Commitment contracts*

Philip Bond[†] Gustav Sigurdsson[‡]

July 29, 2009

Abstract

A large number of recent papers employ preferences with hyperbolic time discounting to study borrowing, saving, and investment decisions. The time inconsistency induced by hyperbolic discounting gives rise to a commitment problem: at any date, an individual would like to commit his future selves to some (state-contingent) course of action. With few exceptions, the existing literature places strong exogenous restrictions on the forms of commitment available. In this paper we analyze the commitment problem in a general contracting framework and characterize the circumstances under which it can—and cannot—be overcome by a suitably designed contract. The fact that an individual's different selves are analogous to multiple individuals who share the same information allows us to exploit insights from implementation theory. When the commitment problem can be solved, the solution entails giving future selves the option to consume early at a high penalty rate. Surprisingly, a hyperbolic individual's ability to save privately makes the commitment problem more difficult to overcome.

*We thank Greg Fischer, Dilip Mookherjee, Debraj Ray, and seminar audiences at El Colegio de Mexico, the London School of Economics, and the University of Pennsylvania for helpful comments. All errors are our own.

[†]University of Pennsylvania.

[‡]University of Pennsylvania.

1 Introduction

Preferences with hyperbolic time discounting, introduced by Strotz (1956),[1] are increasingly used to model individual behavior in a wide variety of settings such as consumer finance (e.g., Laibson 1996 on savings behavior in general; Laibson, Repetto, and Tobacman 1998 on retirement planning; Della Vigna and Malmendier 2004 and Shui and Ausubel 2004 on credit card usage; Skiba and Tobacman 2008 on payday lending; and Jackson 1986 on bankruptcy law), asset pricing (e.g., Luttmer and Mariotti 2003), and procrastination (e.g., O'Donoghue and Rabin 1999a, 1999b, 2001). In his original article, Strotz observed that hyperbolic discounting generates a demand for commitment devices.[2] While some subsequent papers have analyzed the effectiveness of particular forms of commitment (e.g., Laibson 1997 on illiquid assets such as housing wealth), very few papers have analyzed the extent to which commitment is attainable without a priori restrictions on the particular form of commitment device.[3]

In this paper, we analyze a consumption-saving problem in which an individual with hyperbolic time preferences would like to commit at $t = 0$ to a particular consumption plan. To this end, the individual can enter into a contract with an outside party (such as a bank). The key contracting difficulty, however, is that self 0's[4] desired consumption plan may depend on an unverifiable shock that is realized only at $t = 1$, after the contract is signed. Since the shock is unverifiable, the contract cannot directly condition the individual's consumption on its realization.[5] Rather, the contract must provide the individual with sufficient flexibility to choose his consumption in response to the shock's realization. At the same time, however, the contract must provide the individual's future selves with incentives to adhere to self 0's desired consumption plan.

A key insight of our paper is that hyperbolic discounting may actually make it easier to provide an individual with incentives to reveal his private information. The "problem" with hyperbolic discounting is that the individual is essentially a different self at each date. This problem, however, also generates a partially offsetting advantage: any private information held by the individual is

[1] See Frederick, Loewenstein and O'Donoghue (2002) for a recent review of models of time discounting.

[2] See Ariely and Wertenbroch (2002) for direct evidence of a demand for commitment.

[3] See O'Donoghue and Rabin (1999b); Della Vigna and Malmendier (2004); and Amador, Angeletos, and Werning (2006) for models of contracting with a hyperbolic individual. We discuss our relation to these papers in detail below.

[4] We follow the literature and refer to the agent at date t as *self* t.

[5] If the desired consumption plan were fixed or depended solely on verifiable shocks, then the agent's commitment problem would be easy to solve with such a contract: deposit the entire endowment with the bank and instruct the bank to disburse the endowment according to the desired consumption plan.

shared with his future selves. In particular, if self t is hit by a shock, then self $t+1$ also knows that self t was hit by the shock. The fact that the different selves share the same information increases the scope for writing contracts to elicit that information.[6]

In our main result, we characterize exactly when the individual can and cannot fully overcome his commitment problem, in the sense of entering a contract that commits him to self 0's desired consumption plan. Although we do not place any restrictions on the set of possible contracts, the contract that solves the commitment problem takes the familiar form of a long-term savings account with limited early-withdrawal rights. First, the savings account permits penalty-free early withdrawals that provide the individual with just enough flexibility to choose his consumption in response to the shock's realization. Second, in the period following a penalty-free withdrawal, the savings account permits additional withdrawals, but these carry a penalty.

1.1 An example in which commitment is attainable

At date 0, a retired individual with wealth of 3.5 knows that he has three periods to live. He also knows that his child will get married either at date 1 or at date 2, at which time he will incur an expense of 1/2. The individual has log preferences over each period's consumption and his time preferences are quasi-hyperbolic, with a hyperbolic discount factor of $\beta = 1/2$ and no regular time discounting.[7] He would like to commit to self 0's desired consumption plan: consuming exactly 1 at each future date. At the same time, he would also like to have an additional 1/2 available at the wedding date to pay for the wedding. The problem, of course, is that if 1.5 is available to him at date 1, then hyperbolic discounting will lead his future self to consume 1.5 at date 1, even if the wedding is not at date 1 and the wedding expense must rather be incurred at date 2.[8]

Suppose, however, that the individual arranges his financial affairs as follows. At date 0 he deposits 1 each in 1- and 2-period savings accounts and 1.5 in a 3-period savings account. Early withdrawals from 1- and 2-period savings accounts are not permitted, but 1/2 can be withdrawn without penalty from the 3-period account at either date 1 or date 2. Furthermore, if an early

[6] Formally, we use tools from the implementation theory and mechanism design literatures, as we discuss in more detail below. The fact that different selves have different preferences is key.

[7] Formally, his utility at date 0 is $\ln(c_1) + \ln(c_2) + \ln(c_3)$, at date 1 it is $\ln(c_1) + \beta \ln(c_2) + \beta \ln(c_3)$, and at date 2 it is $\ln(c_2) + \beta \ln(c_3)$, where $\beta = 1/2$.

[8] Formally, this commitment problem arises since $\ln(1.5) + \frac{1}{2}\ln\left(1 - \frac{1}{2}\right) + \frac{1}{2}\ln(1) = \ln\left(\frac{3}{2}\sqrt{\frac{1}{2}}\right) > 0 = \ln(1) + \frac{1}{2}\ln(1) + \frac{1}{2}\ln(1)$.

withdrawal is made at date 1, then an additional 1/4 can be withdrawn at date 2, but this second withdrawal carries a penalty of 1/4.

We claim that this arrangement allows the individual to commit to self 0's desired consumption plan while still retaining the flexibility needed to finance the wedding. To see why, first consider self 2's optimal course of action if self 1 withdrew early. In this case, self 2 will withdraw early if and only if the wedding is in fact at date 2.[9] As a result, selves 2 and 3 consume 1 each if self 1 withdrew early to finance the wedding, but 3/4 and 1/2 if self 1 withdrew early to overconsume. It is easy to verify that if self 1 did not take advantage of the penalty-free early withdrawal, then self 2 will, regardless of when the wedding actually occurs.[10]

Next consider self 1's optimal course of action. If the wedding is not at date 1, then self 1 understands that if he were to withdraw early, then self 2 would withdraw early to finance the wedding at date 2, leaving self 3 with very little consumption because of the early-withdrawal penalty. This outcome is unattractive enough to deter self 1 from withdrawing early if the wedding is not at date 1.[11] If, however, the wedding is in fact at date 1, then self 1 will withdraw early the 1/2 he needs, secure in the knowledge that self 2 will not withdraw early and incur the penalty.[12]

A key insight illustrated by the example is that when preferences are hyperbolic, then giving the individual flexibility to overconsume at one date may deter him from overconsuming at an earlier date. In particular, the second, penalty-bearing withdrawal right in the example gives self 2 the flexibility to overconsume. In a world with full commitment, this added flexibility would not be needed, as the first, penalty-free, withdrawal right gives the individual sufficient flexibility to respond to the realization of the shock. In our setting, however, the second withdrawal right serves as a way for self 2 to "punish" self 1 for overconsuming. Loosely speaking, the contract puts the individual in a position where he realizes that if he "slips" at date 1 and overconsumes, then he will "fall off the wagon" at date 2 and overconsume even more.

[9] Formally, $\ln\left(1+\frac{1}{4}-\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}-\frac{1}{2}\right)=\ln\left(\frac{3}{4}\sqrt{\frac{1}{2}}\right)>\ln\left(1-\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)$ (wedding is at date 2) and $\ln(1)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)=0>\ln\left(\frac{5}{4}\sqrt{\frac{1}{2}}\right)=\ln\left(1+\frac{1}{4}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}-\frac{1}{2}\right)$ (wedding is at date 1).

[10] Formally, $\ln\left(1+\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)=\ln\left(\frac{3}{2}\right)>\ln\left(\sqrt{\frac{3}{2}}\right)=\ln(1)+\frac{1}{2}\ln(1.5)$ (wedding is at date 1) and $\ln\left(1+\frac{1}{2}-\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)=0>\ln\left(\frac{1}{2}\sqrt{\frac{3}{2}}\right)=\ln\left(1-\frac{1}{2}\right)+\frac{1}{2}\ln(1.5)$ (wedding is at date 2).

[11] Formally, $\ln(1)+\frac{1}{2}\ln\left(1+\frac{1}{2}-\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)=0>\ln\left(\frac{3}{2}\sqrt{\frac{3}{4}\frac{1}{2}}\right)=\ln\left(1+\frac{1}{2}\right)+\frac{1}{2}\ln\left(1+\frac{1}{4}-\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}-\frac{1}{2}\right)$.

[12] Formally, $\ln\left(1+\frac{1}{2}-\frac{1}{2}\right)+\frac{1}{2}\ln(1)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)=0>\ln\left(\frac{1}{2}\sqrt{\frac{3}{2}}\right)=\ln\left(1-\frac{1}{2}\right)+\frac{1}{2}\ln\left(1+\frac{1}{2}\right)+\frac{1}{2}\ln\left(1.5-\frac{1}{2}\right)$.

1.2 An example in which commitment is not attainable

Consider the following variant of the example above: the wedding either takes place at date 1 or does not take place at all. In this case, the individual would like to consume 1 at each date (withrawing an additional 1/2 at date 1) if the wedding takes place and consume 7/6 at each date if there is no wedding. As we show immediately below, the commitment problem is much more difficult to overcome in this case.

Let $(\hat{c}_2, \hat{c}_3)$ denote the consumption chosen by self 2 when he "punishes" self 1 for withdrawing 1/2 early when the wedding does not take place. Recall, moreover, that self 0 wants self 2 to choose $(1, 1)$ if the wedding does in fact take place. In order for self 2 to prefer $(\hat{c}_2, \hat{c}_3)$ to $(1, 1)$ if and only if the wedding does not take place, we need both $\ln(\hat{c}_2) + \beta \ln(\hat{c}_3) \geq \ln(1) + \beta \ln(1)$ and $\ln(1) + \beta \ln(1) \geq \ln(\hat{c}_2) + \beta \ln(\hat{c}_3)$. Consequently, the solution to the commitment problem is now very delicate: it requires self 2 to break the tie in favor of $(\hat{c}_2, \hat{c}_3)$ if and only if self 1 overconsumed.

The above suggests that it is just possible to overcome the commitment problem in the second example. This, however, does not remain true in the (realistic) case where the individual can save privately from one date to the next. To see this, note that self 1 must prefer $(1, 1, 1)$ to $(1.5, \hat{c}_2, \hat{c}_3)$. Together with self 2's indifference, this implies that $\hat{c}_2 > 1 > \hat{c}_3$, i.e., the "punishment" must entail overconsumption at date 2.[13]

But then it follows that if there is no wedding, then self 1 can withdraw 1/2 early and save a small amount until date 2. Self 2 will then choose $(1, 1)$ since if he is indifferent between $(\hat{c}_2, \hat{c}_3)$ and $(1, 1)$ when self 1 does not save, then he must strictly prefer the latter whenever he inherits any strictly positive level of savings. In essence, the ability to save privately undercuts commitment because it enables self 1 to overconsume and then bribe self 2 to not impose the punishment. Consequently, and perhaps surprisingly, the ability to save privately considerably reduces the range of outcomes that a hyperbolic individal can commit to.

1.3 Discussion

Comparing the two examples, one can see how the nature of the shock determines the difficulty of overcoming the commitment problem. In the former example, self 2's marginal utility of withdraw-

[13]Formally, $\ln(1) + \beta \ln(1) + \beta \ln(1) \geq \ln(1.5) + \beta \ln(\hat{c}_2) + \beta \ln(\hat{c}_3)$ and $\ln(1) + \beta \ln(1) = \ln(\hat{c}_2) + \beta \ln(\hat{c}_3)$ imply $\hat{c}_2 > 1 > \hat{c}_3$. As we show in Section 4, this property of the punishment generalizes beyond this particular example.

ing early is strictly greater when the wedding takes place at date 2 (because of the wedding expense) than when it takes place at date 1. As a result, self 2 strictly prefers withdrawing (respectively, not withdrawing) early if the wedding actually is at date 2 (respectively, date 1). In the latter example, however, self 2's marginal utility of withdrawing early is the same in both states, which makes it difficult to induce him to make a different consumption choice in each state. More generally, commitment is possible only under the following *preference reversal* condition: restricting attention to the case of two states, the state in which self 0's desired date 1 consumption is higher must also be the state in which self 2's marginal utility is lower (relative to marginal utility at date 3).

The above examples illustrate the two most important forces—namely, preference reversal, and the agent's ability to save privately—that affect an individual's ability to overcome his commitment problem. Together, they imply that self 2 must experience preference reversal that is sufficiently strong to induce him to punish if and only if called upon to do so, even if self 1 attempts to bribe him by saving from date 1 to date 2. We give a formal definition of this condition—*strong preference reversal*—in the text below. It is the key necessary condition for overcoming the commitment problem.

Substantively, preference reversal derives from the nature of the shock that the individual faces. When the shock is a *timing shock*, in which an event arrives with certainty but at an uncertain date, commitment is very often possible. Likewise, if uncertainty is over *whether or not* a shock occurs and the individual learns *ahead of time* that a shock will occur (e.g., at date 1 he learns about an essential date 2 expenditure), then commitment is also often possible. By contrast, if there is uncertainty over whether or not a shock occurs and the individual learns about the shock only *contemporaneously*, then commitment is impossible. The two wedding examples above correspond to the first and last case, respectively.

A second and distinct consequence of the individual's ability to save privately is that the second withdrawal right *must* carry a penalty. To see this, return to the example in which the wedding occurs either at date 1 or at date 2 and recall that if self 1 overconsumes, then he is punished by self 2's early withdrawal and overconsumption. When the agent can save privately, this punishment necessarily entails a reduction in total consumption—the penalty rate in the example. By contrast, when the agent cannot save privately, this punishment can be attained without decreasing total consumption.

Our results show that in some settings, an individual can contract to completely overcome his commitment problem, even in the face of uncertainty about his future consumption needs. In these cases, hyperbolic discounting ceases to affect the individual's behavior. Moreover, the contracts that enable an individual to commit are often easy to interpret, as the first example above suggests: at date 0 the individual arranges access to a savings account with limited penalty-free early withdrawal rights, coupled with additional withdrawal rights that carry a penalty. Since there also exist important cases in which commitment is not attainable, however, our results provide partial support for both proponents and antagonists of the importance of hyperbolic discounting.

As the examples illustrate, an individual is only able to commit himself to a state-contingent consumption plan because he is self-aware enough to know that he will discount hyperbolically at future dates. That is, commitment hinges critically on the individual being a different self at each date, and moreover, knowing this in advance. We discuss the extreme case of zero self-awareness, often referred to as naïveté, at the end of the paper.

Finally, although we focus on one particular decision-making bias in this paper—the propensity to overweight current consumption—we believe that our general arguments are widely applicable to many other behavioral biases. For example, consider an individual who is self-aware and understands that he will misinterpret the relevance of a small number of data points in the future. Just as in the current setting, he can potentially commit to a course of action that avoids this bias, while at the same time maintaining flexibility to respond to shocks.

2 Related literature

The closest antecedent to our paper is Amador, Angeletos, and Werning (2006). Like us, they study a hyperbolic individual who is hit by unverifiable shocks, but consider only a two-period version of the problem. This seemingly small difference has large consequences, because with two periods there is effectively only one decision maker (self 1) and therefore any information private to the individual is also private to self 1. Consequently, there is no way to elicit this information without distorting consumption in at least some states. The authors characterize the least-costly way to distort consumption.

Della Vigna and Malmendier (2004) study optimal contracts with hyperbolic individuals where, again, the individuals possess some private information. Like Amador, Angeletos, and Werning they restrict attention to a two-period model, with the same consequences. Moreover, the individual makes only a binary decision in their model (e.g., go to gym, don't go to the gym), so there is little scope to distort outcomes. Instead, the main focus of their paper is on flat versus per-usage fees.

O'Donoghue and Rabin (1999b) study optimal contracts for procrastinators. In their environment, the socially efficient date at which a task should be performed is random. They focus on contracts that induce the individual to perform a task at the efficient date. The contracts they study typically involve a reallocation of resources across states. As they observe, this makes the first-best unattainable if only the individual knows the distribution of shocks. By contrast, the contracts in our paper make no use of the distribution of shocks and are therefore robust in this sense. Indeed, in our setting the individual would voluntarily provide all the information required for the contract at date 0.

In our setting, an individual may be able to commit to a consumption plan by having future selves punish him if he deviates. In our paper, we construct contracts that induce the future selves to punish. Bernheim, Ray, and Yeltikin (1999) and Krusell and Smith (2003) consider models in which an individual is infinitely lived. They show that Markov-perfect equilibria exist in which an individual gains some commitment ability from the fact that deviations will cause future selves to punish him. These papers, however, do not consider the scope for using a contract to increase commitment and do not allow for uncertainty.

Finally, our paper builds on the implementation theory literature[14]—see footnote 18 below—and also belongs to the growing literature on contracting problems with hidden savings (e.g., Kocherlakota 2004, Doepke and Townsend 2006, and He 2008).

[14]See Maskin and Sjöström (2002), Palfrey (2002), and Serrano (2004), for surveys of the implementation theory literature.

3 Model

A single agent consumes over three periods $t = 1, 2, 3$. His preferences are time-inconsistent, with self t's preferences represented by the quasi-hyperbolic separable utility function

$$U^t(c;\phi) = u_t(c_t;\phi) + \beta \sum_{s=t+1}^{3} u_s(c_s;\phi)$$

for $t = 0, 1, 2, 3$ where $c = (c_1, c_2, c_3) \in \mathbb{R}^3_+$ is a *consumption plan* ($u_0(c_0;\phi) = 0$ so date 0 consumption does not matter to the agent); $\phi \in \{\theta, \theta'\}$ is one of two *states* that determine the agent's preferences; u_t is strictly increasing and strictly concave in c_t for every t and every ϕ; and $\beta \in (0,1)$ is the hyperbolic discount factor. There is no regular time discounting and the risk free rate in the economy is zero.

The state is revealed to the agent at $t = 1$ and is unverifiable. Therefore, any contract that the agent enters into with an outside party cannot be made directly contingent on the state. Our main results concern the (realistic) case in which the agent can privately save from one period to the next and his consumption-saving decisions are unverifiable. We also analyze the case of verifiable consumption-saving decisions, which is of independent interest, in Section 4. Finally, we assume that self 2's preferences satisfy the following standard *single crossing property* (see Milgrom and Shannon 1994).

Assumption SC (Single crossing) *Fix ϕ, c, c^a, and c^b such that $c_2 < c_2^a \leq c_2^b$, $U^2(c^a;\phi) \geq U^2(c;\phi) \geq U^2(c^b;\phi)$, and $U^2(c;\phi') \geq U^2(c^a;\phi')$. Then $U^2(c;\phi') \geq U^2(c^b;\phi')$, where $\phi' \neq \phi$.*

In words, single crossing says that self 2's indifference curves in the two states can cross at most once. Two examples of more specific functional forms for utility that satisfy single crossing are *multiplicative shocks*, $\phi \in \mathbb{R}^3_+$ and $u_t(c_t;\phi) = \phi_t u_t(c_t)$; and *additive shocks*, $\phi \in \mathbb{R}^3$, where $\phi_3 = \phi_3'$, and $u_t(c_t;\phi) = u_t(c_t + \phi_t)$.[15] In addition, we make the very mild assumption that if, at any consumption plan, self 2's indifference curves have the same slope in both states, then they must coincide.[16] While both the multiplicative and additive parameterizations can be interpreted

[15]More generally, a sufficient condition for single crossing is that the sign of $\frac{u_2'(x_2;\phi)}{u_3'(x_3;\phi)} - \frac{u_2'(x_2;\phi')}{u_3'(x_3;\phi')}$ is independent of (x_2, x_3), i.e., if shocks affect the ratio of marginal utility across dates 2 and 3 in relatively uniform way across consumption bundles.

[16]We use this assumption only to establish equivalence of the two conditions stated in Proposition 1. It is satisfied

in a large number ways, several interpretations deserve particular discussion:

1. Under the widely used parameterization of utility functions in which consumption and leisure enter multiplicatively, multiplicative shocks can be interpreted (among other ways) as shocks to time endowments. For example, if date t is a vacation day for the agent in state ϕ, his marginal utility of consumption is high.

2. Additive shocks $\phi_t \leq 0$ can be interpreted as essential expenditures. For example, if at date t an individual is sick in state ϕ but not state ϕ', and must pay \$100 for treatment in state ϕ, his utility from spending a total of c_t in state ϕ is the same as spending $c_t - 100$ in state ϕ'. The wedding example of the introduction entails shocks of this type.

3. Symmetrically, additive shocks $\phi_t > 0$ can be interpreted as (unverifiable) increases in income. With slight abuse of language, we continue to refer to c_t as consumption, even though the agent's true consumption in this case is $c_t + \phi_t$.

Finally, we assume throughout that the period utility functions satisfy an Inada condition in all states and at dates. To reflect the two interpretations of additive shocks above, we allow the Inada condition to depend on the state: there exists $\underline{c}_t(\phi)$ such that $u_t'(c_t;\phi) \to \infty$ as $c_t \to \underline{c}_t(\phi)$.[17]

3.1 Commitment problems

At date 0, self 0 would like to commit to his desired (i.e., most-preferred) state-contingent consumption plan $\{c(\phi)\}_\phi$. Let W denote the sum of the agent's initial endowment and verifiable income at each date; then, for $\phi \in \{\theta, \theta'\}$,

$$c(\phi) \equiv \arg\max_{\tilde{c}} \left\{U^0(\tilde{c};\phi)\right\} \text{ s.t. } \sum_s \tilde{c}_s \leq W.$$

If the state were revealed at $t = 0$, then commitment could be easily attained by entering a contract with an outside party (hereafter, "bank") under which the agent gives the bank his entire endowment and instructs the bank to return it over time according to the desired consumption plan in that state.

for the two classes of preferences mentioned above.

[17]So in the two additive shock interpretations discussed, $\underline{c}_t(\phi) = -\phi_t$.

Since the state is not revealed until $t = 1$, however, such a contract will generally not work as self 1 might be tempted to misreport the state to the bank. We say that the agent faces a *commitment problem in state* ϕ if self 0's preferences over c and c' are reversed at $t = 1$ in state ϕ, i.e.,

$$U^1 (c(\phi); \phi) < U^1 (c(\phi'); \phi)$$

where $\phi' \neq \phi$. As the following result shows, a commitment problem in state ϕ rules out a corresponding commitment problem in state $\phi' \neq \phi$ and can only arise when $c(\phi')$ offers more consumption than $c(\phi)$ at $t = 1$.

Lemma 1 *If there is a commitment problem in state* ϕ *then* $c_1(\phi') > c_1(\phi)$*, and so in particular there is no commitment problem in state* $\phi' \neq \phi$*.*

3.2 Alternative interpretations

Our model can be interpreted more broadly than the retired individual's consumption and saving problem from the introduction suggests. In one example, the individual faces a *procrastination* problem, where W is his total endowment of leisure (the time he has left after completing a referee report) and c_t is his leisure in period t. Shocks may be either multiplicative (in a favorite O'Donoghue and Rabin example, a Johnny Depp film is showing) or additive (the individual must take his child to the doctor). In another example, the individual is a *myopic manager*, where c_t is investment in period t, and u is a production function. In this case, hyperbolic discounting captures the manager's present-bias. Finally, although our formal model has three periods, one can interpret utility at date 3—$u_3(c_3; \phi)$—as the agent's expected future utility over a large number of future periods.

4 Commitment with verifiable savings

When does a contract exist that enables the agent to overcome his commitment problem, in the sense of commiting him to self 0's desired consumption plan $\{c(\phi)\}_\phi$? To build intuition, we analyze in this section the special case in which the agent's consumption-saving decisions are verifiable. To simplify notation, we will refer to $c(\theta)$ and $c(\theta')$ as c and c', respectively. We assume, without loss, that $c_1' \geq c_1$; Lemma 1 then implies that if there is a commitment problem, it is in state θ.

By standard revelation principle arguments (see Myerson 1981), we can restrict attention to contracts that give each of selves 1, 2, and 3 a menu of two consumption choices (corresponding to the two states), where each self's menu possibly depends on the previous selves' choices. Furthermore, since $t = 3$ is the final period, self 3 will always choose the highest consumption level on his menu, regardless of the true state. As a result, we can restrict attention to contracts in which self 2's consumption choice dictates self 3's consumption. Finally, by a similar argument (see, e.g., Cole and Kocherlakota 2001), we can restrict attention to contracts that do not allow the agent to save.

The consumption plans c and c' must be among the choices offered. Therefore, self 1's menu must include c_1 and c_1', and self 2's menu after self 1 chooses c_1 (respectively, c_1') must include (c_2, c_3) (respectively, (c_2', c_3')). As a result, self 1's menu must be the set $\{c_1, c_1'\}$ and self 2's menu must be the set $\{(c_2, c_3), (\hat{c}_2', \hat{c}_3')\}$ (if self 1 chooses c_1) or $\{(\hat{c}_2, \hat{c}_3), (c_2', c_3')\}$ (if self 1 chooses c_1'), where $(\hat{c}_2', \hat{c}_3')$ and $(\hat{c}_2, \hat{c}_3)$ are "punishments" chosen by self 2 if self 1 deviates from the desired consumption plan by choosing c_1 in state θ' or c_1' in state θ, respectively.

What are the conditions for incentive compatibility? First, self 1 must be better off choosing c_1 in state θ and c_1' in state θ', anticipating a punishment if he were to deviate. This is the case if

$$\begin{aligned} U^1(c;\theta) &\geq U^1(\hat{c};\theta) \\ U^1(c';\theta') &\geq U^1(\hat{c}';\theta') \end{aligned}$$

where $\hat{c} = (c_1', \hat{c}_2, \hat{c}_3)$ and $\hat{c}' = (c_1, \hat{c}_2', \hat{c}_3')$. Second, self 2 must be better off choosing the punishment in each state if and only if self 1 deviates. This is the case if

$$\begin{aligned} &U^2(\hat{c};\theta) \geq U^2(c';\theta) \text{ and } U^2(c';\theta') \geq U^2(\hat{c};\theta') \\ &U^2(c;\theta) \geq U^2(\hat{c}';\theta) \text{ and } U^2(\hat{c}';\theta') \geq U^2(c;\theta'). \end{aligned}$$

First note that when there is no commitment problem in either state, then the entire consumption decision can be delegated to self 1: $\hat{c} = c'$ and $\hat{c}' = c$ satisfy all the above constraints. Second, note that even when there is a commitment problem in state θ, then there is no need for self 2 to punish self 1 for choosing c_1 in state θ': $\hat{c}' = c$ satisfies all the above constraints in which $\hat{c}'$ appears. As a result, the interesting part of the contracting problem is the choice of $\hat{c}$, the punishment for choosing c_1' in state θ.

As the constraints above indicate, $\hat{c}$ must satisfy three somewhat conflicting criteria. The first two relate to preferences in state θ. First, $\hat{c}$ must be sufficiently *unattractive* relative to c' to deter self 1 from deviating. Second, $\hat{c}$ must be sufficiently *attractive* relative to c' to induce self 2 to punish. These two criteria are satisfied only if $\hat{c}_2 > c_2'$ and $\hat{c}_3 < c_3'$, i.e., if $\hat{c}$ is strictly *front-loaded* relative to c'. This follows from hyperbolic discounting: relative to consumption at date 3, self 2 values consumption at date 2 more than self 1 does. Therefore, if c' were instead front-loaded relative to $\hat{c}$ and self 1 preferred c' to $\hat{c}$, then self 2 would also prefer c' to $\hat{c}$ and would therefore not punish. We use this property, summarized in the following result, throughout the paper.

Lemma 2 *Fix ϕ, c^a, and c^b such that $c_1^a = c_1^b$, $U^1(c^a;\phi) \geq U^1(c^b;\phi)$, and $U^2(c^a;\phi) \leq U^2(c^b;\phi)$. Then $c_2^a \leq c_2^b$ and $c_3^a \geq c_3^b$, both with strict inequality if either $U^1(c^a;\phi) > U^1(c^b;\phi)$ or $U^2(c^a;\phi) < U^2(c^b;\phi)$.*

The third criterion that $\hat{c}$ must satisfy relates to preferences in state θ'. Here, $\hat{c}$ must be sufficiently unattractive relative to c' to deter self 2 from punishing. Given that $\hat{c}$ must be strictly front-loaded relative to c' and self 2 must prefer $\hat{c}$ to c' in state θ, we have established that a particular form of *preference reversal* is a necessary condition for commitment.

Condition PR (Preference reversal) *There exists $\tilde{c}$ such that $\tilde{c}_2 > c_2'$, $U^2(\tilde{c};\theta) \geq U^2(c';\theta)$, and $U^2(c';\theta') \geq U^2(\tilde{c};\theta')$.*

A commitment problem arises because, relative to consumption at future dates, self 1 values current consumption *more* in state θ' than in state θ (and so $c_1' > c_1$); Condition PR, however, says that, relative to consumption at future dates, self 2 values current consumption *less* in state θ' than in state θ.[18]

[18]Condition PR may remind readers of Maskin's (1999) monotonicity condition. However, while PR may fail in our setting, monotonicity is trivially satisfied as long as some self's preferences differ across the two states. In our setting, the social choice rule of interest is $F(\theta) = c$ and $F(\theta') = c'$, where the domain of alternatives from which c and c' are drawn is $\mathbb{R}_+^3$. This social choice rule is monotonic if and only if $U^t(c;\theta) \geq U^t(x;\theta)$ and $U^t(c;\theta') < U^t(x;\theta')$ for some self $t \in \{1,2,3\}$ (self 0 is non-strategic) and some $x \in \mathbb{R}_+^3$, and $U^t(c';\theta') \geq U^t(x;\theta')$ and $U^t(c;\theta) < U^t(x;\theta)$ for some self $t \in \{1,2,3\}$ and some $x \in \mathbb{R}_+^3$. As long as some self's preferences differ across the two states, this condition is satisfied. In particular, the second condition is satisfied by $t = 1$ and $x = c$; this follows directly from the fact that there is a commitment problem in state θ but not in state θ'.

Condition PR is easily checked by comparing the slopes of self 2's indifference curves through c' in states θ and θ'. Summarizing:

Proposition 1 *When savings are verifiable, commitment is possible only if Condition PR holds, or equivalently,* $U_2^2(c';\theta)/U_3^2(c';\theta) \geq U_2^2(c';\theta')/U_3^2(c';\theta')$.

Finally, it is straightforward to show that if the period utility functions u_2 and u_3 are unbounded above and below then Condition PR is sufficient as well as necessary for commitment.

5 Commitment with unverifiable savings

In the remainder of the paper we analyze the case in which the agent's consumption-saving decisions are unverifiable. We denote savings carried over from date t to date $t+1$ by $s_t \geq 0$ and since consumption and withdrawals need not coincide when the agent can save, we denote withdrawals by x to prevent confusion between the two. With slight abuse of notation, we denote the consumption plan $(-s_1, s_1, 0) + x + (0, -s_2, s_2)$ by $s_1 + x - s_2$ (the reader may find it useful to note that savings in this expression are written with respect to self 2, who is the key strategic actor).

The single crossing assumption above only applies to indifference curves (over consumption at dates 2 and 3) in the two states when the savings inherited from date 1 are the same in both states. In the analysis that follows, however, we must repeatedly compare indifference curves when the savings inherited from date 1 differ across the two states. To this end, we extend the single crossing assumption as follows.

Assumption SCB (Single crossing from below) *Fix* $s_1 \geq 0$, ϕ, x, x^a, *and* x^b *such that* $x_2 < x_2^a \leq x_2^b$, $U^2(s_1 + x^a;\phi) \geq U^2(s_1 + x;\phi) \geq U^2(s_1 + x^b;\phi)$, *and* $U^2(x;\phi') \geq U^2(x^a;\phi')$. *Then* $U^2(x;\phi') \geq U^2(x^b;\phi')$, *where* $\phi' \neq \phi$.

In words, single crossing from below says that once self 2's indifference curve in one state with no savings crosses his indifference curve in the other state with positive savings from below, they cannot cross again at higher levels of x_2. Whereas single crossing requires that indifference curves with the same savings cross at most once, single crossing from below (SCB) allows indifference curves with different savings levels to cross twice. In the special case of $s_1 = 0$, the two assumptions are equivalent; therefore, SCB implies single crossing.

SCB is satisfied by preferences with additive shocks and $\theta_3 = \theta_3'$, and by multiplicative shocks under the mild assumption that date 2 utility exhibits either constant or decreasing absolute risk aversion (i.e., $-u_2''/u_2'$ either constant or decreasing).

5.1 The formal contracting problem

As in the case of verifiable consumption-saving decisions, we can restrict attention to the direct revelation mechanism in which self 1 reports the state ϕ, and then self 2 reports the state ϕ and self 1's saving decision. As we show in the appendix, we can further restrict attention to contracts that give self 2 (c_2, c_3) after self 1 chooses c_1, with self 2 then deciding how much to save; this follows from the fact that there is no need to punish self 1 for incorrectly choosing c_1 in state θ', just as in the case of verifiable consumption and saving decisions.

As a result, a contract in our setting is as follows. Self 1 chooses from the set $\{c_1, c_1'\}$ and decides how much to save. If self 1 chooses c_1, then self 2 gets (c_2, c_3) and decides how much to save. If self 1 chooses c_1', then self 2 chooses a point on one of two consumption *schedules* $\{X(s_1)\}_{s_1 \geq 0}$ and $\{\hat{X}(s_1)\}_{s_1 \geq 0}$. We restrict attention to contracts that are finite, in the sense of having at most finitely many points of discontinuity.[19] If self 1 chooses c_1' in state θ' as intended and saves s_1, then self 2 should choose $X(s_1)$; if self 1 deviates and chooses c_1' in state θ and saves s_1, then self 2 should choose the punishment $\hat{X}(s_1)$. Self 2 cannot retroactively alter the choice made by self 1; therefore $X_1(s_1) = \hat{X}_1(s_1) = c_1'$ for every $s_1 \geq 0$. Furthermore, as in the case of verifiable consumption-saving decisions, we can restrict attention to contracts in which the agent does not save on the equilibrium path of any subgame. Finally, since the consumption plans c and c' must be among the choices offered, self 2's menu after self 1 chooses c_1' must include (c_2', c_3'); therefore $X(0) = c'$.

What are the conditions for incentive compatibility? First, self 1 should choose c_1 in state θ and c_1' in state θ' and not save in either state, anticipating a punishment if he were to deviate. This is the case if and only if

$$U^1(c;\theta) \geq U^1(s_1 + \hat{X}(s_1);\theta) \text{ for all } s_1 \geq 0, \tag{IC$_1$}$$

[19]This restriction is used only in proving the necessity half of Theorem 1. It can be relaxed, though only at the cost of introducing economically uninteresting mathematical complexity. Details are available from the authors upon request.

$$U^1\left(c';\theta'\right) \geq U^1\left(s_1 + X\left(s_1\right);\theta'\right) \text{ for all } s_1 \geq 0. \tag{IC$_1'$}$$

Second, if self 1 chooses c_1' and saves s_1, then self 2 should choose the punishment $\hat{X}(s_1)$ if the state is θ, the non-punishment $X(s_1)$ if the state is θ', and in either case not save. This is the case if and only if

$$U^2\left(s_1 + \hat{X}\left(s_1\right);\theta\right) \geq \begin{cases} U^2(s_1 + \hat{X}\left(\tilde{s}_1\right) - s_2;\theta) \\ U^2\left(s_1 + X\left(\tilde{s}_1\right) - s_2;\theta\right) \end{cases} \text{for all } s_1, \tilde{s}_1, s_2 \geq 0, \tag{IC$_2$}$$

$$U^2\left(s_1 + X\left(s_1\right);\theta'\right) \geq \begin{cases} U^2\left(s_1 + X\left(\tilde{s}_1\right) - s_2;\theta'\right) \\ U^2(s_1 + \hat{X}\left(\tilde{s}_1\right) - s_2;\theta') \end{cases} \text{for all } s_1, \tilde{s}_1, s_2 \geq 0. \tag{IC$_2'$}$$

As we show in the appendix, we can restrict attention to the X-schedule that results from giving self 2 (c_2', c_3') and letting him save whatever he likes. To see why, note that self 1 has no incentive to save after correctly choosing c_1' in state θ': c' is self 0's most preferred consumption plan and self 1 values date 1 consumption more than self 0; as a result, (IC$_1'$) is satisfied. Second, note that since $X(0) = c'$ by definition, self 2 can always choose (c_2', c_3') under *any* X-schedule; therefore this restriction does not expand the set of self 2's possible deviations.

It remains to determine the $\hat{X}$-schedule—self 1's punishment for incorrectly choosing c_1' in state θ. For s_1 large enough, there is no need to punish self 1 for choosing c_1' and saving s_1. Define

$$s_1^* \equiv \sup\left\{s_1 : U^1\left(s_1 + c' - s_2;\theta\right) > U^1\left(c;\theta\right) \text{ where } s_2 \in \arg\max_{\tilde{s}_2 \geq 0} U^2\left(s_1 + c' - \tilde{s}_2;\theta\right)\right\}.$$

In words, s_1^* is the highest level of date 1 savings s_1 such that in state θ, self 1 prefers the consumption plan $s_1 + c' - s_2$ to the consumption plan c, anticipating self 2's optimal choice of s_2 if he were to choose the former. In this sense s_1^* is the highest savings level at which there is a commitment problem. As we show in the appendix, for all $s_1 \geq s_1^*$, we can therefore restrict attention to the $\hat{X}$-schedule that results from giving self 2 (c_2', c_3') and letting him save whatever he likes.

To summarize the contract, self 1 chooses c_1 or c_1' and savings s_1. If self 1 chooses c_1, then self 2 gets (c_2, c_3) and chooses savings s_2. If self 1 chooses c_1', then self 2 can choose either (c_2', c_3') (if self 1 saved s_1^* or more) or some point on the schedule $\{\hat{X}(s_1)\}_{s_1 \in [0, s_1^*)}$ (if self 1 saved less than s_1^*), and in either case choose savings s_2. As a result, the interesting part of the contracting problem is determining $\hat{X}(s_1)$ for $s_1 \in [0, s_1^*)$.

Given these simplifying observations, we refer to a *contract* simply by $\hat{X}$ in what follows. When a contract exists that satisfies all the incentive compatibility constraints (with X as defined above), we say that *commitment is possible* and call such a contract a *commitment contract*.

5.2 General properties of commitment contracts

As we will show, the agent's ability to save significantly reduces his ability to overcome his commitment problem. At the most basic level, self 2's ability to save makes it more difficult to punish self 1 for deviating: just as in the case of verifiable consumption-saving decisions, the punishment $\hat{X}(s_1)$ must be strictly front-loaded relative to c'; as before, this follows from Lemma 2 and the fact that there is a commitment problem. When self 2 can save, however, he may be tempted to make the punishment less front-loaded by saving; this limits the severity of punishments that can be imposed on self 1 for deviating.

Self 2's ability to save also implies that the punishment $\hat{X}(s_1)$ must offer strictly less total consumption at dates 2 and 3 than c'. Consequently, the feature of the example that self 2 forfeits some of his endowment when he punishes self 1 is a general one.[20] To see this, note that since $\hat{X}(s_1)$ is strictly front-loaded relative to c', self 2 would never choose c' in state θ' if $\hat{X}(s_1)$ offered at least as much total consumption as c'; this follows from the fact that at c' in state θ', self 2 strictly prefers to borrow a small amount.[21]

Finally, self 1's ability to save implies that self 2 may arrive at date 2 with inherited savings s_1; this makes him less willing to choose a front-loaded punishment. Therefore, as s_1 increases, the punishment $\hat{X}(s_1)$ must become less front-loaded. If instead $\hat{X}(s_1)$ became more front-loaded as s_1 increases, then self 2 with high inherited savings would never choose the point on the schedule intended for him; after all, even self 2 with low savings did not choose that point. Moreover, to avoid one point on the schedule completely dominating another, total consumption $\hat{X}_2(s_1)+\hat{X}_3(s_1)$ must move in the opposite direction from $\hat{X}_2(s_1)$. Hence the punishment $\hat{X}$ becomes less punitive as s_1 increases. These properties are formally established by:

[20] By way of contrast, note that the punishment $(\hat{c}_2, \hat{c}_3)$ in the verifiable savings case may potentially increase total consumption relative to (c_2', c_3'), with the punishment stemming solely from forcing the agent to consume very little at date 3. The possibility of this form of punishment is eliminated by the agent's ability to save in an unverifiable way.

[21] Formally, suppose to the contrary that $\hat{X}_2(s_1)+\hat{X}_3(s_1) \geq c_2'+c_3'$ for some $s_1 \in [0, s_1^*)$ and note that $\hat{X}_2(s_1) > c_2'$. At c' in state θ', self 2 would like to borrow at least a small amount, i.e, $U_2^2(c';\theta') > U_3^2(c';\theta')$ (this follows from $U_2^1(c';\theta') = U_3^1(c';\theta')$ and $\beta < 1$). Hence $U^2(\hat{X}(s_1) - s_2;\theta') > U^2(c';\theta')$ for some $s_2 > 0$, contradicting (IC_2').

Proposition 2 *In any commitment contract, (i)* $\hat{X}_2(s_1) > c'_2$ *and* $\hat{X}_2(s_1) + \hat{X}_3(s_1) < c'_2 + c'_3$ *for all* $s_1 \in [0, s_1^*)$ *and (ii)* $\hat{X}_2(s_1)$ *is weakly decreasing and* $\hat{X}_2(s_1) + \hat{X}_3(s_1)$ *is weakly increasing in* s_1.

5.3 Strong preference reversal

Just as in the case of verifiable consumption-saving decisions, a key condition for the existence of an incentive compatible contract is a form of preference reversal for self 2 across states θ and θ': in order to ensure that self 1 is punished for choosing c'_1 in state θ but not punished for doing so in state θ', there must exist a front-loaded punishment that self 2 will choose in state θ but not in state θ'.

Here self 1's ability to save leads to a more stringent condition than in the case of verifiable consumption-saving decisions: first recall that as self 2 inherits a higher level of savings s_1, he grows less willing to choose a relatively front-loaded punishment. Therefore, self 2 with inherited savings s_1^* is the *least* willing to choose any front-loaded punishment and self 2 with no inherited savings is the *most* willing to choose that punishment. Therefore, to effectively deter self 1 from deviating, there must exist a front-loaded punishment that self 2 with savings s_1^* is willing to choose in state θ but self 2 with no savings will not choose in state θ'. Such a punishment exists only if the following condition—*strong preference reversal*—is satisfied.

Condition SPR (Strong preference reversal) *There exists* $\tilde{x}$ *such that* $\tilde{x}_2 > c'_2$,

$$U^2(s_1^* + \tilde{x}; \theta) \geq U^2\left(s_1^* + c'; \theta\right), \tag{1}$$

$$U^2\left(c'; \theta'\right) \geq U^2\left(\tilde{x} - s_2; \theta'\right) \text{ for all } s_2 \geq 0. \tag{2}$$

Condition SPR extends Condition PR—the preference reversal condition from before—to the case where savings levels differ. As the following result shows, Condition SPR is necessary for commitment:

Proposition 3 *Commitment is possible only if Condition SPR is satisfied.*

Although Proposition 3 is simple to state, its proof is less trivial than one might imagine, since while it is clear that commitment is possible only if for any $s_1 < s_1^*$ there exists some punishment $\tilde{x}$ such that $\tilde{x}_2 > c_2'$ and the analogues of (1) and (2) hold, this is not enough to imply the existence of $\tilde{x}$ with $\tilde{x}_2 > c_2'$ satisfying (1) and (2) at s_1^*.

Just as for Condition PR in the case where savings levels are the same, Condition SPR can be checked by comparing the slopes of two indifference curves through c', although here the comparison yields only a sufficient condition.

Lemma 3 *Condition SPR is satisfied if* $U_2^2\left(s_1^* + c'; \theta\right) / U_3^2\left(s_1^* + c'; \theta\right) \geq U_2^2\left(c'; \theta'\right) / U_3^2\left(c'; \theta'\right)$.

Why is this condition not necessary as well as sufficient for Condition SPR? As the discussion of SCB above suggests, indifference curves may cross more than once when savings levels differ. Therefore, even if the indifference curves cross the wrong way at c', they may cross again at some $\tilde{c}$ such that $\tilde{c}_2 > c_2'$. Allowing for this possibility, Condition SPR can be checked as follows: define $\hat{x}^0$ as the solution to

$$\min_{\tilde{x}} \left\{\tilde{x}_2 + \tilde{x}_3\right\} \text{ s.t. } \tilde{x}_1 = c_1',\ \tilde{x}_2 \geq c_2', \text{ and } U^2\left(s_1^* + \tilde{x}; \theta\right) \geq U^2\left(s_1^* + c'; \theta\right).$$

In words, $\hat{x}^0$ is the most severe (in the sense of minimizing total consumption) front-loaded punishment that self 2 with inherited savings s_1^* would ever choose in state θ. As the following result shows, it is then both necessary and sufficient for Condition SPR that self 2 with no inherited savings does not choose $\hat{x}^0$ in state θ', even if he can save, and self 2 with inherited savings s_1^* strictly prefers to borrow at c' in state θ.

Lemma 4 *Condition SPR is satisfied if and only if*

$$U^2\left(c'; \theta'\right) \geq U^2\left(\hat{x}^0 - s_2; \theta'\right) \text{ for all } s_2 \geq 0, \text{ and } \frac{U_2^2\left(s_1^* + c'; \theta\right)}{U_3^2\left(s_1^* + c'; \theta\right)} > 1. \qquad \text{(SPR}'\text{)}$$

5.4 Main result: necessary and sufficient conditions for commitment

Self 1's ability to save necessitates a stronger preference reversal condition than in the case of verifiable consumption-saving decisions (i.e., Condition PR is replaced by SPR). Self 2's ability to save also imposes a significant restriction on the contracting problem: the punishment must be front-loaded, but self 2 may be tempted to make the punishment less front-loaded by saving; this limits the severity of punishments that can be imposed on self 1 for deviating. We next construct a particular contract, denoted $\hat{X}^*$, with the crucial property that if any commitment contract exists, then $\hat{X}^*$ is the *least front-loaded* among them. This means that $\hat{X}^*$ is the commitment contract under which self 2 is least tempted to save.

Suppose that Condition SPR is satisfied. Starting at $s_1 = s_1^*$ (the highest savings level at which there is a commitment problem), we choose the least front-loaded punishment $\hat{X}^*(s_1^*)$ that is severe enough to deter self 2 with no inherited savings from choosing it in state θ', but also mild enough to induce self 2 with inherited savings s_1^* to choose it in state θ. Formally, $\hat{X}_2^*(s_1^*)$ is defined as the infimum value of $\tilde{x}_2$ such that $\tilde{x}_2 > c_2'$, (1), and (2) are satisfied, while $\hat{X}_3^*(s_1^*)$ is defined by setting (1) to equality. Note that, thus defined, $\hat{X}^*(s_1^*)$ also satisfies (2) with equality.[22] The crucial property of $\hat{X}^*(s_1^*)$ is that it provides a lower bound on the date 2 consumption offered by *any* commitment contract.

Proposition 4 *If $\hat{X}$ is part of a commitment contract, then $\hat{X}_2(s_1) \geq \hat{X}_2^*(s_1^*)$ for all $s_1 \in [0, s_1^*)$. Moreover, if Condition SPR is satisfied, then $\hat{X}^*(s_1^*)$ exists and is unique.*

In common with Proposition 3, the proof of Proposition 4 is more complicated than its simple statement suggests, and for similiar reasons.

As the following result shows, whenever Condition SPR can be verified using the simple derivative condition of Lemma 3, then $\hat{X}^*(s_1^*) = c'$.

Lemma 5 *If $U_2^2(s_1^* + c'; \theta) / U_3^2(s_1^* + c'; \theta) \geq U_2^2(c'; \theta') / U_3^2(c'; \theta')$ then $\hat{X}^*(s_1^*) = c'$.*

As we show in the proof of Theorem 1, if $\hat{X}_2^*(s_1^*) > c_2'$ then the punishment $\hat{X}^*(s_1^*)$ is sufficiently severe to deter self 1 from deviating and saving s_1 for s_1 sufficiently close to s_1^*. Consequently, the least front-loaded contract calls for keeping $\hat{X}^*(s_1)$ fixed at $\hat{X}^*(s_1^*)$ until self 1 is just indifferent towards deviating. Formally, we set $\hat{X}^*(s_1) = \hat{X}^*(s_1^*)$ for all $s_1 \in [s_1^{**}, s_1^*]$, where

$$s_1^{**} = \sup\left\{s_1 \in [0, s_1^*] : U^1(s_1 + \hat{X}^*(s_1^*); \theta) > U^1(c; \theta)\right\}$$

[22] If $\hat{X}_2^*(s_1^*) = c_2'$ then $\hat{X}_3^*(s_1^*) = c_3'$, and since self 2 does not want to save given c' in state θ', (2) holds with equality. For the case $\hat{X}_2^*(s_1^*) > c_2'$, note that if (2) were instead satisfied strictly then $\tilde{x} = \left(\hat{X}_2^*(s_1^*) - \varepsilon_2, \hat{X}_3^*(s_1^*) + \varepsilon_3\right)$ would also satisfy (1) and (2) for small $\varepsilon_2, \varepsilon_3 > 0$ chosen so that (1) holds with equality; this, however, contradicts the definition of $\hat{X}_2^*(s_1^*)$.

and $s_1^{**} = s_1^*$ if the above set is empty. In particular, note that $s_1^{**} = s_1^*$ in the special case of $\hat{X}^*(s_1^*) = c'$.

For $s_1 < s_1^{**}$, the punishment must grow more severe as s_1 decreases in order to deter self 1 from deviating. From Proposition 2 we know that as s_1 decreases and the punishment grows more severe, it must also grow more front-loaded. As we will show, the least front-loaded contract $\hat{X}^*$ calls for punishing self 1 as mildly as possible for deviating. This boils down to keeping him indifferent towards deviating for $s_1 < s_1^{**}$, and is accomplished by the following two differential equations, which together define the contract for all $s_1 < s_1^{**}$.

$$d\hat{X}_2^* = \min\left\{\frac{U_1^1(s_1+\hat{X}^*;\theta) - U_2^1(s_1+\hat{X}^*;\theta)}{U_3^1(s_1+\hat{X}^*;\theta)\left(\frac{U_2^1(s_1+\hat{X}^*;\theta)}{U_3^1(s_1+\hat{X}^*;\theta)} - \frac{U_2^2(s_1+\hat{X}^*;\theta)}{U_3^2(s_1+\hat{X}^*;\theta)}\right)}, 0\right\} ds_1 \tag{3}$$

$$d\hat{X}_3^* = -\frac{U_2^2(s_1+\hat{X}^*;\theta)}{U_3^2(s_1+\hat{X}^*;\theta)} d\hat{X}_2^*. \tag{4}$$

As we will show, $\hat{X}^*$ satisfies all the incentive compatibility constraints except one: self 2 with inherited savings s_1 who correctly chooses $\hat{X}^*(s_1)$ in state θ should not want to save (i.e., $\hat{X}^*$ satisfies the first half of IC_2 for all $s_1, \tilde{s}_1 \in [0, s_1^*)$, $\tilde{s}_1 = s_1$, and $s_2 > 0$). Formally,[23]

$$U_2^2(s_1 + \hat{X}^*(s_1);\theta) \geq U_3^2(s_1 + \hat{X}^*(s_1);\theta) \text{ for all } s_1 \in [0, s_1^*). \tag{NS}$$

Our main result is that a necessary and sufficient condition for commitment is that Conditions SPR and NS are both satisfied.

Theorem 1 *Commitment is possible if and only if Conditions SPR and NS are both satisfied. When both conditions are satisfied, $\hat{X}^*$ is a commitment contract.*

5.5 Sketch proof of Theorem 1

In the previous subsection, we motivated the construction of $\hat{X}^*$ as the least front-loaded commitment contract possible. This feature of the contract is desirable because it ensures that $\hat{X}^*$ is the commitment contract under which self 2 is least tempted to undo a front-loaded punishment by

[23] Note that (NS) is satisfied by construction for $s_1 \geq s_1^*$.

saving. In this subsection, we explain why the contract $\hat{X}^*$ is such that self 2 punishes if and only if he should; and we give more details on why $\hat{X}^*$ is indeed the least front-loaded contract possible.

Self 1's ability to save greatly complicates the contracting problem because he can now choose from a continuum of deviations: he can choose c_1' in state θ, and then save any amount s_1 he wishes. As his deviation savings s_1 increase, it becomes harder to induce self 2 to punish him, since the attractiveness of front-loaded punishment decreases; but at the same time, the severity of the punishment required also decreases. Consequently, we have a continuum of distinct contracting problems, indexed by date 1 savings s_1. Moreover, these problems cannot be tackled in isolation, since we need to make sure that when self 2 inherits savings s_1 he chooses the punishment $\hat{X}(s_1)$ intended for that savings level. Of course, the contract cannot explicitly depend on the level of private savings, and so self 2 must find it in his best interest to choose the punishment intended for each level of savings. That is, $\hat{X}$ must satisfy the self-selection condition,

$$U^2(s_1 + \hat{X}(s_1); \theta) \geq U^2(s_1 + \hat{X}(\tilde{s}_1); \theta) \text{ for all } s_1, \tilde{s}_1 \geq 0. \tag{5}$$

The key step in overcoming these analytical difficulties is to observe that the self-selection constraint (5) implies a strong restriction on how the contract $\hat{X}$ can vary over different savings s_1. Moreover, the restriction is tractable because it requires one only to keep track of self 2's utility, a one-dimensional object. The argument is analytically similar to one used in auction theory (see Lemma 2 in Myerson 1981).

Lemma 6 *If $\hat{X}$ is part of a commitment contract then $U^2(s_1 + \hat{X}(s_1); \theta)$ is strictly increasing and continuous. Moreover, if $\hat{X}$ is continuous at s_1 then $U^2(s_1 + \hat{X}(s_1); \theta)$ is differentiable at s_1 with derivative*

$$\frac{d}{ds_1} U^2(s_1 + \hat{X}(s_1); \theta) = U_2^2(s_1 + \hat{X}(s_1); \theta). \tag{6}$$

Conversely, suppose that $\hat{X}$ is continuous and differentiable over some interval $[s_1^a, s_1^b]$; $\hat{X}_2$ is weakly decreasing; and (6) holds. Then $\hat{X}$ satisfies the self-selection property (5) for s_1, $\tilde{s}_1 \in [s_1^a, s_1^b]$.

Lemma 6 helps explain the *sufficiency* half of Theorem 1. First, it is readily verified that the differential equations (3) and (4) are equivalent to (6) and

$$\frac{d}{ds_1}U^1(s_1+\hat{X}\left(s_1\right);\theta)=0. \tag{7}$$

So $\hat{X}^*$ provides sufficient punishment to deter self 1 from all his continuum of possible deviations, and moreover, by Lemma 6 it satisfies the self-selection property.

Second, self 2 with inherited savings s_1 picks the punishment $\hat{X}^*\left(s_1\right)$ over c' in state θ, i.e.,

$$U^2(s_1+\hat{X}^*\left(s_1\right);\theta)\geq\max_{s_2\geq 0}U^2\left(s_1+c'-s_2;\theta\right). \tag{8}$$

By construction, this is the case for $s_1=s_1^*$. Perhaps surprisingly, self-selection implies this is also the case for all lower savings levels: this follows from the self-selection property (6), together with the fact that $\hat{X}^*$ is front-loaded relative to c'.

Third, self 2 *does not* pick the punishment $\hat{X}^*\left(s_1\right)$ in state θ', i.e., for all s_1 and $\tilde{s}_1$,

$$\max_{s_2\geq 0}U^2(s_1+c'-s_2;\theta')\geq\max_{s_2\geq 0}U^2(s_1+\hat{X}^*\left(\tilde{s}_1\right)-s_2;\theta'). \tag{9}$$

By construction, this is the case for no inherited savings ($s_1=0$) and the specific punishment $\hat{X}^*\left(s_1^*\right)$. Self-selection again implies — together with SCB — that (9) holds for all points on the punishment schedule, i.e., all $\tilde{s}_1$. Moreover, because the punishment $\hat{X}^*$ is front-loaded relative to c', and entails less total consumption, this preference for c' over $\hat{X}^*\left(s_1\right)$ in state θ' is only strengthened as self 2 inherits more savings, and so (9) holds for all s_1 and $\tilde{s}_1$, as required.

Next, we turn to the *necessity* half of Theorem 1, and provide more details for our claim that $\hat{X}^*$ is the least front-loaded contract possible. Consider a specific level of date 1 savings s_1. If the contract is to provide sufficient punishment for cheating by self 1, then $U^1(s_1+\hat{X}\left(s_1\right);\theta)\leq U^1\left(s_1+c';\theta\right)$. Moreover, self 2 must have no desire to save given the allocation $\hat{X}\left(s_1\right)$, and so $U_2^2(s_1+\hat{X}\left(s_1\right);\theta)\geq U_3^2(s_1+\hat{X}\left(s_1\right);\theta)$. Together, these two conditions imply that given any value of $\hat{X}_3\left(s_1\right)$, $\hat{X}_2\left(s_1\right)$ must lie below some cutoff.

Suppose for a moment that $\hat{X}$ has already been defined for all $\tilde{s}_1$ above s_1. From Lemma 6, the self-selection property (5) tells us what level of utility $U^2(s_1+\hat{X}\left(s_1\right);\theta)$ the contract must deliver. From the two conditions discussed in the previous paragraph, we know that $\hat{X}_2\left(s_1\right)$ must be sufficiently low given $\hat{X}_3(s_1)$. It is possible that no choice of $\hat{X}(s_1)$ can satisfy these conditions, while giving self 2 the target level of utility. In this case, commitment is not possible. The key observation is that this is *less likely* to happen if the target level of utility is low.[24]

What contract features help reduce the target utility $U^2(s_1 + \hat{X}(s_1); \theta)$? From Lemma 6, it helps to keep $\hat{X}_2(\tilde{s}_1)$ as low as possible for $\tilde{s}_1$. By the same argument, selecting $\hat{X}(s_1)$ with the lowest date 2 allocation makes it more likely that the incentive compatibility conditions can be satisfied for even lower levels of savings. This is achieved by setting the "sufficient punishment" constraint $U^1(s_1 + \hat{X}(s_1); \theta) \leq U^1(s_1 + c'; \theta)$ to hold with equality: if instead it holds with strict inequality, the date 3 allocation can be raised and the date 2 allocation even further reduced.

This argument implies that the most front-loaded contract possible is defined by keeping $U^1(s_1 + \hat{X}(s_1); \theta)$ constant (and equal to $U^1(s_1 + c'; \theta)$), subject to the self-selection constraint (5). As we have already noted, these two conditions are equivalent to the differential equations (3) and (4) that define $\hat{X}^*$.

6 Interpreting the strong preference reversal (SPR) condition

Condition SPR gives a necessary condition for the individual's commitment problem to be solvable. The condition makes clear how self 2's prefences must differ across states θ and θ', and casts light on the key economic conditions required for commitment. However, it is not easy to relate the condition to specific cases in which self 0 can commit to exactly what he wants. To this end, in this section we restate condition SPR in terms of self 0's most preferred consumption plans, c and c'. Throughout the section we assume that the shock hitting the individual has passed by date 3, i.e., $u_3(\cdot; \theta) = u_3(\cdot; \theta')$.

In the wedding example we gave in the introduction, self 0's most preferred consumption plans in states θ (wedding at date 2) and θ' (wedding at date 1) are $c = \left(1, \frac{3}{2}, 1\right)$ and $c' = \left(\frac{3}{2}, 1, 1\right)$ respectively. One way to think about the difference between c' and c is that, in c', at date 1 the individual "borrows" an extra 1/2 relative to c, but then "repays" the 1/2 at date 2. In other words, the increase in date 1 consumption afforded by c' comes solely at the expense of date 2

[24] This argument depends on the fact that the indifference curves of self 2 in (x_2, x_3)-space are steeper than those of self 1. We aim to keep the discussion in the text as intuitive as possible; all formal details are in the proofs in the appendix.

consumption (relative to c), without negatively impacting date 3 consumption.

Whenever shocks are such that self 0's desired consumptions c and c' satisfy this property, condition SPR is satisfied:

Lemma 7 *Suppose that $u_3(\cdot;\theta) \equiv u_3(\cdot;\theta')$ and there is a commitment problem. Condition SPR is satisfied if $c_3' > c_3$; or if $c_3' = c_3$ and the shock is either additive or multiplicative.*

What sort of shocks have the property that $c_3' \geq c_3$? The wedding example of the introduction is an example of a *timing shock*, in which the agent's utility in dates 1 and 2 in state θ is the same as utility in dates 2 and 1 in state θ'. In other words, self 0 knows he will be hit be shock in the future, but does not know when the shock will arrive. In addition to the wedding example, good examples include the replacement of an old car when it breaks down; and the need to make a downpayment on a house purchase. Formally, the timing shock is defined by $u_1'(\cdot;\theta) = u_2'(\cdot;\theta') < u_2'(\cdot;\theta) = u_1'(\cdot;\theta')$. For timing shocks, if self 0's most preferred consumption in state θ is $c = (c_1, c_2, c_3)$, then his most preferred consumption in state θ' is simply $c' = (c_2, c_1, c_3)$, and so date 3 consumption is clearly the same in c and c'.

A second class of shocks for which $c_3' \geq c_3$ are *one-period-ahead shocks* in which the individual learns about a change in his utility one period in advance, i.e., the shock only affects date 2 utility. For example, an individual might learn today that, at some future date, he will receive vacation, boosting his marginal utility of consumption at that date; or he will receive a bonus; or he will lose his job;[25] or, in the procrastination interpretation of our setting, that his favorite film is showing. Formally, a one-period ahead shock is defined by $u_1(\cdot;\theta) = u_1(\cdot;\theta')$, $u_2'(\cdot;\theta) > u_2'(\cdot;\theta')$. One-period ahead shocks lower the marginal utility of consumption at date 2 in state θ', relative to state θ. This means that, in state θ', self 0 wishes to reallocate some his consumption away from date 2 and to date 1—the source of the commitment problem—and also to date 3, so that $c_3' > c_3$.[26]

The opposite of a one-period-ahead shock is a *contemporaneous shock*, in which an individual learns about a change his utility only contemporaneously, i.e., the shock only affects date 1 util-

[25] That is, at date 1 he learns he will lose his job at date 2, but anticipates finding a new job by date 3.

[26] Formally, since c and c' are self 0's most preferred consumptions, $u_1(c_1;\theta) = u_3(c_3;\theta)$ and $u_1(c_1';\theta') = u_3(c_3';\theta')$, implying

$$u_3(c_3';\theta') - u_3(c_3;\theta) = u_1(c_1';\theta') - u_1(c_1;\theta).$$

Since $u_1(\cdot;\theta) = u_1(\cdot;\theta')$ and $u_3(\cdot;\theta) = u_3(\cdot;\theta')$ for one-period ahead shocks, it follows from $c_1' \geq c_1$ that either $c_3' > c_3$, as claimed; or that $c_t' = c_t$ for $t = 1, 2, 3$, which contradicts either $u_2(c_2;\theta) = u_3(c_3;\theta)$ or $u_2(c_2';\theta') = u_3(c_3';\theta')$.

ity. Formally, a contemporaneous shock is defined by $u_1'(\cdot;\theta) < u_1'(\cdot;\theta')$, $u_2(\cdot;\theta) = u_2(\cdot;\theta')$. In contrast to the cases of one-period ahead and timing shocks, self 0 is never able to overcome commitment problems associated with contemporaneous shocks. This follows from the facts that (A) the key savings level s_1^* is strictly positive whenever a commitment problem exists; and (B) for contemporaneous shocks, self 2's preferences are identical across the two states. To see that SPR fails, consider any $\tilde{x}$ that is strictly front-loaded relative to c', i.e., $\tilde{x}_2 > c_2'$, and such that self 2 prefers c' to $\tilde{x}$ in state θ' when he has no savings, i.e., $u_2(\tilde{x}_2;\theta') - u_2(c_2';\theta') + \beta(u_3(\tilde{x}_3;\theta') - u_3(c_3';\theta')) \leq 0$. Facts (A) and (B) above, together with strict concavity of u_2, imply that self 2 must then strictly prefer c' to $\tilde{x}$ in state θ when he has savings s_1^*, i.e., $u_2(s_1^* + \tilde{x}_2;\theta') - u_2(s_1^* + c_2';\theta') + \beta(u_3(\tilde{x}_3;\theta') - u_3(c_3';\theta')) < 0$. In words, there is no punishment $\tilde{x}$ that self 2 would pick in state θ but not in state θ'.

Contemporaneous shocks raise the marginal utility of consumption at date 1 in state θ', relative to state θ. This means that, in state θ', self 0 wishes to reallocate some his consumption away from both dates 2 and 3 towards date 1, so that $c_3' < c_3$. More generally, and given the discussion above, one might conjecture that condition SPR fails whenever $c_3' < c_3$, i.e., whenever the extra consumption afforded at date 1 under c' is not fully repaid by date 2. For the case of additive shocks, we are able to confirm that this is indeed the case:[27]

Lemma 8 *Suppose shocks are additive. If $c_3' < c_3$ then condition SPR is not satisfied, and commitment is impossible.*

Finally, in this section we have focused exclusively on condition SPR. When this condition is satisfied, it is potentially possible for self 0 to overcome his commitment problem. To confirm that this is indeed possible, one must also check condition NS. Unfortunately, we have been unable to achieve any useful analytic characterization of when NS is satisfied. However, for the case of logarithmic period utility we have checked a very wide class of possible parameterizations of both timing and one-period ahead shocks,[28] and found that NS is always satisfied.

[27]When shocks are not additive, we have been unable to either prove that $c_3' < c_3$ implies that condition SPR fails, or to provide a counterexample.

[28]Specifically, for the additive shock specification we have have checked all combinations the hyperbolic discount rate β, and all specifications of a timing or one-period shocks, in all combinations, for a range of endowment values W. We have performed a similar exercise for the multiplicative shock specification, but with the restrictions that the timing shocks are of the form $(\theta,\theta') = ((\zeta,\psi,\zeta),(\psi,\zeta,\zeta))$, and that one-period ahead shocks are of the form $(\theta,\theta') = ((\zeta,\zeta,\zeta),(\zeta,\psi,\zeta))$ or $((\zeta,\psi,\zeta),(\zeta,\zeta,\zeta))$.

7 Concluding remarks

Our analysis characterizes situations in which an individual can contract to completely overcome his commitment problem, even in the face of uncertainty about his future consumption needs. In these situations, hyperbolic discounting ceases to affect the individual's behavior. At the same time, there also exist important cases in which such commitment is not attainable, even though we have placed absolutely no restriction on the class of allowable commitment devices. Overall our results provide partial support for both proponents and antagonists of the importance of hyperbolic discounting. Moreover, although we focus on one particular decision-making bias in this paper—the propensity to overweight current consumption—we believe that our general arguments are widely applicable to many other behavioral biases. We leave an exploration of this last point for future research.

Contractual implementation

When commitment is possible, commitment contracts take an easy-to-interpret form. At date 1, the individual is given flexibility to meet his consumption needs via an option to access savings ("early withdrawal") and/or borrow at a penalty-free rate. After exercising this option, any further early withdrawals and/or borrowing carries a penalty. This pair of features resembles the withdrawal rights observed in many real-world savings products, as well as borrowing fee schedules in many credit contracts.

The exact form of a commitment contract depends on the parameters of the problem, such as the degree of impatience (β) and the shocks (θ and θ'). A notable feature of our setting relative to much of the contract literature is that the agent is happy to truthfully report these details at date 0. In essence, self 0 is the principal. In this sense, a nice feature of our contracting problem is that it is considerably less informationally demanding than those between a distinct principal and agent(s).

In common with much of the contract literature, we have conducted our analysis under an assumption of *exclusivity*: after writing the original contract at date 0, at dates 1 and 2 the agent cannot write additional contracts with other counterparties. Three points are worth making here.

First, our assumption that the individual can privately save represents a partial relaxation of the exclusivity assumption.

Second, in situations in which borrowing markets are thin or nonexistent, our contractual solution is robust to relaxing the exclusivity assumption. To see this, consider a setting in which self 0's most-preferred consumption plans c and c' do not require any borrowing (e.g., at date 0 the agent already has his entire endowment W).[29] In this case, the original contract can be viewed as a savings contract specifying withdrawal rights. Exclusivity simply requires that self 1 cannot himself enter more savings contracts. But this restriction can be easily included in the initial self 0 contract by including an agreement that self 2 can borrow interest-free against any funds deposited in savings vehicles by self 1. Note that this gives self 2 the incentive to voluntarily report the existence of these funds. Moreover, note that the inclusion of this clause does not violate the assumption of missing borrowing markets, since it relates to an out-of-equilibrium event and so the original date-0 counterparty would be prepared to take a loss on this part of the contract.

Third, and related, when borrowing markets exist, enforcement of exclusivity requires that at date 0 the agent can pledge not to borrow more from other creditors at dates 1 and 2—but *does not require* any restrictions on his future savings options. Since many real-world mechanisms exist to restrict future borrowing (e.g., credit registries, collateral registries, prepayment penalties, debt covenants), this degree of exclusivity may indeed be attainable in reality.

Generality

Throughout, we have restricted attention to the case of three consumption dates and just two shocks. These simplifying assumptions help keep the economic intuition of our analysis clear. However, they are not essential to our main insight: the agent's different selves share the same information and this increases the scope for writing contracts to elicit that information. Moreover, and as noted previously, one can interpret $u_3(c_3; \phi)$ as the sum of expected future utility in state ϕ given wealth c_3. In this sense, our analysis is considerably more general than it might at first appear.

We have also focused throughout on the question of whether it is possible to achieve full commitment, i.e., self 0's most preferred consumption plans c and c'. This allows us to present our

[29] If this condition is not satisfied, and borrowing is impossible, then it is clearly impossible for the agent to attain consumption c and c'.

results in the starkest way—in some situations, hyperbolic discounting has no effect on outcomes once commitment contracts are allowed for. When commitment is not possible, however, our analysis can also be used as an input into a constrained-optimization problem, i.e., what consumption plans c and c' maximize self 0's utility subject to the constraint that self 0 can commit to c and c'?

Naïveté

As we note in the introduction, our analysis relies on some sophistication on the part of the agent regarding his own biases—specifically, selves 0 and 1 must be self-aware enough to know that they will discount hyperbolically at future dates. Although for expositional transparency we have focused on the case of complete sophistication, our analysis is also applicable to partially naïve agents who underestimate—but do not completely dismiss—the extent of their own future impatience.

Clearly, at the opposite extreme of complete naïveté—in which each self believes his future selves will discount exponentially—commitment is completely impossible. Moreover, and more interestingly, offering a commitment contract designed for a sophisticated agent to a complete naïf can make the naïf worse off, even compared to the alternative of a contract that makes no attempt to tackle the commitment problem and simply allows self 1 to choose between c and c'.

This last point is easily demonstrated in the case of verifiable savings. First, note that in any commitment contract, the "punishment" $(\hat{c}_2, \hat{c}_3)$ must satisfy

$$u_2(\hat{c}_2;\theta) + u_3(\hat{c}_3;\theta) < u_2\left(c_2';\theta\right) + u_3\left(c_3';\theta\right), \tag{10}$$

since otherwise the punishment would not be enough to deter self 1 from overconsuming in state θ.[30] Consequently, at date 1 a naïf will pick c_1' in state θ, believing that self 2 will pick c_2'. However, after self 1 picks c_1' self 2 in fact picks the punishment $(\hat{c}_2, \hat{c}_3)$. Self 0's equilibrium utility in state θ is hence $U^0\left(\left(c_1', \hat{c}_2, \hat{c}_3\right);\theta\right)$. But by (10), this is strictly less than the utility self 0 would get from a contract allowing self 1 to simply choose freely between c and c', namely $U^0(c';\theta)$. We leave for future research the complicated question of how to balance the welfare gain a commitment contract delivers to sophisticated agents with the welfare loss stemming from its misuse at the hands of naïve agents.

[30] Formally, if $u_2(\hat{c}_2;\theta) + u_3(\hat{c}_3;\theta) \geq u_2(c_2';\theta) + u_3(c_3';\theta)$ then $U^1\left(\left(c_1', \hat{c}_2, \hat{c}_3\right);\theta\right) \geq U^1(c';\theta) > U^1(c;\theta)$.

References

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

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

[3] Bernheim, B.D., D. Ray, and S. Yeltekin (1999), "Self-Control, Saving, and the Low Asset Trap," working paper.

[4] Cole, H.L., and N.R. Kocherlakota (2001), "Efficient Allocations with Hidden Income and Hidden Storage," *Review of Economic Studies* 68, 523-542.

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

[6] Doepke, M., and R.M. Townsend (2006), "Dynamic mechanism design with hidden income and hidden actions," *Journal of Economic Theory* 126, 235-285.

[7] Frederick, S., G. Loewenstein, and T. O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature* 40, 351-401.

[8] He, Zhiguo (2008), "Dynamic Compensation Contracts with Private Savings," working paper.

[9] Jackson, Thomas H. (1986), *The Logic and Limits of Bankruptcy Law.* Cambridge, MA: Harvard University Press.

[10] Kocherlakota, Naryana R. (2004), "Figuring out the impact of hidden savings on optimal unemployment insurance," *Review of Economic Dynamics* 7, 541-554.

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

[12] Laibson, D. (1996), "Hyperbolic Discount Functions, Undersaving and Savings Policy," working paper.

[13] Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics* 112, 443-477.

[14] Laibson, D., A. Repetto, and J. Tobacman (1998), "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity* 57, 91-196.

[15] Luttmer, E.G.J, and T. Mariotti (2003), "Subjective Discounting in an Exchange Economy," *Journal of Political Economy* 111, 959-989.

[16] Maskin, E. and T. Sjöström (2002), "Implementation Theory," in K. Arrow, A. Sen, and K. Suzumara, eds., *Handbook of Social Choice and Welfare*, vol. I. New York: Elsevier Science.

[17] Milgrom, P. and C. Shannon (1994), "Monotone Comparative Statics," *Econometrica* 62, 157-180.

[18] Myerson, R.B. (1981), "Optimal Auction Design," *Mathematics of Operations Research* 6, 58-73.

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

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

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

[22] Palfrey, T. (2002), "Implementation Theory," in R. Aumann and S. Hart, eds., *Handbook of Game Theory with Economic Applications*, vol. III. New York: Elsevier Science.

[23] Serrano, R. (2004), "The Theory of Implementation of Social Choice Rules," *SIAM Review* 46, 377-414.

[24] Shui H. and L. Ausubel, (2004), "Consumer Time Inconsistency: Evidence from an Experiment in the Credit Card Market," working paper.

[25] Skiba, P.M., and J. Tobacman (2008), "Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default," working paper.

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

Appendix

Results omitted from main text

Lemma A-1 *Fix a state ϕ. Let x^a and x^b be allocations such that $x_2^a \leq x_2^b$ and $x_2^a + x_3^a \geq x_2^b + x_3^b$. Then if $\max_{s_2 \geq 0} U^2(s_1 + x^a - s_2; \phi) \geq (\leq) \max_{s_2 \geq 0} U^2(s_1 + x^b - s_2; \phi)$ for some s_1, the same is true for all higher (lower) levels of s_1.*

Proof. Define

$$f(s_1) = \max_{s_2 \geq 0} U^2(s_1 + x^a - s_2; \phi) - \max_{s_2 \geq 0} U^2\left(s_1 + x^b - s_2; \phi\right).$$

If $x^a = x^b$, then $f(s_1) = 0$ for all s_1; the result is immediate. If $x_2^a + x_3^a > x_2^b + x_3^b$ and $x_2^a = x_2^b$, then $f(s_1) > 0$ for all s_1; the result is immediate. If $x_2^a + x_3^a = x_2^b + x_3^b$ and $x_2^a < x_2^b$, then there exists $\hat{s}_1$ such that for all $s_1 \geq \hat{s}_1$, $f(s_1) = 0$ and for all $s_1 < \hat{s}_1$, $f(s_1) < 0$; the result is immediate. Finally, consider the case where $x_2^a < x_2^b$ and $x_2^a + x_3^a > x_2^b + x_3^b$.

By standard arguments, f is differentiable. To establish the result, we show that $f'(s_1) > 0$ whenever $f(s_1) = 0$. Fix an s_1 such that $f(s_1) = 0$. Writing explicitly,

$$f'(s_1) = U_2^2(s_1 + x^a - s_2^a; \phi) - U_2^2\left(s_1 + x^b - s_2^b; \phi\right),$$

where $s_2^a \in \arg\max_{s_2 \geq 0} U^2(s_1 + x^a - s_2; \phi)$ and $s_2^b \in \arg\max_{s_2 \geq 0} U^2\left(s_1 + x^b - s_2; \phi\right)$. For the case $x_2^a - s_2^a < x_2^b - s_2^b$ the result is immediate.

The remainder of the proof deals with the case $x_2^a - s_2^a \geq x_2^b - s_2^b$. Since $x_2^a < x_2^b$, we must have $s_2^b > 0$. Suppose that, contrary to the claimed result, $f'(s_1) \leq 0$. So

$$U_3^2(s_1 + x^a - s_2^a; \phi) \leq U_2^2(s_1 + x^a - s_2^a; \phi) \leq U_2^2\left(s_1 + x^b - s_2^b; \phi\right) = U_3^2\left(s_1 + x^b - s_2^b; \phi\right),$$

where the final equality follows from the fact that date 2 savings s_2^b are at an interior optimum. Hence $x_3^a + s_2^a \geq x_3^b + s_2^b$. So the allocation $s_1 + x^b - s_2^b$ gives more consumption at both dates 2 and 3 than the allocation $s_1 + x^a - s_2^a$, and (since $x_2^a + x_3^a > x_2^b + x_3^b$) the comparison is strict at at least one of the two dates. But then $f(s_1) < 0$, giving a contradiction and completing the proof.

Lemma A-2 *If a commitment contract exists, then* $U_2^2(s_1^* + c'; \theta) \geq U_3^2(s_1^* + c'; \theta)$.

Proof. Let $\hat{X}$ be part of a commitment contract, suppose to the contrary that $U_2^2(s_1^* + c'; \theta) < U_3^2(s_1^* + c'; \theta)$, and note that by continuity of preferences and the definition of s_1^*, there exists $s_1 < s_1^*$ such that $U_2^2(s_1 + c'; \theta) < U_3^2(s_1 + c'; \theta)$ and $U^1(s_1 + c' - s_2; \theta) > U^1(s_1 + c; \theta)$ where, just as in the definition of s_1^*, $s_2 \in \arg\max_{\tilde{s}_2 \geq 0} U^2(s_1 + c' - \tilde{s}_2; \theta)$. From (IC$_1$), $U^1(s_1 + c; \theta) \geq U^1(s_1 + \hat{X}(s_1); \theta)$ and therefore $U^1(s_1 + c' - s_2; \theta) > U^1(s_1 + \hat{X}(s_1); \theta)$; from (IC$_2$) and the fact that $X(0) = c'$, $U^2(s_1 + \hat{X}(s_1); \theta) \geq U^2(s_1 + c' - s_2; \theta)$; it then follows from Lemma 2 that $\hat{X}_2(s_1) > c'_2 - s_2$ and $\hat{X}_3(s_1) < s_2 + c'_3$. From the definition of s_2 and the fact that $U_2^2(s_1 + c'; \theta) < U_3^2(s_1 + c'; \theta)$, we have $U_2^2(s_1 + c' - s_2; \theta) = U_3^2(s_1 + c' - s_2; \theta)$. But then $U_2^2(s_1 + \hat{X}(s_1); \theta) < U_3^2(s_1 + \hat{X}(s_1); \theta)$, violating (IC$_2$) and contradicting our supposition that $\hat{X}$ is part of an incentive compatible contract.

Proofs of results stated in main text

Proof of Proposition 1. The main text establishes the necessity of Condition PR. The equivalence of the two stated conditions is immediate from single-crossing and our mild assumption that if, at any consumption plan, self 2's indifference curves have the same slope in both states, then they completely coincide.

Proof of Lemma 1. Since $c(\phi)$ is self 0's most-preferred consumption in state ϕ, $U^0(c(\phi); \phi) - U^0(c(\phi'); \phi) \geq 0$. Since there is a commitment problem in state ϕ, $U^1(c(\phi); \phi) - U^1(c(\phi'); \phi) < 0$. Substracting the second inequality from the first implies $-(1-\beta)(u_1(c_1(\phi); \phi) - u_1(c_1(\phi'); \phi)) > 0$, and hence $c_1(\phi') > c_1(\phi)$.

Proof of Lemma 2. Immediate from $U^1(x^a; \phi) - U^1(x^b; \phi) = -(1-\beta)(u_2(x_2^a; \phi) - u_2(x_2^b; \phi)) + U^2(x^a; \phi) - U^2(x^b; \phi)$.

Proof of Proposition 2. **Claim 1** For all $s_1 \in [0, s_1^*)$, $\hat{X}_2(s_1) > c'_2$.

Proof of Claim 1. Note that from Lemma A-2 and concavity of preferences, $0 \in \arg\max_{\tilde{s}_2 \geq 0} U^2(s_1 + c' - \tilde{s}_2; \theta)$ for all $s_1 \in [0, s_1^*]$. It then follows from continuity of preferences and the definition of s_1^* that

$U^1(s_1^* + c'; \theta) = U^1(c; \theta)$. Together with $U^1(c'; \theta) > U^1(c; \theta)$ and the concavity of $U^1(s_1 + c'; \theta)$ in s_1, this implies that for all $s_1 \in [0, s_1^*)$, $U^1(s_1 + c'; \theta) > U^1(c; \theta)$.

Fix $s_1 \in [0, s_1^*)$. From (IC_1), $U^1(s_1 + c; \theta) \geq U^1(s_1 + \hat{X}(s_1); \theta)$ and therefore $U^1(s_1 + c'; \theta) > U^1(s_1 + \hat{X}(s_1); \theta)$; from ($\text{IC}_2$) and the fact that $X(0) = c'$, $U^2(s_1 + \hat{X}(s_1); \theta) \geq U^2(s_1 + c'; \theta)$; it then follows from Lemma 2 that $\hat{X}_2(s_1) > c_2'$.

Claim 2 For all $s_1 \in [0, s_1^*)$, $\hat{X}_2(s_1) + \hat{X}_3(s_1) < c_2' + c_3'$.

Proof of Claim 2. Suppose to the contrary that for some $s_1 \in [0, s_1^*)$, $\hat{X}_2(s_1) + \hat{X}_3(s_1) \geq c_2' + c_3'$ and note that from Claim 1, $\hat{X}_2(s_1) > c_2'$. By the definition of c', $U_2^1(c'; \theta') = U_3^1(c'; \theta')$; therefore $U_2^2(c'; \theta') > U_3^2(c'; \theta')$ (since $\beta < 1$) and $U^2(\hat{X}(s_1) - s_2; \theta') > U^2(c'; \theta')$ for some $s_2 > 0$, violating (IC_2') and contradicting our supposition that $\hat{X}$ is part of an incentive compatible contract.

Claim 3 For all $s_1 \geq 0$, $\hat{X}_2(s_1)$ is weakly decrasing in s_1 and $\hat{X}_2(s_1) + \hat{X}_3(s_1)$ is weakly increasing in s_1.

Proof of Claim 3. Fix s_1 and $\tilde{s}_1 > s_1$. First note that, by (IC_2), $U^2(s_1 + \hat{X}(s_1); \theta) \geq U^2(s_1 + \hat{X}(\tilde{s}_1); \theta)$ and $U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1); \theta) \geq U^2(\tilde{s}_1 + \hat{X}(s_1); \theta)$, implying

$$U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1); \theta) - U^2(s_1 + \hat{X}(\tilde{s}_1); \theta) \geq U^2(\tilde{s}_1 + \hat{X}(s_1); \theta) - U^2(s_1 + \hat{X}(s_1); \theta).$$

Since date 3 consumption is the same under $\tilde{s}_1 + \hat{X}(\tilde{s}_1)$ and $s_1 + \hat{X}(\tilde{s}_1)$ on the one hand, and under $\tilde{s}_1 + \hat{X}(s_1)$ and $s_1 + \hat{X}(s_1)$ on the other, it follows from concavity of preferences that $\hat{X}_2(\tilde{s}_1) \leq \hat{X}_2(s_1)$. Second note that, by ($\text{IC}_2$), $U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1); \theta) \geq U^2(\tilde{s}_1 + \hat{X}(s_1) - s_2; \theta)$ for $s_2 = \hat{X}_2(s_1) - \hat{X}_2(\tilde{s}_1) \geq 0$. Since date 2 consumption is the same under $\tilde{s}_1 + \hat{X}(\tilde{s}_1)$ and $\tilde{s}_1 + \hat{X}(s_1) - s_2$, it follows that date 3 consumption is greater under the former, i.e., $\hat{X}_3(\tilde{s}_1) \geq \hat{X}_3(s_1) + s_2$. Substituting in for s_2 implies $\hat{X}_2(\tilde{s}_1) + \hat{X}_3(\tilde{s}_1) \geq \hat{X}_2(s_1) + \hat{X}_3(s_1)$.

Proof of Proposition 3. The proof is by contradiction: suppose that Condition SPR does not hold but commitment is nonetheless possible.

Given Lemma 4 (proved below), Condition SPR$'$ does not hold. We start by showing that failure of SPR$'$ implies that there exists some $s_2^+ \in \mathbb{R}$ such that $U^2(c'; \theta') < U^2(\hat{x}^0 - s_2^+; \theta')$. This is trivially true if $\frac{U_2^2(s_1^* + c'; \theta)}{U_3^2(s_1^* + c'; \theta)} > 1$. For the case $\frac{U_2^2(s_1^* + c'; \theta)}{U_3^2(s_1^* + c'; \theta)} \leq 1$, note first that this implies $\hat{x}^0 = c'$. The definition of c' implies that $U_2^1(c'; \theta') = U_3^1(c'; \theta')$ and hence (since $\beta < 1$) $U_2^2(c'; \theta') > U_3^2(c'; \theta')$, so for some $s_2^+ < 0$, $U^2(c'; \theta') < U^2(c' - s_2^+; \theta') = U^2(\hat{x}^0 - s_2^+; \theta')$.

Parallel to $\hat{x}^0$, let $\hat{x}^0(s_1)$ be the consumption plan $\tilde{x}$ that minimizes $\tilde{x}_2+\tilde{x}_3$ subject to $\tilde{x}_1 = c_1'$, $\tilde{x}_2 \geq c_2'$, and $U^2(s_1+\tilde{x};\theta) \geq U^2(s_1+c';\theta)$. Observe that (i) in the limit as s_1 approaches s_1^*, $\hat{x}^0(s_1)$ approaches $\hat{x}^0$ and (ii) $U^2\left(s_1+\hat{x}^0(s_1);\theta\right) = U^2(s_1+c';\theta) \geq U^2\left(s_1+\hat{x}^0(s_1)-s_2;\theta\right)$ for all s_2 if $\hat{x}_2^0(s_1) > c_2'$ and for all $s_2<0$ if $\hat{x}_2^0(s_1) = c_2'$.

Let $(X,\hat{X})$ be a commitment contract (at least one exists by supposition). The following two claims, which contradict each other, prove the proposition.

Claim 1 *For all* $s_1 < s_1^*$, $\hat{X}_2(s_1) \geq \hat{x}_2^0 - s_2^+$.

Proof of Claim 1. Suppose to the contrary that for some $s_1 < s_1^*$, $\hat{X}_2(s_1) < \hat{x}_2^0 - s_2^+$. By Proposition 2, $\hat{X}_2(s_1)$ is weakly decreasing in s_1 and $\hat{X}_2(s_1) > c_2'$ for all $s_1 < s_1^*$. Therefore, $\hat{X}_2(\tilde{s}_1) \in \left(c_2', \hat{x}_2^0 - s_2^+\right)$ for all $\tilde{s}_1 \in (s_1, s_1^*)$. By (IC$_2$) and (IC$_2'$),

$$\begin{aligned} U^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta) &\geq U^2\left(\tilde{s}_1+c';\theta\right) \\ U^2\left(c';\theta'\right) &\geq U^2(\hat{X}(\tilde{s}_1);\theta') \end{aligned}$$

for all $\tilde{s}_1$, in particular for all $\tilde{s}_1 \in (s_1, s_1^*)$. Let $s_2(\tilde{s}_1)$ be such that $\hat{x}_2^0(\tilde{s}_1) - s_2(\tilde{s}_1) = \hat{x}_2^0 - s_2^+$ and note that $c_2' < \hat{x}_2^0 - s_2^+$ then implies $\hat{x}_2^0(\tilde{s}_1) - s_2(\tilde{s}_1) > c_2'$. It follows that either $\hat{x}_2^0(\tilde{s}_1) > c_2'$ or $s_2(\tilde{s}_1) < 0$. From observation (ii) above,

$$U^2\left(\tilde{s}_1+c';\theta\right) \geq U^2\left(\tilde{s}_1+\hat{x}^0(\tilde{s}_1)-s_2(\tilde{s}_1);\theta\right).$$

Condition SCB implies

$$U^2\left(c';\theta'\right) \geq U^2\left(\hat{x}^0(\tilde{s}_1)-s_2(\tilde{s}_1);\theta'\right)$$

which, in the limit as $\tilde{s}_1$ approaches s_1^*, contradicts our supposition that $U^2\left(c';\theta'\right) < U^2\left(\hat{x}^0 - s_2^+;\theta'\right)$ and completes the proof of Claim 1.

Claim 2 *For some* $s_1 < s_1^*$, $\hat{X}_2(s_1) < \hat{x}_2^0 - s_2^+$.

Proof of Claim 2. Suppose to the contrary that for all $s_1 < s_1^*$, $\hat{X}_2(s_1) \geq \hat{x}_2^0 - s_2^+$. By (IC$_2$),

$$U^2(s_1+\hat{X}(s_1);\theta) \geq U^2\left(s_1+c';\theta\right)$$

which, by the definition of $\hat{x}^0(s_1)$, implies that $\hat{X}_2(s_1)+\hat{X}_3(s_1) \geq \hat{x}_2^0(s_1)+\hat{x}_3^0(s_1)$ for all $s_1 < s_1^*$.

Let $s_2(s_1)$ be such that $\hat{X}_2(s_1) - s_2(s_1) = \hat{x}_2^0 - s_2^+$ and note that by supposition, $s_2(s_1) \geq 0$. Therefore, for all $s_1 < s_1^*$, (IC_2') implies

$$U^2\left(c';\theta'\right) \geq U^2(\hat{X}(s_1) - s_2(s_1);\theta').$$

For all $s_1 < s_1^*$, date 2 consumption under $\hat{X}(s_1) - s_2(s_1)$ is $\hat{x}_2^0 - s_2^+$, while date 3 consumption is

$$\hat{X}_3(s_1) + s_2(s_1) \geq \hat{x}_2^0(s_1) + \hat{x}_3^0(s_1) - \hat{X}_2(s_1) + s_2(s_1) = \hat{x}_2^0(s_1) + \hat{x}_3^0(s_1) - \hat{x}_2^0 + s_2^+.$$

Therefore,

$$U^2\left(c';\theta'\right) \geq U^2((\hat{X}_1(s_1), \hat{x}_2^0 - s_2^+, \hat{x}_2^0(s_1) + \hat{x}_3^0(s_1) - \hat{x}_2^0 + s_2^+);\theta')$$

which, in the limit as s_1 approaches s_1^*, contradicts our supposition that $U^2\left(c';\theta'\right) < U^2\left(\hat{x}^0 - s_2^+;\theta'\right)$, completing the proof of Claim 2 and thus of Proposition 3.

Proof of Lemma 3. Suppose that $U_2^2(s_1^* + c';\theta)/U_3^2(s_1^* + c';\theta) \geq U_2^2\left(c';\theta'\right)/U_3^2\left(c';\theta'\right)$ but, to the contrary, Condition SPR is violated. Since c' is self 0's most preferred consumption, $U_2^2\left(c';\theta'\right)/U_3^2\left(c';\theta'\right) > 1$. So there exists $\tilde{x}$ such that $\tilde{x}_2 > c_2'$, $U^2\left(c';\theta'\right) = U^2\left(\tilde{x};\theta'\right) \geq U^2\left(\tilde{x} - s_2;\theta'\right)$ for all $s_2 \geq 0$. Since by supposition SPR is violated, $U^2(s_1^* + c';\theta) > U^2(s_1^* + \tilde{x};\theta)$. So there exists $s_1 \in (0, s_1^*)$ such that $U^2(s_1 + c';\theta) > U^2(s_1 + \tilde{x};\theta)$ (note that $s_1^* > 0$ whenever there is a commitment problem). By adding date 3 consumption to $\tilde{x}$, one obtains $\tilde{x}'$ such that $\tilde{x}_2' > c_2'$, $U^2(s_1 + c';\theta) = U^2(s_1 + \tilde{x}';\theta)$, and $U^2\left(c';\theta'\right) < U^2\left(\tilde{x}';\theta'\right)$. Since $U_2^2(s_1^* + c';\theta)/U_3^2(s_1^* + c';\theta) \geq U_2^2\left(c';\theta'\right)/U_3^2\left(c';\theta'\right)$ and $s_1 < s_1^*$, then $U_2^2(s_1 + c';\theta)/U_3^2(s_1 + c';\theta) > U_2^2\left(c';\theta'\right)/U_3^2\left(c';\theta'\right)$. But then $U^2(s_1 + c';\theta) = U^2(s_1 + \tilde{x}';\theta)$ and SCB imply that $U^2\left(c';\theta'\right) \geq U^2\left(\tilde{x}';\theta'\right)$, a contradiction.

Proof of Lemma 4. Note that $\frac{U_2^2\left(s_1^* + c';\theta\right)}{U_3^2\left(s_1^* + c';\theta\right)} > 1$ if and only if $\hat{x}_2^0 > c_2'$. Given this, it is immediate that SPR′ implies SPR. Suppose now that SPR′ is not satisfied. We show that SPR is not satisfied either.

First, suppose that $\hat{x}_2^0 \leq c_2'$. By the definition of $\hat{x}^0$, this implies that for all $\tilde{x}$ such that $\tilde{x}_2 > c_2'$ and $U^2(s_1^* + \tilde{x};\theta) \geq U^2(s_1^* + c';\theta)$, $\tilde{x}_2 + \tilde{x}_3 > c_2' + c_3'$. As a result, for every such $\tilde{x}$, $U^2\left(c';\theta'\right) < U^2\left(\tilde{x} - s_2;\theta'\right)$ for $s_2 = \tilde{x}_2 - c_2'$. So SPR does not hold.

Second, suppose that $\hat{x}_2^0 > c_2'$ but that $U^2\left(c';\theta'\right) < U^2\left(\hat{x}^0 - s_2^+;\theta'\right)$ for some $s_2^+ \geq 0$. Note that by the definition of $\hat{x}^0$, $\tilde{x}_2 + \tilde{x}_3 \geq \hat{x}_2^0 + \hat{x}_3^0$ for all $\tilde{x}$ such that $U^2(s_1^* + \tilde{x};\theta) \geq U^2(s_1^* + c';\theta)$.

On the one hand, if $\tilde{x}$ is such that $\tilde{x}_2 \geq \hat{x}_2^0 - s_2^+$ and $U^2(s_1^* + \tilde{x}; \theta) \geq U^2(s_1^* + c'; \theta)$, we have $U^2(c'; \theta') < U^2(\tilde{x} - s_2; \theta')$ when $s_2 \geq 0$ is set such that $\tilde{x}_2 - s_2 = \hat{x}_2^0 - s_2^+$.

On the other hand, if $\tilde{x}$ is such that $c_2' < \tilde{x}_2 < \hat{x}_2^0 - s_2^+$ and $U^2(s_1^* + \tilde{x}; \theta) \geq U^2(s_1^* + c'; \theta)$, we have $U^2(c'; \theta') < U^2(\tilde{x}; \theta')$, as follows. Suppose to the contrary that $U^2(c'; \theta') \geq U^2(\tilde{x}; \theta')$. By the definition of $\hat{x}^0$, $U^2(s_1^* + \hat{x}^0 - s_2^+; \theta) \leq U^2(s_1^* + c'; \theta)$. So SCB implies that $U^2(c'; \theta') \geq U^2(\hat{x}^0 - s_2^+; \theta')$, contradicting our supposition that $U^2(c'; \theta') < U^2(\hat{x}^0 - s_2^+; \theta')$.

So SPR does not hold, completing the proof.

Proof of Proposition 4. **Claim 1** If $\hat{X}_2^*(s_1^*) > c_2'$ then $U_2^2(s_1^* + \hat{X}^*(s_1^*); \theta) \geq U_3^2(s_1^* + \hat{X}^*(s_1^*); \theta)$.

Proof of Claim 1. Suppose to the contrary that $U_2^2(s_1^* + \hat{X}^*(s_1^*); \theta) < U_3^2(s_1^* + \hat{X}^*(s_1^*); \theta)$ and define $\tilde{x}(\tilde{x}_2)$ by $\tilde{x}_1(\tilde{x}_2) = c_1'$, $\tilde{x}_2(\tilde{x}_2) = \tilde{x}_2$, and $U^2(s_1^* + \tilde{x}(\tilde{x}_2); \theta) = U^2(s_1^* + \hat{X}^*(s_1^*); \theta) = U^2(s_1^* + c'; \theta)$. There then exists $\varepsilon > 0$ such that $\hat{X}_2^*(s_1^*) - \varepsilon > c_2'$ and $\tilde{x}_2(\hat{X}_2^*(s_1^*) - \varepsilon) + \tilde{x}_3(\hat{X}_2^*(s_1^*) - \varepsilon) < \hat{X}_2^*(s_1^*) + \hat{X}_3^*(s_1^*)$. Hence $U^2(c'; \theta') > U^2(\tilde{x}(\hat{X}_2^*(s_1^*) - \varepsilon) - s_2; \theta')$ for all $s_2 \geq 0$, contradicting the definition of $\hat{X}^*(s_1^*)$.

Claim 2 Fix s_1 and x such that $x_1 = c_1'$, $x_2 > c_2'$, $U^2(s_1 + x; \theta) \geq U^2(s_1 + c'; \theta)$, and $U^2(c'; \theta') \geq U^2(x - s_2; \theta')$ for all $s_2 \geq 0$. Fix $\tilde{x}$ such that $\tilde{x}_1 = c_1'$, $\tilde{x}_2 \geq x_2$, $U^2(s_1 + \tilde{x}; \theta) = U^2(s_1 + c'; \theta)$, and $U_2^2(s_1 + \tilde{x}; \theta) \geq U_3^2(s_1 + \tilde{x}; \theta)$. Then $U^2(c'; \theta') \geq U^2(\tilde{x} - s_2; \theta')$ for all $s_2 \geq 0$.

Proof of Claim 2. We first show that $U^2(c'; \theta') \geq U^2(\tilde{x} - s_2; \theta')$ for all $s_2 \in [0, \tilde{x}_2 - x_2]$. Note that $U^2(s_1 + \tilde{x}; \theta) = U^2(s_1 + c'; \theta)$ and $U_2^2(s_1 + \tilde{x}; \theta) \geq U_3^2(s_1 + \tilde{x}; \theta)$ imply that $U^2(s_1 + c'; \theta) \geq U^2(s_1 + \tilde{x} - s_2; \theta)$ for all $s_2 \geq 0$, in particular for all $s_2 \in [0, \tilde{x}_2 - x_2]$. Therefore, since $\tilde{x}_2 - s_2 \geq x_2$ for all $s_2 \in [0, \tilde{x}_2 - x_2]$, it follows from SCB that $U^2(c'; \theta') \geq U^2(\tilde{x} - s_2; \theta')$ for all $s_2 \in [0, \tilde{x}_2 - x_2]$.

We next show that $U^2(c'; \theta') \geq U^2(\tilde{x} - s_2; \theta')$ for all $s_2 \geq \tilde{x}_2 - x_2$. Note that $\tilde{x}_2 \geq x_2$, $U^2(s_1 + x; \theta) \geq U^2(s_1 + c'; \theta) = U^2(s_1 + \tilde{x}; \theta)$, and $U_2^2(s_1 + \tilde{x}; \theta) \geq U_3^2(s_1 + \tilde{x}; \theta)$ imply that $x_2 + x_3 \geq \tilde{x}_2 + \tilde{x}_3$. Therefore it follows from $U^2(c'; \theta') \geq U^2(x - s_2; \theta')$ for all $s_2 \geq 0$ that $U^2(c'; \theta') \geq U^2(\tilde{x} - s_2; \theta')$ for all $s_2 \geq \tilde{x}_2 - x_2$.

Final step of proof. Suppose to the contrary that for some $s_1^+ < s_1^*$, $\hat{X}_2(s_1^+) < \hat{X}_2^*(s_1^*)$. So $\hat{X}_2^*(s_1^*) > c_2'$ by Proposition 2. For any $\tilde{s}_1$ define $\tilde{x}(\tilde{s}_1)$ by $\tilde{x}_1 = c_1'$, $\tilde{x}_2(\tilde{s}_1) = \hat{X}_2(s_1^+)$, and $U^2(s_1 + \tilde{x}(\tilde{s}_1)) = U^2(\tilde{s}_1 + c'; \theta)$. By the definition of $\hat{X}^*(s_1^*)$, $U^2(c'; \theta') < U^2(\tilde{x}(s_1^*) - s_2; \theta')$ for some $s_2 \geq 0$. By continuity of preferences, there thus exist $s_1 \in (s_1^+, s_1^*)$ and $s_2 \geq 0$ such that

$$U^2(c'; \theta') < U^2(\tilde{x}(s_1) - s_2; \theta'). \tag{A-1}$$

Since $s_1 > s_1^+$, Proposition 2 implies that $c_2' < \hat{X}_2(s_1) \leq \hat{X}_2\left(s_1^+\right) = \tilde{x}_2(s_1)$. By (IC$_2$), $U^2(s_1 + \hat{X}(s_1);\theta) \geq U^2(s_1 + c';\theta)$, and by (IC$_2'$), $U^2\left(c';\theta'\right) \geq U^2(\hat{X}(s_1) - s_2;\theta')$ for all $s_2 \geq 0$. By Claim 1, $U_2^2(s_1^* + \hat{X}^*(s_1^*);\theta) \geq U_3^2(s_1^* + \hat{X}^*(s_1^*);\theta)$. Together with $\tilde{x}_2(s_1^*) = \hat{X}_2\left(s_1^+\right) < \hat{X}_2^*(s_1^*)$ and $U^2(s_1^* + \tilde{x}(s_1^*)) = U^2(s_1^* + c';\theta) = U^2(s_1^* + \hat{X}^*(s_1^*);\theta)$, this implies $U_2^2(s_1^* + \tilde{x}(s_1^*);\theta) \geq U_3^2(s_1^* + \tilde{x}(s_1^*);\theta)$. Finally note that since for any $\tilde{s}_1$, $U^2(\tilde{s}_1 + \tilde{x}(\tilde{s}_1)) = U^2(\tilde{s}_1 + c';\theta)$ and $\tilde{x}_2(\tilde{s}_1) = \tilde{x}_2(s_1^*) > c_2'$, we must have $U_2^2(s_1 + \tilde{x}(s_1);\theta) \geq U_3^2(s_1 + \tilde{x}(s_1);\theta)$ for $s_1 < s_1^*$. (This is due to the clockwise rotation of the indifference curve through c' as $\tilde{s}_1$ is decreased.) Since $\tilde{x}(s_1)$ satisfies all the conditions of Claim 2 with $x = \hat{X}(s_1)$, it follows that $U^2\left(c';\theta'\right) \geq U^2\left(\tilde{x}(s_1) - s_2;\theta'\right)$ for all $s_2 \geq 0$, a contradiction to (A-1).

Proof of Lemma 5. We show that for all $\varepsilon > 0$, there exists $\tilde{x}$ such that $c_2' < \tilde{x}_2 < c_2' + \varepsilon$ and (1)-(2) are both satisfied. Suppose to the contrary that there exists $\varepsilon > 0$ such that this is not the case. An easy adaptation of the proof of Lemma 3 gives a contradiction.

Proof of Lemma 6. For the first half of the result, recall that the incentive compatibility conditions imply the self-selection condition (5). It follows immediately that for every s_1, $\tilde{s}_1 \geq 0$,

$$\begin{aligned} U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1);\theta) - U^2(s_1 + \hat{X}(\tilde{s}_1);\theta) &\geq U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1);\theta) - U^2(s_1 + \hat{X}(s_1);\theta) \\ &\geq U^2(\tilde{s}_1 + \hat{X}(s_1);\theta) - U^2(s_1 + \hat{X}(s_1);\theta). \quad \text{(A-2)} \end{aligned}$$

From inequality (A-2), $U^2(s_1 + \hat{X}(s_1;\theta);\theta)$ is strictly increasing, and continuous. (For continuity, simply note that both $U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1;\theta);\theta) - U^2(s_1 + \hat{X}(\tilde{s}_1;\theta);\theta)$ and $U^2(\tilde{s}_1 + \hat{X}(s_1;\theta);\theta) - U^2(s_1 + \hat{X}(s_1;\theta);\theta)$ converge to 0 as $\tilde{s}_1 \to s_1$.) To establish differentiability, divide (A-2) everywhere by $\tilde{s}_1 - s_1$, yielding

$$\begin{aligned} \frac{U^2(\tilde{s}_1 + X(\tilde{s}_1);\theta) - U^2(s_1 + X(\tilde{s}_1);\theta)}{\tilde{s}_1 - s_1} &\geq \frac{U^2(\tilde{s}_1 + \hat{X}(\tilde{s}_1);\theta) - U^2(s_1 + \hat{X}(s_1);\theta)}{\tilde{s}_1 - s_1} \\ &\geq \frac{U^2(\tilde{s}_1 + X(s_1);\theta) - U^2(s_1 + X(s_1);\theta)}{\tilde{s}_1 - s_1}. \end{aligned}$$

If X is continuous at s_1, then the upper bound and the lower bound both converge to $U_2^2(s_1 + \hat{X}(s_1);\theta)$ as $\tilde{s}_1 \to s_1$, establishing differentiability.

For the converse, fix s_1, and consider the function defined by $f(\tilde{s}_1) \equiv U^2(s_1 + \hat{X}(\tilde{s}_1);\theta)$.

Observe that

$$\begin{aligned}
f'(\tilde{s}_1) &= U_2^2(s_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_2(\tilde{s}_1)}{d\tilde{s}_1} + U_3^2(s_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_3(\tilde{s}_1)}{d\tilde{s}_1} \\
&= \frac{U_2^2(s_1+\hat{X}(\tilde{s}_1);\theta)}{U_2^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)}U_2^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_2(\tilde{s}_1)}{d\tilde{s}_1} \\
&\quad + \frac{U_3^2(s_1+\hat{X}(\tilde{s}_1);\theta)}{U_3^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)}U_3^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_3(\tilde{s}_1)}{d\tilde{s}_1}.
\end{aligned}$$

From (6),

$$U_2^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_2(\tilde{s}_1)}{d\tilde{s}_1} + U_3^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_3(\tilde{s}_1)}{d\tilde{s}_1} = 0,$$

and so

$$f'(\tilde{s}_1) = U_2^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)\frac{d\hat{X}_2(\tilde{s}_1)}{d\tilde{s}_1}\left(\frac{U_2^2(s_1+\hat{X}(\tilde{s}_1);\theta)}{U_2^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)} - \frac{U_3^2(s_1+\hat{X}(\tilde{s}_1);\theta)}{U_3^2(\tilde{s}_1+\hat{X}(\tilde{s}_1);\theta)}\right).$$

Clearly $f'(\tilde{s}_1 = s_1) = 0$. From $\frac{d\hat{X}_2(\tilde{s}_1)}{d\tilde{s}_1} \leq 0$, $U_{22}^2 < 0$, and the independence of U_3^2 from s_1, it follows that $s_1'(\tilde{s}_1) \leq 0$ if $\tilde{s}_1 \geq s_1$ and $f'(\tilde{s}_1) \geq 0$ if $\tilde{s}_1 \leq s_1$. Hence f attains its maximum at $\tilde{s}_1 = s_1$, implying the result.

Proof of Theorem 1 (sufficiency)
Define $X^*(s_1)$ as the solution to

$$\max_{\tilde{c}} \left\{U^2\left(s_1+\tilde{c};\theta'\right)\right\} \text{ s.t. } \tilde{c}_1 = c_1',\ \tilde{c}_2 \leq c_2', \text{ and } \tilde{c}_2+\tilde{c}_3 = c_2'+c_3'.$$

In words, X^* results from giving self 2 (c_2', c_3') and letting him save whatever he likes.

Claim 1 *If* $(X, \hat{X})$ *satisfies the incentive constraints* (IC_1), (IC_1'), (IC_2) *and* (IC_2'), *then commitment is possible using a contract in which self* 2 *picks from* X *and* $\hat{X}$ *when self* 1 *chooses* c_1', *and self* 2 *receives* (c_2, c_3) *and saves whatever he likes after self* 1 *chooses* c_1.

Proof of Claim 1. Under such a contract, self 2's only choice is how much to save. Since his choice of savings in either state is optimal by definition, incentive compatibility for self 2 is immediate.

We next show that $U_2^2(c;\theta) \geq U_3^2(c;\theta)$, i.e., after self 1 chooses c_1 and saves nothing in state θ (as he should in equilibrium), self 2 does not save either, thus ensuring that self 0's preferred consumption plan c is chosen in equilibrium. To see why, note that it follows from the definition of c that $U_2^1(c;\theta) = U_3^1(c;\theta)$; since $\beta < 1$ it then follows that $U_2^2(c;\theta) > U_3^2(c;\theta)$.

We next show that, for all $s_1, s_2 \geq 0$, $U^1(c;\theta) \geq U^1(s_1 + c - s_2;\theta)$, i.e., we show that self 1 does not want to save after correctly choosing c_1 in state θ, no matter how much self 2 subsequently saves. To see why, first note that for $\beta = 1$, it follows from the definition of c that $U^1(c;\theta) \geq U^1(s_1 + c - s_2;\theta)$ for all $s_1, s_2 \geq 0$, or

$$u_1(c_1;\theta) - u_1(c_1 - s_1;\theta) \geq \beta\left(u_2(s_1 + c_2 - s_2;\theta) - u_2(c_2;\theta) + u_3(s_2 + c_3;\theta) - u_3(c_3;\theta)\right).$$

Since $s_1 \geq 0$, the left hand side is positive; it follows that the condition is also satisfied for $\beta < 1$.

We finally show that, for all $s_1, s_2 \geq 0$, $U^1\left(c';\theta'\right) \geq U^1\left(s_1 + c - s_2;\theta'\right)$, i.e., we show that self 1 does not want to choose c_1 in state θ', no matter how much self 2 subsequently saves. To see why, first note that for $\beta = 1$, it follows from the definition of c' that $U^1\left(c';\theta'\right) \geq U^1\left(s_1 + c - s_2;\theta'\right)$ for all $s_1, s_2 \geq 0$, or

$$u_1\left(c_1';\theta'\right) - u_1\left(c_1 - s_1;\theta'\right) \geq \beta\left(u_2\left(s_1 + c_2 - s_2;\theta'\right) - u_2\left(c_2';\theta'\right) + u_3\left(s_2 + c_3;\theta'\right) - u_3\left(c_3';\theta'\right)\right).$$

Since $s_1 \geq 0$ and $c_1 \leq c_1'$, the left hand side is positive; it follows that the condition is also satisfied for $\beta < 1$.

Claim 2 *For all $s_1 \geq 0$, $\hat{X}_2^*(s_1) + \hat{X}_3^*(s_1) \leq c_2' + c_3'$ and $\hat{X}_2^*(s_1) + \hat{X}_3^*(s_1)$ is nondecreasing in s_1.*

Proof of Claim 2. For $s_1 > s_1^*$, $\hat{X}_2^*(s_1) + \hat{X}_3^*(s_1) = c_2' + c_3'$ by the definition of $\hat{X}^*$. For $s_1 = s_1^*$, $\hat{X}_2^*(s_1) + \hat{X}_3^*(s_1) \leq c_2' + c_3'$ by the definition of $\hat{X}^*(s_1^*)$. For $s_1 \in [0, s_1^*]$, $\hat{X}^*$ is continuous and $U_3^2(s_1 + \hat{X}^*(s_1);\theta)d\hat{X}_3^*(s_1) = -U_2^2(s_1 + \hat{X}^*(s_1);\theta)d\hat{X}_2^*(s_1)$. By NS, $U_3^2(s_1 + \hat{X}^*(s_1);\theta) \geq U_2^2(s_1 + \hat{X}^*(s_1);\theta)$ and therefore $d\hat{X}_2^*(s_1) + d\hat{X}_3^*(s_1) \geq 0$.

Claim 3 *$\hat{X}^*$ satisfies (IC$_1$).*

Proof of Claim 3. Recall from page 15 that, for $s_1 \geq s_1^*$, $\hat{X}^*(s_1)$ solves $\max_{\tilde{c}}\left\{U^2(s_1 + \tilde{c};\theta)\right\}$ s.t. $\tilde{c}_1 = c_1'$, $\tilde{c}_2 \leq c_2'$, and $\tilde{c}_2 + \tilde{c}_3 = c_2' + c_3'$. It follows from the definition of s_1^* that (IC$_1$) is satisfied for all $s_1 \geq s_1^*$.

For $s_1 \in [s_1^{**}, s_1^*]$, $\hat{X}^*(s_1) = \hat{X}^*(s_1^*)$. From the definition of $\hat{X}^*(s_1^*)$, $U^2(s_1^* + \hat{X}^*(s_1^*);\theta) = U^2(s_1^* + c';\theta)$ and $\hat{X}_2^*(s_1^*) \geq c_2'$; it follows from Lemma 2 that $U^1(s_1^* + \hat{X}^*(s_1^*);\theta) \leq U^1(s_1^* + c';\theta) =$

$U^1(c;\theta)$, where the last equality follows from the definition of s_1^* and Lemma A-2. As a result, (IC$_1$) is satisfied for $s_1 = s_1^*$; it follows from the definition of s_1^{**} that it is satisfied for all $s_1 \in [s_1^{**}, s_1^*]$.

For $s_1 \in [0, s_1^{**}]$, $U^1(s_1+\hat{X}^*(s_1);\theta)$ is nondecreasing in s_1. To see why, take the total differential

$$-\left(U_1^1(s_1+\hat{X}^*;\theta) - U_2^1(s_1+\hat{X}^*;\theta)\right) ds_1 + U_2^1(s_1+\hat{X}^*;\theta) d\hat{X}_2^* + U_3^1(s_1+\hat{X}^*;\theta) d\hat{X}_3^*$$

which, using the definition of $d\hat{X}_3^*$, we can rewrite as

$$\begin{aligned}&-\left(U_1^1(s_1+\hat{X}^*;\theta) - U_2^1(s_1+\hat{X}^*;\theta)\right) ds_1 \\ &+U_3^1\left(s_1+\hat{X}^*;\theta\right)\left(\frac{U_2^1(s_1+\hat{X}^*;\theta)}{U_3^1(s_1+\hat{X}^*;\theta)} - \frac{U_2^2(s_1+\hat{X}^*;\theta)}{U_3^2(s_1+\hat{X}^*;\theta)}\right) d\hat{X}_2^*\end{aligned}$$

which, using the definition of $d\hat{X}_2^*$, simplifies to

$$\max\left\{0, -\left(U_1^1(s_1+\hat{X}^*;\theta) - U_2^1(s_1+\hat{X}^*;\theta)\right)\right\} ds_1.$$

By the definition of s_1^{**}, $U^1(s_1^{**}+\hat{X}^*(s_1^{**});\theta) = U^1(c;\theta)$; it follows that (IC$_1$) is satisfied for all $s_1 \in [0, s_1^{**}]$.

Claim 4 *X^* satisfies (IC$'_1$).*

Proof of Claim 4. Parallel to the final step of Claim 1.

Claim 5 *The contract $(\hat{X}^*, X^*)$ satisfies (IC$_2$).*

Proof of Claim 5. Step A. We show that for all $s_1, s_1' \in [0, s_1^*]$ and $s_2 \geq 0$, $U^2(s_1+\hat{X}^*(s_1);\theta) \geq U^2(s_1+\hat{X}^*(s_1') - s_2;\theta)$. Consider the function $f(s_1') \equiv \max_{s_2 \geq 0} U^2(s_1+\hat{X}^*(s_1') - s_2;\theta)$ and note that by condition (NS), $f(s_1' = s_1) = U^2(s_1+\hat{X}^*(s_1);\theta)$; we will therefore show that $f(s_1')$ has a global maximum at $s_1' = s_1$. By standard envelope arguments,

$$\frac{df(s_1')}{ds_1'} = U_2^2(s_1+\hat{X}^*(s_1') - s_2';\theta)\frac{d\hat{X}_2^*(s_1')}{ds_1'} + U_3^2(s_1+\hat{X}^*(s_1') - s_2';\theta)\frac{d\hat{X}_3^*(s_1')}{ds_1'},$$

where $s_2' \in \arg\max_{s_2 \geq 0} U^2(s_1+\hat{X}^*(s_1') - s_2;\theta)$. By the definition of $\hat{X}^*$,

$$\frac{df(s_1')}{ds_1'} = U_3^2(s_1+\hat{X}^*(s_1') - s_2';\theta)\left(\frac{U_2^2(s_1+\hat{X}^*(s_1') - s_2';\theta)}{U_3^2(s_1+\hat{X}^*(s_1') - s_2';\theta)} - \frac{U_2^2(s_1'+\hat{X}^*(s_1');\theta)}{U_3^2(s_1'+\hat{X}^*(s_1');\theta)}\right)\frac{d\hat{X}_2^*(s_1')}{ds_1'}.$$

Observe that $\left.\frac{df(s_1')}{ds_1'}\right|_{s_1'=s_1} = 0$. Given that $\frac{d\hat{X}_2^*(s_1')}{ds_1'} \leq 0$, it suffices to show that

$$\frac{U_2^2(s_1 + \hat{X}^*(s_1') - s_2'; \theta)}{U_3^2(s_1 + \hat{X}^*(s_1') - s_2'; \theta)} \geq (\leq) \frac{U_2^2(s_1' + \hat{X}^*(s_1'); \theta)}{U_3^2(s_1' + \hat{X}^*(s_1'); \theta)} \text{ if } s_1' > (<) s_1,$$

since doing so establishes global concavity of f.

Consider the case of $s_1' > s_1$. If $U_2^2(s_1' + \hat{X}^*(s_1'); \theta) = U_3^2(s_1' + \hat{X}^*(s_1'); \theta)$ then the required inequality is immediate. If instead $U_2^2(s_1' + \hat{X}^*(s_1'); \theta) > U_3^2(s_1' + \hat{X}^*(s_1'); \theta)$, then $U_2^2(s_1 + \hat{X}^*(s_1'); \theta) > U_3^2(s_1 + \hat{X}^*(s_1'); \theta)$ also, implying that $s_2' = 0$. The required inequality then follows. Finally, consider the case of $s_1' < s_1$. If $U_2^2(s_1 + \hat{X}^*(s_1') - s_2'; \theta) = U_3^2(s_1 + \hat{X}^*(s_1') - s_2'; \theta)$ the required inequality is immediate. If instead $U_2^2(s_1 + \hat{X}^*(s_1') - s_2'; \theta) > U_3^2(s_1 + \hat{X}^*(s_1') - s_2'; \theta)$ then $s_2' = 0$, and the required inequality follows as before.

Step B. We show that for all $s_1 > s_1^*$, $s_1' \in [0, s_1^*]$, and $s_2 \geq 0$, $U^2(s_1 + \hat{X}^*(s_1); \theta) \geq U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta)$. By the definition of $\hat{X}^*$, $U^2(s_1 + \hat{X}^*(s_1); \theta) = \max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta)$ for all $s_1 > s_1^*$; we will therefore show that for all $s_1 > s_1^*$ and $s_1' \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta) \geq \max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta)$. By condition (NS), $U^2(s_1^* + \hat{X}^*(s_1^*); \theta) = \max_{s_2 \geq 0} U^2(s_1^* + \hat{X}^*(s_1^*) - s_2; \theta)$; by the definition of $\hat{X}^*(s_1^*)$, $U^2(s_1^* + c'; \theta) = U^2(s_1^* + \hat{X}^*(s_1^*); \theta)$; it then follows from $\hat{X}_2^*(s_1^*) \geq c_2'$ that $U^2(s_1^* + c'; \theta) = \max_{s_2 \geq 0} U^2(s_1^* + c' - s_2; \theta)$ which, together with Step A, implies that for all $s_1' \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2(s_1^* + c' - s_2; \theta) \geq \max_{s_2 \geq 0} U^2(s_1^* + \hat{X}^*(s_1') - s_2; \theta)$. Since for all $s_1' \in [0, s_1^*]$, $\hat{X}_2^*(s_1') \geq c_2'$, it then follows from Lemma A-1 that for all $s_1 > s_1^*$ and $s_1' \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta) \geq \max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta)$.

Step C. We show that for all $s_1 \in [0, s_1^*]$, $s_1' > s_1^*$, and $s_2 \geq 0$, $U^2(s_1 + \hat{X}^*(s_1); \theta) \geq U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta)$. By the definition of $\hat{X}^*$, $\max_{s' > s_1^*, s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta) = \max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta)$ for all $s_1 > s_1^*$; as before, $U^2(s_1 + \hat{X}^*(s_1); \theta) = \max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1) - s_2; \theta)$; we will therefore show that for all $s_1 \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1) - s_2; \theta) \geq \max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta)$. As before, $\max_{s_2 \geq 0} U^2(s_1^* + \hat{X}^*(s_1^*) - s_2; \theta) = \max_{s_2 \geq 0} U^2(s_1^* + c' - s_2; \theta)$; since $\hat{X}_2^*(s_1^*) \geq c_2'$ it follows from Lemma A-1 that for all $s_1 \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1^*) - s_2; \theta) \geq \max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta)$ which, together with the first step of the proof establishes that for all $s_1 > s_1^*$ and $s_1' \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta) \geq \max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta)$.

Step D. It follows directly from the definitions of $\hat{X}^*$ and X^* that for all $s_1, s_1' > s_1^*$ and $s_2 \geq 0$, $U^2(s_1 + \hat{X}^*(s_1); \theta) \geq U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta)$

Step E. By the definition of X^*, Step C establishes that for all $s_1 \in [0, s_1^*]$, $s_1' \geq 0$, and $s_2 \geq 0$, $U^2(s_1 + \hat{X}^*(s_1); \theta) \geq U^2(s_1 + X^*(s_1') - s_2; \theta)$ and Step D establishes that for all $s_1 > s_1^*$, $s_1' \geq 0$, and $s_2 \geq 0$, $U^2(s_1 + \hat{X}^*(s_1); \theta) \geq U^2(s_1 + X^*(s_1') - s_2; \theta)$.

Claim 6 *The contract $(\hat{X}^*, X^*)$ satisfies (IC_2').*

Proof of Claim 6. We first show that for all $s_1 \geq 0$, $s_1' \in [0, s_1^*]$, and $s_2 \geq 0$, $U^2\left(s_1 + X^*(s_1); \theta'\right) \geq U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta')$. By the definition of X^*, $U^2\left(s_1 + X^*(s_1); \theta'\right) = \max_{s_2 \geq 0} U^2\left(s_1 + c' - s_2; \theta'\right)$; we will therefore show that for all $s_1 \geq 0$ and $s_1' \in [0, s_1^*]$, $\max_{s_2 \geq 0} U^2\left(s_1 + c' - s_2; \theta'\right) \geq \max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta')$. Consider the case of $s_1 = 0$: since c' is self 0's desired allocation in state θ', $\max_{s_2 \geq 0} U^2\left(c' - s_2; \theta'\right) = U^2\left(c'; \theta'\right)$; by the definition of $\hat{X}^*(s_1^*)$, $U^2\left(c'; \theta'\right) = \max_{s_2 \geq 0} U^2(\hat{X}^*(s_1^*) - s_2; \theta')$. By the previous claim, $U^2(s_1^* + \hat{X}^*(s_1^*); \theta) \geq U^2(s_1^* + \hat{X}^*(s_1') - s_2; \theta)$ for all $s_1' \in [0, s_1^*]$ and $s_2 \geq 0$. If $\hat{X}_2^*(s_1') - s_2 > \hat{X}_2^*(s_1^*)$, then it follows from the definition of $\hat{X}^*(s_1^*)$ and SCB that $\max_{s_2 \geq 0} U^2(\hat{X}^*(s_1^*) - s_2; \theta') \geq U^2(\hat{X}^*(s_1') - s_2; \theta)$. If $\hat{X}_2^*(s_1') - s_2 \leq \hat{X}_2^*(s_1^*)$, then it follows from the fact that $\hat{X}_2^*(s_1') + \hat{X}_3^*(s_1') \leq \hat{X}_2^*(s_1^*) + \hat{X}_3^*(s_1^*)$ that $\max_{s_2 \geq 0} U^2(\hat{X}^*(s_1^*) - s_2; \theta') \geq U^2(\hat{X}^*(s_1') - s_2; \theta)$. As a result, $\max_{s_2 \geq 0} U^2\left(s_1 + c' - s_2; \theta'\right) \geq \max_{s_2 \geq 0} U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta')$ for $s_1 = 0$ and all $s_1' \in [0, s_1^*]$; it then follows from Lemma A-1 that the same is true for all $s_1 \geq 0$ and $s_1' \in [0, s_1^*]$.

Finally, it follows directly from the definitions of $\hat{X}^*$ and X^* that for all $s_1 \geq 0$ and $s_1' \geq s_1^*$, $U^2\left(s_1 + X^*(s_1); \theta'\right) \geq U^2(s_1 + \hat{X}^*(s_1') - s_2; \theta')$ and for all $s_1, s_1' \geq 0$, $U^2\left(s_1 + X^*(s_1); \theta'\right) \geq U^2\left(s_1 + X^*(s_1') - s_2; \theta'\right)$.

Proof of Theorem 1 (necessity)

Suppose that SPR is satisfied and NS is violated but, to the contrary, commitment is possible. Let $\hat{X}$ be part of an commitment contract.

Claim 1 *If $\hat{X}^*$ violates (NS) for some $s_1 \in [0, s_1^*]$ then it violates (NS) for some $s_1^+ \in [0, s_1^*]$ such that $U^1(s_1^+ + \hat{X}^*\left(s_1^+\right); \theta) = U(c; \theta)$.*

Proof of Claim 1. We first show that there exists $s_1^{***} \in [0, s_1^{**}]$ such that $U^1(s_1 + \hat{X}^*(s_1); \theta) = U^1(c; \theta)$ if $s_1 \in [s_1^{***}, s_1^{**}]$, and $\hat{X}^*$ is constant over $[0, s_1^{***}]$. If $U_1^1(s_1 + \hat{X}^*(s_1); \theta) - U_2^1(s_1 + \hat{X}^*(s_1); \theta) \leq 0$ then it follows from the definition of $\hat{X}^*$ that the same is true for all lower s_1.

Define

$$s_1^{***} = \sup \left\{ s_1 \in [0, s_1^{**}] : U_1^1(s_1 + \hat{X}^* (s_1) ; \theta) - U_2^1(s_1 + \hat{X}^* (s_1) ; \theta) \leq 0 \right\},$$

with $s_1^{***} = 0$ if the above set is empty; $\hat{X}^*$ is then constant on $[0, s_1^{***}]$. Moreover, from the total differential of $U^1(s_1 + \hat{X}^* (s_1) ; \theta)$ evaluated in Claim 3 of the sufficiency proof, $U^1(s_1 + \hat{X}^* (s_1) ; \theta) = U^1 (c; \theta)$ on $[s_1^{***}, s_1^{**}]$.

We next show that (NS) is violated for some $s_1^+ \in [s_1^{***}, s_1^{**}]$. From the definition of $\hat{X}^*$ and Lemma A-2, (NS) is satisfied at s_1^*. Since $\hat{X}^*$ is constant on $[s_1^{**}, s_1^*]$, it follows that (NS) is satisfied for all $s_1 \in [s_1^{**}, s_1^*]$. From the above, $\hat{X}^*$ is constant on $[0, s_1^{***}]$; it follows that if (NS) is violated at some $s_1 \in [0, s_1^{***})$, then it must be violated at s_1^{***} too.

Claim 2 $U^2(s_1^+ + \hat{X} \left(s_1^+\right) ; \theta) < U^2(s_1^+ + \hat{X}^* \left(s_1^+\right) ; \theta)$ *and* $U^2(s_1^* + \hat{X} (s_1^*) ; \theta) \geq U^2(s_1^* + \hat{X}^* (s_1^*) ; \theta)$.

Proof of Claim 2. Suppose to the contrary that $U^2(s_1^+ + \hat{X} \left(s_1^+\right) ; \theta) \geq U^2(s_1^+ + \hat{X}^* \left(s_1^+\right) ; \theta)$. From the previous claim, $U^1(s_1^+ + \hat{X}^* \left(s_1^+\right) ; \theta) = U (c; \theta)$; therefore $U^1(s_1^+ + \hat{X} \left(s_1^+\right) ; \theta) \leq U^1(s_1^+ + \hat{X}^* \left(s_1^+\right) ; \theta)$. Together with Lemma 2, these two inequalities imply that $\hat{X}_2 \left(s_1^+\right) \geq \hat{X}_2^* \left(s_1^+\right)$ and $\hat{X}_3 \left(s_1^+\right) \leq \hat{X}_3^* \left(s_1^+\right)$ which, in turn, implies that $U_2^2(s_1^+ + \hat{X} \left(s_1^+\right) ; \theta) < U_3^2(s_1^+ + \hat{X} \left(s_1^+\right) ; \theta)$, contradicting the supposition that $\hat{X}$ is part of a commitment contract.

From the definition of $\hat{X}^* (s_1^*)$, $U^2(s_1^* + \hat{X}^* (s_1^*) ; \theta) = U^2 (s_1^* + c'; \theta)$; it follows that $U^2(s_1^* + \hat{X} (s_1^*) ; \theta) \geq U^2(s_1^* + \hat{X}^* (s_1^*) ; \theta)$.

Claim 3 *There exists* $s_1 \in \left(s_1^+, s_1^*\right)$ *such that either* $U^1(s_1 + \hat{X} (s_1) ; \theta) > U^1 (c'; \theta)$ *or* $\hat{X}_2 (s_1) < \hat{X}_2^* (s_1^*)$.

Proof of Claim 3. To show this, we define $\tilde{X}^*$—a perturbation of $\hat{X}^*$. Define $\tilde{x}^*$ by $\tilde{x}_2^* < \hat{X}_2^* (s_1^*)$, and $U^1 (s_1^* + \tilde{x}^*; \theta) = U^1 \left(s_1^* + \hat{X}^* (s_1^*) ; \theta\right)$; it follows from Lemma 2 that $U^2 (s_1^* + \tilde{x}^*; \theta) < U^2 \left(s_1^* + \hat{X}^* (s_1^*) ; \theta\right)$. Given $\tilde{x}^*$, define $\tilde{X}^*$ in a parallel way to $\hat{X}^*$. That is, let

$$\tilde{s}_1^{**} = \sup \left\{ s_1 \in [0, s_1^*] : U^1 (s_1 + \tilde{x}^*; \theta) > U^1 (c; \theta) \right\},$$

with $\tilde{s}_1^{**} = s_1^*$ if the above set is empty. Then, for every $s_1 \in [\tilde{s}_1^{**}, s_1^*]$, let $\tilde{X}^*(s_1) = \tilde{x}^*$ and for every $s_1 \in [0, \tilde{s}_1^{**}]$, define $\tilde{X}^*(s_1)$ by the differential equations

$$d\tilde{X}_2^* = \frac{U_1^1(s_1 + \tilde{X}^*; \theta) - U_2^1\left(s_1 + \tilde{X}^*; \theta\right)}{U_3^1\left(s_1 + \tilde{X}^*; \theta\right)\left(\frac{U_2^1\left(s_1 + \tilde{X}^*; \theta\right)}{U_3^1\left(s_1 + \tilde{X}^*; \theta\right)} - \frac{U_2^2\left(s_1 + \tilde{X}^*; \theta\right)}{U_3^2\left(s_1 + \tilde{X}^*; \theta\right)}\right)} ds_1$$

$$d\hat{X}_3^* = -\frac{U_2^2\left(s_1 + \tilde{X}^*; \theta\right)}{U_3^2\left(s_1 + \tilde{X}^*; \theta\right)} d\hat{X}_2^*.$$

As long as $\tilde{x}^*$ is chosen close enough to $\hat{X}^*(s_1^*)$, $U^2(s_1^+ + \hat{X}_2(s_1^+); \theta) < U^2\left(s_1^+ + \tilde{X}_2^*(s_1^+); \theta\right)$. By the definition of $\tilde{x}^*$, $U^2(s_1^* + \hat{X}(s_1^*); \theta) > U^2\left(s_1^* + \tilde{X}^*(s_1^*); \theta\right)$. Since $\hat{X}$ is continuous at all but at most a finite number of points, it then follows that there exists some $s_1 \in (s_1^+, s_1^*)$ such that $U^2(s_1 + \hat{X}(s_1); \theta) > U^2\left(s_1 + \tilde{X}^*(s_1); \theta\right)$ and $\frac{d}{ds_1}U^2(s_1 + \hat{X}(s_1); \theta) \geq \frac{d}{ds_1}U^2\left(s_1 + \tilde{X}^*(s_1); \theta\right)$ (differentiability follows from Lemma 6 and from continuity of $\hat{X}$ and $\tilde{X}^*$). From Lemma 6, $\frac{d}{ds_1}U^2(s_1 + \hat{X}(s_1); \theta) = U_2^2(s_1 + \hat{X}(s_1); \theta)$ and $\frac{d}{ds_1}U^2\left(s_1 + \tilde{X}^*(s_1); \theta\right) = U_2^2\left(s_1 + \tilde{X}^*(s_1); \theta\right)$; therefore $\hat{X}_2(s_1) \leq \tilde{X}_2^*(s_1)$ and $\hat{X}_3(s_1) > \tilde{X}_3^*(s_1)$. Lemma 2 then implies that $U^1(s_1 + \hat{X}(s_1); \theta) > U^1\left(s_1 + \tilde{X}^*(s_1); \theta\right)$.

If $U^1\left(s_1 + \tilde{X}^*(s_1); \theta\right) = U^1(c'; \theta)$, the proof is complete. If $U^1\left(s_1 + \tilde{X}^*(s_1); \theta\right) < U^1(c'; \theta)$ (the opposite inequality is ruled out by the definition of $\tilde{X}^*$) then, by the definition of $\tilde{X}^*$, $\tilde{X}^*(s_1) = \tilde{x}^*$ and $\hat{X}_2(s_1) \leq \tilde{X}_2^*(s_1) = \tilde{x}_2^* < \hat{X}_2^*(s_1^*)$.

Final step of proof. It follows from Claims 1-3 that either $U^1(s_1 + \hat{X}(s_1); \theta) > U^1(c'; \theta)$ or $\hat{X}_2(s_1) < \hat{X}_2^*(s_1^*)$. The former contradicts our supposition that $\hat{X}$ is part of a commitment contract and the latter contradicts Proposition 4.

Proof of Lemma 7. Throughout, we use $s_2(s_1) \equiv \arg\max_{s_2 \geq 0} U^2(s_1 + c' - s_2; \theta)$. We start by establishing:

Claim $s_1^* \leq c_1' - c_1$.

Proof of Claim. From Lemma 1, $c_1' > c_1$. Recall that s_1^* is the supremum value of s_1 such that

$$\begin{aligned} & u_1(c_1' - s_1; \theta) + \beta u_2(s_1 + c_2' - s_2(s_1); \theta) + \beta u_3(s_2(s_1) + c_3'; \theta) \\ > \; & u_1(c_1; \theta) + \beta u_2(c_2; \theta) + \beta u_3(c_3; \theta). \end{aligned} \tag{A-3}$$

Using the definition of c as self-0's most preferred consumption in state θ, along with $c_1+c_2+c_3 = c_1'+c_2'+c_3'$, for any $\tilde{s}_1$, s_2, at $\beta=1$,

$$\begin{aligned} & u_1(c_1;\theta) - u_1\left(c_1' - \left(c_1' - c_1 + \tilde{s}_1\right);\theta\right) \\ \geq \; & \beta u_2\left(c_1' - c_1 + \tilde{s}_1 + c_2' - s_2;\theta\right) + \beta u_3\left(c_3' + s_2;\theta\right) - \beta u_2(c_2;\theta) - \beta u_3(c_3;\theta). \end{aligned} \tag{A-4}$$

Consequently, (A-4) holds strictly if $\tilde{s}_1 > 0$ and $\beta < 1$. It follows that (A-3) cannot hold for any $s_1 > c_1' - c_1$. Hence $s_1^* \leq c_1' - c_1$, completing the proof of the claim.

When $c_1' + c_2' < c_1 + c_2$, the claim implies that

$$u_2'\left(s_1^* + c_2';\theta\right) \geq u_2'\left(c_1' - c_1 + c_2';\theta\right) > u_2'(c_2;\theta) = u_3'(c_3;\theta),$$

where the final equality follows from the definition of c as self 0's most preferred consumption. Moreover, $c_1' + c_2' < c_1 + c_2$ is equivalent to $c_3 < c_3'$, and so

$$u_2'\left(s_1^* + c_2';\theta\right) > u_3'\left(c_3';\theta\right) = u_3'\left(c_3';\theta'\right) = u_2'\left(c_2';\theta'\right),$$

where the equalities follow from the assumption that that the shock has no effect on date 3 preferences, and from the definition of c'. Since, moreover, $u_3'(c_3';\theta) = u_3'\left(c_3';\theta'\right)$, we have shown that

$$\frac{u_2'\left(s_1^* + c_2';\theta\right)}{\beta u_3'\left(c_3';\theta\right)} > \frac{u_2'\left(c_2';\theta'\right)}{\beta u_3'\left(c_3';\theta'\right)},$$

which by Lemma 3 completes the proof for the case $c_1' + c_2' < c_1 + c_2$.

The remainder of the proof deals with the case in which $c_1' + c_2' = c_1 + c_2$ (and so $c_3' = c_3$). If in state θ self 2 anticipates consumption $(c_2 - c_2') + c_2' = c_2$ at date 2 and $c_3' = c_3$, he does not save (by the definition of c). Hence $s_2(s_1 = c_2 - c_2') = 0$. Since $c_2 - c_2' = c_1' - c_1$, it follows that (A-3) holds with equality at $s_1 = c_2 - c_2'$. Hence (combined with the above claim) $s_1^* = c_2 - c_2'$.

For additive shocks, observe that by the definition of c and c', along with $c_3 = c_3'$,

$$u_2'(c_2 + \theta_2) = u_3'(c_3 + \theta_3) = u_3'\left(c_3' + \theta_3'\right) = u_2'\left(c_2' + \theta_2'\right),$$

and hence $c_2 + \theta_2 = c_2' + \theta_2'$. Consequently, $s_1^* + \theta_2 = \theta_2'$, and so self 2 has exactly the same preferences in state θ with savings s_1^* and state θ' with no savings. Hence SPR is satisfied.

Finally, for multiplicative shocks, define the state θ' indifference curve through c' by $U^2\left(x;\theta'\right) = U^2\left(c';\theta'\right)$. We show that for $x_2 \geq c_2'$ this indifference curve lies above the state θ savings s_1^* indifference curve through c', i.e., $U^2\left(s_1^* + x;\theta\right) \geq U^2\left(s_1^* + c';\theta\right)$. Substituting out date 3, we must show that

$$u_2\left(s_1^* + x_2;\theta\right) - u_2\left(x_2;\theta'\right) \geq u_2\left(s_1^* + c_2';\theta\right) - u_2\left(c_2';\theta'\right). \tag{A-5}$$

The LHS of (A-5) is zero at $x_2 = c_2'$; moreover, its deriviative with respect to x_2 is also zero at $x_2 = c_2'$ since (using $c_3 = c_3'$)

$$u_2'\left(s_1^* + c_2';\theta\right) = u_2'\left(c_2;\theta\right) = u_3'\left(c_3;\theta\right) = u_3'\left(c_3';\theta'\right) = u_2\left(c_2';\theta'\right). \tag{A-6}$$

Finally, the second derivative of the LHS of (A-5) can be written as

$$\left|\frac{u_2''\left(x_2;\theta'\right)}{u_2'\left(x_2;\theta'\right)}\right| u_2'\left(x_2;\theta'\right) - \left|\frac{u_2''\left(s_1^* + x_2;\theta\right)}{u_2'\left(s_1^* + x_2;\theta\right)}\right| u_2'\left(s_1^* + x_2;\theta\right),$$

which is strictly positive at $x_2 = c_2'$ by (A-6), multiplicative shocks, and DARA.[31] By SCB, this completes the proof.

Proof of Lemma 8. When shocks are additive, $c_t = \frac{W+\theta_1+\theta_2+\theta_3}{3} - \theta_t$ and $c_t' = \frac{W+\theta_1'+\theta_2'+\theta_3'}{3} - \theta_t'$. Hence $c_3' < c_3$ is equivalent to $\theta_1' + \theta_2' < \theta_1 + \theta_2$ (recall both here and throughout the proof that we assume $\theta_3' = \theta_3$).

When shocks are additive, indifference curves cross at most once (even with different saving levels), so SPR fails if and only if $u'\left(s_1^* + c_2' + \theta_2\right) < u'\left(c_2' + \theta_2'\right)$, that is, if and only if $s_1^* + \theta_2 > \theta_2'$.

Recall that s_1^* is defined by

$$s_1^* \equiv \sup\left\{s_1 : U^1\left(s_1 + c' - s_2;\theta\right) > U^1\left(c;\theta\right) \text{ where } s_2 \in \arg\max_{\tilde{s}_2 \geq 0} U^2\left(s_1 + c' - \tilde{s}_2;\theta\right)\right\}.$$

The existence of a commitment problem guarantees that $s_1^* > 0$. However, with the extra stipulation that $s_1^* = 0$ if the above set is empty, then s_1^* is well-defined regardless of whether or not there

[31] For CARA preferences, multiplicative shocks are the same as additive shocks: see above.

is a commitment problem.

Claim If $s_1^* > 0$ and $s_1^* + \theta_2 = \theta_2'$ then $\frac{ds_1^*}{d\theta_2} > -1$.

Proof of Claim. First, note that if in state θ self 2 inherits savings $s_1^* = \theta_2' - \theta_2$ and the consumption plan c', he does not save, since he strictly discounts consumption at date 3 and $s_1^* + c_2' + \theta_2 = c_2' + \theta_2' = c_3' + \theta_3' = c_3' + \theta_3$. Second, note that $\beta u'(s_1^* + c_2' + \theta_2) < u'(-s_1^* + c_1' + \theta_1)$, since otherwise the definition of s_1^* is contradicted.

Given these two observations, and substituting for c and c', it follows that s_1^*, θ and θ' satisfy $s_1^* = \theta_2' - \theta_2$,

$$\begin{aligned} & u\left(-s_1^* + \frac{W + \theta_1' + \theta_2' + \theta_3'}{3} - \theta_1' + \theta_1\right) \\ & + \beta u\left(s_1^* + \frac{W + \theta_1' + \theta_2' + \theta_3'}{3} - \theta_2' + \theta_2\right) + \beta u\left(\frac{W + \theta_1' + \theta_2' + \theta_3'}{3}\right) \\ = \ & u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3}\right) + \beta u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3}\right) + \beta u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3}\right) \end{aligned} \quad \text{(A-7)}$$

and

$$\beta u'\left(s_1^* + \frac{W + \theta_1' + \theta_2' + \theta_3'}{3} - \theta_2' + \theta_2\right) < u'\left(-s_1^* + \frac{W + \theta_1' + \theta_2' + \theta_3'}{3} - \theta_1' + \theta_1\right).$$

Define $Y = \theta_1' + \theta_2' - \theta_1 - \theta_2$. Then substituting in for Y and $s_1^* = \theta_2' - \theta_2$ implies that

$$\begin{aligned} & u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3} + \frac{Y}{3} - Y\right) + \beta u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3} + \frac{Y}{3}\right) + \beta u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3} + \frac{Y}{3}\right) \\ = \ & u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3}\right) + \beta u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3}\right) + \beta u\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3}\right) \end{aligned}$$

and

$$\beta u'\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3} + \frac{Y}{3}\right) < u'\left(\frac{W + \theta_1 + \theta_2 + \theta_3}{3} + \frac{Y}{3} - Y\right).$$

The only value of Y that satisfies both equations is $Y = 0$.

So far we have shown that $s_1^* > 0$ and $s_1^* + \theta_2 = \theta_2'$ together imply $\theta_1' + \theta_2' = \theta_1 + \theta_2$. Differentiation of (A-7) and substitution of $s_1^* + \theta_2 = \theta_2'$ and $\theta_1' + \theta_2' = \theta_1 + \theta_2$ imply $-\frac{ds_1^*}{d\theta_2} + \beta\left(\frac{ds_1^*}{d\theta_2} + 1\right) = \frac{1+2\beta}{3}$, or equivalently, $\frac{ds_1^*}{d\theta_2} + \frac{1}{3} = 0$, completing the proof of the claim.

Now, suppose that contrary to the claimed result, there exist shocks, $\bar{\theta}$ and $\bar{\theta}'$ say, such that there is a commitment problem, $\bar{\theta}_1' + \bar{\theta}_2' < \bar{\theta}_1 + \bar{\theta}_2$, and $s_1^* + \bar{\theta}_2 \leq \bar{\theta}_2'$. Because there is a commitment problem, $s_1^* > 0$. Hence $\bar{\theta}_1' - \bar{\theta}_1 < \bar{\theta}_2 - \bar{\theta}_2' \leq -s_1^* < 0$.

Next, note that if $\bar{\theta}_2$ is replaced with $\tilde{\theta}_2 < \bar{\theta}_2$ such that $\bar{\theta}_1 + \tilde{\theta}_2 = \bar{\theta}_1' + \bar{\theta}_2'$, then from the proof of Lemma 7, $s_1^* = \bar{\theta}_2' - \tilde{\theta}_2 > 0$. (Note that $\bar{\theta}_1' < \bar{\theta}_1$ is used for this inference, since it implies $c_1' > c_1$).

The savings level s_1^* is continuous as a function of θ_2. Consequently, there exists some $\theta_2 \in \left[\tilde{\theta}_2, \bar{\theta}_2\right]$ such that $s_1^* + \theta_2 = \bar{\theta}_2'$ and $\frac{ds_1^*}{d\theta_2} + 1 \leq 0$. Note that at this θ_2, $s_1^* = \bar{\theta}_2' - \theta_2 > 0$. But this contradicts the above claim, completing the proof.