (1999): "Payoff Information and Self-Confirming Equilibrium," *Journal of Economic Theory*, 89, 165–185.

——— (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282–303.

Dekel, E. and B. L. Lipman (2010): "Costly Self Control and Random Self Indulgence," Mimeo, Boston University.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47, 315–372.

DellaVigna, S. and U. Malmendier (2006): "Paying Not to Go to the Gym," *The American Economic Review*, 96, 694–719.

Easley, D. and N. M. Kiefer (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045–1064.

Fudenberg, D. (2006): "Advancing Beyond Advances in Behavioral Economics," *Journal of Economic Literature*, 44, 694–711.

Fudenberg, D. and D. K. Levine (1993a): "Self-Confirming Equilibrium," *Econometrica*, 61, 523–545.

——— (1993b): "Steady State Learning and Nash Equilibrium," *Econometrica*, 61, 547–573.

——— (2006): "A Dual Self Model of Impulse Control," *American Economic Review*, 96, 1449–1476.

——— (2009): "Learning and Equilibrium," *Annual Review of Economics*, 1, 385–420.

Giné, X., D. Karlan, and J. Zinman (2009): "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation," Mimeo, Yale University.

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

——— (2005): "The Revealed Preferences Theory of Changing Tastes," *Review of Economic Studies*, 72, 429–448.

Harris, C. and D. Laibson (2003): "Hyberbolic Discounting and Consumption," in *Advances in Economics and Econometrics: Theory and Applications: Eighth World Congress*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, Econometric Society, 258–297.

Kahneman, D., P. P. Wakker, and R. Sarin (1997): "Back to Bentham? Explorations of Experienced Utility," *The Quarterly Journal of Economics*, 112, 375–405.

Kaper, J., E. Wagena, M. Willemsen, and C. Van Schayck (2005): "Reimbursement for smoking cessation treatment may double the abstinence rate: results of a randomized trial." *Addiction*, 100, 1012.

Kőszegi, B. (2006): "Ego utility, overconfidence, and task choice," *Journal of the European Economic Association*, 4, 673–707.

Kreps, D. (1979): "A Representation Theorem for "Preference for Flexibility"," *Econometrica*, 47, 565–578.

Krusell, P. and A. A. Smith, Jr (2003): "Consumption-Savings Decisions with Quasi-Geometric Discounting," *Econometrica*, 71, 365–375.

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

Loewenstein, G. and T. O'Donoghue (2007): "The Heat of the Moment: Modeling Interactions Between Affect and Deliberation," Mimeo, Cornell University.

Noor, J. (2007): "Commitment and Self-Control," *Journal of Economic Theory*, 135, 1–34.

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

Pollak, R. A. (1968): "Consistent planning," *The Review of Economic Studies*, 35, 201–208.

Rubinstein, A. (2005): "Discussion of 'Behavioral Economics'," in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, ed. by R. Blundell, W. Newey, and T. Persson, Econometric Society, 246–254.

Shui, H. and L. M. Ausubel (2005): "Time Inconsistency in the Credit Market," Mimeo, University of Maryland–College Park.

Strotz, R. H. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23, 165–180.

Thaler, R. H. and S. Benartzi (2004): "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 112, 164–182.

Thaler, R. H. and H. M. Shefrin (1981): "An Economic Theory of Self-Control," *Journal of Political Economy*, 89, 392–406.

Wertenbroch, K. (1998): "Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice," *Marketing Science*, 17, 317–337.

A Online Appendix

A.1 Section 2

So as to make the analysis of the example self-contained, I offer direct proofs of these assertions in the text. Throughout, I let c_t denote the commitment choice made in period t.

Proof of Proposition 1

I first establish that V exists, satisfies Equation 1, is continuous and non-decreasing in μ. Define the function $V^t(\mu)$ recursively as follows:

$$\begin{aligned} V^0(\mu) &= 0 \\ V^t(\mu) &= \max\left\{ \begin{array}{c} b - \frac{1}{2} + \delta V^{t-1}(\mu), \\ (1-\mu)\left(\underline{\theta} b\left(b - \frac{\underline{\theta} b}{2} + \delta V^{t-1}(\mu^+)\right) + \delta(1-\underline{\theta} b) V^{t-1}(\mu^-)\right) + \\ \mu\left(\bar{\theta} b\left(b - \frac{\bar{\theta} b}{2} + \delta V^{t-1}(\mu^+)\right) + \delta\left(1-\bar{\theta} b\right) V^{t-1}(\mu^-)\right). \end{array} \right\} \end{aligned}$$

where μ^+ and μ^- are the posteriors conditional on the Planner choosing flexibility, and the Doer choosing $a_t = 1$ and $a_t = -1$ respectively, given a prior μ. To simplify notation, let

$$W^t(\mu) = \begin{array}{c} (1-\mu)\left(\underline{\theta} b\left(b - \frac{\underline{\theta} b}{2} + \delta V^{t-1}(\mu^+)\right) + \delta(1-\underline{\theta} b) V^{t-1}(\mu^-)\right) + \\ \mu\left(\bar{\theta} b\left(b - \frac{\bar{\theta} b}{2} + \delta V^{t-1}(\mu^+)\right) + \delta\left(1-\bar{\theta} b\right) V^{t-1}(\mu^-)\right). \end{array}$$

I verify that V^t satisfies continuity and non-increasing. I then define V as the point-wise limit of V^t and show that convergence is uniform. This verifies the existence, continuity, and monotonicity of V.

Lemma 1. *For each $t \geq 0$, $V^t(\mu)$ is continuous and non-decreasing in μ.*

Proof. Clearly $V^0(\mu)$ is continuous and non-decreasing in μ. By induction, if these properties hold for $t \leq \tau$, then they hold for $\tau + 1$ since $V^{\tau+1}$ is the upper-envelope of sums of continuous functions. Q.E.D.

Lemma 2. *For all $\mu \in [0, 1]$, $V^t(\mu)$ is non-decreasing in t.*

Proof. Note that $V^1(\mu) > 0 = V^0(\mu)$ for all $\mu \in [0, 1]$. Let $V^t(\mu) \geq V^{t-1}(\mu)$ for all $\mu \in [0, 1]$, and $t \leq \tau$. Then $b - \frac{1}{2} + \delta V^\tau(\mu) \geq b - \frac{1}{2} + \delta V^{\tau-1}(\mu)$, and $W^{\tau+1}(\mu) \geq W^\tau(\mu)$. Therefore, $V^{\tau+1}(\mu) \geq V^\tau(\mu)$. Q.E.D.

Given that $V^t(\mu)$ is non-decreasing in t, and $V^t(\mu)$ is uniformly bounded above by $\frac{b}{1-\delta}$, by the Bolzano-Weierstrass Theorem, there exists a point-wise limit $V(\mu) =$

$\lim_{t\to\infty} V^t(\mu)$. In fact, since $V^t(\mu) \leq V(\mu) \leq V^t(\mu) + \delta^t \frac{b}{1-\delta}$, for every μ, $|V(\mu) - V^t(\mu)| \leq \delta^t \frac{b}{1-\delta}$. Consequently, the convergence is uniform leading to the following corollary.

Corollary 1. *$V(\mu)$ satisfies Equation 1, and is continuous and non-decreasing in μ.*

Let $W(\mu) = \lim_{t\to\infty} W^t(\mu)$. Observe that $V(0) = \frac{b-1/2}{1-\delta} > \underline{\theta} b \left(b - \frac{\underline{\theta} b}{2}\right) + \delta V(0) = W(0)$, and $\frac{b-1/2}{1-\delta} < \frac{\bar{\theta} b \left(b - \frac{\bar{\theta} b}{2}\right)}{1-\delta} = V(1) = W(1)$. Also observe that if $V(\mu) > W(\mu)$, then for all $\tilde{\mu} < \mu$, $V(\tilde{\mu}) = V(\mu) > W(\mu) \geq W(\tilde{\mu})$, since W is non-decreasing in μ. Therefore, the set, $\{\mu \in [0,1] : V(\mu) > W(\mu)\}$ is an interval. By continuity of V and W, there exists μ^* such that for $\mu \geq \mu^*$, $V(\mu) = W(\mu)$, and for $\mu \leq \mu^*$, $V(\mu) = \frac{b-1/2}{1-\delta}$.

Proof of Theorem 1

Define the public likelihood ratio, $\lambda(h^t) \equiv \frac{\Pr(\theta = \bar{\theta} | h^t)}{\Pr(\theta = \underline{\theta} | h^t)}$, and denote the random variable with realizations $\lambda(h^t)$ by λ_t. Let $\lambda^* = \frac{\mu^*}{1-\mu^*}$. Consider the case where $\theta = \underline{\theta}$. Observe that $\langle \lambda_t \rangle$ is a super-martingale:

$$E_t[\lambda_{t+1} | \theta = \underline{\theta}] = \begin{cases} \lambda_t & \text{if } \underline{\theta} > 0, \\ \left(1 - \bar{\theta} b\right) \lambda_t & \text{if } \underline{\theta} = 0. \end{cases}$$

By Doob's Supermartingale Convergence Theorem for non-negative random variables (Billingsley, 1995, p. 468-469), there exists a non-negative random variable λ_∞ with support $[0, \infty)$ such that $\lambda_t \to_{a.s.} \lambda_\infty$.

Lemma 3. *$Supp(\lambda_\infty) \subseteq [0, \lambda^*)$*

Proof. Suppose that for some fixed $x \in (\lambda^*, \infty)$ that $x \in Supp(\lambda_\infty)$. Consider the ball $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$ where $\varepsilon < \min\left\{\left(\frac{\bar{\theta} - \underline{\theta}}{\bar{\theta} + \underline{\theta}}\right) x, \left(\frac{b(\bar{\theta} - \underline{\theta})}{2 - b(\bar{\theta} + \underline{\theta})}\right) x, x - \lambda^*\right\}$. Since x is in the support of λ_∞, then λ_i eventually enters and perpetually remains in $B_\varepsilon(x)$ with positive probability. Observe that for all $\tilde{x} \in B_\varepsilon(x)$, if $\lambda_i = \tilde{x}$, the Planner chooses to remain flexible since $\tilde{x} > \lambda^*$. The Planner then observes either the activity not being undertaken or the activity being undertaken. Consequently, $\lambda_{i+1} \in \left\{\left(\frac{1-\bar{\theta} b}{1-\underline{\theta} b}\right) \tilde{x}, \left(\frac{\bar{\theta}}{\underline{\theta}}\right) \tilde{x}\right\}$ if $\underline{\theta} \neq 0$ and $\lambda_{i+1} = \left(1 - \bar{\theta} b\right) \tilde{x}$ if $\underline{\theta} = 0$; neither value is in $B_\varepsilon(x)$. Therefore, $\lambda_i \in B_\varepsilon(x) \Rightarrow \lambda_{i+1} \notin B_\varepsilon(x)$; thus, $x \notin Supp(\lambda_\infty)$, leading to a contradiction.

To prove $\lambda^* \notin Supp(\lambda_\infty)$: suppose the opposite. It must be that for all ε, λ_i eventually enters and perpetually remains in $B_\varepsilon(\lambda^*)$ with positive probability. Let $\varepsilon < \min\left\{\lambda^* \left(\frac{\bar{\theta} - \underline{\theta}}{\bar{\theta} + \underline{\theta}}\right), \lambda^* \left(\frac{b(\bar{\theta} - \underline{\theta})}{2 - b(\bar{\theta} + \underline{\theta})}\right)\right\}$, and consider the set $B = [\lambda^*, \lambda^* + \varepsilon]$. Observe that along any sample-path, if $\lambda_i < \lambda^*$, then for every $k > i$, $\lambda_k = \lambda_i < \lambda^*$. Consequently, the only sample-paths along which $\langle \lambda_i \rangle$ converges to λ^* are those in which

$\lambda_i \geq \lambda^*$ for every i, and $\lim_{i\to\infty} \lambda_i = \lambda^*$. For $\tilde{x}$ in B, if $\lambda_i = \tilde{x}$, the Planner remains flexible since $\tilde{x} \geq \lambda^*$. Then, as before, $\lambda_{i+1} \in \left\{ \tilde{x} \left(\frac{1-\bar{\theta}b}{1-\underline{\theta}b} \right), \tilde{x} \left(\frac{\bar{\theta}}{\underline{\theta}} \right) \right\}$, neither of which is in B. Q.E.D.

The above establishes that $\Pr\left(\lim_{t\to\infty} \lambda_t \in [0, \lambda^*) \,|\theta = \underline{\theta}\right) = 1$. Since the Planner chooses $c = \{1\}$ for beliefs in that range, the first part of the result is established.

Consider $\rho\left(h^t\right) = \frac{\Pr\left(\theta=\underline{\theta}|h^t\right)}{\Pr\left(\theta=\bar{\theta}|h^t\right)}$, and let ρ_t be the random variable which has realization $\rho\left(h^t\right)$. Let $\rho^* = \frac{1-\mu^*}{\mu^*}$. Consider the case where $\theta = \bar{\theta}$ and observe that $\langle \rho_t \rangle$ is a martingale. By the Martingale Convergence Theorem, there exists a random variable ρ_∞ with $Supp\left(\rho_\infty\right) \subseteq [0, \infty)$ such that $\rho_t \to_{a.s.} \rho_\infty$.

Lemma 4. *$Supp\left(\rho_\infty\right) \subseteq \{0\} \cup (\rho^*, \infty)$*

Proof. The proof is analogous to that of Lemma 3. Q.E.D.

The above Lemma establishes that conditional on $\theta = \bar{\theta}$, $\mu_\infty \in [0, \mu^*) \cup \{1\}$ with probability 1. Moreover, observe that for a given μ_0 and $\rho_0 = \frac{1-\mu_0}{\mu_0}$, $\Pr\left(\lim_{t\to\infty} c_t = \{1\} \,|\theta = \bar{\theta}\right) \geq \left(1-\bar{\theta}b\right)^{\left\lceil \frac{\log p^* - \log \rho_0}{\log(1-\underline{\theta}b) - \log\left(1-\bar{\theta}b\right)} \right\rceil} > 0$. To see that $\Pr\left(\lim_{t\to\infty} c_t = \{-1, 1\} \,|\theta = \bar{\theta}\right) > 0$: observe that since $\langle \rho_t \rangle$ is a martingale, $E\left[\rho_t\right] = \rho_0 \leq \rho^*$ for every t where the inequality follows from $\mu_0 \geq \mu^*$. By Fatou's Lemma, $E\left[\rho_\infty\right] \leq \lim_{t\to\infty} E\left[\rho_t\right] \leq \rho^*$, and as $(0, \rho^*] \cap Supp\left(\rho_\infty\right)$ is non-empty, $\Pr\left(\rho_\infty = 0\right) > 0$. Therefore $\Pr\left(\lim_{t\to\infty} c_t = \{-1, 1\} \,|\theta = \bar{\theta}\right) > 0$.

A.2 Section 3

Before proving the results, I make some useful preliminary observations about the full information environment.

Lemma 5. *In Model 1, the following holds:*

1. *π is continuous in (θ, c).*
2. *If $\pi(\theta, c) \geq \hat{\pi}$, then $\pi(\theta', c) \geq \hat{\pi}$ for all $\theta' > \theta$.*
3. *FCD holds if and only if $\hat{\Theta}$ is of zero measure or there exists some interval of the form $\left[\theta, \bar{\theta}\right] \subseteq \hat{\Theta}$ such that $\mu_0\left(\hat{\Theta} \backslash \left[\theta, \bar{\theta}\right]\right) = 0$.*
4. *If there exists $\hat{\theta} < \bar{\theta}$ in $\hat{\Theta}$ such that for some $\bar{\varepsilon} > 0$, $\left[\hat{\theta}, \hat{\theta} + \varepsilon\right] \nsubseteq \hat{\Theta}$ for all $\varepsilon \in (0, \bar{\varepsilon})$, then there exists $c \notin \hat{C}\left(\hat{\theta}\right)$ such that $\pi\left(\hat{\theta}, c\right) = \hat{\pi}$.*
5. *If there exists $\hat{\theta}$ in $\hat{\Theta}$ such that for some $\bar{\varepsilon} > 0$, $\left[\hat{\theta}, \hat{\theta} + \varepsilon\right] \subseteq \hat{\Theta}$ for all $\varepsilon \in (0, \bar{\varepsilon})$, then there exists $\gamma > 0$ such that $\pi\left(\hat{\theta}, c\right) < \hat{\pi} - \gamma$ for all $c \notin \hat{C}\left(\hat{\theta}\right)$.*

Proof. (1) For any commitment, $c = (A, \eta)$, $\pi(\theta, c) = \int_{\{s:W(s,\theta,\eta)\geq 0\}} U(s)\, d\nu$. Continuity of π follows from the continuity of U and W (in their respective arguments) and from ν being atomless.

(2) Consider, as a function of s', the following, $\int_{s'}^{\bar{s}} U(s)\, d\nu$: it is continuous in s', takes on a value of π when $s = \underline{s}$, a value of 0 when $s' = \bar{s}$, is positively sloped for all s' such that $U(s') < 0$, and is negatively sloped when $U(s) > 0$. By continuity, there exists $\tilde{s}$ such that $U(\tilde{s}) > 0$ and $\hat{\pi} \leq \int_{s'}^{\bar{s}} U(s)\, d\nu$ if and only if $s' \leq \tilde{s}$. Observe that $\pi(\theta, c) = \int_{s(\theta,c)}^{\bar{s}} U(s)\, d\nu$ where $s(\theta, c) = \min(\{s \in S : a_D(s, \theta, c) = 1\})$. Accordingly, $\pi(\theta, c) \geq \hat{\pi}$ implies that $s(\theta, c) \leq \tilde{s}$. By the strict single-crossing condition, $a_D(s, \theta', c) \geq a_D(s, \theta, c)$ for any $\theta' > \theta$, and therefore, $s(\theta', c) \leq s(\theta, c)$. Hence, $\pi(\theta', c) \geq \hat{\pi}$.

(3) For any θ, and $\theta' > \theta$, $\hat{C}(\theta') \subseteq \hat{C}(\theta)$, and therefore, for almost every θ in $\hat{\Theta}$, FCD is satisfied if and only if $\left[\theta, \bar{\theta}\right] \subseteq \hat{\Theta}$.

(4) There exists a sequence $\{\theta_n\}_{n=1}^{\infty} \searrow \hat{\theta}$ such that $\theta_n \notin \hat{\Theta}$. For each θ_n, $C^*(\theta_n)$ is single-valued; denote the associated optimal nudge by $\eta^*(\theta_n)$. Since W is continuous, the sequence $\{\eta^*(\theta_n)\}_{n=1}^{\infty}$ has a limit, and because $\mathcal{N}$ is closed, the limit $\eta = \lim_{n\to\infty} \eta^*(\theta_n)$ is feasible. Consider the commitment $c = (A, \eta)$: if $c \in \hat{C}\left(\hat{\theta}\right)$, then $c \in \hat{C}(\theta_n)$ for each n. Because $\theta_n \notin \hat{\Theta}$, $\eta^*(\theta_n) < \eta$ for each n, and therefore, there exists a sequence of nudges $\{\eta_m\}_{m=1}^{\infty} \nearrow \eta$. Thus, if $c \in \hat{C}\left(\hat{\theta}\right)$, there exist commitments $c_m = (A, \eta_m)$ such that $\pi\left(\hat{\theta}, c_m\right) > \hat{\pi}$ for sufficiently large m, contradicting $\hat{\theta}$ being in $\hat{\Theta}$. It follows that c is in $C^*\left(\hat{\theta}\right) \backslash \hat{C}\left(\hat{\theta}\right)$.

(5) Suppose towards a contradiction that there exists a sequence of commitments $\{c_m\}_{m=1}^{\infty}$ $\notin \hat{C}\left(\hat{\theta}\right)$ such that for all $\gamma > 0$, there exists M such that $\pi\left(\hat{\theta}, c_m\right) \geq \hat{\pi} - \gamma$ for every $m > M$. Consider the sequence $\{s_m\}_{m=1}^{\infty}$ such that $s_m = s\left(\hat{\theta}, c_m\right)$. By above, it must be that $\lim_{m\to\infty} s_m = \tilde{s}$. Since $\mathcal{C}$ is closed, there exists a commitment $c = \lim_{m\to\infty} c_m$ such that $\pi(\theta, c) = \hat{\pi}$ and $c \notin \hat{C}\left(\hat{\theta}\right)$. By monotonicity, $s(\theta, c) \in (\underline{s}, \tilde{s})$ for all for θ in $\left(\hat{\theta}, \hat{\theta} + \varepsilon\right)$ for every sufficiently small $\varepsilon' > 0$, and therefore, $\pi(\theta, c) > \hat{\pi}$, contradicting the assumption that $\left[\hat{\theta}, \hat{\theta} + \varepsilon\right] \subseteq \hat{\Theta}$ for all $\varepsilon \in (0, \bar{\varepsilon})$. Q.E.D.

Notice that Observation 1 is subsumed by Lemma 5.

In Model 2, define $\hat{a}_P(s,c)$ as the action that the Planner would select when restricted to a commitment set c, and let $\pi_P(c) = \int_S u(\hat{a}_P(s,c), s)\, d\nu$.

Lemma 6. *In Model 2, the following holds:*

1. *$\pi(\theta, c) \leq \pi(\theta', c)$ for all $c \in \mathcal{C}$ and $\theta < \theta'$. If $\nu(s : a_D(s, \theta', c) \neq a_D(s, \theta, c)) > 0$, then the inequality is strict.*
2. *If $\theta' \in \hat{\Theta}$, then $\theta \in \hat{\Theta}$ for all $\theta < \theta'$.*
3. *If θ and $\theta' > \theta$ both are in $\hat{\Theta}$, then $\hat{C}(\theta) \subseteq \hat{C}(\theta')$.*
4. *π is continuous in (θ, c).*
5. *If FCD holds, and $\mu_0\left(\hat{\Theta}\right) > 0$, then for all $\hat{\theta}$ in $\hat{\Theta}$ and θ not in $\hat{\Theta}$, $\hat{C}\left(\hat{\theta}\right) \nsubseteq \hat{C}(\theta)$.*

Proof. (1) Let M be the menu associated to commitment c. For any θ and $\theta' > \theta$, it follows from the strict single-crossing condition that $a_D(s, \theta, c) \leq a_D(s, \theta', c) \leq \hat{a}_P(s, c)$ for every s. Because $u(a, s)$ is strictly quasi-concave for each s, and $\hat{a}_P(s, c)$ uniquely attains $\max_{a \in M} u(a, s)$ it follows that $u(a_D(s, \theta', c), s) \geq u(a_D(s, \theta, c), s)$ for every s and therefore, $\pi(\theta', c) \geq \pi(\theta, c)$. For any state s in which $a_D(s, \theta', c) \neq a_D(s, \theta, c)$, it follows from strict quasi-concavity that $u(a_D(s, \theta', c), s) > u(a_D(s, \theta, c), s)$, and therefore, $\pi(\theta', c) > \pi(\theta, c)$ if $\nu(s : a_D(s, \theta', c) \neq a_D(s, \theta, c)) > 0$.

(2) For any type $\theta \notin \hat{\Theta}$ and commitment $c \in C^*(\theta)$, it follows from the definition of $\hat{\Theta}$ that $\pi(\theta, c) > \hat{\pi}$. By above, $\pi(\theta', c) \geq \pi(\theta, c)$ for any $\theta' > \theta$, and therefore, $\theta' \notin \hat{\Theta}$. Taking the contrapositive yields the result.

(3) Suppose that $c \in \hat{C}(\theta)$, and let M be the menu associated to c. Since $a_D(s, \theta, c) = \hat{a}$, the strict single-crossing property implies that $a_D(s, \theta', c) \geq \hat{a}$ for all s. Consider the set $A_{\theta', c} = \{s : a_D(s, \theta', c) > \hat{a}\}$; I prove that $\nu(A_{\theta', c}) = 0$. Suppose otherwise: by Assumption 1, $\hat{a}_P(s, c) \geq a_D(s, \theta', c)$, and since u is strictly quasi-concave, $u(a_D(s, \theta', c), s) > u(\hat{a}, s)$ for any $s \in A_{\theta', c}$. Then, $\pi(\theta', c) > \hat{\pi}$, but since $\theta' \in \hat{\Theta}$, $\pi^*(\theta') = \hat{\pi}$ leading to a contradiction. Therefore, $c \in \hat{C}(\theta')$.

(4) Consider a sequence $\{\theta_n, c_n\}_{n=1}^{\infty}$ that converges to (θ, c). Since W is continuous, $a_D(s, \theta_n, c_n) \to a_D(s, \theta, c)$ for ν-a.e.; by the continuity of u, $U(a_D(s, \theta_n, c_n), s) \to U(a_D(s, \theta, c), s)$ for ν-a.e. Therefore, $\pi(\theta_n, c_n) \to \pi(\theta, c)$.

(5) By above, if $\hat{C}\left(\hat{\theta}\right) \subseteq \hat{C}(\theta)$, then $\hat{C}(\theta') \subseteq \hat{C}(\theta)$ for every $\theta' < \hat{\theta}$. Therefore if FCD holds, then the only type for which the condition may not be satisfied is $\underline{\theta}$. Consider a sequence of types $\{\theta_n\}_{n=1}^{\infty} \searrow \underline{\theta}$. Consider N such that for all $n > N$, $W(\underline{a}, \underline{s}, \theta_n, 0) > W(\hat{a}, \underline{s}, \theta_n, 0)$; because $\underline{a} = a_D(\underline{s}, \underline{\theta}, A, 0)$, and W is continuous, there exists such an N. Consider some $n > N$. By strict quasi-concavity of W, $W(a, \underline{s}, \theta_n, 0) > W(\hat{a}, \underline{s}, \theta_n, 0)$ for every $a \in [\underline{a}, \hat{a})$. Therefore, for every commitment $c \in \hat{C}(\theta_n)$, the associated menu M

must not include any actions strictly below $\hat{a}$. Since $a_D(s, \underline{\theta}, c) \leq a_D(s, \theta_n, c)$ for every c, it follows that $\hat{C}(\underline{\theta}) = \hat{C}(\theta_n)$. Since the condition is satisfied for θ_n it follows that for every $\theta \notin \hat{\Theta}$, $\hat{C}(\underline{\theta}) \nsubseteq \hat{C}(\theta)$. Q.E.D.

Proof of Proposition 2

The first claim follows from Lemma 6. Suppose that $\mu_0\left(\hat{\Theta}\right) > 0$. If for every $\theta \notin \hat{\Theta}$, $\hat{C}(\underline{\theta}) \nsubseteq \hat{C}(\theta)$, then because $\hat{C}(\underline{\theta}) \subseteq \hat{C}(\theta')$ for every $\hat{\theta} \in \hat{\Theta}$, it follows that for every $\theta' \in \hat{\Theta}$ and every $\theta \notin \hat{\Theta}$, $\hat{C}\left(\hat{\theta}\right) \nsubseteq \hat{C}(\theta)$. On the other hand, if for some $\theta \notin \hat{\Theta}$, it is the case that $\hat{C}(\underline{\theta}) \subseteq \hat{C}(\theta)$, then this poses a contradiction to Lemma 6.

Let $\Theta^L = \{\theta : a_D(\bar{s}, \theta, A, 0) \leq \hat{a}\}$. By the argument in the text, $\hat{\Theta} \subseteq \Theta^L$. Suppose that $\hat{\Theta} = \Theta^L$: then for all $\hat{\theta}$ in $\hat{\Theta}$ and θ not in $\hat{\Theta}$, the commitment set $[\hat{a}, \bar{a}] \in \hat{C}(\theta) \backslash \hat{C}(\theta)$. Therefore, FCD is satisfied. Now, suppose that $\hat{\Theta}$ is a strict subset of Θ^L and has strictly positive measure. Let $\theta \in \Theta^L \backslash \hat{\Theta}$, and let $\hat{\theta} \in \hat{\Theta}$. For any $c \in \hat{C}\left(\hat{\theta}\right)$, let $\hat{S} = \left\{s \in S : a_D\left(s, \hat{\theta}, c\right) = \hat{a}\right\}$; by definition, $\nu\left(\hat{S}\right) = 1$. By Lemma 6, $\theta > \hat{\theta}$, and therefore by the strict single crossing assumption, $a_D(s, \theta, c) \geq a_D\left(s, \hat{\theta}, c\right)$ for every s in $\hat{S}$. Moreover, since θ is in Θ^L, it follows that $a_D(s, \theta, A, 0) \leq \hat{a}$ and because $W(a, s, \theta, 0)$ is strictly quasi-concave in a, it follows that for every s in $\hat{S}$, $a_D(s, \theta', c) = \hat{a}$. Thus, $c \in \hat{C}(\theta')$ violating FCD.

For the solution when the Planner is partially informed, I draw on techniques and results from Easley and Kiefer (1988). An application of their Theorems 1-3 establish the existence and uniqueness of the value function, $V(\mu; \delta)$, and that the correspondence of maximizers is u.h.c. Let $c(\mu)$ be a selection of the maximizers from this correspondence. The standard martingale property on beliefs holds.

Observation 3. *Beliefs satisfy the martingale property: for A, a Borel subset of Θ, $E[\mu_{t+1}(A) | \mu_t] = \mu_t(A)$.*

Accordingly, the Martingale Convergence Theorem applies, and there is a $P(\Theta)$-valued random variable μ_∞ such that $\mu_t \to_{a.s.} \mu_\infty$. I describe possible limit-beliefs in the support of μ_∞. Suppose that $\tilde{\mu}_\infty$ is a *generic* limit-belief on a sample-path for which beliefs converge, and for this sample-path, let $\tilde{c}_\infty$ be the set of limit points of the sequence of optimal commitment choices $\{c(\mu_t)\}_{t=1}^{\infty}$. Formally,

$$\tilde{c}_\infty = \{c \in \mathcal{C} : \text{for all } \varepsilon > 0, \, T < \infty, \; |c(\tilde{\mu}_t) - c| \leq \varepsilon \text{ for some } t \geq T\};$$

since the correspondence of maximizers is u.h.c., $\tilde{c}_\infty$ exists. An implication of Lemma 3 of Easley and Kiefer (1988) is the following.

Lemma 7. *For all $\theta, \theta' \in Supp(\tilde{\mu}_\infty)$, $a_D(s, \theta, c) = a_D(s, \theta', c)$ for almost all s, and for all c in $\tilde{c}_\infty$.*

Let ∂_θ denote the Dirac measure that places probability 1 on type θ. Since W satisfies the strict single-crossing property on (a, θ), a corollary of the above lemma is:

Corollary 2. *For $\theta \in Supp(\tilde{\mu}_\infty)$, and $c \in \tilde{c}_\infty$, if for all $a \in A$, $\nu(\{s \in S : a_D(s, \theta, c) = a\}) < 1$, then $\tilde{\mu}_\infty = \partial_\theta$. If θ, $\theta' \in Supp(\hat{u}_\infty)$ and $\theta \neq \theta'$ then there exists $a \in A$ such that $\nu(\{s \in S : a_D(s, \theta, c) = a\}) = 1$.*

Recall that $m(\mu, c) = \int_\theta \pi(\theta, c)\, d\mu$; let $m^*(\mu) = \max_{c \in \mathcal{C}} m(\mu, c)$. As is standard in experimentation contexts, the dynamically optimal commitment choice in response to some limit belief $\tilde{\mu}_\infty$ must attain the maximal stage payoff $m^*(\tilde{\mu}_\infty)$ since at the limiting belief, the Planner is no longer learning and all option value is exhausted. By Lemma 4 of Easley and Kiefer (1988), the following holds.

Lemma 8. *$c \in \tilde{c}_\infty$ implies that $m(\tilde{\mu}_\infty, c) = m^*(\tilde{\mu}_\infty)$.*

Corollary 2 and Lemma 8 together imply that if for some $c \in \tilde{c}_\infty$, $a_D(s, \theta, c) = a$ for almost all s and for almost all $\theta \in Supp(\tilde{\mu}_\infty)$, then necessarily, $a = \hat{a}$, the optimal full commitment.

Proof of Theorem 2

It is straightforward to establish the result when $\mu_0\left(\hat{\Theta}\right) = 0$. Suppose that $\mu_0\left(\hat{\theta}\right) > 0$. It suffices to show that

$$Supp(\mu_\infty) \subset \left\{\tilde{\mu}_\infty \in P(\Theta) : Supp(\tilde{\mu}_\infty) \subseteq \hat{\Theta} \text{ or } \tilde{\mu}_\infty = \partial_\theta \text{ for some } \theta \notin \hat{\Theta}\right\}. \tag{3}$$

If Equation 3 is satisfied: for $\theta \in \hat{\Theta}$, then almost-surely, $Supp(\tilde{\mu}_\infty) \subset \hat{\Theta}$, and so $V(\mu_t; \delta) \to_{a.s.} \hat{\pi} = \pi^*(\theta)$; for $\theta \notin \hat{\Theta}$, then almost-surely, $\tilde{\mu}_\infty = \partial_\theta$, and therefore $V(\mu_t; \delta) \to_{a.s} \pi^*(\theta)$.

To establish that Equation 3 is true when FCD is satisfied, suppose towards a contradiction that it were false. Then there exists a limit belief $\tilde{\mu}_\infty$ that is not an element of the RHS. Let $\tilde{c} \in \tilde{c}_\infty$ be a commitment that maximizes expected payoffs. By Corollary 2 and Lemma 8, $a_D(s, \theta, \tilde{c}) = \hat{a}$ for almost all s, for almost all θ in $Supp(\tilde{\mu}_\infty)$. Thus, $m^*(\tilde{\mu}_\infty) = \hat{\pi}$. I demonstrate below that there exists some commitment $c \in \mathcal{C}$ such that for all $\theta \in Supp(\tilde{\mu}_\infty)$, $\pi(\theta, c) \geq \hat{\pi}$ and $\tilde{\mu}_\infty(\{\theta : \pi(\theta, c) > \hat{\pi}\}) > 0$. Thus, $m(\tilde{\mu}_\infty, c) > \hat{\pi}$, leading to a contradiction.

Model 1: let $\check{\theta} = \inf(Supp(\tilde{\mu}_\infty))$. By Lemma 5, $Supp(\tilde{\mu}_\infty) \nsubseteq \hat{\Theta}$ implies that either $\check{\theta} \notin \hat{\Theta}$ or there exists a sequence $\{\theta_n\}_{n=1}^\infty \searrow \check{\theta}$ such that $\theta_n \notin \hat{\Theta}$. In the former case, consider a commitment $c \in C^*\left(\check{\theta}\right)$: then $\pi\left(\check{\theta}, c\right) > \hat{\pi}$, and by continuity, $\pi(\theta, c) > \hat{\pi}$ for all θ in $\left(\check{\theta}, \check{\theta} + \varepsilon\right)$ for a sufficiently small ε. By Lemma 5, $\pi(\theta, c) \geq \hat{\pi}$ for all $\theta \in Supp(\tilde{\mu}_\infty)$. In the latter case, consider the commitment $c \in C^*\left(\check{\theta}\right) \backslash \hat{C}\left(\check{\theta}\right)$, which exists by Lemma 5. By continuity, $\pi(\theta, c) > \hat{\pi}$ for all θ in $\left(\check{\theta}, \check{\theta} + \varepsilon\right)$ for a sufficiently small ε, and $\pi(\theta, c) \geq \hat{\pi}$ for all $\theta \in Supp(\tilde{\mu}_\infty)$. Thus, in either case, $m(\tilde{\mu}_\infty, c) > \hat{\pi}$.

Model 2: let $\check{\theta} = \inf(Supp(\tilde{\mu}_\infty))$, and let $\theta^* \in Supp(\tilde{\mu}_\infty) \backslash \hat{\Theta}$ such that $\theta^* < \sup(Supp(\tilde{\mu}_\infty))$. If $Supp(\hat{\mu}_\infty) \cap \hat{\Theta}$ is empty, then consider a commitment $c \in C^*\left(\check{\theta}\right)$; because $\check{\theta} \notin \hat{\Theta}$, $\pi\left(\check{\theta}, c\right) > \hat{\pi}$, and therefore, by Lemma 6, $\pi(\theta, c) > \hat{\pi}$ for all $\theta \in Supp(\mu_\infty)$. If $Supp(\tilde{\mu}_\infty) \cap \hat{\Theta}$ is non-empty, then by Lemma 6, we know that $\check{\theta}$ is in $\hat{\Theta}$. By FCD, there exists $c \in \hat{C}\left(\check{\theta}\right) \backslash \hat{C}(\theta^*)$: Lemma 6 establishes that $\pi(\theta, c) \geq \hat{\pi}$ for all $\theta \in Supp(\tilde{\mu}_\infty)$, and $\pi(\theta, c) > \hat{\pi}$ for all $\theta > \theta^*$. Since $\tilde{\mu}_\infty\left(\left[\theta^*, \bar{\theta}\right]\right) > 0$, it follows that $m(\tilde{\mu}_\infty, c) > \hat{\pi}$.

Proof of Theorem 3

Suppose that FCD fails.

In Model 1, consider $\hat{\theta} \in \hat{\Theta}$ such that $\left[\hat{\theta}, \hat{\theta} + \varepsilon\right] \subseteq \hat{\Theta}$ for sufficiently small ε, and $\left[\hat{\theta}, \bar{\theta}\right] \nsubseteq \hat{\Theta}$. By Lemma 5, such a type $\hat{\theta}$ exists and there exists $\gamma > 0$ such that for any commitment $c \notin \hat{C}\left(\hat{\theta}\right)$, $\pi\left(\hat{\theta}, c\right) < \hat{\pi} - \gamma$. Let $\hat{c} \in \hat{C}\left(\hat{\theta}\right)$: by continuity of π, there exists ε' and γ' such that $\pi(\theta, c) < \hat{\pi} - \gamma'$ for all $\theta \in \left[\hat{\theta}, \hat{\theta} + \varepsilon'\right]$. I establish that for each $\delta > 0$, there exists $\rho \in (0, 1)$ such that $c(\mu) \in \hat{C}\left(\hat{\theta}\right)$ for all μ such that $\mu\left(\left[\hat{\theta}, \hat{\theta} + \varepsilon'\right]\right) > 1 - \rho$. Observe that for any commitment $c \notin \hat{C}\left(\hat{\theta}\right)$, the difference in expected static payoffs between commitment choices $\hat{c}$ and c is

$$m(\mu, \hat{c}) - m(\mu, c) > \gamma' \mu\left(\left[\hat{\theta}, \hat{\theta} + \varepsilon'\right]\right) + \left(1 - \mu\left(\left[\hat{\theta}, \hat{\theta} + \varepsilon'\right]\right)\right)\left(\int_{U(s)<0} U(s)\, d\nu\right).$$

Since the Planner can always fully commit in the future, it follows that

$$\int_{P(\Theta)} V(\tilde{\mu}; \delta)\, dQ(\mu, \hat{c}) \geq \hat{\pi}.$$

Also, since an alternative commitment c can at best lead to full information payoffs in one period, and because $\hat{c} \in \hat{C}(\theta)$ for $\theta \in \left[\tilde{\theta}, \tilde{\theta} + \varepsilon'\right]$,

$$\int_{P(\Theta)} V(\tilde{\mu}; \delta)\, dQ(\mu, c) \leq \hat{\pi}\mu\left(\left[\hat{\theta}, \hat{\theta} + \varepsilon'\right]\right) + \left(\int_{U(s)>0} U(s)\, d\nu\right)\left(1 - \mu\left(\left[\hat{\theta}, \hat{\theta} + \varepsilon'\right]\right)\right).$$

Combining the above inequalities, the payoff difference between choosing commitment $\hat{c}$

and c is

$$\begin{aligned}&(1-\delta)\,m(\mu,\hat{c})+\delta\int_{P(\Theta)}V(\tilde{\mu};\delta)\,dQ(\mu,\hat{c})-\left((1-\delta)\,m(\mu,c)+\delta\int_{P(\Theta)}V(\tilde{\mu};\delta)\,dQ(\mu,c)\right)\\ > \;&(1-\delta)\,\gamma'\mu\left(\left[\hat{\theta},\hat{\theta}+\varepsilon'\right]\right)+\left(1-\mu\left(\left[\hat{\theta},\hat{\theta}+\varepsilon'\right]\right)\right)\left(\int_{U(s)<0}U(s)\,d\nu\right).\end{aligned}$$

Since γ' is strictly positive and independent of δ, it follows that for each δ, there exists $\rho\in(0,1)$ such that $\mu\left(\left[\hat{\theta},\hat{\theta}+\varepsilon'\right]\right)>1-\rho$ ensures that the RHS of the above inequality is strictly positive. Thus, for any such belief μ, $c(\mu)\in\hat{C}\left(\hat{\theta}\right)\subseteq\hat{C}(\theta)$ for every $\theta>\hat{\theta}$. Consider a prior for the Planner such that $\mu_0\left(\left[\hat{\theta},\hat{\theta}+\varepsilon'\right]\right)$ is at least $1-\rho$. Accordingly the Planner chooses a commitment that forces every type higher than $\hat{\theta}$ to fully commit. For any type in $\left[\hat{\theta},\bar{\theta}\right]\backslash\hat{\Theta}$, the Planner's posterior attributes at least $1-\rho$ to $\left[\hat{\theta},\hat{\theta}+\varepsilon'\right]$, and therefore, the Planner continues to fully commit these types. Because $\mu_0\left(\left[\hat{\theta},\bar{\theta}\right]\backslash\hat{\Theta}\right)>0$, it follows that learning is inadequate.

Consider Models 2*a*-*b*: by Proposition 2, FCD is violated if $\hat{\Theta}=\left[\underline{\theta},\tilde{\theta}\right]$ for some $\tilde{\theta}<\bar{\theta}$ and $a_D\left(\bar{s},\tilde{\theta},A,0\right)<\hat{a}$. Consider $\check{\theta}>\tilde{\theta}$ such that $a_D\left(\bar{s},\check{\theta},A,0\right)<\hat{a}$; denote $\check{a}$ for $a_D\left(\bar{s},\check{\theta},A,0\right)$. I establish that for every $\delta>0$, there exists an open set of beliefs that induce a Planner to never select a commitment c that would permit an action below $\check{a}$. Accordingly, the Planner is not able to distinguish types in $\left(\tilde{\theta},\check{\theta}\right]$ (which are not in $\hat{\Theta}$) from types in $\hat{\Theta}$.

First, I establish a preliminary lemma. For two commitments c and c', let $d(c,c')$ denote the Hausdorff distance between their associated menus, and let $N_\partial(c)$ denote the (relativized) open neighborhood of c. Let $C_\varepsilon(\theta)=\{c\in\mathcal{C}:\pi(\theta,c)\geq\pi^*(\theta)-\varepsilon\}$ denote the set of commitments that approximate the full commitment payoffs for type θ. The following lemma states that such commitments must be close to those that are optimal for type θ.

Lemma 9. *For every $\partial>0$, there exists $\bar{\varepsilon}>0$ such that for every $\varepsilon<\bar{\varepsilon}$, $C_\varepsilon(\theta)\subseteq\cup_{c\in C^*(\theta)}N_\partial(c)$.*

Proof. Suppose otherwise towards a contradiction: then there exists $\partial>0$ such that for every $\varepsilon>0$, $C_\varepsilon(\theta)\nsubseteq\cup_{c\in C^*(\theta)}N_\partial(c)$ for every n. Consider the subset of commitments $\mathcal{C}'=\mathcal{C}\backslash\cup_{c\in C^*(\theta)}N_\partial(c)$; by hypothesis, $\mathcal{C}'$ is non-empty and as the closed subset of a compact set, $\mathcal{C}'$ is compact. Because $C_\varepsilon(\theta)\cap\mathcal{C}'\neq\emptyset$ for every $\varepsilon>0$, $\sup_{c\in\mathcal{C}'}\pi(\theta,c)=\pi^*(\theta)$, and yet, because $C^*(\theta)\cap\mathcal{C}'=\emptyset$, $\pi(\theta,c)$ does not attain its maximum on $\mathcal{C}'$. Because π is continuous, this contradicts the compactness of $\mathcal{C}'$. Q.E.D.

Let $\mathcal{C}''$ denote the set of commitments that permit some action below $\check{a}$ to be chosen by the Doer. The intersection of $\mathcal{C}''$ and $C^*(\underline{\theta})$ is empty because whenever an action below $\check{a}$ is permitted, the Doer of type $\underline{\theta}$ does not choose action $\hat{a}$ with probability 1. Moreover, there exists $\partial > 0$ sufficiently small such that $\mathcal{C}'' \nsubseteq \cup_{c \in C^*(\theta)} N_\partial(c)$. Therefore, as an implication of Lemma 9, there exists $\varepsilon > 0$ such that $\sup_{c \in \mathcal{C}''} \pi(\underline{\theta}, c) < \hat{\pi} - \varepsilon$. By continuity of π, there exists $\varepsilon', \gamma' > 0$ such that $\pi(\theta, c) < \hat{\pi} - \gamma'$ for all θ in $[\underline{\theta}, \underline{\theta} + \varepsilon']$. Let $\hat{c} \in \hat{C}(\underline{\theta})$, and let $\check{c} \in \mathcal{C}''$. Using an argument analogous to that above, it follows that:

$$\begin{aligned} & (1-\delta) m(\mu, \hat{c}) + \delta \int_{P(\Theta)} V(\tilde{\mu}; \delta) \, dQ(\mu, \hat{c}) - \left((1-\delta) m(\mu, \check{c}) + \delta \int_{P(\Theta)} V(\tilde{\mu}; \delta) \, dQ(\mu, \check{c}) \right) \\ > \ & (1-\delta) \gamma' \mu([\underline{\theta}, \underline{\theta} + \varepsilon']) - (1 - \gamma' \mu([\underline{\theta}, \underline{\theta} + \varepsilon'])) \left(\int_S (u(a_P(s), s) - u(\hat{a}, s)) \, d\nu \right). \end{aligned}$$

Since γ' is strictly positive and independent of δ, it follows that for each δ, there exists $\rho \in (0,1)$ such that $\mu\left[\left(\hat{\theta}, \hat{\theta} + \varepsilon'\right)\right] > 1 - \rho$ ensures that the Planner prefers to select a commitment $\hat{c}$ over $\check{c}$. To establish the claim in the text, consider μ_0 such that $\mu_0([\underline{\theta}, \underline{\theta} + \varepsilon']) > 1 - \rho$: observe that when the Doer is of type $\left[\underline{\theta}, \check{\theta}\right]$, the Planner never learns the Doer's type because he never chooses commitments that permit actions below $\check{a}$ and therefore, his posterior beliefs after every history ascribe at least probability $1 - \rho$ to $[\underline{\theta}, \underline{\theta} + \varepsilon']$. Since $\mu_0\left(\left[\underline{\theta}, \check{\theta}\right] \backslash \hat{\Theta}\right) > 0$, learning is inadequate.

Proof of Theorem 4

The argument proceeds in two steps, similar to those of Theorems 3.3 and 4.3 of Aghion et al. (1991). Fix $\varepsilon > 0$.

Step 1: For a finite set of commitments $\mathcal{C}_\mathcal{N} = \{c_1, ..., c_N\}$, let $\pi^*_{\mathcal{C}_\mathcal{N}}(\theta) = \max_{c_n \in \mathcal{C}_\mathcal{N}} \pi(\theta, c)$. First I establish that there exists $\mathcal{C}_\mathcal{N}$ such that

$$\sup_{\theta \in \Theta} \left(\pi^*(\theta) - \pi^*_{\mathcal{C}_\mathcal{N}}(\theta) \right) \leq \frac{\varepsilon}{2}. \tag{4}$$

For two commitments, c and c', let $d(c, c')$ denote their distance; in Model 1, this corresponds to the standard (Euclidean) norm on nudges $\mathcal{N}$, and in Model 2, this corresponds to the Hausdorff metric on menus $\mathcal{M}$, and let $B_\partial(c) = \{c' \in \mathcal{C} : d(c', c) \leq \delta \}$. By Lemma 5 and Lemma 6, π is continuous in (θ, c), and because $\Theta \times \mathcal{C}$ is compact, π is uniformly continuous. Therefore, there exists ∂ such that

$$\sup_{\theta \in \Theta} (\pi^*(\theta) - \pi(\theta, c)) \leq \frac{\varepsilon}{2} \text{ for every } c \in \cup_{c' \in C^*(\theta)} B_\partial(c').$$

Since $\mathcal{C}$ is compact, there exists a finite sub-cover $\mathcal{C}_\mathcal{N} = \{c_n\}_{n=1}^N$ such that for every $c \in \mathcal{C}$, there exists $c_n \in \mathcal{C}_{\mathcal{N}}$ such that $d(c, c_n) < \partial$. Therefore, Inequality 4 is satisfied.

Step 2: To consider the setting that yields the minimal information in each period, I suppose that σ_t is completely uninformative; therefore, the Planner learns about θ only through the action frequencies. For any subset $A' \subseteq A$, define $\varphi(A', \theta, c) = \nu(\{s : a_D(s, \theta, c) \in A'\})$. I begin by establishing that for any commitment c_n, the Planner can use the empirical frequency of actions to accurately estimate his payoffs from choosing c_n repeatedly. First, consider the case in which for every action $a \in A$, $\varphi(\{a\}, \theta, c_n) < 1$. Denote the posterior mean after observing history h^t by $\hat{\theta}_t$ and denote the estimate of expected payoffs by $\hat{\pi}_t(c_n) = \pi\left(\hat{\theta}_t, c_n\right)$. By Corollary 2, $\Pr\left(\lim_{t\to\infty} \hat{\theta}_t = \theta\right) = 1$, and by the Continuous Mapping Theorem, $\Pr(\lim_{t\to\infty} \hat{\pi}_t(c_n) = \pi(\theta, c_n))$. In the complementary case in which there exists some $a \in A$ such that $\varphi(\{a\}, \theta, c_n) = 1$, then it trivially follows that $\Pr(\lim_{t\to\infty} \hat{\pi}_t(c_n) = \pi(\theta, c_n)) = 1$. Therefore, for every $\varepsilon > 0$, there exists T_N such that for every $t \geq T_N$,

$$\Pr\left(\sup_{n\leq N} |\hat{\pi}_t(c_n) - \pi(\theta, c_n)| \geq \frac{\varepsilon}{2}\right) \leq \varepsilon.$$

Consider a strategy in which the Planner chooses each commitment $c_n \in \mathcal{C}_{\mathcal{N}}$ a total of T_N times and then chooses from the commitment that yields the highest estimate of payoffs. Let θ denote the type of the Doer. Using K to denote a bound on $|u(a, s)|$, it follows that for every δ:

$$V(\mu_0; \delta) \geq \left(1 - \delta^{NT_N}\right)(-K) + \delta^{NT_N}\left((1-\varepsilon)\left(\int_\Theta \pi^*_{\mathcal{C}_{\mathcal{N}}}(\theta)\, d\mu_0 - \frac{\varepsilon}{2}\right) - K\varepsilon\right).$$

Therefore, taking the limit as the Planner becomes arbitrarily patient and applying Inequality 4,

$$\lim_{\delta\to 1} V(\mu_0; \delta) \geq (1-\varepsilon)\left(\int_\Theta \pi^*(\theta)\, d\mu_0 - \varepsilon\right) - K\varepsilon.$$

Since ε is arbitrary, the conclusion follows.

Proof of Theorem 5

Suppose that FCD is satisfied. Trivially, there exists AC-RSCE in which the Planner chooses a commitment $C^*(\theta)$ and has beliefs that ascribe probability 1 to type θ. In any other AC-RSCE, it must be the case that the Planner is fully committing; i.e., the commitment choice c is rationalized by some belief μ such that $m(\mu, c) = m^*(\mu) = \hat{\pi}$. By the argument of Theorem 2, $Supp(\mu) \subseteq \hat{\Theta}$, and therefore, such AC-RSCE can arise if and only if type θ is in $\hat{\Theta}$.

When FCD fails, in Model 1, consider types and beliefs from the proof of Theorem 3 which would suffice for $\delta = 0$; such beliefs rationalize full commitment while including the true type in its support. In Models 2a-b, referring to proof of Theorem 3, consider a ρ that corresponds to $\delta = 0$. Consider a belief μ_0 that ascribes of at least probability $(1-\rho)$ to $[\underline{\theta}, \underline{\theta} + \varepsilon']$, and the remaining measure to $\left[\underline{\theta} + \varepsilon', \check{\theta}\right]$. By the argument of Theorem 3, the Planner will never choose an action below $\check{a}$, and because the Planner ascribes no probability to types above $\check{\theta}$, there is no reason to choose an action below $\hat{a}$. Thus, for types in $\left[\tilde{\theta}, \check{\theta}\right]$, there exists AC-RSCE in which the Planner chooses to fully commit although for all such types, a fully informed Planner would attain strictly higher payoffs.

A.3 Section 4.1

I first prove the result when the Planner has access to merely illiquid assets and then extend it to the case in which the Planner also has access to singleton menus. I begin by describing properties of the savings problem that prove useful in later analysis. Let $U(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$, when $\sigma \neq 1$, and $U(c) = \log(c)$ if $\sigma = 1$. Consider a recursive formulation of the problem in the absence of self-control problems:

$$V(u, y) = Max_{c \in [0,y]} uU(c) + \delta \int_{\mathcal{U}} V(u', R(y-c))\, dF. \tag{5}$$

Restricting attention to a linear consumption path, let $c_P(u)\, y$ denote the optimal policy when the wealth is y and value of consumption is u. The stochastic Euler equation is

$$\frac{u}{(c_P(u)\, y)^{\sigma}} = \delta R \int_{\mathcal{U}} \frac{u'}{(c_P(u')\, R(1-c_P(u))\, y)^{\sigma}} dF. \tag{6}$$

A guess-and-verify approach establishes that

$$c_P(u) = \frac{Cu^{1/\sigma}}{Cu^{1/\sigma} + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}},$$

where C is the unique solution to $E\left[\left(Cu^{1/\sigma} + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}\right)^{\sigma}\right] = 1$ satisfies the Euler equation.

Consider the solution to the optimization problem when the Planner is restricted to having no flexibility:

$$\begin{aligned} \hat{V}(y) &= \max_{c \in [0,y]} \left[\int_{\mathcal{U}} uU(c)\, dF + \delta \hat{V}(R(y-c))\right] \\ &= \max_{c \in [0,y]} \left[U(c)\, dF + \delta \hat{V}(R(y-c))\right], \end{aligned} \tag{7}$$

where the equality follows from $E[u] = 1$. A guess and verify approach yields that the optimal policy is a *consumption rate* $\hat{c} = 1 - \delta^{1/\sigma} R^{(1-\sigma)/\sigma}$, and holding fraction $\hat{s} = 1 - \hat{c}$ of illiquid wealth. This fraction $\hat{s}$ corresponds to $\hat{a}$ from Section 3, the optimal action when there is no ability to respond to shocks. Moreover, the value satisfies

$$\hat{V}(y) = \begin{cases} \dfrac{y^{1-\sigma}}{(1-\sigma)\hat{c}^{\sigma}} - \dfrac{1}{(1-\sigma)(1-\delta)} & \text{if } \sigma > 1, \\ \dfrac{\log(y) + \log(1-\delta)}{\hat{c}} + \dfrac{\delta \log(R\delta)}{(1-\delta)^2} & \text{if } \sigma = 1 \end{cases},$$

and therefore, $\hat{V}'(y) = \left(\dfrac{1}{\hat{c}y}\right)^{\sigma}$.

Lemma 10. *$\hat{c} \in (c_P(\underline{u}), c_P(\bar{u}))$ for all y.*

Proof. If $c_P(\underline{u}) > \hat{c}$, then since c is strictly increasing, for all $u \in [\underline{u}, \bar{u}]$,

$$\frac{Cu^{1/\sigma}}{Cu^{1/\sigma} + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}} > 1 - \delta^{1/\sigma} R^{(1-\sigma)/\sigma}$$

yields

$$\left(Cu^{1/\sigma} + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}\right)^{\sigma} > 1,$$

and thus, $E\left[\left(Cu^{1/\sigma} + \delta^{1/\sigma} R^{(1-\sigma)/\sigma}\right)^{\sigma}\right] > 1$, contradicting the definition of C. An analogous argument establishes that $c_P(\bar{u}) > \hat{c}$. Q.E.D.

Now consider the full information benchmark in which the Planner has beliefs that ascribe probability 1 to type θ. For any $\bar{c}$, let $\hat{u}(\theta, \bar{c}, y)$ denote the bunching point:

$$\hat{u}(\theta, \bar{c}, y) = \begin{cases} \max\{u \in \mathcal{U} : c(u, \theta)\, y \leq \bar{c}\} & \text{if } c(\underline{u}, \theta)\, y \leq \bar{c}, \\ \underline{u} & \text{otherwise.} \end{cases}$$

The dynamic program that a fully informed Planner faces is

$$\bar{V}^{\theta}(y) = Max_{\bar{c}} \left(\begin{array}{c} \int_{u < \hat{u}(\theta,\bar{c},y)} \left(uU(yc(u,\theta)) + \delta \bar{V}^{\theta}(Ry(1 - c(u,\theta)))\right) dF \\ + \int_{u \geq \hat{u}(\theta,\bar{c},y)} \left(uU(\bar{c}) + \delta \bar{V}^{\theta}(R(y - \bar{c}))\right) dF \end{array} \right).$$

Lemma 11. *For each Doer type θ, there exists a unique optimal bunching point $\hat{u}(\theta)$ independent of y.*

Proof. Suppose that $\hat{u}(\theta, \bar{c}, y) \in (\underline{u}, \bar{u})$ for some y: then the optimization problem can be rewritten as selecting some $\hat{u}$ and with $\bar{c} = yc(\hat{u}, \theta)$. The above program can be re-written as

$$\bar{V}^{\theta}(y) = Max_{\hat{u}} \left(\begin{array}{c} \int_{u<\hat{u}(y)} \left(uU\left(yc\left(u,\theta\right)\right) + \delta \bar{V}^{\theta}\left(Ry\left(1-c\left(u,\theta\right)\right)\right)\right) dF \\ + \int_{u \geq \hat{u}(y)} \left(uU\left(yc\left(\hat{u},\theta\right)\right) + \delta \bar{V}^{\theta}\left(R\left(y-\bar{c}\right)\right)\right) dF \end{array} \right).$$

Suppose that the solution were a function $\hat{u}(\theta, y)$ which depended on y. Let $y_0 \neq 1$ be two exogenous starting wealths in which $\hat{u}(\theta, y_0) \neq \hat{u}(\theta, 1)$. For any sequence $\{u_t\}_{t=0}^{\infty}$, consider the induced choice of bunching points $\{\hat{u}_t\}_{t=0}^{\infty}$ when the wealth is y_0 and the choice of bunching points $\{\hat{u}'_t\}_{t=0}^{\infty}$ when the wealth is 1. Observe that if the Planner implemented the choice of bunching points $\{\hat{u}_t\}_{t=0}^{\infty}$ when the wealth is 1, his utility would be $\left(\frac{1}{y_0}\right)^{1-\sigma}$ multiplied by the utility he gets when the initial wealth is y_0, modulo some constant. Since this holds for every sequence of realizations $\{u_t\}_{t=0}^{\infty}$, it follows that

$$\left(\bar{V}^{\theta}(y_0) + C\right) \leq (y_0)^{1-\sigma} \left(\bar{V}^{\theta}(1) + C\right),$$

where C is a constant. The analogous logic implies that

$$\left(\bar{V}^{\theta}(y_0) + C\right) \geq (y_0)^{1-\sigma} \left(\bar{V}^{\theta}(1) + C\right),$$

and therefore that $\bar{V}^{\theta}(y_0) = (y_0)^{1-\sigma} \left(\bar{V}^{\theta}(1) + C\right) - C$. It immediately follows that the optimum $\hat{u}(\theta, y)$ is independent of y and can be written as $\hat{u}(\theta)$. Since the stage payoff is strictly concave in $\hat{u}$, it follows from standard arguments that $\hat{u}$ is unique. Q.E.D.

Lemma 12. *The Planner fully commits the Doer if and only if* $c(\underline{u}, \theta) \geq \hat{c}$.

Proof. Applying Leibniz's Rule, the derivative of the objective function with respect to $\hat{u}$ is

$$\left(\int_{u>\hat{u}} \frac{u}{\left(yc\left(\hat{u},\theta\right)\right)^{\sigma}} dF - \delta R\left(1-F\left(\hat{u}\right)\right) \frac{d\bar{V}^{\theta}}{dy'}\Big|_{y'=Ry(1-c(\hat{u},\theta))} \right) \left(y \left. \frac{dc\left(u,\theta\right)}{du} \right|_{u=\hat{u}} \right). \qquad (8)$$

If the Planner chooses to fully commit in every future period, it follows that

$$\frac{d\bar{V}^{\theta}}{dy} = \frac{d\hat{V}}{dy} = \left(\frac{1}{\hat{c}y}\right)^{\sigma}.$$

For full commitment to be optimal, the sign of the derivative from Equation 8 at $\underline{u}$ must be non-positive implying that

$$\frac{1}{\left(c\left(\underline{u},\theta\right)\right)^{\sigma}} - \frac{\delta R^{1-\sigma}}{\left(1-c\left(\underline{u},\theta\right)\right)^{\sigma}} \left(\frac{1}{\hat{c}}\right)^{\sigma} \leq 0.$$

Substituting $(1-\hat{c})^{\sigma}$ for $\delta R^{1-\sigma}$ implies that $\hat{c} \leq c(\underline{u}, \theta)$. Thus, $\hat{\Theta} = [\theta : c(\underline{u}, \theta) \geq \hat{c}]$. Q.E.D.

Proof of Theorem 6

Observe that when the Planner purchases $\hat{s}y_t$ level of illiquid assets and offers the consumption menu $[0, \hat{c}y_t]$, it fully commits all types $\hat{\theta}$ in $\hat{\Theta}$ and partially commits all types θ not in $\hat{\Theta}$. This immediately implies that Equation 3 is satisfied, and the conclusion of globally adequate learning follows.

Now, I address the inclusion of singleton sets; in particular, suppose that the Planner could commit by offering a consumption menu $\{\hat{c}y_t\}$ when he has wealth y_t (any other singleton menu is *ex ante* inferior). The extra analysis that is necessary in this case is to establish that for any type $\theta \notin \hat{\Theta}$, the Planner prefers to offer the menu $[0, \hat{c}y_t]$ rather than fully commit using the singleton $\{\hat{c}y_t\}$. While this may be true more generally, I establish it only for $\sigma \geq 1$.

Lemma 13. *When facing a Doer of type $\theta < \frac{\hat{c}}{c(\underline{u})}$, the Planner strictly prefers offering the menu $[0, \hat{c}y_t]$ to the singleton $\{\hat{c}y_t\}$ in each period if $\sigma \geq 1$.*

Proof. The argument proceeds through several steps. I use a one-stage deviation principle argument whereby even if the Planner is using the singleton in each future period, he prefers to use the larger menu presently.

Begin by considering the auxiliary problem of finding $\tilde{c}(u)$ that solves for each u,

$$Max_c uU(cy) + \delta\hat{V}(R(1-c)y).$$

The first-order condition yields

$$\frac{u}{(\tilde{c}(u)y)^{\sigma}} = \frac{\delta R}{(\hat{c}Ry(1-\tilde{c}(u)))^{\sigma}},$$

and therefore, that

$$\tilde{c}(u) = \frac{\hat{c}u^{1/\sigma}}{\hat{c}u^{1/\sigma} + \delta^{1/\sigma}R^{(1-\sigma)/\sigma}}.$$

Thus, when $u < 1$, then the Planner chooses consumption $\tilde{c}(u)$ less than $\hat{c}$. Given the concavity of U, when $u \leq 1$, the Planner solving this auxiliary problem strictly prefers a consumption $[\tilde{c}(u), \hat{c})$ to the consumption $\hat{c}$. Next, I establish that for every u, $c_P(u) \geq \tilde{c}(u)$ so long as $\sigma \geq 1$. Recall that

$$c_P(u) = \frac{Cu^{1/\sigma}}{Cu^{1/\sigma} + \delta^{1/\sigma}R^{(1-\sigma)/\sigma}},$$

in which C solves

$$E\left[\left(Cu^{1/\sigma}+1-\hat{c}\right)^{\sigma}\right]=1.$$

Because $\sigma \geq 1$, it follows from Minkowski's Inequality that

$$\begin{aligned} 1 &= \left(E\left[\left(Cu^{1/\sigma}+1-\hat{c}\right)^{\sigma}\right]\right)^{1/\sigma} \\ &\leq \left(E\left[\left(Cu^{1/\sigma}\right)^{\sigma}\right]\right)^{1/\sigma}+\left(E\left[(1-\hat{c})^{\sigma}\right]\right)^{1/\sigma} \\ &= C+1-\hat{c}, \end{aligned}$$

and therefore that $C \geq \hat{c}$ and that $c_P(u) \geq \tilde{c}(u)$.

Now suppose that the Planner faces a Doer of type θ that satisfies $c(\underline{u},\theta) < \hat{c}$. Let $\hat{u}$ solve $c(u,\theta)=\hat{c}$; because $c(u,\theta) \geq c_P(u) \geq \tilde{c}(u)$, and $\tilde{c}(1)=\hat{c}$, it follows that $\hat{u} \leq 1$. For all $u \geq \hat{u}$, the Doer chooses the same consumption from the menu $\{\hat{c}y\}$ and $[0,\hat{c}y]$. However, if $u < \hat{u}$, the Doer chooses $c(u,\theta) \in [\tilde{c}(u),\hat{c})$ when offered the menu $[0,\hat{c}y]$. Because the Planner strictly prefers the choice of $c(u,\theta)$ over $\hat{c}$ for these taste shocks, and $\Pr(u < \hat{u}) > 0$, it follows that the Planner strictly prefers offering the menu $[0,\hat{c}y]$ to the menu $\{\hat{c}y\}$ presently, even when offering the singleton menu in the future. Thus, the Planner prefers the former to the latter in each period. Q.E.D.

Theorem 8. *Consider the setting in which the Planner has access to illiquid assets and singleton menus. Then if $\sigma \geq 1$, learning is globally adequate.*

Proof. By the analysis in Lemma 12, it follows that $\hat{\Theta} \supseteq \{\theta : c(\underline{u},\theta) \geq \hat{c}\}$. If $c(\underline{u},\theta) < \hat{c}$, then by Lemma 13, the Planner prefers the menu $[0,\hat{c}y]$ to $\{\hat{c}y\}$. Therefore, $\hat{\Theta} = \{\theta : c(\underline{u},\theta) \geq \hat{c}\}$. Parallel to the argument for Model 2 in Theorem 2, Equation 3 holds thereby establishing that learning is globally adequate. Q.E.D.

A.4 Section 4.2

Proof of Theorem 7

Suppose that the Doer's type is given by θ. Let

$$\pi(\theta,\Gamma)=\int_{\left\{s:\Gamma(s)\geq\frac{s}{\theta}\right\}}(b-\Gamma(s))\,ds+\int_{\left\{s:\Gamma(s)<\frac{s}{\theta}\right\}}(s-\Gamma(s))\,ds$$

denote the Planner's payoff when he chooses to remain flexible and exert self-control $\Gamma(s)$

in state s. Naturally, an optimal rule is a cutoff of the form

$$\Gamma(s) = \begin{cases} s/\theta & \text{if } s < s^*, \\ 0 & \text{otherwise.} \end{cases}$$

At the cutoff, the Planner is indifferent between exerting self-control to complete the task and not doing so and thus, the cutoff value solves

$$b - \frac{s^*}{\theta} = s^*,$$

thereby equaling $\frac{\theta b}{\theta+1} \in (0,1)$. The value from flexibility when the Doer is of type θ is therefore, $\frac{1}{2}\left(\frac{\theta b^2}{\theta+1} + 1\right)$. In contrast, the value from commitment is simply b, and therefore, the Planner prefers to be flexible and exercise self-control if and only if

$$\frac{1}{2}\left(\frac{\theta b^2}{\theta+1} + 1\right) \geq b,$$

which reduces to $\theta > \theta^* = \frac{(1-b)^2}{2b-1}$.

Suppose that $\underline{\theta} > \theta^*$: then $\hat{\Theta}$ is empty because the Planner chooses the menu $\{-1, 1\}$ and exert costly self-control for each type of the Doer. Thus, FCD trivially holds. Since the distribution of σ is increasing in $\Gamma - \frac{s}{\theta}$ (in terms of FOSD), it follows from Lemma 3 of Easley and Kiefer (1988) that for any limit choice in which the Planner chooses the menu $\{-1, 1\}$, and applies self-control according to some function $\Gamma(s)$, the Planner's limit belief almost-surely is ∂_θ. Therefore, learning can be impeded only if the Planner chooses the menu $\{1\}$. Notice that because $\underline{\theta} > \theta^*$, the payoff from choosing the menu $\{-1, 1\}$, and exerting self-control $s/\underline{\theta}$ when s is less than $\frac{\theta b}{\underline{\theta}+1}$ is greater than b, regardless of the Doer's actual type θ. Therefore, by Lemma 8, for any possible limit belief μ_∞, the Planner never chooses the menu $\{1\}$. Globally adequate learning follows.

Now suppose that $\underline{\theta} < \theta^*$: then the Planner prefers to remain flexible for $\theta > \theta^*$ but fully commit for $\theta \leq \theta^*$; i.e., $\hat{\Theta} = [\underline{\theta}, \theta^*]$. FCD is necessarily violated: for all $\hat{\theta} \in \hat{\Theta}$, full commitment comes from choosing the singleton menu and not exerting any costly self-control, which also fully commits all types not in $\hat{\Theta}$. That for every $\delta > 0$, learning is inadequate for some open set of priors follows from an argument identical to that in Theorem 3: there exists $\varepsilon', \gamma' > 0$ such that $\pi(\theta, \Gamma) < \hat{\pi} - \gamma'$ for all θ in $[\underline{\theta}, \underline{\theta} + \varepsilon']$, and for all functions $\Gamma : [0,1] \to \Re_+$. The conclusion then follows that for every $\delta > 0$, for any full support prior that attributes sufficient mass to $[\underline{\theta}, \underline{\theta} + \varepsilon']$, the Planner chooses the menu $\{1\}$ and thus, learning is inadequate.

A.5 Section 5

Proof of Proposition 3

The precise condition that is sufficient is that $\frac{\mu_0\lambda_0}{\mu_0\lambda_0+(1-\mu_0)(1-\lambda_0)} > \frac{2b-1-\underline{\theta}b}{2b\left(\bar{\theta}-\underline{\theta}\right)}$; notice that the RHS is strictly in $(0,1)$ since $b \in \left(\frac{1}{2-\underline{\theta}}, \frac{1}{2-\bar{\theta}}\right)$, and therefore, if μ_0 and λ_0 are sufficiently high, the condition is satisfied.

To trace out how this condition implies undercommitment: it can be shown through calculation that of the four possible cases $\left\{\left(\bar{\theta},0\right), \left(\bar{\theta},\left(\bar{\theta}-\underline{\theta}\right)b\right), \left(\underline{\theta},0\right), \left(\underline{\theta},\left(\bar{\theta}-\underline{\theta}\right)b\right)\right\}$, only if $(\underline{\theta},0)$ is realized does a fully informed Planner (who knows (θ,ξ)) choose to commit. Suppose that this is the realized state of the world. Along any history h^i in which the Planner has been flexible so far, the ratio of posteriors satisfies:

$$\frac{\Pr\left(\bar{\theta},\left(\bar{\theta}-\underline{\theta}\right)b|h^i\right)}{\Pr\left(\left(\underline{\theta},0\right)|h^i\right)} = \frac{\mu_0\lambda_0}{\left(1-\mu_0\right)\left(1-\lambda_0\right)}, \tag{9}$$

because the state-dependent action probabilities are identical. Consider a sample-path in which the Planner has been flexible at all stages up to this point (I later verify that this is optimal). By the Strong Law of Large Numbers, almost-surely, the limiting frequency with which $a=1$ is $\underline{\theta}b$, and therefore, the Planner's posterior ascribe zero probability to the states $\left\{\left(\bar{\theta},0\right), \left(\underline{\theta},\left(\bar{\theta}-\underline{\theta}\right)b\right)\right\}$. Let $\tilde{\mu}$ be the Planner's posterior belief that $(\theta,\xi) = \left(\bar{\theta},\left(\bar{\theta}-\underline{\theta}\right)b\right)$, and observe that such a Planner prefers flexibility to commitment if:

$$b-\frac{1}{2}-\tilde{\mu}\left(\bar{\theta}-\underline{\theta}\right)b < \tilde{\mu}\underline{\theta}b\left(b-\left(\frac{\left(\bar{\theta}-\underline{\theta}\right)b+\bar{\theta}b}{2}\right)\right)+\left(1-\tilde{\mu}\right)\underline{\theta}b\left(b-\frac{\underline{\theta}b}{2}\right) \tag{10}$$

which simplifies into $\tilde{\mu} > \frac{2b-1-\underline{\theta}b}{2b\left(\bar{\theta}-\underline{\theta}\right)}$. Observe that since $\tilde{\mu} = \frac{\mu_0\lambda_0}{\mu_0\lambda_0+(1-\mu_0)(1-\lambda_0)}$, the Planner prefers flexibility to commitment at the limit of such a sample-path.

It is straightforward to prove that the Planner would also remain flexible at every time period: the Planner ascribes the ratio of posteriors as in Equation 9 and ascribes positive probability to the two other states in which he strictly prefers to remain flexible. Therefore, Inequality 10 holds for any belief μ_t that the Planner has. Thus, the Planner perpetually remains flexible when $(\underline{\theta},0)$ is the realized state although he would fully commit were he to learn the state.

Now suppose that the Planner is informed that $\xi=0$; this coincides with Section 2 and so the results of Theorem 1 apply. If the Planner is instead informed that $\underline{\theta}=0$, then along any sample-path in which he remains flexible, his beliefs converge to putting probability 1 on $\xi=0$, and therefore, again, such a Planner chooses to commit.