

Long-Term Contracting with Time-Inconsistent Agents*

Daniel Gottlieb[†] and Xingtan Zhang[‡]

October 10, 2017

Abstract

In many contractual relationships, consumers are allowed to terminate agreements at will. We study how removing commitment power from consumers affects equilibrium contracts and welfare when consumers have self-control problems. We show that removing commitment power is welfare improving when consumers are sufficiently dynamically inconsistent. Controlling for impatience, it is easier to sustain long-term contracts when consumers are time-inconsistent than when they are dynamically consistent. Moreover, because naive consumers overestimate the surplus from keeping a long-term contract, it is easier to sustain long-term contracts with them than with sophisticates. In fact, we show that the welfare loss from time-inconsistency of naifs vanishes as the number of periods grows (with or without commitment). As a result, naive agents get a higher welfare than sophisticates if the contracting horizon is large. We also show that limiting the fees that companies can charge actually hurts naive consumers.

JEL: D03, D81, D86

*Preliminary version; comments are especially welcome. We thank Marina Halac, George Mailath, Philipp Strack, Jeremy Tobacman, and Edward Van Wesep for helpful comments.

[†]Washington University in St. Louis, dgottlieb@wustl.edu

[‡]University of Colorado at Boulder, xingtan.zhang@colorado.edu.

1 Introduction

In many contractual relationships, consumers are allowed to terminate the agreements they have with a firm and start new agreements with another firm, but firms are contractually forced to serve consumers for the duration of the agreement. For example, in mortgages and other credit markets, borrowers can prepay their debt but debtors cannot force them to repay before the contract is due. Similarly, in long-term insurance policies (such as life insurance, long-term care insurance, or annuities), policyholders are allowed to cancel their policies at all times but firms cannot drop them.

Regulations that allow consumers to terminate agreements at will are often motivated by attempts to protect them. However, in standard models with rational consumers, removing their ability to commit to long-term contracts can only hurt them. Thus, either these regulations are misguided or they are aimed at consumers who are not perfectly rational. In this paper, we study the effect of allowing consumers to terminate long-term agreements when consumers are not perfectly rational, but, instead, have self-control problems.

Several papers have shown that firms can mitigate lack of commitment on the customer side by offering front-loaded payment schedules, which lock (rational) consumers into their contracts.¹ Similarly, several researchers have shown that mortgages are front loaded to mitigate prepayment risk.² For example, Hendel and Lizzeri (2003) theoretically and empirically examine how life insurers mitigate reclassification risk by offering front-loaded policies. More recently, Handel et al. (2017) show that front-loaded long-term health insurance contracts can produce large welfare gains by insuring policyholders against reclassification risk. But, while this literature assumes that individuals are time consistent, in practice many individuals are present biased.³ In particular, there is evidence that present bias is an important feature of mortgage markets.⁴ Therefore, it is

¹This literature originates with Harris and Holmstrom (1982) who present a theory of wage rigidity based on the assumption that firms can make binding contracts with workers but workers are always allowed to switch to better jobs. See also Dionne and Doherty (1994), Pauly et al. (1995), Cochrane (1995), and Krueger and Uhlig (2006).

²See Brueckner (1994) and Stanton and Wallace (1998).

³See Laibson (1997), O'Donoghue and Rabin (1999, 2001), and Augenblick et al. (2015)

⁴Schlafmann (2016), for example, empirically studies self control in mortgage markets and shows that requiring higher down payments and restricting prepayment can help customers. Similarly, Ghent (2011) argues that providing

important to understand how the ability to terminate long-term agreements affects contracts and welfare when consumers are present biased.

Because present-biased individuals are tempted to over-consume, one might think that this bias would exacerbate the problem of lack of commitment. However, we show that the opposite is true: controlling for impatience, it is easier to sustain long-term contracts when consumers are time-inconsistent than when they are time-consistent. This is because, when considering a time-inconsistent and a time-consistent agents with the same “average impatience,” the time-consistent agent discounts periods further in the future by more than a time-inconsistent agent. As a result, the time-inconsistent agent is less hurt by front-loading payments from the periods sufficiently far in the future, which helps supporting long-term contracts.

Moreover, it is easier to sustain long-term contracts when consumers are not fully aware of their bias (i.e., they are naive) than when they are aware of their bias (i.e., they are sophisticated). Naive consumers think that they will not be as tempted to lapse in the future as they are today, thereby overestimating the benefit from keeping a long-term contract. This makes them more reluctant to lapse than sophisticates, who fully understand that their temptation to lapse will persist. Therefore, somewhat paradoxically, consumers who are unaware of their self control problems are more likely to keep their contracts. In particular, we show that as the number of periods grows, the welfare cost from time inconsistency converges to zero for naifs but not for sophisticates.

Comparing the outcomes with one- and two-sided commitment, we find that sufficiently time-inconsistent consumers are hurt by having commitment power. This is because being able to commit to long-term contracts allows them to over-borrow. When consumers can walk away from contracts, firms do not allow them to borrow, which is welfare improving if over-borrowing is sufficiently problematic. While this result is in line with regulation, which often mentions consumer

access to mortgages with lower initial payments decreases savings due to time inconsistency. Gathergood and Weber (2017) study mortgage choices in the UK and find that present bias substantially raises the likelihood of choosing alternative mortgage products. And Atlas et al. (2017) study data from a nationally-representative sample of US households and find that present-biased individuals are more likely to choose mortgages with lower up-front costs. They also find that present-biased individuals are less likely to refinance their mortgages, which is consistent with our results in the case of partial naiveté. See also Bar-Gill (2009) for a description of behavioral aspects of the subprime mortgage market.

protection as the reason for allowing them to terminate agreements at will, it contrasts with the idea that the provision of commitment devices is necessarily welfare improving.⁵

Next, we evaluate the welfare consequences of imposing an upper bound on the fees that can be charged in each period. This type of policy has been popular among regulators as a way to protect consumers who would otherwise be tempted to over-borrow. For example, the Credit CARD Act of 2009 limits the amount of interest and fees that credit card companies can charge. Similarly, Dodd-Frank (Title XIV) has many provisions that restrict penalties or fees that can be charged in mortgage contracts. In insurance, nonforfeiture laws specify minimum payments that must be made to customers who surrender their permanent life insurance policies or annuities.⁶ And, in long-term care insurance, the Department of Financial Services specifies minimum benefits that must be provided as well as minimum cash benefits that must be paid to those who lapse.

We show that, when consumers are naive, limiting the fee that firms can charge is never welfare improving. The intuition is that, by preventing back-loaded fees, one-sided commitment limits the amount of over-borrowing that customers may engage in. When interest rates are low so customers would like to borrow, one-sided commitment already prevents them from obtaining long-term contracts and imposing a maximum fee does not affect the equilibrium. The only case where a maximum fee can affect the equilibrium is when interest rates are high so that, in equilibrium, customers would like to save. But, in this case, such a policy reduces savings, moving the equilibrium further away from the optimum.⁷

It is often suggested that educating behavioral agents about their biases would increase their welfare.⁸ In our model, however, making individuals aware of their dynamic consistency increases their temptation to lapse, reducing savings and welfare. Therefore, in our model, educating behavioral agents about their naiveté is actually harmful!

⁵See Bryan et al. (2010) for a survey of the literature on commitment devices.

⁶Each states follows slight variations of the general guidelines from the National Association of Insurance Commissioners (NAIC).

⁷One-sided commitment is key for this result. With two-sided commitment, imposing a maximum fee sometimes helps naive consumers (Heidhues and Kőszegi, 2010).

⁸See, for example, Bar-Gill (2009) for a discussion of this type of policy in the context of subprime mortgage markets.

Our paper fits into a recent literature on contracting with behavioral agents, summarized in Kőszegi (2014) and Grubb (2015). In particular, we build on the credit card model of Heidhues and Kőszegi (2010) by introducing the assumption that consumers cannot commit to long-term contracts, considering more than two consumption periods, and allowing for uncertainty.⁹

Our paper is also related to a literature that studies commitment contracts with time-inconsistent agents (c.f. Amador et al. (2006); Halac and Yared (2014); Galberti (2015); Bond and Sigurdsson (2017)). This literature studies the trade-off between commitment and flexibility: agents are able to commit to any contract but, because they face an unverifiable taste shock, their contract must induce them to report their shock correctly (i.e., they value the flexibility to adjust to different taste shocks). There are two differences between these papers and ours. First, we study a different incentive aspect: the agent's incentive to lapse and re-contract with other firms. Second, these papers focus on sophisticated agents, whereas we also consider partially naive agents.

Finally, our paper belongs to a literature on dynamic risk-sharing, where we are most closely related to Krueger and Uhlig (2006) and Makarov and Plantin (2013). Krueger and Uhlig (2006) show that, in the presence of one-sided commitment, risk-sharing contracts can arise in equilibrium because front-loaded payments effectively lock in consumers with firms. Makarov and Plantin (2013) introduce the assumption that households can terminate loan contracts at any point in time and show that the model accounts for several features of the subprime market. The main difference between these models and ours is that we assume that consumers are dynamically inconsistent.¹⁰

The paper proceeds as follows. We first consider a simple version of the model, in which the agent has a constant deterministic income. We start by describing the framework (Subsection 2.1) and the equilibrium concept (2.2). We then derive conditions for the market to break down (2.3) and for the agent's welfare to be higher without commitment power (2.4). Then, we show that

⁹Several other papers also study contracting with behavioral agents. DellaVigna and Malmendier (2004) study a two-period model in which a firm sells an indivisible good to time-inconsistent customers. Eliaz and Spiegler (2006) study a two-period model in which a monopolist contracts with consumers with different abilities to predict their future tastes. Grubb (2009) studies a model in which consumers overestimate the precision of their demand forecasts. Parsons and Van Wesep (2013) study the equilibrium timing of pay for firms that hire workers with self-control problems.

¹⁰Our paper is also related to Bernheim et al. (2015), who show that poverty traps naturally emerge in intertemporal allocation problems with credit constraints and sophisticated time-inconsistent individuals.

the welfare loss of naive agents vanishes as the contracting horizon grows (2.5) and establish that imposing a maximum fee is welfare decreasing when agents are naive (2.6). In Section 3, we turn to arbitrary (possibly stochastic) income paths, generalizing the main results obtained in the case of a constant income. Section 4 concludes. Proofs are presented in the appendix.

2 Model with Constant Income

2.1 Environment

We study a market in which at least two principals (“firms”) offer contracts to an agent (“consumer”). Time is discrete and finite: $t \in \{1, \dots, T\}$ with $T \geq 3$. The agent has a constant deterministic endowment of w in each period. This assumption is only made to simplify the exposition: we allow for arbitrary (possibly stochastic) endowment paths in Section 3.

Principals are risk neutral can freely save or borrow at an interest rate $R \geq 1$. A principal who collects a stream of payments $\{p_t\}_{t \geq s}$ starting at time s gets payoff

$$\sum_{t=s}^T \frac{p_t}{R^{t-s}}.$$

The agent is risk averse and can only transfer consumption across time by transacting with a principal. Following O’Donoghue and Rabin (1999), we consider three possible preferences for the agent: time consistent, (partially) naive, and sophisticated. All of them have instantaneous utility functions $u : \mathbb{R}_+ \rightarrow \mathbb{R}$, where u is strictly increasing, strictly concave, twice continuously differentiable, and bounded from below. At time s , the agent evaluates the stream of consumption $\{c_t\}_{t \geq s}$ according to

$$u(c_s) + \beta \sum_{t>s} \delta^{t-s} u(c_t). \tag{1}$$

The agent is *time consistent* if $\beta = 1$ and *time inconsistent* if $\beta < 1$. A time-inconsistent agent can be either naive or sophisticated. A *naif* has true time-consistency parameter β but believes that, in the future, he will behave like an agent with time-consistency parameter $\hat{\beta} \in (\beta, 1]$. A *sophisticate*

perfectly knows his time-consistency parameter: $\hat{\beta} = \beta$.

Competition is modeled by allowing each principal to offer a contract to the agent in each period. Principals can always commit to long-term contracts, meaning that, after a contract has been accepted, the principal must honor its terms for as long as the agent wishes to keep it. We consider two different scenarios. In the first scenario, the agent is also able to commit to long-term contracts (“two-sided commitment”). In the second scenario, the agent is allowed to drop his contract in any period and replace it with a new one (“one-sided commitment”). As usual, we assume that there are no direct costs of signing contracts, so the agent’s outside option is endogenously determined by the best possible contract that he can obtain. To simplify the exposition, we will use an equilibrium notion that suppresses explicit strategic considerations. In the appendix, we show that this notion is equivalent to the Subgame Perfect Nash Equilibrium of the game in which firms simultaneously offer contracts to the agent.

Because of the Bertrand nature of competition between firms, a competitive equilibrium will maximize the agent’s perceived utility subject to subject to zero profits as well as some additional constraints. The first constraint requires the agent’s lapsing decision at each period to maximize the perceived utility of that period’s self. For the purposes of characterizing the equilibrium consumption and welfare, there is no loss of generality in restricting attention to contracts in which the agent never lapses (“renegotiation proofness”). To see this, consider a contract in which the agent lapses in some period, replacing it with a contract from another principal. Since the other principal cannot lose money by offering this new contract, the old principal could have offered a contract that substituted the terms of the old contract from this period on with the terms of the new contract, and the agent would have accepted to remain with the old principal. Therefore, for the purpose of understanding the consumption profiles that can be supported in equilibrium, we can impose a *non-lapsing constraint* that requires contracts to be renegotiation proof. This constraint requires that the agent’s outside option does not exceed the value from staying with the current principal, where the outside option is the best possible contract that other principals can provide.¹¹

¹¹When the non-lapsing constraints bind, there are also equilibria in which the agent lapses and recontracts with another principal. But these equilibria are equivalent in the sense that the agent gets the same consumption stream and

The second constraint arises from the fact that a naive agent disagrees with the principal about the preferences of his future selves. Then, instead of offering a fixed consumption stream, a principal can obtain higher profits by offering a contract with different options: a *baseline* option corresponding to what the agent thinks that his future selves will choose (B), and an *alternative* option corresponding to what the principal knows that they will choose (A). More precisely, as in Heidhues and Kőszegi (2010), there is no loss of generality in restricting attention to contracts in which, after each non-terminal history, at least one party believes that each of the terms offered will be chosen (“non-redundancy”).

When the agent is either time-consistent or sophisticated, his beliefs about which contracts his future actions coincide with the beliefs of the principal. Then, non-redundancy means that we can restrict attention to contracts that specify a single consumption stream. On the other hand, because a naive agent may disagree with the principal regarding which contract term he will choose, non-redundancy means that the agent is given two options after any non-terminal history: a baseline and an alternative option. Those contracts must be such that, at each point, the naif must prefer to choose the alternative option (“incentive compatibility”) but, when predicting his future choices, he must believe that he will choose the baseline option (“perceived choice”).

A time- t history h^t is a list of possible options chosen by the agent up to time t : $h^1 = \emptyset$, $h^2 \in \{A, B\}$, $h^3 \in \{AA, AB, BA, BB\}$, etc. Notice that $|h^t| = 2^{t-1}$ for any $t < T$. However, since there are no actions after the last period, there is no space for disagreement at $t = T$, so that $h^T = h^{T-1}$. Figure 1 depicts the possible histories if there are four periods and the agent is naive.

On the equilibrium path, the agent always picks option A , but believes that, in all future periods, he will pick B . For example, in the credit card model studied in Heidhues and Kőszegi (2010), the borrower is given two repayment options: in the baseline option, he repays his debt immediately; in the alternative option, he pays a fee to postpone the payment. We refer to $c^E = (c_1, c_2(A), c_3(AA), \dots, c_T(A \dots A))$ as the naive agent’s consumption on the equilibrium path and to $c_t^E = c_t(A \dots A)$ as his time- t consumption on the equilibrium path.

all principals make the same profits.

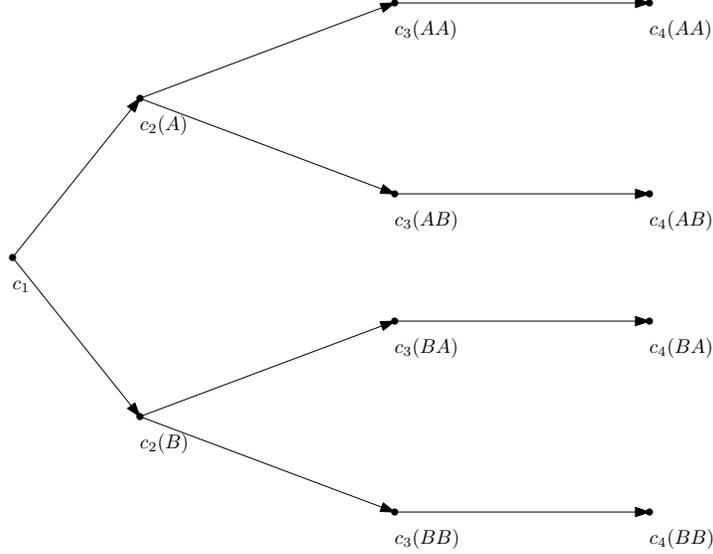


Figure 1: **Contracting with a naive agent.** The figure represents possible histories of a contract offered to a naive agent when $T = 4$. The agent thinks he will choose the baseline (B) option in each node, but he always ends up choosing the alternative (A) option. In the initial period, the agent thinks he will receive the consumption stream $(c_1, c_2(B), c_3(BB), c_4(BB))$. In period 2, however, he ends up deviating and obtaining $c_2(A)$, while thinking that he will receive $c_3(AB)$ and $c_4(AB)$ in periods 3 and 4. Then, he deviates again in period 3, obtaining consumption $c_3(AA)$ and $c_4(AA)$ in periods 3 and 4.

A *competitive equilibrium with two-sided commitment* is a set of non-redundant contracts that maximizes the agent's expected utility in period 1 subject to the zero profits, incentive compatibility, and perceived choice constraints. A *competitive equilibrium with one-sided commitment* is a set of non-redundant, renegotiation-proof contracts that maximizes the agent's expected utility in period 1 subject to the zero profits, incentive compatibility, and perceived choice constraints.

We say that the *market breaks down* if the agent gets the same consumption as the endowment in all periods along the equilibrium path. If the market does not break down, we say that the equilibrium features a *long-term contract*.

Following most of the literature (c.f. DellaVigna and Malmendier (2004); O'Donoghue and Rabin (1999, 2001)), we take the agent's long-run preferences as a measure of welfare. Formally, for a consumption stream $\{c_t\}$, we let

$$W(\{c_t\}) \equiv \sum_{t=1}^T \delta^{t-1} u(c_t). \quad (2)$$

denote the welfare associated with it.

2.2 Equilibrium

In this section, we study the equilibria with one-sided commitment for a time-consistent agent, a sophisticated agent, and a naive agent separately.

2.2.1 Time-Consistent Agent

We start with the benchmark case of a time-consistent agent. Let V_τ denote the agent's utility from the best possible contract that can be obtained at time τ , defined recursively as:

$$V_\tau := \max_{(c_\tau, \dots, c_T)} \sum_{t=\tau}^T \delta^{t-\tau} u(c_t), \quad (3)$$

subject to

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-\tau}} = \sum_{t=\tau}^T \frac{w}{R^{t-\tau}}, \quad (4)$$

$$\sum_{t=s}^T \delta^{t-s} u(c_t) \geq V_s, \quad \text{for } s = \tau + 1, \dots, T. \quad (5)$$

Equation (4) is the zero-profit condition, and (5) are the non-lapsing constraints.

Since, by definition, the equilibrium contract maximizes the agent's utility in period 1 subject to the zero-profit condition and the non-lapsing constraints, it solves the program above for $\tau = 1$. While one could use backward induction to determine the outside option in each future period and solve this program directly, it is helpful to reformulate the program as follows:

$$\max_{\{c_t\}} \sum_{t=1}^T \delta^{t-1} u(c_t), \quad (6)$$

subject to

$$\sum_{t=1}^T \frac{c_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}}, \quad (7)$$

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-\tau}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-\tau}}, \quad \tau = 2, \dots, T. \quad (8)$$

This program replaces the non-lapsing constraints by the requirement that, in each period, the present discounted value of future income must exceed the the present value of future consumption. In other words, discounted profit from every continuation contract cannot be positive.

In general, (8) is a relaxation of the non-lapsing constraints (5): if the continuation contract gave positive profits at some period, the agent would be able to increase his utility in that period by replacing it with another contract that gave zero profits. However, the two programs are equivalent if the agent is dynamically consistent. By zero profits, (8) requires payments to be front-loaded, in the sense that, at each point in time, the accumulated profits from the contract must be (weakly) positive. Suppose a solution to program (6)-(8) did not satisfy (5), so that there exists a continuation contract that gives zero profits while increasing the agent's continuation utility. Then, substituting the original continuation contract by this new one increases the agent's utility and gives non-negative profits at $t = 1$, which contradicts the optimality of the original contract.

Solving program (6)-(8), gives the following result:

Lemma 1. *Suppose the agent is time consistent. There exists a unique equilibrium. Moreover:*

1. *if $R \leq \frac{1}{\delta}$, the market breaks down;*
2. *if $R > \frac{1}{\delta}$, the equilibrium features a long-term contract and the agent's equilibrium consumption grows over time.*

The lemma shows that long-term contracts can be supported if and only if the interest rate is high enough that, in the absence of commitment issues, the consumer would choose to save ($R > \frac{1}{\delta}$). In this case, the agent initially consumes less than his income, effectively making irreversible up-front payments that prevent him from lapsing in the future. If the agent would

prefer to borrow at the prevailing interest rate ($R \leq \frac{1}{\delta}$), the market breaks down since he cannot commit not to drop any contract.

2.2.2 Sophisticated Agent

Next, consider the sophisticated agent. Let V_τ^S denote the agent's utility from the best possible contract that can be obtained at time τ , defined recursively as:

$$V_\tau^S := \max_{(c_\tau, \dots, c_T)} u(c_\tau) + \beta \sum_{t=\tau+1}^T \delta^{t-1} u(c_t), \quad (9)$$

subject to

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-1}} = \sum_{t=\tau}^T \frac{w}{R^{t-1}}, \quad (10)$$

$$u(c_s) + \beta \sum_{t=s+1}^T \delta^{t-s} u(c_t) \geq V_s^S, \forall s = \tau + 1, \dots, T. \quad (11)$$

The equilibrium contract solves the program above for $\tau = 1$. Because the agent is dynamically inconsistent, we cannot replace the non-lapsing constraints by the net present value condition (8), since his temptation to lapse depends not only on the amount of front-loaded resources but also on the time-inconsistency parameter β . To obtain the solution, we apply backward induction to calculate the outside option in each future period:

Lemma 2. *Suppose the agent is sophisticated. There exists a unique equilibrium. Moreover:*

1. *the market breaks down if and only if $R \leq r_T(\beta, \delta)$ for some $r_T(\beta, \delta) > \frac{1}{\delta}$,*
2. *$r_T(\beta, \delta)$ is decreasing in β and in T , and*
3. *$\lim_{\beta \nearrow 1} r_T(\beta, \delta) = \lim_{T \nearrow \infty} r_T(\beta, \delta) = \frac{1}{\delta}$.*

When the interest rate is low, the agent would like to borrow. However, because he cannot commit to repay his debt, no firm would accept to agree to lend him money. So, in equilibrium, the agent consumes his endowment in each period – i.e., the market breaks down. Claim 1 states that

the cutoff for the market to break down $r_T(\beta, \delta)$ is higher than the cutoff when there is no present bias $\frac{1}{\delta}$. We will return to this comparison in Section 2.3.¹²

Claims 2 and 3 examine how the market breakdown condition varies with the degree of time inconsistency and with the horizon of the contracting problem. Claim 2 states that market breaks down “less often” when the agent has a higher time-consistency parameter β and when there are more periods T . Intuitively, as β increases, the need for immediate gratification decreases, reducing the need to front-load payments in order to support long-term contracting. Moreover, when the market breaks down with T periods, the non-lapsing constraints (11) bind in each of the T periods, forcing sophisticates to consume their endowments. Since the reservation utility in the second period, V_2^S , is the same as the equilibrium program in an economy with $T - 1$ periods, the market must also break down in the $(T - 1)$ -period economy. In other words, with a longer horizon, principals have more instruments to try to prevent the market from breaking down.

Claim 3 states that, as the agent becomes close to time consistent or as the number of periods goes to infinity, the cutoff for the market to break down approaches the cutoff with a time-consistent agent (obtained in Lemma 1). Therefore, market breakdown conditions are approximately the same for an “almost time-consistent” sophisticate as for a time-consistent agent. They are also approximately the same for a time-consistent agent and for a sophisticate when the horizon is “sufficiently long.”

While Lemma 2 gives conditions under which the market breaks down, it says nothing about the equilibrium *consumption*. We now examine the equilibrium consumption for an agent who is “almost time-consistent” or has “sufficiently long horizons.”

Corollary 1. *Suppose the agent is sophisticated. The equilibrium consumption is a continuous function of $\beta \in [0, 1]$.*

¹²When the interest rate is high enough ($R > R_T^S(\beta, \delta)$ for some $R_T^S(\beta, \delta) > r_T(\beta, \delta)$), sophisticates find it optimal to save in every period. Then, no non-lapsing constraint binds and the equilibrium consumption is same as in a model with two-sided commitment. That is, although the current self would like to consume more than what was previously agreed, the irreversible up-front payment is large enough to ensure that they never find it beneficial to lapse. For intermediate interest rates ($r_T(\beta, \delta) < R < R_T^S(\beta, \delta)$), a time-consistent agent receives the same long-term contract as when there is full commitment but a sophisticated agent does not.

The corollary establishes that the equilibrium contract of a sophisticated agent is continuous in β , implying that an “almost time-consistent” agent gets approximately the same contract as a time-consistent agent does. This continuity result is not obvious: as we will see below, an “almost sophisticated naif” gets a very different contract than a sophisticate.

As the example below shows, the result for “sufficiently long horizons” is remarkably different than the one when the agent is “almost time-consistent.” Although in both cases the cutoff for the market to breakdown converges to the cutoff of a time-consistent agent, the difference between the equilibrium consumption of a sophisticated agent and a time-consistent agent does not vanish as the number of periods T grows. As a result, the welfare of sophisticate remains bounded away from the welfare of his long-term selves as the number of periods grows:

Example 1. Let $u(c) = 1 - \exp(-c)$. As we show in the appendix, there exists $\bar{R} > 1$ such that when $R > \bar{R}$, the equilibrium consumption of time-consistent and sophisticated agents equal:

$$c_t^C = w - A + (T - 1) \log(\delta R),$$

and

$$c_t^S(\beta) = \begin{cases} w - A - \log(\beta) \left(\frac{\frac{1}{R} + \dots + \frac{1}{R^{T-1}}}{1 + \frac{1}{R} + \dots + \frac{1}{R^{T-1}}} \right) + (T - 1) \log(\delta R) & \text{for } t = 1, \\ w - A + \log(\beta) \frac{1}{1 + \frac{1}{R} + \dots + \frac{1}{R^{T-1}}} + (T - 1) \log(\delta R) & \text{for } t > 1, \end{cases}$$

where $A \equiv \frac{\frac{\log(\delta R)}{R} + \dots + \frac{(T-1)\log(\delta R)}{R^{T-1}}}{1 + \frac{1}{R} + \dots + \frac{1}{R^{T-1}}}$. Notice that, while the difference in their equilibrium consumption vanishes as β approaches 1, it persists as the contracting horizon grows:

$$\lim_{T \rightarrow \infty} (c_t^C - c_t^S(\beta)) = \begin{cases} \log(\beta) \frac{1}{R} < 0 & \text{for } t = 1, \\ -\log(\beta) \frac{R-1}{R} > 0 & \text{for } t > 1. \end{cases}$$

2.2.3 Naive Agent

We now turn to the naive agent, who has a time-consistency parameter $\beta < 1$ but believes that, in the future, he will behave like an agent with parameter $\hat{\beta} \in (\beta, 1]$. Recall that, by non-redundancy,

the contract specifies, for each non-terminal history, a baseline option (what the agent thinks that he will choose), which we will index by B , and an alternative option (what the principal knows that the agent will choose), which we will index by A .

To understand the constraints of the equilibrium program, imagine a principal designing a contract to a naive agent. The contract must satisfy the following constraints. First, the agent must believe that he will choose the baseline option (B) in all future periods. This is the perceived choice constraint (“PCC”). Second, the agent must end up picking the alternative (A) instead of the baseline option. This is the incentive constraint (“IC”). And third, since the agent is not able to commit not to lapse, the contract must also satisfy the non-lapsing constraints.

Instead of solving the naive agent’s program directly, it is helpful to consider the contracting problem of a dynamically consistent agent who differs from the time-consistent agent from Subsection 2.2.1 in that he discounts consumption in the last period by an additional factor β . The equilibrium with this agent solves:

$$\max_{\{c_t\}} \sum_{t=1}^{T-1} \delta^{t-1} u(c_t) + \beta \delta^{T-1} u(c_T), \quad (12)$$

subject to

$$\sum_{t=1}^T \frac{c_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}}. \quad (13)$$

We refer to this program as the *auxiliary program with two-sided commitment*. With one-sided commitment, the equilibrium must also solve the non-lapsing constraints. Since this agent is dynamically consistent, as shown in Section 2.2.1, we can rewrite the non-lapsing constraints in terms of the expected net present value of consumption and income for all continuation contracts:

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-\tau}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-\tau}}, \quad \tau = 2, \dots, T. \quad (14)$$

We refer to this program as the *auxiliary program with one-sided commitment*.

Let c_N^{1E} and c_N^{2E} denote consumptions on the equilibrium path of a naive agent with one- and

two-sided commitment, respectively, and let c^{1A} and c^{2A} be solutions to these auxiliary programs. The following lemma establishes that these objects coincide:

Lemma 3. *Suppose the agent is naive. The consumption on the equilibrium path coincides with the solution of the auxiliary problem: $c_N^{1E} = c^{1A}$ and $c_N^{2E} = c^{2A}$.*

The auxiliary programs highlight the main source of distortion with a naive agent: underweighting the last period. With one-sided commitment, the auxiliary program is significantly more tractable than the original equilibrium program, since it does not require a backward induction algorithm and the constraints do not depend on the agent's preferences. In fact, both auxiliary programs are independent of the agent's perceived time-inconsistency parameter $\hat{\beta}$, implying that consumption on the equilibrium path does not depend on the agent's degree of naiveté.

Solving Program (12)-(14), we can characterize the equilibrium contract, obtaining the following result:

Lemma 4. *Suppose the agent is naive ($\beta < \hat{\beta} \leq 1$) and there is one-sided commitment. There exists an equilibrium. Moreover:*

1. *All equilibria have the same consumption on the equilibrium path,*
2. *Consumption on the equilibrium path is a continuous function of $\beta \in (0, 1]$, and*
3. *The market breaks down if and only if $R \leq \frac{1}{\delta}$.*

Lemma 4 shows that the condition for long-term contracting to be possible with a naive agent is the same as with a time-consistent agent. In particular, unlike in the case of sophisticates, the interest rates under which the market breaks down do not depend on the agent's time-consistency parameter β . We will study this in greater detail in the next section, when we compare conditions for the market to break down for different types of agents.

Since the constraints of the auxiliary program do not depend on β and $\hat{\beta}$, it is straightforward to derive comparative statics for naive consumers. Consumption in the last period is weakly increasing in the agent's time-consistency parameter β , and consumption in all other periods is weakly

decreasing in β . Moreover, since the objective function does not depend on the perceived time-consistency parameter $\hat{\beta}$, consumption on the equilibrium path does not depend on the agent's degree of the naiveté. Introducing any small amount of naiveté discontinuously shifts the equilibrium from the one described in Lemma 2 to the one described in Lemma 4.¹³ In particular, the equilibrium path consumption would remain unchanged if firms faced a population of agents with heterogeneous $\hat{\beta}$ parameters.

When $R > \frac{1}{\delta}$ (so that a long-term contract is provided), the equilibrium consumption is increasing over time for $t = 1, \dots, T - 1$. By the non-lapsing constraint, consumption in the last period cannot be lower than the endowment w . In some cases, it is equal to w , leaving the last-period self indifferent between lapsing or remaining with the original contract. When $R > \frac{1}{\beta\delta}$ even the last-period self strictly prefers not to lapse.

In each period, the current self believes that the next self will pick a lower immediate consumption in exchange for a higher consumption in the future. However, each future self ends up picking an option with a higher immediate consumption and a lower future consumption. That is, the agent's perceived consumption stream is more front-loaded than the actual consumption stream.

2.3 Market Breakdown

Fix a discount parameter δ . Combining the conditions for the market to break down, we find that, when $\frac{1}{\delta} < R < r_T(\beta, \delta)$, the equilibrium features long-term contracts if the agent is naive but the market breaks down if the agent is sophisticated. Outside of this interval, either both of them receive long-term contracts or the market breaks down for both. Therefore, *naiveté helps the provision of long-term contracts*. The intuition is as follows. Front-loaded payments are key to sustain long-term contracts. Sophisticates fully understand how front-loaded payments will hurt their future selves. Naive agents, however, believe that their future selves will be less hurt by front-

¹³The discontinuity of the equilibrium consumption in the agent's naiveté is not unique to our setting (c.f. Heidhues and Kőszegi (2010)).

loaded contracts than they actually will. Therefore, they overestimate the surplus from remaining with a contract, which makes them more willing to accept a long-term contract and to keep it.

Because the condition for long-term contracts to be provided is the same for a naive and for a time-consistent agent, one might be tempted to also conclude that it is easier to provide long-term contracts for a time-consistent agent than for a sophisticate. However, this interpretation is not warranted, as holding δ fixed also makes time-inconsistent agents more impatient than time-consistent agents (because of the additional discount factor β).

To allow us to compare time-consistent and time-inconsistent agents while holding their “average impatience” fixed, we introduce the notion of a weighted level of impatience. Formally, consider a vector of weights $\alpha = (\alpha_1, \dots, \alpha_T)$, where $\alpha_i > 0$ and $\sum \alpha_i = 1$. An agent’s α -weighted measure of impatience is:

$$\alpha_1 + \alpha_2\beta\delta + \dots + \alpha_T\beta\delta^{T-1}. \quad (15)$$

That is, if a time-consistent agent has discount parameter δ_C and a time-inconsistent agent has discount parameters (δ_I, β) , they have the same α -weighted impatience if:

$$\alpha_1 + \alpha_2\delta_C + \dots + \alpha_T\delta_C^{T-1} = \alpha_1 + \alpha_2\beta\delta_I + \dots + \alpha_T\beta\delta_I^{T-1}. \quad (16)$$

Intuitively, both agents discount the stream of utils α in the same way. Of course, they still discount other streams differently. One example of α -weighted impatience is the effective discount factor introduced by Chade et al. (2008), which corresponds to an α -weighted impatience with uniform weights ($\alpha_i = \frac{1}{T}$ for all i). Simple algebraic manipulations show that, for any fixed vector of weights, $\beta\delta_I < \delta_C$ and $\beta\delta_I^{N-1} > \delta_C^{N-1}$. That is, because both agents have the same average impatience, present-biased individuals discount the immediate future by more and later periods by less than time-consistent individuals.

We now compare when the market breaks down for each type of agents while fixing an α -weighted impatience. Recall the conditions for market breakdown: (1) $R \leq \frac{1}{\delta_C}$ for a time-

consistent agent; (2) $R \leq r_T(\beta, \delta_I)$ for a sophisticate; and (3) $R \leq \frac{1}{\delta_I}$ for a partial naif. Since $\frac{1}{\delta_I} < \frac{1}{\delta_C}$ and $\frac{1}{\delta_I} < r_T(\beta, \delta_I)$, it is easier to sustain long-term contracting with naifs than with both sophisticates and time-consistent consumers.

We claim that it is easier to sustain long-term contracting with sophisticates than with time-consistent agents. The reason is that, while present-biased individuals discount the immediate future by more than time-consistent individuals, they discount later periods by less (holding their weighted impatience constant), which makes them more willing to save further into the future. Therefore, sophisticates have a higher demand for instruments that cannot be liquidated in the immediate future, which relaxes the non-lapsing constraints.

Formally, recall that the equilibrium contract for a sophisticate maximizes the utility of the period-1 self subject to zero profits and non-lapsing constraints. Starting from $c_1 = c_2 = \dots = c_T = w$, suppose we shift consumption from period 1 to period T by $\epsilon > 0$: $c_1 = w - \epsilon$, $c_N = w + \epsilon R^{T-1}$. This transfer keeps the non-lapsing and zero profits constraints satisfied and changes the agent's utility by

$$(\beta \delta_I^{T-1} R^{T-1} - 1) u'(w) \epsilon. \quad (17)$$

For the market to break down, shifting consumption to the last period cannot increase the agent's utility, so we must have

$$\beta \delta_I^{T-1} R^{T-1} \leq 1. \quad (18)$$

Using the fact that the time-inconsistent agent discounts the last period by less ($\beta \delta_I^{T-1} \geq \delta_C^{T-1}$), we find that (18) implies $\delta_C R \leq 1$. That is, whenever the market breaks down for sophisticates, it also breaks down for time-consistent agents. We have, therefore, established the following results:

Proposition 1. *Fix a vector of weights α and consider time-consistent, sophisticated, and naive agents with the same α -weighted impatience.*

1. *If the market breaks down for naive agents, it breaks down for sophisticates.*
2. *If the market breaks down for sophisticates, it breaks down for time-consistent agents.*

The prediction from Proposition 1 is consistent with the results from Atlas et al. (2017), who find that present-biased individuals are less likely to refinance their mortgages than time-consistent ones.¹⁴

Notice that whether we control for impatience is key for the predictions of the model. In Lemma 2, we found that sophisticates were less likely than time-consistent agents to obtain long-term contracts with the same “long-term discount factor” δ . However, holding the long-term discount factor fixed conflates the effects of discounting and time inconsistency. Proposition 1 shows that, controlling for discounting, sophisticates are actually more likely to obtain long-term contracts.

2.4 Removing Commitment Power

As mentioned in the introduction, governments often regulate how easy it is for consumers to terminate agreements and switch firms. We now study how giving consumers the right to terminate agreements affects contracts and welfare when agents are time-inconsistent. Allowing agents to commit to long-term contracts removes the non-lapsing constraint, which makes borrowing possible. While allowing agents to borrow can only help time-consistent agents, it can hurt time-inconsistent agents, who are tempted to “over-borrow.” Hence, removing commitment power from time-inconsistent agents can increase their welfare.

Suppose the agent is naive. If the interest rate is sufficiently high, the non-lapsing constraints do not bind and the agent gets a long-term contract in which he saves, obtaining the same welfare with and without commitment power. On the other hand, if $R \leq \frac{1}{\delta}$, the market breaks down under one-sided commitment. In that case, we need to compare the welfare loss from not being able to borrow against the welfare gain from preventing time-inconsistent agents from “over-borrowing.” In particular, if the time inconsistency problem is severe enough, the agent is better off without commitment power.¹⁵

¹⁴Note that our results are true even though there are no immediate transaction costs in refinancing. Introducing those costs would further accentuate our results, since time-inconsistent agents are more averse to immediate costs.

¹⁵In our analysis, there is no contracting stage before any consumption is possible. Adding a contracting stage does

Proposition 2. 1. *Suppose the agent is naive. There exists $\underline{R} < \frac{1}{\delta}$ and $\bar{\beta}_N \in (0, 1)$ such that the welfare with one-sided commitment is greater than the welfare with two-sided commitment when $\underline{R} < R < \frac{1}{\delta}$ and $\beta < \bar{\beta}_N$.*

2. *Suppose the agent is sophisticated. There exist \underline{R}_1 and \underline{R}_2 with $\underline{R}_1 < \frac{1}{\delta} < \underline{R}_2$ and $\bar{\beta}_S \in (0, 1)$ such that the welfare with one-sided commitment is greater than the welfare with two-sided commitment when $\underline{R}_1 < R < \underline{R}_2$ and $\beta < \bar{\beta}_S$.*

Of course, one-sided commitment does not always yield a higher welfare than under two-sided commitment. In particular, if agents are close enough to time consistent (β close to 1) and the interest rate is low enough, “over-borrowing” is not as bad as not being able to borrow at all, the welfare with two-sided commitment is higher.

2.5 Asymptotic Welfare Loss: Naiveté is Bliss

We now examine how time-inconsistency affects the agent’s welfare. Let c_C^1 and c_S^1 denote the equilibrium consumption of the time-consistent and sophisticated agent with one-sided commitment, and let c_C^2 and c_S^2 denote their equilibrium consumption with two-sided commitment. Recall that c_C^{1E} and c_C^{2E} are the consumption on the equilibrium path of the naive agent with one- and two-sided commitment. Note that these expressions are functions of the parameters of the model, although we omit this dependence for notational simplicity. Since the welfare function coincides with the utility of time-consistent agents, they obtain the highest possible welfare:

$$W(c_C^i) \geq W(c_N^{iE}) \text{ and } W(c_C^i) \geq W(c_S^{iE}), \quad i = 1, 2.$$

Consider the case with one-sided commitment. When $R\delta \leq 1$, the market shuts down for all agents and they all get the same welfare. When $R\delta > 1$, these inequalities are strict (time-inconsistent agents always over-consume in the initial period and under-consume in the last period relative to the welfare-maximizing consumption stream).

not affect our results for naifs. However, sophisticates would get the same contract as time-consistent agents, obtaining a higher welfare with two-sided commitment than with one-sided commitment.

Our main result shows that, as the contracting horizon grows, the welfare loss of the naive agent vanishes. To ensure that the sum of discounted utility converges, so that the limit of the welfare function as the horizon goes to infinity is well defined, we make the standard assumptions of a bounded utility function and discounting.¹⁶

Proposition 3. *Suppose u is bounded and $\delta < 1$. Then,*

$$\lim_{T \nearrow +\infty} [W(c_C^1) - W(c_N^{1E})] = 0 \text{ and } \lim_{T \nearrow +\infty} [W(c_C^2) - W(c_N^{2E})] = 0.$$

The intuition for the result is as follows. The consumption of a naive agent coincides with the solution of the auxiliary problem, which only distorts the last period by β . As number of periods grows, the effect of distorting a single period becomes small. In the limit, the naive agent obtain the same consumption stream as time-consistent agents. Therefore, if the contracting length is large enough, the welfare loss of a naive agent is small. In particular, the welfare of naive agents is (weakly) higher than the welfare sophisticates when the contracting length is large. Notice that this result holds for any fixed discount factors β and δ , and so it is unrelated to the folk theorem literature from repeated games.

Proposition 3 differs sharply from existing results. In particular, DellaVigna and Malmendier (2004) and Heidhues and Kőszegi (2010) find that, with two-sided commitment, sophisticates receive same consumption as time-consistent agents, while naive agents obtain a distorted consumption. The reason for this apparent discrepancy is that the proposition applies to the case of a sufficiently large contracting length, whereas in their models there are only two periods of consumption.

Proposition 3 also implies that if the contracting length is large enough, removing commitment power from naive agents cannot increase welfare, since it is suboptimal to remove commitment power from time-consistent agents.

¹⁶The assumption of bounded utility can be substantially generalized. The proof of the proposition remains unchanged if, instead, we make the weaker assumption that the game is continuous at infinity – see Fudenberg and Tirole (1991, pp. 110).

For a fixed T , it is possible to construct examples where the sophisticated agent gets either a higher or a lower welfare than the naive agent.¹⁷ However, if the market breaks down for the sophisticated agent ($R \leq r_T(\beta, \delta)$), algebraic manipulations show that the naive agent must still obtain a higher welfare than the sophisticated agent (see the online appendix). The intuition for this result is that if the market also breaks down for the naive agent, they both obtain the same consumption in all periods. But if the market does not break down for the naive agent, the ability to smooth consumption gives them a strictly higher welfare than if they consumed their endowment, which is what happens to sophisticates.

2.6 Maximum Fees

We now consider the effects of imposing a maximum fee in each period. As previously, it is convenient to write the contract in terms of the agent's consumption instead of in terms of payments to the principal. The fee paid in each state corresponds to the difference between the endowment and the consumption: $w - c_t(h^t)$. Therefore, specifying a maximum fee is equivalent to mandating a consumption floor \underline{c} . Notice that if the consumption floor exceeded the agent's endowment, any contract would give negative profits to the firm. Therefore, the consumption floor cannot exceed the agent's endowment, $0 < \underline{c} < w$, or, equivalently, the maximum fee cannot be negative (since, otherwise, no firm would ever break even).

Of course, if the agent is time-consistent, imposing a consumption floor introduces additional constraints in the agent's welfare maximization program, which cannot increase welfare. Suppose, instead, that the agent is naive. Recall that the equilibrium contract solves the auxiliary program, which coincides with the equilibrium program of a time-consistent agent that discounts the last period more heavily. By the non-lapsing constraint, the consumption in the last period cannot be lower than the agent's income, so the consumption floor never binds in the last period.

¹⁷For example, consider a 3-period model with $w = 1$ and $R = 1.5$. Let $u(c) = 1 - e^{-c}$, $\delta = 1$, and $\beta = 0.5$. Solving the auxiliary program, we find that the naive agent's consumption on the equilibrium path is $(0.8378, 1.2433, 1)$, which gives a welfare of 1.9110. The sophisticate's equilibrium consumption equals $(0.9675, 0.9105, 1.2073)$, giving him a welfare of 1.9187, which is greater than the welfare of the naive agent.

If $R\delta \leq 1$, the agent would like to borrow, so the market breaks down and the consumption floor is not binding in any period. If $R\delta > 1$, the agent prefers to save in the initial periods, and consumption is increasing along the first $T - 1$ periods. Therefore, whenever the consumption floor binds, it must bind in the initial periods, reducing saving. Since a naive agent already under-saves relative to the welfare-maximizing amount, this policy hurts them whenever it is binding:

Proposition 4. *Suppose the agent is naive and has no commitment power. Then, mandating a minimum consumption weakly decreases welfare.*

Unlike with naive agents, a consumption floor may increase savings for sophisticated agents, making the welfare effect ambiguous (see Appendix B). This is because the equilibrium with a sophisticated agent may have very low consumption in intermediate periods. In those cases, a consumption floor would increase initial savings, which can increase welfare.

3 General Income Distributions

We now generalize the main results to arbitrary income distributions. In order to have a well-defined notion of what it means for the contracting length to grow in a potentially non-stationary setting, we consider T -period truncations of a setting with infinitely many periods.

Formally, for each $t \in \mathbb{N}$, let \mathbb{S}_t be the set of possible spaces of the world. Let $p(s_t|s_\tau)$ denote the probability of reaching state $s_t \in \mathbb{S}_t$ conditional on state $s_\tau \in \mathbb{S}_\tau$. A state of the world describes all uncertainty realized up to that period, so it is not possible to reach the same state from different states of the world (i.e., if $s_\tau \neq \hat{s}_\tau$ and $p(s_t|s_\tau) > 0$ then $p(s_t|\hat{s}_\tau) = 0$). Without loss of generality, we assume that no uncertainty is realized before the initial period: $\mathbb{S}_1 = \{\emptyset\}$. A T -period truncation is a model with these states and conditional probabilities up to period T , at which point the game ends (or, equivalently, the agent is forced to consume his endowment in all periods greater than T). At state s_t , the agent earns income $w(s_t)$. Notice that, by taking degenerate distributions, this framework incorporates environments with deterministic incomes. Moreover, since the probabilities of reaching future states depend on the current state, the framework also

allows for persistent shocks. Therefore, this general framework allows the agent to be subject to reclassification risk.

As usual, stochastic payments are evaluated by taking expectations. Thus, the expected discounted profit at state s_τ of a principal who collects state-dependent payments $\{\pi(s_t)\}$ is

$$\sum_{t=\tau}^T \sum_{s_t \in \mathbb{S}_t} p(s_t | s_\tau) \frac{\pi(s_t)}{R^{t-\tau}}.$$

Similarly, at state s_τ , the agent's expected discounted utility from state-dependent consumption $\{c(s_t)\}$ is

$$u(c(s_\tau)) + \beta \sum_{t=\tau+1}^T \sum_{s_t \in \mathbb{S}_t} p(s_t | s_\tau) \delta^{t-s} u(c(s_t)).$$

The definitions of histories and contracts are analogous to the ones from Section 2, except that they are now functions of the state of the world. Therefore, a contract at state s_t specifies a payment to the principal in all future states of the world (and, if the agent is naive, conditional on whether the agent picks A or B in each period up to t).

3.1 Removing Commitment Power

We now generalize the result from Section 2.4, where we showed that removing commitment power from the agent can make him better off. We start with the case of a sophisticated agent:

Proposition 5. *Suppose the agent is sophisticated. There exists $\bar{\beta} > 0$ and $\bar{\delta} < 1$ such that, if $\beta < \bar{\beta}$ and $\delta > \bar{\delta}$, the welfare with one-sided commitment is greater than the welfare with two-sided commitment.*

The intuition for this result is that giving commitment power to the agent allows him to borrow, and time-inconsistent agents are tempted to over-borrow. Thus, the welfare effect of removing commitment depends whether the welfare gain from being able to borrow is outweighed by the welfare loss from over-borrowing. The over-borrowing problem is more severe if the agent is very time inconsistent (β is small) and saving is important for welfare (δ is large). In that case, removing

commitment power improves the agent's welfare.

Next, consider a naive agent. Recall that, on the equilibrium path, a naive agent gets the same consumption as a dynamically consistent agent who under-weights the last period by an additional factor β . Therefore, giving commitment power to a naive agent allows him to smooth consumption in the first $T - 1$ periods (where the objective function coincides with the welfare function), while leaving too little consumption for the last period. This distortion in the last period is large when the agent is sufficiently time inconsistent (β is low), in which case the last period consumption will be close to zero. So, if consuming zero in the last period hurts the agent enough and β is low, the agent is better off without commitment.

To formalize this argument, let \mathcal{V}_S denote the agent's welfare from smoothing consumption perfectly in the first $T - 1$ periods and consuming zero in the last period:

$$\mathcal{V}_S \equiv \max_{\{c(s_t)\}} \sum_{t=1}^{T-1} \sum_{s_t \in S_t(s_1)} \delta^{t-\tau} p(s_t|s_1) u(c(s_t)) + \delta^{T-1} u(0),$$

subject to

$$\sum_{t=1}^{T-1} \sum_{s_t \in S_t(s_1)} p(s_t|s_1) \frac{w(s_t) - c(s_t)}{R^{t-1}} = 0.$$

Let \mathcal{V}_{NS} denote the agent's welfare from consuming the endowment in each state (no consumption smoothing):

$$V_{NS} \equiv \sum_{t=1}^T \sum_{s_t \in S_t(s_1)} \delta^{t-\tau} p(s_t|s_1) u(w(s_t)).$$

Proposition 6. *Suppose the agent is naive and $\mathcal{V}_{NS} > \mathcal{V}_S$. There exists $\bar{\beta}_N$ such that if $\beta < \bar{\beta}_N$, the welfare with one-sided commitment is greater than the welfare with two-sided commitment.*

3.2 Asymptotic Welfare Loss

We now generalize the result from Section 2.5, where we showed that the welfare loss of the naive agent vanishes as the contracting horizon grows.

Proposition 7. *Suppose u is bounded and $\delta < 1$. Then,*

$$\lim_{T \nearrow +\infty} [W(c_C^1) - W(c_N^{1E})] = 0 \text{ and } \lim_{T \nearrow +\infty} [W(c_C^2) - W(c_N^{2E})] = 0.$$

Proposition 7 relies on showing that equivalence between the naive agent's consumption on the equilibrium path and the equilibrium with a time-consistent agent who discounts the last period by an additional factor β (see the online appendix). Then, as in Proposition 3 the result follows from the fact that, as number of periods grows, the effect of distorting a single period becomes small.

As with Proposition 3, Proposition 7 implies that when the contracting period is large, the welfare of naive agents is a (weakly) higher than the welfare of sophisticates, and that, if the contracting horizon is large enough, a naive agent is always better off with commitment power (since a time-consistent agent is always better off with commitment).

4 Conclusion

Consumer protection laws usually allow consumers to terminate certain contracts at will. In a rational framework, these are puzzling laws since allowing consumers to terminate contracts introduces renegotiation proofness constraints, which can only make the contracting problem less efficient. In this paper, we show that when consumers have self-control problems, removing their commitment power can be welfare improving. Somewhat surprisingly, we show that the welfare loss from dynamic inconsistency vanishes as the contracting when agents are naive. This result suggests that enforcing long-term contracts may be an effective way to restore efficiency in competitive markets when agents are naive.

We focus on one particular deviation from rationality – dynamic inconsistency – for two reasons. First, the decision to leave a previous contract is fundamentally an intertemporal decision, and dynamic inconsistency is the most well-studied bias in intertemporal decisions. Second, there is evidence that this bias is an important feature of credit markets where consumers are allowed to leave previous agreements, such as mortgages or credit cards.

Nevertheless, we think that the policies that remove commitment power are also important in settings with other biases. For example, most developed countries have regulations that specify “cooling-off” periods, in which firms must allow consumers to return goods. Allowing consumers to return recently purchased products is an effective policy when consumers suffer from “projection bias,” that is, they mispredict their future tastes, overestimating how much it will resemble their current tastes.¹⁸ Individuals who mis-predict their future tastes will buy objects and then regret it. For example, although they would voluntarily choose a non-refundable ticket to an expensive trip while craving for a vacation, they later regret it.

¹⁸See Loewenstein et al. (2003); Camerer et al. (2003); Conlin et al. (2007).

Appendix A: Proofs

Proof of Lemma 1. The equilibrium contract solves:

$$\max_{\{c_t\}} \sum_{t=1}^T \delta^{t-1} u(c_t), \quad (\text{A1})$$

subject to

$$\sum_{t=1}^T \frac{c_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}}, \quad (\text{A2})$$

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-\tau}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-\tau}}, \forall 2 \leq \tau \leq T. \quad (\text{A3})$$

Existence follows from the fact that the objective is continuous and the the set of feasible contracts is non-empty ($c_t = w$ for all t satisfies the constraints) and compact. Since the objective function is strictly concave and the set of feasible contracts is convex, the solution is unique.

The Lagrangian is

$$\mathcal{L} = \sum_{t=1}^T \delta^{t-1} u(c_t) - \sum_{\tau=1}^T \lambda_{\tau} \left(\sum_{t=\tau}^T \frac{c_t}{R^{t-1}} - \sum_{t=\tau}^T \frac{w}{R^{t-1}} \right), \quad (\text{A4})$$

where $\lambda_{\tau} \geq 0$. Then $\delta^{t-1} u'(c_t) = \frac{\sum_{\tau=1}^t \lambda_{\tau}}{R^{t-1}}$, or equivalently, $u'(c_t) = \frac{\sum_{\tau=1}^t \lambda_{\tau}}{(\delta R)^{t-1}}$.

First, consider $\delta R \leq 1$. Then

$$u'(c_1) \leq u'(c_2) \leq \dots \leq u'(c_T),$$

therefore, $c_1 \geq c_2 \geq \dots \geq c_T$. From the zero profit condition, we then would have $c_T \leq w$. We also must have $c_T \geq w$ to prevent the agent to leave the contract at the last period. So it must be the case that $c_T = w$. Now we have $c_1 \geq \dots \geq c_{T-1} \geq w$. From the zero profit condition $c_{T-1} \leq w$. By the non-lapsing condition, we need to have $c_{T-1} \geq w$. Similarly we conclude $c_{T-1} = w$. Using the same argument, we find $c_1 = \dots = c_T = w$. In other words, the market breaks down.

Second, consider $\delta R > 1$. We can solve the problem with the zero-profit condition and then verify that the resource constraint (A3) holds automatically. Solving the problem with only the zero-profit condition gives $c_1 < c_2 < \dots < c_T$. Notice that $c_1 < w$ because otherwise, $w \leq c_1 < c_2 < \dots < c_T$, contradicted to the zero profit condition. Similarly we have $c_T > w$. Let ξ be the smallest index such that

$$c_1 < \dots < c_\xi < w \leq c_{\xi+1} < \dots < c_T.$$

It is clear that (A3) holds strictly for $\tau \geq \xi+1$. Now consider $\tau \leq \xi$, we have

$$\begin{aligned} \sum_{t=\tau}^T \frac{c_t}{R^{t-1}} &= \sum_{t=1}^T \frac{c_t}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{c_t}{R^{t-1}} \\ &= \sum_{t=1}^T \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{c_t}{R^{t-1}} \\ &> \sum_{t=1}^T \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{w}{R^{t-1}} \\ &= \sum_{t=\tau}^T \frac{w}{R^{t-1}}. \end{aligned} \tag{A5}$$

So the resource constraint (A3) holds strictly for $\tau \leq \xi$. In all, the equilibrium contract features growing consumption and the long-term contract is supported in this market. \square

Proof of Lemma 2. We show the existence by induction. For $T = 1$, the only feasible consumption is $c_1 = w$, so that $V_1^S = u(w)$. Suppose the maximization program has a solution for any period that is smaller than T . Notice that the program defining V_2^S is equivalent to a $(T - 1)$ -period problem. Then V_τ^S exists, for any $\tau = 2, \dots, T$.

We now show that V_1^S exists. First, notice that the objective function is a continuous function in $c = (c_1, \dots, c_T)$. Second, we show that the feasible set is non-empty. Starting with the autarky consumption (w, w, \dots, w) , we construct a consumption profile that satisfies all the constraints as follows. Suppose the non-lapsing constraint is first violated at τ , then we replace the last $(T - \tau + 1)$ periods consumption (w, \dots, w) by the contract (c'_τ, \dots, c'_T) , which solves V_τ . By design, $(w, \dots, w, c'_\tau, \dots, c'_T)$ is a consumption profile that satisfies all the constraints, im-

plying that the feasible set is non-empty. Lastly, from the zero-profit condition, it follows that $c_t \in [0, R^{t-1} \sum_{t=1}^T \frac{w}{R^{t-1}}]$, implying the feasible set for the program V_1^S is a compact subset of \mathbb{R}^T . Since a continuous function always has a maximum over a non-empty compact set, a maximizer exists in our setting. Then, an equilibrium exists.

We now show that the equilibrium is unique. Letting $x_t = u(c_t)$, we can write the dual program as follows.

$$\max_{x_1, \dots, x_T} \sum_{t=1}^T \frac{w}{R^{t-1}} - \sum_{t=1}^T \frac{u^{-1}(x_t)}{R^{t-1}}, \quad (\text{A6})$$

subject to

$$x_s + \beta \sum_{t=s+1}^T \delta^{t-s} x_t \geq V_s^S, \forall s = 1, \dots, T. \quad (\text{A7})$$

Note the objective function in the above program is concave and the feasible set is a linear set. Therefore there exists a unique (x_1, \dots, x_T) , which in turn implies the uniqueness of equilibrium.

We now show claim (i). Let λ_1 denote the Lagrangian multiplier associated with the zero-profit constraint, and let λ_τ denote the Lagrangian multiplier associated with the non-lapsing constraints.

The Lagrangian optimality conditions imply:

$$\begin{aligned} u'(c_1) &= \lambda_1, \\ \beta \delta u'(c_2) &= \frac{\lambda_1}{R} - \lambda_2 u'(c_2), \\ &\dots \\ \beta \delta^{T-1} u'(c_T) &= \frac{\lambda_1}{R^{T-1}} - (\beta \delta^{T-2} \lambda_2 + \dots + \beta \delta \lambda_{T-1} + \lambda_T) u'(c_T). \end{aligned}$$

Let $r \equiv \delta R$ and $x_i \equiv \lambda_{i+1} R^{i+1}$ for $i = 1, \dots, T-1$.

We first examine the conditions for the market to break down: $c_1 = \dots = c_T = w$. We need to have $\lambda_\tau \geq 0, \forall \tau \geq 2$, or equivalently, $x_i \geq 0, \forall 1 \leq i \leq T-1$. We can rewrite the conditions in a

matrix form as follows.

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta r & 1 & 0 & \cdots & 0 \\ \beta r^2 & \beta & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \beta r^{T-2} & \beta r^{T-3} & & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{T-1} \end{pmatrix} = \begin{pmatrix} 1 - \beta r \\ 1 - \beta r^2 \\ 1 - \beta r^3 \\ \vdots \\ 1 - \beta r^{T-1} \end{pmatrix}.$$

Inverting the lower-triangular matrix, we can find that the necessary and sufficient condition for $x_i \geq 0, \forall 1 \leq i \leq T - 1$ is:

$$1 \geq \beta r + \beta(1 - \beta)r^2 + \cdots + \beta(1 - \beta)^{T-2}r^{T-1}. \quad (\text{A8})$$

The right-hand-side is an increasing function of r . If $r = 0$, LHS > RHS. If $r \rightarrow \infty$, LHS < RHS. So there exists a unique $r_T(\beta)$ such that the (A8) becomes an equality. Since the the RHS evaluated at $r = 1$ is strictly less than 1:

$$\sum_{t=0}^{T-2} \beta(1 - \beta)^t < \sum_{t=0}^{\infty} \beta(1 - \beta)^t = \beta \frac{1}{1 - (1 - \beta)} = 1,$$

we must have $r_T(\beta) > 1$. Let $r_T(\beta, \delta) := \frac{r_T(\beta)}{\delta} > \frac{1}{\delta}$. The market breaks down when $R \leq r_T(\beta, \delta)$.

We now turn to the properties of $r_T(\beta, \delta)$. We first show that $r_T(\beta, \delta)$ is decreasing in β . Recall that $r_T(\beta, \delta) = \frac{r_T(\beta)}{\delta}$, where $r_T(\beta)$ solves the equation (A8). It is sufficient to show that $r_T(\beta)$ is decreasing in β . The right hand side of (A8) is a geometric series, implying

$$1 = \beta r \frac{1 - (1 - \beta)^{T-1}r^{T-1}}{1 - (1 - \beta)r}. \quad (\text{A9})$$

Rearranging terms leads to

$$1 - r + \beta(1 - \beta)^{T-1}r^T = 0. \quad (\text{A10})$$

Taking derivative with respect to β , we obtain

$$\begin{aligned}
r'_T(\beta) &= \frac{(1 - T\beta)(1 - \beta)^{T-2}r_T^T(\beta)}{1 - \beta(1 - \beta)^{T-1}r_T^{T-1}T} \\
&= \frac{(1 - T\beta)(1 - \beta)^{T-2}r_T^{T+1}(\beta)}{r_T - \beta(1 - \beta)^{T-1}r_T^T(\beta)T} \\
&= \frac{(1 - T\beta)(1 - \beta)^{T-2}r_T(\beta)^{T+1}}{r_T(\beta) - (r_T(\beta) - 1)T} \\
&= \frac{(1 - T\beta)(1 - \beta)^{T-2}r^{T+1}}{T - (T - 1)r_T(\beta)}. \tag{A11}
\end{aligned}$$

We can rewrite equation (A10) as

$$r^T = \frac{1}{\beta(1 - \beta)^{T-1}}(r - 1). \tag{A12}$$

This equation has two real positive roots, one of which is $\frac{1}{1-\beta}$ and the other one is $r_T(\beta)$. The left-hand-side of equation (A12) is a line with a slope $\frac{1}{\beta(1-\beta)^{T-1}}$. We can verify that when $\beta = \frac{1}{T}$. The left-hand-side and the right-hand-side of equation (A12) are tangent at $r = \frac{T}{T-1}$. Thus, we have $r_T(\beta) > \frac{T}{T-1}$ if $\beta < \frac{1}{T}$ and $r_T(\beta) < \frac{T}{T-1}$ if $\beta > \frac{1}{T}$. Finally, from equation (A11), we know that $r'(\beta) < 0$. So $r_T(\beta, \delta)$ is a decreasing function of β .

We then show that $r_T(\beta, \delta)$ is decreasing in T . Suppose the interest rate is such that the market breaks down with T periods, i.e., (w, w, \dots, w) is the solution to the sophisticates' program (9). Consider the non-lapsing constraint (11) at time 2. First, since (w, \dots, w) must satisfy the constraints, $u(w) + \beta \sum_{i=2}^T \delta^{i-1}u(w) \geq V_2^S$. On the other hand, by definition, $V_2^S \geq u(w) + \beta \sum_{i=2}^T \delta^{i-1}u(w)$. So $V_2^S = u(w) + \beta \sum_{i=2}^T \delta^{i-1}u(w)$, implying the market breaks down with $(T - 1)$ period. So we have shown that if the interest rate is such that the market breaks down with T periods, the market also breaks down with $(T - 1)$ periods. Put differently, the cutoff $r_T(\beta, \delta)$ must be decreasing in T .

Finally, we show claim (3). As $\beta \rightarrow 1$, the right-hand-side of equation (A8) becomes βr . So, $\lim_{\beta \nearrow 1} r_T(\beta)$. Then it follows that $\lim_{\beta \nearrow 1} r_T(\beta, \delta) = \frac{\lim_{\beta \nearrow 1} r_T(\beta)}{\delta} = \frac{1}{\delta}$. As $T \rightarrow +\infty$, the right-hand-side of equation (A8) becomes $\frac{\beta r}{1 - (1 - \beta)r}$. Solving $\frac{\beta r}{1 - (1 - \beta)r} = 1$ gives $r = 1$. Thus,

$$\lim_{T \nearrow \infty} r_T(\beta, \delta) = \frac{\lim_{T \nearrow \infty} r_T(\beta)}{\delta} = \frac{1}{\delta}. \quad \square$$

Proof of Corollary 1. The proof follows directly from applying Theorem of Maximum on the dual program (A6) in the proof of Lemma 2. \square

Proof of Lemma 3. We first focus on the one-sided commitment. We show that all IC constraints are binding. To simplify the exposition, we focus on the case of $T = 4$ here (the proof for general T is essentially the same except for additional notation). There are two ICs:

$$u(c_2(A)) + \beta[\delta u(c_3(AB)) + \delta^2 u(c_4(AB))] \geq u(c_2(B)) + \beta[\delta u(c_3(BB)) + \delta^2 u(c_4(BB))], \quad (\text{A13})$$

$$u(c_3(AA)) + \beta \delta u(c_4(AA)) \geq u(c_3(AB)) + \beta \delta u(c_4(AB)). \quad (\text{A14})$$

First, notice that these two ICs give upper bounds on $c_4(BB)$ and $c_4(AB)$. Since no other constraints restrict $c_4(BB)$ and $c_4(AB)$ from above, (A13) must be binding at an optimum (otherwise, we can raise $c_4(BB)$, giving the agent a higher utility). Substitute the binding (A13) in the objective to eliminate $c_4(BB)$:

$$\begin{aligned} & u(c_1) + \beta[\delta u(c_2(B)) + \delta^2 u(c_3(BB)) + \delta^3 u(c_4(BB))] \\ &= u(c_1) + \delta u(c_2(A)) + \beta[\delta^2 u(c_3(AB)) + \delta^3 u(c_4(AB))] + (\beta - 1)\delta u(c_2(B)). \end{aligned}$$

By the same argument, (A14) must bind (otherwise, we can raise $c_4(AB)$, increasing the agent's utility). Substituting the binding (A14) in the objective, gives:

$$u(c_1) + \delta u(c_2) + \delta^2 u(c_3) + \beta \delta^3 u(c_4) + (\beta - 1)[\delta u(c_2(B)) + \delta^2 u(c_3(AB))].$$

Since $\beta < 1$, we want to pick $c_2(B)$, $c_3(AB)$ as small as possible (subject to the constraints). We now show that all the perceived non-lapsing constraints hold if we set them at their lowest possible values (zero, by our normalization): $c_2(B) = c_3(AB) = 0$.

Let the contract \hat{c} denote the maximizer to the perceived outside option program, \hat{V}_2^N . Suppose $\hat{V}_2^N = u(\hat{c}_2) + \hat{\beta}(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4))$. We obtain

$$\begin{aligned} & u(c_2(B)) + \hat{\beta}(\delta u(c_3(BB)) + \delta^2 u(c_4(BB))) \\ &= \frac{\hat{\beta}}{\beta} \beta (\delta u(c_3(BB)) + \delta^2 u(c_4(BB))) \\ &= \frac{\hat{\beta}}{\beta} (u(c_2(A)) + \beta(\delta u(c_3(AB)) + \delta^2 u(c_4(AB))), \end{aligned}$$

where the first equality follows from $c_2(B) = 0$ and the second uses the binding IC constraint. From the non-lapsing constraint at time 2, we know that $u(c_2(A)) + \beta(\delta u(c_3(AB)) + \delta^2 u(c_4(AB))) \geq V_2^N$. Since V_2^N is the best possible outside option at time 2, in particular, it is greater than or equal to the utility provided by the contract \hat{c} , implying

$$\begin{aligned} u(c_2(B)) + \hat{\beta}(\delta u(c_3(BB)) + \delta^2 u(c_4(BB))) &\geq \frac{\hat{\beta}}{\beta} V_2^N \\ &\geq \frac{\hat{\beta}}{\beta} [u(\hat{c}_2) + \beta(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4))] \\ &> u(\hat{c}_2) + \hat{\beta}(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4)) = \hat{V}_2^N, \end{aligned}$$

where the first line follows from the non-lapsing constraint at time 2, the second uses the revealed preference, and the last line uses $\hat{\beta} \geq \beta$. This shows that all the perceived non-lapsing constraints hold.

We next verify that all the perceived choice constraints hold. Notice that

$$\begin{aligned} u(c_3(AB)) + \hat{\beta} \delta u(c_4(AB)) &= \hat{\beta} \delta u(c_4(AB)) \\ &= \frac{\hat{\beta}}{\beta} \left(u(c_3(AA)) + \hat{\beta} \delta u(c_4(AA)) \right) \\ &\geq u(c_3(AA)) + \hat{\beta} \delta u(c_4(AA)), \end{aligned} \tag{A15}$$

and

$$\begin{aligned}
& u(c_2(B)) + \hat{\beta}[\delta u(c_3(BB)) + \delta^2 u(c_4(BB))] \\
&= \hat{\beta}[\delta u(c_3(BB)) + \delta^2 u(c_4(BB))] \\
&= \frac{\hat{\beta}}{\beta} [u(c_2(A)) + \beta[\delta u(c_3(AB)) + \delta^2 u(c_4(AB))]] \\
&\geq u(c_2(A)) + \hat{\beta}[\delta u(c_3(AB)) + \delta^2 u(c_4(AB))].
\end{aligned} \tag{A16}$$

So the perceived choice constraints hold.

So far, we have shown that $c_2(B) = c_3(AB) = 0$ under the equilibrium contract. We also showed that we can disregard the perceived choice constraints and perceived non-lapsing constraints. Recall that c_t^E denotes the consumption on the equilibrium path at time t . Substituting the binding ICs, the non-lapsing constraints on the equilibrium path can be simplified to $u(c_t^E) + \delta u(c_{t+1}^E) + \dots + \beta\delta^{4-t}u(c_4^E) \geq V_t^N$.

Therefore, the original program reduces to the auxiliary program:

$$\max_{(c_1, \dots, c_T)} u(c_1) + \delta u(c_2) + \dots + \delta^{T-2}u(c_{T-1}) + \beta\delta^{T-1}u(c_T), \tag{A17}$$

subject to

$$\sum_{t=1}^T \frac{c_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}}, \tag{A18}$$

$$u(c_t) + \delta u(c_{t+1}) + \dots + \beta\delta^{T-t}u(c_T) \geq V_t^A, \forall 2 \leq t \leq T. \tag{A19}$$

If there is two-sided commitment, the above argument go through except that we can omit the non-lapsing constraints (A19). In all, the consumption on the equilibrium path coincides with the solution of the auxiliary problem: $c_N^{1E} = c^{1A}$ and $c_N^{2E} = c^{2A}$.

Proof of Lemma 4. We can rewrite the non-lapsing constraints in the auxiliary problem as

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-1}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-1}}, \forall 2 \leq \tau \leq T. \quad (\text{A20})$$

Consider the following program.

$$\max_{c_t} u(c_1) + \delta u(c_2) + \cdots + \delta^{T-2} u(c_{T-1}) + \beta \delta^{T-1} u(c_T), \quad (\text{A21})$$

subject to

$$\sum_{t=1}^T \frac{c_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}}, \quad (\text{A22})$$

$$\sum_{t=\tau}^T \frac{c_t}{R^{t-1}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-1}}, \forall 2 \leq \tau \leq T. \quad (\text{A23})$$

This program has a concave objective function and the feasible set is a non-empty linear set, so there exists a solution. Since the program does not depend on β , it is clear that all equilibria have the same consumption on the equilibrium path. By Theorem of Maximum, the consumption on the equilibrium path is a continuous function of $\beta \in (0, 1]$. We note that it may not be right-continuous at $\beta = 0$ because the step showing binding incentive constraints in the proof of Lemma 3 requires $\beta > 0$.

The Lagrangian is

$$\mathcal{L} = \sum_{t=1}^{T-1} \delta^{t-1} u(c_t) + \beta \delta^{T-1} u(c_T) - \sum_{\tau=1}^T \lambda_{\tau} \left(\sum_{t=\tau}^T \frac{c_t}{R^{t-1}} - \sum_{t=\tau}^T \frac{w}{R^{t-1}} \right),$$

where $\lambda_{\tau} \geq 0$. If $1 \leq t \leq T - 1$, we have $\delta^{t-1} u'(c_t) = \frac{\sum_{\tau=1}^t \lambda_{\tau}}{R^{t-1}}$, or equivalently, $u'(c_t) = \frac{\sum_{\tau=1}^t \lambda_{\tau}}{(\delta R)^{t-1}}$. If $t = T$, $u'(c_t) = \frac{\sum_{\tau=1}^t \lambda_{\tau}}{\beta (\delta R)^{t-1}} > \frac{\sum_{\tau=1}^t \lambda_{\tau}}{(\delta R)^{t-1}}$.

If $\delta R \leq 1$, then $u'(c_1) \leq u'(c_2) \leq \cdots \leq u'(c_T)$, therefore $c_1 \geq c_2 \geq \cdots \geq c_T$. From the zero-profit condition, we then have $c_T \leq w$. We also have $c_T \geq w$ from self T's non-lapsing

constraint. So it must be the case that $c_T = w$. Now we have $c_1 \geq \dots \geq c_{T-1} \geq w$. From the zero profit condition $c_{T-1} \leq w$. By the non-lapsing condition, we need to have $c_{T-1} \geq w$. Similarly we conclude $c_{T-1} = w$. Applying the same argument, we have $c_1 = \dots = c_T = w$.

Now if $\delta R > 1$, consider the problem with the same objective function and the zero profit condition and the last period non-lapsing constraint:

$$\tilde{c} = (\tilde{c}_1, \dots, \tilde{c}_T) = \arg \max_{c_t} u(c_1) + \delta u(c_2) + \dots + \delta^{T-2} u(c_{T-1}) + \beta \delta^{T-1} u(c_T),$$

subject to

$$\sum_{t=1}^T \frac{c_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}},$$

$$c_T \geq w.$$

Applying Lagrangian condition gives $u'(\tilde{c}_1) = \dots = (\delta R)^{T-2} u'(\tilde{c}_{T-1}) \leq \beta (\delta R)^{T-1} u'(\tilde{c}_T)$. Since $\delta R > 1$, we have $\tilde{c}_1 < \tilde{c}_2 < \dots < \tilde{c}_{T-1}$. We next verify that \tilde{c} satisfies all the non-lapsing constraints: $\sum_{t=\tau}^T \frac{c_t}{R^{t-1}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-1}}, \forall 2 \leq \tau \leq T$, in which case, \tilde{c} would be the optimal solution for the original problem and the long-term contract is supported in this market. To see that, let ξ be the largest index such that $\tilde{c}_\xi < w$. If $\tau \geq \xi + 1$, $\sum_{t=\tau}^T \frac{\tilde{c}_t}{R^{t-1}} \geq \sum_{t=\tau}^T \frac{w}{R^{t-1}}$. If $\tau \leq \xi$, we have

$$\begin{aligned} \sum_{t=\tau}^T \frac{\tilde{c}_t}{R^{t-1}} &= \sum_{t=1}^T \frac{\tilde{c}_t}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{\tilde{c}_t}{R^{t-1}} \\ &= \sum_{t=1}^T \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{\tilde{c}_t}{R^{t-1}} \\ &> \sum_{t=1}^T \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{w}{R^{t-1}} \\ &= \sum_{t=\tau}^T \frac{w}{R^{t-1}}. \end{aligned} \tag{A24}$$

If $\tilde{c}_T > w$, then the solution is given by $u'(\tilde{c}_1) = \dots = (\delta R)^{T-2}u'(\tilde{c}_{T-1}) = \beta(\delta R)^{T-1}u'(\tilde{c}_T)$ and $\sum_{t=1}^T \frac{\tilde{c}_t}{R^{t-1}} = \sum_{t=1}^T \frac{w}{R^{t-1}}$.

If $\tilde{c}_T = w$, the problem can be further reduced to

$$\max_{c_t} u(c_1) + \delta u(c_2) + \dots + \delta^{T-2}u(c_{T-1}),$$

subject to

$$\sum_{t=1}^{T-1} \frac{c_t}{R^{t-1}} = \sum_{t=1}^{T-1} \frac{w}{R^{t-1}}.$$

Then the solution is determined by $u'(\tilde{c}_1) = \dots = (\delta R)^{T-2}u'(\tilde{c}_{T-1})$, $\sum_{t=1}^{T-1} \frac{\tilde{c}_t}{R^{t-1}} = \sum_{t=1}^{T-1} \frac{w}{R^{t-1}}$, and $\tilde{c}_T = w$.

In summary, the market breaks down if and only if $R \leq \frac{1}{R}$. □

Proof of Proposition 1. Presented in the text. □

Proof of Proposition 2.

Part 1. Since $R \leq \frac{1}{\delta}$, the market breaks down under one-sided commitment. Then we know that the welfare under one-sided commitment is $U^1 = u(w) \sum_{i=0}^{T-1} \delta^i$.

Under two-sided commitment, the equilibrium-path consumption solves the following program:

$$\max_{c_1, \dots, c_T} \sum_{i=t}^{T-1} \delta^{t-1} u(c_i) + \beta \delta^{T-1} u(c_T) \tag{A25}$$

subject to

$$\sum_{t=1}^T \frac{c_t - w}{R^{t-1}} = 0 \tag{A26}$$

The solution is given by $u'(c_1^*) = \delta R u'(c_2^*) = \dots = \beta(\delta R)^{T-1}u'(c_T^*)$ and $\sum \frac{c_i^* - w}{R^{i-1}} = 0$. Then the welfare under two-sided commitment is $U^{N2} = \sum_{i=1}^T \delta^{i-1} u(c_i^*)$.

Taking derivative w.r.t. β on $u'(c_1^*) = \delta R u'(c_2^*) = \dots = \beta(\delta R)^{T-1} u'(c_T^*)$ gives

$$u''(c_1^*) \frac{\partial c_1^*}{\partial \beta} = \delta R u''(c_2^*) \frac{\partial c_2^*}{\partial \beta} = \dots = (\delta R)^{T-1} u'(c_T^*) + \beta(\delta R)^{T-1} u''(c_T^*) \frac{\partial c_T^*}{\partial \beta} \quad (\text{A27})$$

So $\frac{\partial c_i^*}{\partial \beta}$ have same signs for all $1 \leq i \leq T-1$. We also know that $\frac{\partial c_1^*}{\partial \beta} + \frac{\partial c_2^*}{\partial \beta} \frac{1}{R} + \dots + \frac{\partial c_T^*}{\partial \beta} \frac{1}{R^{T-1}} = 0$.

So $\frac{\partial c_T^*}{\partial \beta}$ has a different sign. One can verify that it must be the case that

$$\frac{\partial c_1^*}{\partial \beta} < 0, \frac{\partial c_2^*}{\partial \beta} < 0, \dots, \frac{\partial c_{T-1}^*}{\partial \beta} < 0, \frac{\partial c_T^*}{\partial \beta} > 0 \quad (\text{A28})$$

Then,

$$\begin{aligned} \frac{\partial U^{N2}}{\partial \beta} &= u'(c_1^*) \frac{\partial c_1^*}{\partial \beta} + \delta u'(c_2^*) \frac{\partial c_2^*}{\partial \beta} + \dots + \delta^{T-1} u'(c_T^*) \frac{\partial c_T^*}{\partial \beta} \\ &= u'(c_1^*) \left(- \sum_{i=2}^T \frac{\partial c_i^*}{\partial \beta} \frac{1}{R^{i-1}} \right) + \delta u'(c_2^*) \frac{\partial c_2^*}{\partial \beta} + \dots + \delta^{T-1} u'(c_T^*) \frac{\partial c_T^*}{\partial \beta} \\ &= (1 - \beta) \delta^{T-1} u'(c_T^*) \frac{\partial c_T^*}{\partial \beta} > 0. \end{aligned} \quad (\text{A29})$$

So the welfare U^{N2} is a strictly increasing function of β . We also know that if $\beta \rightarrow 1$, the equilibrium consumption also maximizes the welfare, so $U^{N2} > U^{N1}$. We next consider the case that $\beta \rightarrow 0$. We first consider the case that $\delta R = 1$, it then follows that

$$\begin{aligned} U^{N2} &= u(c_1^*) + \dots + \delta^{T-2} u(c_{T-1}^*) + \delta^{T-1} u(0) \\ &= (1 + \delta + \dots + \delta^{T-1}) \frac{u(c_1^*) + \dots + \delta^{T-2} u(c_{T-1}^*) + \delta^{T-1} u(0)}{1 + \delta + \dots + \delta^{T-1}} \\ &= (1 + \delta + \dots + \delta^{T-1}) \frac{u(c_1^*) + \dots + \left(\frac{1}{R}\right)^{T-2} u(c_{T-1}^*) + \left(\frac{1}{R}\right)^{T-1} u(0)}{1 + \frac{1}{R} + \dots + \left(\frac{1}{R}\right)^{T-1}} \\ &< (1 + \delta + \dots + \delta^{T-1}) u \left(\frac{c_1^* + \dots + \frac{c_{T-1}^*}{R^{T-2}} + \frac{0}{R^{T-1}}}{1 + \frac{1}{R} + \dots + \frac{1}{R^{T-1}}} \right) \\ &= (1 + \delta + \dots + \delta^{T-1}) u(w) = U^{N1}, \end{aligned} \quad (\text{A30})$$

where the inequality comes from the fact that u is concave and Jensen's inequality. By continu-

ity, there exists \underline{R} and $\bar{\beta}_N$ such that if $\underline{R} < R \leq \frac{1}{\delta}$ and $\beta < \bar{\beta}_N$, the welfare under one-sided commitment is higher than that under two-sided commitment: $U^1 > U^{N2}$.

Part 2. If $R \leq r_T(\beta, \delta)$, the market breaks down under one-sided commitment. Then we know that the welfare under one-sided commitment is $U^1 = u(w) \sum_{i=0}^{T-1} \delta^i$.

Under two-sided commitment, the equilibrium consumption solves the following program:

$$\max_{c_1, \dots, c_T} u(c_1) + \beta \sum_{t=2}^T \delta^{t-1} u(c_t) \quad (\text{A31})$$

subject to

$$\sum_{t=1}^T \frac{c_t - w}{R^{t-1}} = 0 \quad (\text{A32})$$

The solution is given by $u'(c_1) = \beta \delta R u'(c_2) = \dots = \beta (\delta R)^{T-1} u'(c_T)$ and $\sum \frac{c_i - w}{R^{i-1}} = 0$. Then the welfare under two-sided commitment is $U^{S2} = \sum_{i=1}^T \delta^{i-1} u(c_i)$.

Taking derivative w.r.t β on $\frac{u'(c_1)}{\beta} = \delta R u'(c_2) = \dots = (\delta R)^{T-1} u'(c_T)$ gives

$$\frac{u''(c_1) \frac{\partial c_1}{\partial \beta} \beta - u'(c_1)}{\beta^2} = \delta R u''(c_2) \frac{\partial c_2}{\partial \beta} = \dots = (\delta R)^{T-1} u''(c_T) \frac{\partial c_T}{\partial \beta} \quad (\text{A33})$$

So $\frac{\partial c_i}{\partial \beta}$ have same signs for all $2 \leq i \leq T$. We also know that $\frac{\partial c_1}{\partial \beta} + \frac{\partial c_2}{\partial \beta} \frac{1}{R} + \dots + \frac{\partial c_T}{\partial \beta} \frac{1}{R^{T-1}} = 0$. So $\frac{\partial c_1}{\partial \beta}$ has a different sign. One can verify that it must be the case that

$$\frac{\partial c_1}{\partial \beta} < 0, \frac{\partial c_2}{\partial \beta} > 0, \dots, \frac{\partial c_{T-1}}{\partial \beta} > 0, \frac{\partial c_T}{\partial \beta} > 0 \quad (\text{A34})$$

Then,

$$\begin{aligned} \frac{\partial U^{S2}}{\partial \beta} &= u'(c_1) \frac{\partial c_1}{\partial \beta} + \delta u'(c_2) \frac{\partial c_2}{\partial \beta} + \dots + \delta^{T-1} u'(c_T) \frac{\partial c_T}{\partial \beta} \\ &= u'(c_1) \frac{\partial c_1}{\partial \beta} + \delta u'(c_2) \frac{\partial c_2}{\partial \beta} + \dots + \delta^{T-1} u'(c_T) R^{T-1} \left(- \sum_{i=1}^{T-1} \frac{\partial c_i}{\partial \beta} \frac{1}{R^{i-1}} \right) \\ &= \left(1 - \frac{1}{\beta}\right) u'(c_1) \frac{\partial c_1}{\partial \beta} > 0. \end{aligned} \quad (\text{A35})$$

So the welfare U^{S2} is a strictly increasing function of β . We also know that if $\beta \rightarrow 1$, the equilibrium consumption also maximizes the welfare, so $U^{S2} > U^1$. At the other extreme, if $\beta \rightarrow 0$, $U^{S2} = u(W)$, where $W = w \sum_{i=1}^T \frac{1}{R^{i-1}}$. Note that when $R \geq \frac{1}{\delta}$,

$$\begin{aligned}
U^{S2} &= u\left(w \sum_{i=1}^T \frac{1}{R^{i-1}}\right) \\
&\leq u\left(w \sum_{i=1}^T \delta^{i-1}\right) \\
&= u\left(w \sum_{i=1}^T \delta^{i-1}\right) + \sum_{i=1}^{T-1} \delta^i u(0) \\
&< u(w) \sum_{i=0}^{T-1} \delta^i = U^1,
\end{aligned} \tag{A36}$$

where the first inequality comes from $R \geq \frac{1}{\delta}$ and the second inequality is because of the concavity of u . By continuity, there exist $\underline{R}_1, \underline{R}_2$, and $\bar{\beta}_S$ such that when $\underline{R}_1 < R < \underline{R}_2$ and $\beta < \bar{\beta}_S$, the welfare under one-sided commitment is higher than that under two-sided commitment. \square

Proof of Proposition 3. We show the result for one-sided commitment. The proof with two-sided commitment is similar and omitted here. If $\delta R \leq 1$, the market breaks down for both time-consistent agents and naifs, implying that they consume their incomes at each period. Thus, $W(c_C^1) = W(c_N^{1E})$.

Denote $c_{C,t}^1$ and $c_{N,t}^{1E}$ the equilibrium period- t consumption for a time-consistent agents and a naif, respectively. Suppose $\delta R > 1$. First we show that $c_{C,1}^1 \leq c_{N,1}^{1E}$. Notice that $u'(c_{C,1}^1) = \dots = (\delta R)^{T-2} u'(c_{C,T-1}^1) = (\delta R)^{T-1} u'(c_{C,T}^1)$. It follows from the proof of Lemma 4 that $u'(c_{N,1}^{1E}) = \dots = (\delta R)^{T-2} u'(c_{N,T-1}^{1E}) \leq \beta (\delta R)^{T-1} u'(c_{N,T}^{1E})$. If $c_{C,1}^1 > c_{N,1}^{1E}$, $u'(c_{C,1}^1) < u'(c_{N,1}^{1E})$, which implies that $u'(c_{C,i}^1) < u'(c_{N,i}^{1E})$ for all $1 \leq i \leq T-1$, i.e. $c_{C,i}^1 - c_{N,i}^{1E} > 0$. Since $\beta < 1$, we can similarly find that $c_{C,T}^1 - c_{N,T}^{1E} > 0$. But this can't be true as it violates the zero-profit condition. So we must

have $c_{C,1}^1 \leq c_{N,1}^{1E}$, which implies that $c_{C,i}^1 \leq c_{N,i}^{1E}$ for all $1 \leq i \leq T-1$. Note that

$$\begin{aligned}
& \lim_{T \nearrow +\infty} (W(c_C^1) - W(c_N^{1E})) \\
&= \lim_{T \nearrow +\infty} \sum_{i=1}^{T-1} \delta^{i-1} [u(c_{C,i}^1) - u(c_{N,i}^{1E})] + \delta^{T-1} [u(c_{C,T}^1) - u(c_{N,T}^{1E})] \\
&\leq \lim_{T \nearrow +\infty} 0 + \delta^{T-1} [u(c_{C,T}^1) - u(c_{N,T}^{1E})] \\
&= 0,
\end{aligned}$$

where the last step uses the fact that u is bounded and $\delta < 1$. Therefore, we have shown that as the contracting horizon grows, the welfare loss of the naive agent vanishes.

$$\lim_{T \nearrow +\infty} (W(c_C^1) - W(c_N^{1E})) = 0.$$

□

Proof of Proposition 4. If $\delta R \leq 1$, the market breaks down for naifs, so welfare is unchanged with mandating a minimum consumption. Now suppose $\delta R > 1$. Denote $(c_1^1, c_2^1, \dots, c_T^1)$ the maximizer to the program V_1^A , and $(c_1^2, c_2^2, \dots, c_T^2)$ the maximizer to the program with the manage, denoted as V_1^{PA} .

We first claim that $c_T^1 \geq c_T^2$. We know that $c_1^1 < c_2^1 < \dots < c_{T-1}^1$ and $c_T^1 \geq w$. If $c_1^1 \geq \underline{c}$, then the welfare is unchanged with the mandate since none of the consumption is affected. If $c_1^1 < \underline{c}$, then c_1^2 would hit the lowest possible consumption level, \underline{c} . So $c_1^1 < c_1^2$. Assume that k is the largest index such that $c_k^1 < \underline{c}$. Since $\sum_{t=1}^T \frac{c_t^1}{R^{t-1}} = \sum_{t=1}^T \frac{c_t^2}{R^{t-1}}$, it is clear that $\sum_{t=k+1}^T \frac{c_t^1}{R^{t-1}} > \sum_{t=k+1}^T \frac{c_t^2}{R^{t-1}}$. Since the auxiliary program is dynamically consistent, both $(c_{k+1}^1, \dots, c_T^1)$ and $(c_{k+1}^2, \dots, c_T^2)$ maximize the time $(k+1)$ auxiliary program but subject to different resource constraints. More resources must lead to a weakly higher last period consumption, implying $c_T^1 \geq c_T^2$.

Note that

$$\begin{aligned}
& u(c_1^1) + \delta u(c_2^1) + \cdots + \delta^{T-1} u(c_T^1) \\
&= u(c_1^1) + \delta u(c_2^1) + \cdots + \beta \delta^{T-1} u(c_T^1) + (1 - \beta) \delta^{T-1} u(c_T^1) \\
&\geq u(c_1^2) + \delta u(c_2^2) + \cdots + \beta \delta^{T-1} u(c_T^2) + (1 - \beta) \delta^{T-1} u(c_T^1) \\
&\geq u(c_1^2) + \delta u(c_2^2) + \cdots + \beta \delta^{T-1} u(c_T^2) + (1 - \beta) \delta^{T-1} u(c_T^2) \\
&= u(c_1^2) + \delta u(c_2^2) + \cdots + \delta^{T-1} u(c_T^2),
\end{aligned} \tag{A37}$$

where the first inequality is from the fact that $(c_1^1, c_2^1, \dots, c_T^1)$ maximizes the program V_1^A , and the second inequality follows from that $c_T^1 \geq c_T^2$. So the mandate weakly decreases welfare. \square

Proof of Proposition 5. Fix an equilibrium with two-sided commitment. Let $x(s_t) \equiv u(c(s_t))$ denote the agent's utility from the consumption he gets in state s_t in this equilibrium. We claim that there exists some utility level \underline{u} to the agent, for which $\{x(s_t)\}$ solves the program:

$$\max_{\{x(s_t)\}} \sum_{t=1}^T \sum_{s_t \in \mathbb{S}_t} p(s_t | s_1) \frac{w(s_t) - u^{-1}(x(s_t))}{R^{t-1}} \tag{A38}$$

subject to

$$x(s_1) + \beta \sum_{t=2}^T \sum_{s_t \in \mathbb{S}_t(s_1)} \delta^{t-1} p(s_t | s_1) x(s_t) \geq \underline{u}. \tag{A39}$$

That is, the agent's consumption on the equilibrium path maximizes the principal's expected discounted profits subject to providing the agent's initial self a utility of at least \underline{u} . If this were not the case, the principal would be able to increase her profits while giving the agent a higher period-1 utility, so the original contract would not be an equilibrium.

This program corresponds to the maximization of a strictly concave function over a convex set, so that, by the Theorem of the Maximum, the solution is continuous in $\beta \in [0, 1]$. Because when $\beta = 0$, consuming in any period other than in the initial period is costly and does not increase the agent's utility, the agent consumes all expected PDV of income in the first period:

$c(s_1) = \sum_{t=1}^T \sum_{s_t} p(s_t|s_1) \frac{w(s_t)}{R^{t-1}}$ and $c(s_t) = 0$ for all $s_t \neq s_1$.

Next, fix an equilibrium with one-sided commitment. We claim that there exists some utility level \underline{u} to the agent, for which $\{x(s_t)\}$ solves the program:

$$\max_{\{x(s_t)\}} \sum_{t=1}^T \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \frac{w(s_t) - u^{-1}(x(s_t))}{R^{t-1}}, \quad (\text{A40})$$

subject to

$$x(s_1) + \beta \sum_{t=2}^T \sum_{s_t \in \mathbb{S}_t} \delta^{t-1} p(s_t|s_1) x(s_t) \geq \underline{u}, \quad (\text{A41})$$

and

$$x(s_{\tilde{\tau}}) + \beta \sum_{t > \tilde{\tau}} \sum_{s_t \in \mathbb{S}_t} \delta^{t-\tilde{\tau}} p(s_t|s_{\tilde{\tau}}) x(s_t) \geq V^S(s_{\tilde{\tau}}) \quad \forall s_{\tilde{\tau}} \in \mathbb{S}_{\tilde{\tau}}, \forall \tilde{\tau}, \quad (\text{A42})$$

where the outside option $V^S(s_t)$ is the utility of the best contract that the agent can obtain by signing a new contract offered at state s_t :

$$V^S(s_\tau) \equiv \max_{c(s_\tau), \{c(s_t): s_t \in \mathbb{S}_t\}} u(c(s_\tau)) + \beta \sum_{t > \tau} \sum_{s_t \in \mathbb{S}_t} \delta^{t-\tau} p(s_t|s_\tau) u(c(s_t)),$$

subject to

$$\sum_{t \geq \tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \frac{w(s_t) - c(s_t)}{R^{t-\tau}} = 0,$$

$$u(c(s_{\tilde{\tau}})) + \beta \sum_{t > \tilde{\tau}} \sum_{s_t \in \mathbb{S}_t} \delta^{t-\tilde{\tau}} p(s_t|s_{\tilde{\tau}}) u(c(s_t)) \geq V^S(s_{\tilde{\tau}}) \quad \forall s_{\tilde{\tau}} \in S_{\tilde{\tau}}(s_\tau),$$

for $\tilde{\tau} = \tau + 1, \dots, T$.

The optimal solution is continuous in $\beta \in [0, 1]$. When $\beta = 0$, the agent would like to consume in the current period as much as possible. Applying backward induction starting from states in period $T-1$, we find that the renegotiation proofness constraints bind in all continuation programs, so that $c(s_{T-1}) = w(s_{T-1})$ and $c(s_T) = w(s_T)$ for the outside options. Proceeding backwards, it follows that the solution of this program features $c(s_t) = w(s_t)$ in all states. That is, with $\beta = 0$, the agent would like to borrow as much as possible. But firms know that, after borrowing, the

agent would prefer to drop the contract instead of repaying, so they are not willing to lend. So, in equilibrium, the agent consumes his income in all states.

Comparing the solutions under one- and two-sided commitment, we find that the agent's welfare is higher with one-sided commitment consuming the endowment in every state than consuming the expected PDV of all income right away and consuming zero in all future periods if the following condition holds:

$$\sum_{t=1}^T \sum_{s_t} \delta^{t-1} p(s_t|s_1) u(w(s_t)) > u\left(\sum_{t=1}^T \sum_{s_t} p(s_t|s_1) \frac{w(s_t)}{R^{t-1}}\right),$$

By Jensen's inequality, this condition is satisfied if $R\delta \geq 1$. Because of the continuity and $R \geq 1$, there exists $\bar{\beta}_S > 0$ and $\bar{\delta}_S \equiv \frac{1}{R} < 1$ such that, if $\beta < \bar{\beta}_S$ and $\delta > \bar{\delta}_S$, the welfare with one-sided commitment dominates the welfare with two-sided commitment. We have therefore established Proposition 5. □

Proof of Proposition 6. First, the welfare with two-sided commitment approaches to \mathcal{V}_S as β approaches to zero. It suffices to show that the welfare with one-sided commitment is bounded below by \mathcal{V}_{NS} . In the remainder of the proof, we will therefore focus on the equilibrium with one-sided commitment.

We claim that for β close to zero, the equilibrium consumption equals the endowment in all last-period states: $c(s_T) = w(s_T), \forall s_T \in \mathbb{S}_T(s_1)$. To see this, consider a perturbation that shifts consumption from a state in the last period to the preceding state, that is, it increases $c(s_{T-1})$ by $\epsilon > 0$ and reduces $c(s_T)$ by $\frac{\epsilon R}{p(s_T|s_{T-1})}$ for some $s_T \in \mathbb{S}_T$ with $p(s_T|s_{T-1}) > 0$. Let W_{s_T} denote the future value of all income up to state s_T . The amount W_{s_T} is how much the agent would be able to consume at state s_T if he saves all his income from all periods for the last one. It therefore gives an upper bound on how much the agent can consume in the last period. Since there are finitely many states and $W_{s_T} < \infty$ for all s_T , we can take the uniform bound $W \equiv \max_{s_T} W_{s_T}$. This

perturbation affects the LHS of the non-lapsing constraint at state s_t by

$$\begin{aligned} & p(s_{T-1}|s_t) [u'(c(s_{T-1})) - \beta R \delta u'(c(s_T))] \delta^{T-1-t} \epsilon \\ & > p(s_{T-1}|s_t) [u'(0) - \beta R \delta u'(W_{s_T})] \delta^{T-1-t} \epsilon, \end{aligned}$$

which is positive whenever

$$\frac{u'(0)}{R \delta u'(W)} > \beta. \quad (\text{A43})$$

The perturbation has exactly the same effect on the objective function (scaled down by δ^t and multiplied by the probability of reaching state s_{T-1}). Thus, as long as β satisfies (A43), the equilibrium will have the smallest consumption possible in the last period, which is determined by the non-lapsing constraint.

Substituting $c(s_T) = w(s_T)$ in the auxiliary program, it becomes analogous to the program of a time-consistent agent except that the contracting problem ends at period $T - 1$ instead of period T :

$$\max_{\{c(s_t)\}} \sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t(s_1)} \delta^{t-1} p(s_t|s_1) u(c(s_t)),$$

subject to

$$\sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t(s_1)} p(s_t|s_1) \frac{w(s_t) - c(s_t)}{R^{t-1}} = 0,$$

and

$$\sum_{t=\tilde{\tau}}^{T-1} \sum_{s_t \in \mathbb{S}_t(s_{\tilde{\tau}})} \delta^{t-\tilde{\tau}} p(s_t|s_{\tilde{\tau}}) u(c(s_t)) \geq (V')^C(s_{\tilde{\tau}}) \text{ for all } s_{\tilde{\tau}},$$

for all $\tilde{\tau} = 2, \dots, T$, where $(V')^C(s_{\tilde{\tau}})$ denotes the outside option for the time-consistent agent in this $(T - 1)$ -period economy.

It is straightforward to verify that $(V')^C(s_1)$ is bounded below by the utility from consuming the endowment in all states. If the endowment already satisfies the non-lapsing constraints, then the result follows from revealed preference because the endowment also satisfies zero profits. If the endowment does not satisfy the non-lapsing constraints, any renegotiation of the endowment

satisfies the zero-profits condition and gives the time-consistent agent a strictly higher utility conditional on that state. So, replacing the endowment by the solution of the continuation program in all states where the non-lapsing constraints are violated leads to a profile of consumption that satisfies the constraints and gives a utility greater than the utility of consuming the endowment in each period. It thus follows by revealed preference that the solution of the program also gives a higher utility than consuming the endowment in all states.

Since the solution of a naive agent coincides with the solution of this auxiliary program, their welfare is also bounded below by the welfare from consuming their endowment in all periods \mathcal{V}_{NS} when (A43) holds. Therefore, by continuity, if $\mathcal{V}_{NS} > \mathcal{V}_S$, there exists $\bar{\beta}_N$ such that if $\beta < \bar{\beta}_N$, the welfare with one-sided commitment dominates the welfare with two-sided commitment. \square

Proof of Proposition 7. In the following we show the result for one-sided commitment. The proof with two-sided commitment is similar and omitted here. We need to show that $\lim_{T \nearrow +\infty} [W(c_C^1) - W(c_N^{1E})] = 0$.

For each parameter β , let $V^A(\beta)$ denote the maximum value attained by the solution of the auxiliary program with one-sided commitment. Notice that the feasible set is independent of β . When $\beta = 1$, the auxiliary program becomes a time-consistent agent's program, so that $V^A(1) = W(c_C^1)$. Since the objective function is strictly increasing in β , it follows from the Envelope Theorem that $V^A(\beta)$ is strictly increasing in β , so that $V^A(1) \geq V^A(0)$. Note that $W(c_N^{1E}) \geq V^A(\beta)$. To obtain the result, it suffices to show that:

$$\lim_{T \nearrow +\infty} [V^A(1) - V^A(0)] \leq 0.$$

Consider the auxiliary program with one-sided commitment when $\beta = 0$, which attains maximum value $V^A(0)$. Let $\mathbf{c}^0 \equiv \{c^0(s_t) : s_t \in S_t(s_1), 1 \leq t \leq T\}$ denote a solution to this program. Since the objective function does not depend on $c(s_T)$ when $\beta = 0$, the solution has the lowest possible value for $c(s_T)$ that still satisfies the constraints: $\mathbf{c}^0(s_T) = w(s_T)$. Substituting this equality back, we obtain the same program that determines the consumption of a time-consistent agent

with a contracting horizon consisting of the first $(T - 1)$ periods.

Recall that c_C^1 is the equilibrium consumption of a time-consistent agent. Since c_C^1 is in the feasible set, income cannot exceed consumption for any last-period state: $c_C^1(s_T) \geq w(s_T)$. Therefore, by revealed preference ($V^A(0)$ maximizes expected utility in the first $T - 1$ periods and uses weakly higher resources), we must have

$$\begin{aligned} V^A(0) &= \sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t(s_1)} \delta^{t-1} p(s_t | s_1) u(c^0(s_t)) \\ &\geq \sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t(s_1)} \delta^{t-1} p(s_t | s_1) u(c_C^1(s_t)) \\ &= V^A(1) - \delta^{T-1} \sum_{s_T \in \mathbb{S}_T(s_1)} p(s_T | s_1) u(c_C^1(s_T)), \end{aligned}$$

where the first line uses the definition of $V^A(0)$, the second line uses revealed preference, and the third line uses the definition of $V^A(1)$. Since $\delta < 1$ and u is bounded, we have

$$\lim_{T \nearrow +\infty} \delta^{T-1} \sum_{s_T \in \mathbb{S}_T(s_1)} p(s_T | s_1) u(c_C^1(s_T)) = 0,$$

which establishes that $\lim_{T \nearrow +\infty} [V^A(1) - V^A(0)] \leq 0$. □

Appendix B: Additional Results

Equivalence Between Competitive Equilibrium and Subgame Perfect Nash Equilibrium

We show that the notion of competitive equilibrium that suppresses strategic considerations is equivalent to the Subgame Perfect Nash Equilibrium of the game in which firms simultaneously offer arbitrary menus of contracts to the agent.

We first review these two approaches.

Definition 1 (Competitive equilibrium). 1. A competitive equilibrium with one-sided commitment is a profile of non-redundant contracts $\{\mathcal{C}_t\}_{t=1, \dots, T}$ such that \mathcal{C}_t maximizes the agent's

perceived utility subject to the zero profits, incentive compatibility, perceived choice, and non-lapsing constraints for each period $t = 1, \dots, T$.

2. A competitive equilibrium with two-sided commitment is a profile non-redundant contracts at time 1, \mathcal{C}_1 , which maximizes the agent's perceived utility subject to the zero profits, incentive compatibility and perceived choice constraints.

Alternatively, we can use the game theoretic framework, in which the contract is determined by a game played by different selves and principals. We focus on pure strategy. Since principals can offer any contracts at any time, the strategy is a profile of contracts $\{\mathcal{C}'_t\}_{t=1,\dots,T}$. Note that principals can offer a space of messages/menus/actions in each period and a consumption conditional on the history of messages/menus/actions sent by the agent.

Formally, a contract at time t , \mathcal{C}'_t , specify consumption on each possible state for a future time $\tau \geq t$. Denote the set of possible states by $K_{t,\tau}$, in which the first subscript reminds the timing of when the contract is offered and the second subscript reminds the decision-making time τ . The contract can specify a consumption for each different income states, so the entire possible contracting space must be greater than the state for income uncertainty. In addition, $K_{t,\tau}$ can be arbitrary. To keep analysis tractable, we assume that $K_{t,\tau}$ has a product structure. Otherwise, we can always add more states that are not never reached so that it has a product structure. Specifically, suppose $K_{t,\tau} = \mathbb{S}_\tau \times H_t$, in which H_t summaries the income-irrelevant messages/actions that the agent can send at time t . For that reason, we say H_t the income-irrelevant history. Without loss of generality, $H_1 = \emptyset$. Denote h_t a generic element in H_t . We call h_t an income-irrelevant message/action. Denote $H_\tau(h_t)$ the states that can be reached at time τ from a earlier history $h_t \in H_t$ for $\tau > t$.

Consider an agent makes a decision at time τ . Suppose the income-irrelevant messages that has been reached is $h_{\tau-1}$, which is an element in $H_{\tau-1}$. At time τ , the agent learn the income state, i.e., s_τ is realized. The agent needs to decide whether he will lapse or not, and in case that the contract is not lapsed, he also needs to decide which income-irrelevant message/action to send. Formally, the strategy can be summarized by a pair (d_τ, a_τ) , in which $d_\tau \in \{0, 1\}$ indicates whether the agent

stays with the principal or not, and a_τ is a message/action in $H_\tau(h_{\tau-1})$, which he needs to choose if he stays. If $d_\tau = 1$, then the agent stays otherwise the contract is lapsed. If there is two-sided commitment, we can without loss of generality omit d_τ .

Since the agent is time-inconsistent, each self has his own preference. He also needs to predict future self's behavior. The outcome is determined by the game played between different selves and principals. We are interested in the Subgame Perfect Nash Equilibria (SPNE). The SPNE is solved by treating the agent's decisions in each period as if it were taken by a different player (i.e., a different "self"). We now give the formal definition of SPNE.

Definition 2 (Subgame Perfect Nash Equilibria). 1. A SPNE with one-sided commitment is a profile of principals' contracts $\{C'_t\}_{t=1,\dots,T}$ and self- τ agent's decision (d_τ, a_τ) contingent on each possible history such that

- (a) (Principal's maximization) At any time t , the contract C'_t maximizes principals' profit holding the strategies of each self of the agent fixed.
- (b) (Agent's optimization) Time τ self chooses $a_\tau \in H_\tau(h_{\tau-1})$ to maximize his utility V_τ^N . Denote the utility from contract C_τ as V'_τ . The decision to stay or leave must be given by $d_\tau = 1_{V_\tau^N \geq V'_\tau}$.

2. A SPNE with two-sided commitment is a profile of principals' contract C'_1 and the agent's decision a_τ contingent on each possible history such that

- (a) (Principal's maximization) The contract C'_1 maximizes principals' profit holding the strategies of each self of the agent fixed.
- (b) (Agent's optimization) Time τ self chooses $a_\tau \in H_\tau(h_{\tau-1})$ to maximize his perceived utility.

We are now ready to present the equivalent result.

Proposition 8. *For any competitive equilibrium, there exists a SPNE that gives the agent exactly same actual consumption and perceived consumption. For any SPNE, there exists a competitive*

equilibrium that gives the agent exactly same actual consumption and perceived consumption.

Proof. For any competitive equilibrium, consider a candidate SPNE that the contracts offered by principals are identical to the ones in the competitive equilibrium and the agent never lapses and always chooses the alternative options. Since the competitive equilibrium satisfies non-lapsing, perceived choice, and incentive constraint, it is clear to see that the candidate equilibrium is indeed an SPNE. In addition, the SPNE gives the agent exactly same actual consumption and perceived consumption as the competitive equilibrium.

We now show the second part of the proposition. Suppose there is one-sided commitment. For the ease of exposition, we use the following definition.

Definition 3. Two SPNE are said to be equivalent if all selves have same actual consumption and perceived consumption.

We first show the following useful lemmas.

Lemma 5. *For any SPNE, there is an equivalent SPNE in which the agent never lapses, i.e., $d_\tau = 1, \forall \tau$.*

Proof. To prove the lemma, consider a SPNE in which the agent lapses in some period $d_\tau = 0$, replacing it with a contract C'_τ from another principal. Since the other principal cannot lose money by offering this new contract, the old principal could have offered a contract that substituted the terms of the old contract from this period on with the terms of the new contract, and the agent would have accepted to remain with the old principal. The constructed new contracts together with the agent's optimal decision forms a SPNE that is equivalent to the original SPNE. \square

Lemma 6. *For any SPNE, there is an equivalent SPNE that has at most two options following any history, i.e., $\#|H_t(h_{t-1})| \leq 2, \forall h_{t-1} \in H_{t-1}$ and $t \geq 2$.*

Proof. Note that the agent can perfectly predict their future selves' decision. Suppose $t_1 < t_2 < t_3$. Because of the agent's preference, self t_1 's prediction about self t_3 's decision coincides with self t_2 's prediction about self t_3 's decision. Restricting $H_t(h_{t-1})$ to two messages, which either the

agent would actually choose or the agent thought he would choose, does not affect the actual consumption or the perceived consumption. Put differently, if $H_t(h_{t-1})$ has at least three messages, then there is at least one of them that the agent never sends and the agent never believes other selves would send. Therefore, we can restrict the income-irrelevant message space to be at most two: one that the agent actually choose, and one that the agent thought he would choose. \square

From above lemmas, we can assume that non-lapsing constraints (i.e., $d_\tau = 1$) and non-redundancy conditions (i.e., $\#|H_t(h_{t-1})| \leq 2$) hold in any SPNE. Now consider the contracts, denoted by $\{C'_t\}_{t=1, \dots, T}$, offered by principals in the SPNE. The contract C'_t must maximizes the agent's utility subject to the zero profits, incentive compatibility, perceived choice, and non-lapsing constraints. In other words, the contracts $\{C'_t\}_{t=1, \dots, T}$ must form a competitive equilibrium with one-sided commitment. In addition, this competitive equilibrium gives the agent exactly same actual consumption and perceived consumption.

Note that in our above proof, there is nothing special about the one-sided commitment. If there is two-sided commitment, instead of considering all contracts at any time, we just focus on the contract at time 1, and all our proof are same. \square

Welfare Effect of a Minimum Consumption Policy for Sophisticates

In this appendix, we show that the welfare effect of a minimum consumption policy can be ambiguous when the agent are sophisticated. To do that, we present two simple examples, with one showing that the policy increases welfare, and the other one showing that the policy decreases welfare.

Suppose $T = 3, \beta = 0.9, R = 1.1, \delta = 1, w = 1, u(c) = \log(c)$. We first solve the problem without the mandate. We can find the equilibrium contract by solving $u'(c_1) = \beta R u'(c_2) = \beta R^2 u'(c_3)$, which gives $c_1^1 = 0.977, c_2^1 = 0.9672, c_3^1 = 1.0639$. We can verify that all the con-

straints hold:

$$c_1^1 + \frac{c_2^1}{R} + \frac{c_3^1}{R^2} = 2.7355 = 1 + \frac{1}{R} + \frac{1}{R^2}, \quad (\text{B1})$$

$$u(c_2^1) + \beta u(c_3^1) = 0.0224 > 0 = V_2^S. \quad (\text{B2})$$

Thus, the welfare without the mandate is $W_1 = u(c_1^1) + u(c_2^1) + u(c_3^1) = 0.0053$.

Example of the mandate increasing welfare: Now consider the problem with the mandate $\underline{c} = 0.97$. Then $c_2^2 = \underline{c} = 0.97$. Solving $\max u(c_1) + \beta u(c_3)$ gives $c_1^2 = 0.9756$ and $c_3^2 = 1.0625$. We can verify that all the constraints hold:

$$c_1^2 + \frac{c_2^2}{R} + \frac{c_3^2}{R^2} = 2.7355 = 1 + \frac{1}{R} + \frac{1}{R^2}, \quad (\text{B3})$$

$$u(c_2^2) + \beta u(c_3^2) = 0.0241 > 0 = V_2^S. \quad (\text{B4})$$

The welfare with the mandate is $W_2 = u(c_1^2) + u(c_2^2) + u(c_3^2) = 0.0055$. Thus, the mandate strictly increases welfare.

Example of the mandate decreasing welfare: Assume that now $\underline{c} = 0.98$. So $c_1^2 \geq \underline{c} > c_1^1$. We can prove that the mandate strictly decreases welfare by the following:

$$\begin{aligned} u(c_1^1) + u(c_2^1) + u(c_3^1) &= \frac{1}{\beta} (u(c_1^1) + \beta(u(c_2^1) + u(c_3^1))) - \frac{1-\beta}{\beta} u(c_1^1) \\ &\geq \frac{1}{\beta} (u(c_1^2) + \beta(u(c_2^2) + u(c_3^2))) - \frac{1-\beta}{\beta} u(c_1^1) \\ &> \frac{1}{\beta} (u(c_1^2) + \beta(u(c_2^2) + u(c_3^2))) - \frac{1-\beta}{\beta} u(c_1^2) \\ &= u(c_1^2) + u(c_2^2) + u(c_3^2), \end{aligned} \quad (\text{B5})$$

where the first inequality comes from the fact that (c_1^1, c_2^1, c_3^1) is the solution without the mandate, and the second inequality comes from that $c_1^1 < c_1^2$.

References

- Manuel Amador, Iván Werning, and George-Marios Angeletos. Commitment vs. flexibility. *Econometrica*, 74(2):365–396, 2006.
- Stephen A Atlas, Eric J Johnson, and John W Payne. Time preferences and mortgage choice. *Journal of Marketing Research*, pages jmr–14, 2017.
- Ned Augenblick, Muriel Niederle, and Charles Sprenger. Working over time: Dynamic inconsistency in real effort tasks. *Quarterly Journal of Economics*, 130(3):1067–1115, 2015.
- Oren Bar-Gill. The law, economics and psychology of subprime mortgage contracts. *Cornell Law Review*, 94(5):1073–1152, 2009.
- B. Douglas Bernheim, Debraj Ray, and Sevin Yeltekin. Poverty and self-control. *Econometrica*, 83(5):1877–1911, 2015.
- Philip Bond and Gustav Sigurdsson. Commitment contracts. *Review of Economic Studies*, Forthcoming, 2017.
- Jan K. Brueckner. Borrower mobility, adverse selection, and mortgage points. *Journal of Financial Intermediation*, 3(4):416–441, 1994.
- Gharad Bryan, Dean Karlan, and Scott Nelson. Commitment devices. *Annual Review of Economics*, 2(1):671–698, 2010.
- Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O’Donoghue, and Matthew Rabin. Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania law review*, 151(3):1211–1254, 2003.
- Hector Chade, Pavlo Prokopovych, and Lones Smith. Repeated games with present-biased preferences. *Journal of Economic Theory*, 139(1):157–175, 2008.

- John H Cochrane. Time-consistent health insurance. *Journal of Political Economy*, pages 445–473, 1995.
- Michael Conlin, Ted O’Donoghue, and Timothy Vogelsang. Projection bias in catalog orders. *American Economic Review*, 97(4):1–33, 2007.
- Stefano DellaVigna and Ulrike Malmendier. Contract design and self-control: Theory and evidence. *The Quarterly Journal of Economics*, pages 353–402, 2004.
- Georges Dionne and Neil A Doherty. Adverse selection, commitment, and renegotiation: Extension to and evidence from insurance markets. *Journal of Political Economy*, 102(2):209–235, 1994.
- Kfir Eliaz and Ran Spiegler. Contracting with diversely naive agents. *Review of Economic Studies*, 73(3):689–714, 2006.
- Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.
- Simone Galberti. Commitment, flexibility, and optimal screening of time inconsistency. *Econometrica*, 83(4):1425–1465, 2015.
- John Gathergood and Jörg Weber. Financial literacy, present bias and alternative mortgage products. *Journal of Banking & Finance*, 78:58–83, 2017.
- Andra Ghent. Subprime mortgages, mortgage choice, and hyperbolic discounting. *Unpublished paper, Zicklin School of Business*, 2011.
- Michael D Grubb. Selling to overconfident consumers. *American Economic Review*, 99(5):1770–1807, 2009.
- Michael D Grubb. Overconfident consumers in the marketplace. *The Journal of Economic Perspectives*, 29(4):9–35, 2015.

- Marina Halac and Pierre Yared. Fiscal rules and discretion under persistent shocks. *Econometrica*, 82(5):1557–1614, 2014.
- Ben Handel, Igal Hendel, and Michael D Whinston. The welfare effects of long-term health insurance contracts. *Working paper*, 2017.
- Paul Heidhues and Botond Kőszegi. Exploiting naivete about self-control in the credit market. *The American Economic Review*, pages 2279–2303, 2010.
- Igal Hendel and Alessandro Lizzeri. The role of commitment in dynamic contracts: Evidence from life insurance. *Quarterly Journal of Economics*, 118(1):299–327, 2003. ISSN 00335533.
- Botond Kőszegi. Behavioral contract theory. *Journal of Economic Literature*, 52(4):1075–1118, 2014.
- Dirk Krueger and Harald Uhlig. Competitive risk sharing contracts with one-sided commitment. *Journal of Monetary Economics*, 53(7):1661–1691, 2006.
- David Laibson. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–477, 1997.
- George Loewenstein, Ted O’Donoghue, and Matthew Rabin. Projection bias in predicting future utility. *The Quarterly Journal of Economics*, 118(4):1209–1248, 2003.
- Igor Makarov and Guillaume Plantin. Equilibrium subprime lending. *Journal of Finance*, 68(3): 849–879, 2013.
- Ted O’Donoghue and Matthew Rabin. Doing it now or later. *American Economic Review*, 89(1): 103–124, 1999.
- Ted O’Donoghue and Matthew Rabin. Choice and procrastination. *Quarterly Journal of Economics*, 116(1):121–160, 2001.

Christopher A Parsons and Edward D Van Wesep. The timing of pay. *Journal of Financial Economics*, 109(2):373–397, 2013.

Mark V Pauly, Howard Kunreuther, and Richard Hirth. Guaranteed renewability in insurance. *Journal of Risk and Uncertainty*, 10(2):143–156, 1995.

Kathrin Schlafmann. Housing, mortgages, and self control. Technical report, Stockholm University, 2016.

Richard Stanton and Nancy Wallace. Mortgage choice: What’s the point? *Real Estate Economics*, 26(2):173–205, 1998.

Online Appendix - Not For Publication

Welfare comparison when the market breaks down for sophisticates. Letting $c_{N,t}^{1E}$ denote $c_{N,t}^1$ denote the consumption on the equilibrium path of a naif and a sophisticate in period t , we have $c_N^{1E} \geq w = c_{S,T}^1$. Then,

$$\begin{aligned}
 W(c_N^{1E}) &= u(c_{N,1}^{1E}) + \delta u(c_{N,2}^{1E}) + \cdots + \delta^{T-1} u(c_{N,T}^{1E}) \\
 &= u(c_{N,1}^{1E}) + \delta u(c_{N,2}^{1E}) + \cdots + \delta^{T-2} u(c_{N,T-1}^{1E}) + \beta \delta^{T-1} u(c_{N,T}^{1E}) + (1 - \beta) \delta^{T-1} u(c_{N,T}^{1E}) \\
 &\geq u(c_{S,1}^1) + \delta u(c_{S,2}^1) + \cdots + \delta^{T-2} u(c_{S,T-1}^1) + \beta \delta^{T-1} u(c_{S,T}^1) + (1 - \beta) \delta^{T-1} u(c_{S,T}^1) \\
 &= u(c_{S,1}^1) + \delta u(c_{S,2}^1) + \cdots + \delta^{T-1} u(c_{S,T}^1) \\
 &= W(c_S^1),
 \end{aligned} \tag{OA1}$$

where the second line adds and subtracts $\beta \delta^{T-1} u(c_{N,T}^{1E})$, the third line uses the fact that c_N^{1E} solves the auxiliary problem and $c_{N,T}^{1E} \geq c_{S,T}^1$, and the fourth line rearranges terms. Therefore, the naive agent obtains a higher welfare than the sophisticate when the market breaks down for the sophisticated agent.

Proof of Lemma 3 for General Income Distributions. Suppose $h^t = (s_t, a)$, where s_t is the uncertainty state and a is the messages up to time t . For the ease of exposition, we define the projection of state spaces to the uncertainty state and the messages/actions state.

$$h^t|_s = s_t, h^t|_a = a.$$

Denote $A_k(h^t)$ denote a path of actions up to time $(t + k)$ in which the first t actions are given by $h^t|_a$ and the last k actions are given by the alternative option (A). Similarly, denote $B_k(h^t)$ denote a path of actions up to time $(t + k)$ in which the first t actions are given by $h^t|_a$ and the last k actions are given by the baseline option (B).

We are now ready to write down the objective function and constraints. The objective function

is the value of agent's perceived utility:

$$u(c(s_1)) + \beta \sum_{t>1} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \delta^{t-1} u(c(s_t, B_{t-1}(\emptyset))). \quad (\text{OA2})$$

Perceived choice constraints (PCC) at time $\tilde{\tau}$ requires that the agent believes that at time $\tau > \tilde{\tau}$ they would choose the baseline option over the alternative option, using time-consistency parameter $\hat{\beta}$.

$$\begin{aligned} & u(c(s_\tau, B(h^{\tau-1}))) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t+1-\tau}(h^{\tau-1}))) \\ & \geq u(c(s_\tau, A(h^{\tau-1}))) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t-\tau}(A(h^{\tau-1}))))). \end{aligned} \quad (\text{OA3})$$

Incentive constraints (IC) at time $\tau > 1$ requires that the agent decides to choose the alternative option when they use their true time-consistency parameter β .

$$\begin{aligned} & u(c(s_\tau, A(h^{\tau-1}))) + \beta \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t-\tau}(A(h^{\tau-1})))) \\ & \geq u(c(s_\tau, B(h^{\tau-1}))) + \beta \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t+1-\tau}(h^{\tau-1}))). \end{aligned} \quad (\text{OA4})$$

If there is one-sided commitment, then we also have non-lapsing constraints. Depending on the timing of evaluating the outside option, there are two type of non-lapsing constraints: the actual non-lapsing constraint that is evaluated *at* the time of decision-making self and the perceived non-lapsing constraint that is evaluated *before* the time of decision-making self. Suppose $\tau > 1$.

$$u(c(s_\tau, B(h^{\tau-1}))) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t+1-\tau}(h^{\tau-1}))) \geq \hat{V}(s_\tau, h^{\tau-1}), \quad (\text{OA5})$$

$$u(c(s_\tau, A(h^{\tau-1}))) + \beta \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t-\tau}(A(h^{\tau-1})))) \geq V(s_\tau, h^{\tau-1}), \quad (\text{OA6})$$

where $\hat{V}(s_\tau, h^{\tau-1})$ denotes the perceived outside option and $V(s_\tau, h^{\tau-1})$ denotes the actual outside

option.

The feasible contracts associated with one-sided commitment are given by the sets of contracts that satisfy (OA3), (OA4), (OA5), and (OA6). The feasible contracts associated with two-sided commitment are given by the sets of contracts that satisfy (OA3) and (OA4).

Since on the equilibrium path, the agent always chooses option A . We say h^t is on the equilibrium path, if the actions $h^t|_a$ only include option A .

In the following, assume $\beta > 0$. We first note that the incentive constraints (OA4) must be binding on the equilibrium path, because otherwise we can increase $c(s_T, B_{T+1-\tau}(h^{\tau-1}))$ without affecting all other constraints while weakly increase the agent's perceived utility. Given (OA4) are binding, we can simplify (OA3) as

$$u(c(s_\tau, B(h^{\tau-1}))) \leq u(c(s_\tau, A(h^{\tau-1}))). \quad (\text{OA7})$$

Substituting the binding IC constraints in the objective gives

$$\begin{aligned} & \sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \delta^{t-1} u(c(s_t, B_{t-1}(\emptyset))) + \beta \sum_{s_T \in \mathbb{S}_T} p(s_T|s_1) \delta^{T-1} u(c(s_T, B_{T-1}(\emptyset))) \\ & + (\beta - 1) \sum_{t=2}^{T-1} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \delta^{t-1} u(c(s_t, A(B_{t-2}(\emptyset)))) \end{aligned}$$

Since $\beta < 1$, we want to choose $c(s_t, A(B_{t-2}(\emptyset)))$ as small as possible (subject to the constraints).

We now show that we can set them at their lowest possible values (zero, by our normalization): $c(s_t, A(B_{t-2}(\emptyset))) = 0$. First, the PCC (OA3) holds because (OA7) holds.

We then verify that the perceived non-lapsing constraints hold if actual non-lapsing constraints (OA6) hold. Suppose $\{\hat{c}(s_t) : t \geq \tau\}$ solves the perceived outside option program $\hat{V}(s_\tau, h^{\tau-1})$. So we have

$$\hat{V}(s_\tau, h^{\tau-1}) = u(\hat{c}(s_\tau, B(h^{\tau-1}))) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(\hat{c}(s_t, B_{t+1-\tau}(h^{\tau-1}))). \quad (\text{OA8})$$

We next verify the perceived non-lapsing constraints. Note that

$$\begin{aligned} & u(c(s_\tau, B(h^{\tau-1}))) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t+1-\tau}(h^{\tau-1}))) \\ &= u(0) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t+1-\tau}(h^{\tau-1}))) \end{aligned} \quad (\text{OA9})$$

$$= (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta} \left(u(c(s_\tau, A(h^{\tau-1}))) + \beta \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(c(s_t, B_{t-\tau}(A(h^{\tau-1})))) \right) \quad (\text{OA10})$$

$$\geq (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta} V(s_\tau, h^{\tau-1}) \quad (\text{OA11})$$

$$\geq (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta} \left(u(\hat{c}(s_\tau, B(h^{\tau-1}))) + \beta \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(\hat{c}(s_t, B_{t+1-\tau}(h^{\tau-1}))) \right) \quad (\text{OA12})$$

$$= (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta} u(\hat{c}(s_\tau, B(h^{\tau-1}))) + \hat{\beta} \sum_{t>\tau} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_\tau) \delta^{t-\tau} u(\hat{c}(s_t, B_{t+1-\tau}(h^{\tau-1}))) \quad (\text{OA13})$$

$$= (1 - \frac{\hat{\beta}}{\beta})u(0) + (\frac{\hat{\beta}}{\beta} - 1)u(\hat{c}(s_\tau, B(h^{\tau-1}))) + \hat{V}(s_\tau, h^{\tau-1}) \quad (\text{OA14})$$

$$\geq \hat{V}(s_\tau, h^{\tau-1}), \quad (\text{OA15})$$

where (OA9) is from $c(s_\tau, B(h^{\tau-1})) = 0$, (OA10) is from the IC constraints (OA4), (OA11) is from the actual non-lapsing constraints (OA6), (OA12) is due to the revealed preference argument since \hat{c} is also a candidate contract for the program $V(s_\tau, h^{\tau-1})$, (OA13) is a simple algebra, (OA14) comes from the definition of $\hat{V}(s_\tau, h^{\tau-1})$ is (OA8), and finally (OA15) is because of $\hat{c}(s_\tau, B(h^{\tau-1})) \geq 0$.

So far, we have shown that we can drop the perceived choice constraints and the perceived non-lapsing constraints. Consequently, the program reduces to the following auxiliary program.

$$\max \sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \delta^{t-1} u(c(s_t, B_{t-1}(\emptyset))) + \beta \sum_{s_T \in \mathbb{S}_T} p(s_T|s_1) \delta^{T-1} u(c(s_T, B_{T-1}(\emptyset)))$$

subject to the zero-profit condition and the actual non-lapsing constraints. So $c^{1E} = c^{1A}$.

If there is two-sided commitment, using the same argument, we can rewrite the program as the auxiliary program with two-sided commitment. So we have $c^{2E} = c^{2A}$. □