# Long-Term Contracting with Time-Inconsistent Agents*

Daniel Gottlieb† and Xingtan Zhang‡

November 16, 2018

**Abstract**

We study contractual relationships between (partially naive) time-inconsistent consumers and risk neutral firms in settings with one- and two-sided commitment. Our main result is that as the number of periods grows, the welfare loss from time-inconsistency vanishes. We use our results to study two common regulatory interventions: removing commitment power from consumers and imposing limits on the fees that firms can charge. For each fixed contracting horizon, removing commitment power increases welfare when consumers are sufficiently time-inconsistent. However, removing commitment power cannot help if the contracting horizon is long. With one-sided commitment, setting a maximum fee weakly hurts consumers.

JEL: D81, D86, D91, I18

*We thank Douglas Bernheim, David Dillenberger, Drew Fudenberg, Faruk Gul, Marina Halac, Botond Koszegi, Shengwu Li, George Mailath, Pietro Ortoleva, Wolfgang Pesendorfer, Philipp Strack, Dmitry Taubinsky, Jeremy Tobacman, Juuso Toikka, Edward Van Wesep, Leeat Yariv, Muhamet Yildiz, and seminar participants at Harvard-MIT, Princeton, UC Berkeley, University of Colorado Boulder, Wharton, and at the SITE Psychology and Economics Workshop for comments. Gottlieb is grateful for the hospitality of the Economics Department at Harvard University, where part of this research was conducted. Zhang acknowledges financial support from de Soto Capital Markets Program at University of Colorado Boulder.

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

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

# 1 Introduction

A large literature in behavioral economics has studied markets with present-biased consumers who underestimate their bias ("partial naiveté"). An important finding from this literature is that the equilibrium is inefficient and regulation that accounts for internalities can increase welfare.[1] Models in this literature generally assume that there are only three periods, which is the minimum needed for dynamic inconsistency to play a role. But this is an unrealistic assumption since, in these models, periods are thought to be very short, typically no more than a day (O'Donoghue and Rabin, 2015).

This paper considers a general contracting model with present-biased consumers and an arbitrary number of periods. To explore consumer naiveté, equilibrium contracts offer two options at each point in time: a front-loaded and a back-loaded option. Consumers think they will pick the front-loaded option but pick the back-loaded one, effectively postponing payments to the next period. Payments are postponed until the last period, when they cannot be postponed any longer. As a result, the equilibrium has smooth consumption in all but the last period. Because the relative weight on the last period shrinks as the number of periods grows, the consumption path of present-biased consumers converges to the path that maximizes their long-term preferences. Therefore, the welfare loss from present bias vanishes as the contracting horizon grows.

We use our framework to study the effect of removing commitment power from consumers. Many markets have one-sided commitment (e.g., mortgages, car loans, life insurance, long-term care, and annuities).[2] Regulations that allow consumers to terminate agreements at will are often motivated by attempts to protect them. But in standard models, removing a rational consumer's commitment power can only hurt the consumer. Our setting is a natural candidate for studying the effect regulating commitment power, because committing to a future action and lapsing on a previous agreement are inherently inter-temporal decisions, and present bias is the most well-studied bias in intertemporal decision-making. Moreover, there is evidence that present bias is an important feature in some credit markets where regulation prevents consumers from being able to commit to long-term contracts, such as in mortgage or credit card markets.[3]

For a fixed horizon, removing commitment power helps consumers who are sufficiently time inconsistent. This is because time-inconsistent consumers are tempted to overborrow. Since firms would not lend to consumers who can walk away from contracts, removing commitment power restricts their access to savings, which increases welfare when consumers are sufficiently time

[1]See, for example, Gruber and Koszegi (2001); O'Donoghue and Rabin (2003); DellaVigna and Malmendier (2004); Heidhues and Kőszegi (2010).

[2]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 markets – 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.

[3]Schlafmann (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 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. For credit card markets, see Meier and Sprenger (2010).

inconsistent. This result is in line with regulation, which often mentions consumer protection as the reason for allowing them to terminate agreements at will, although it contrasts with the intuition that the provision of commitment devices is necessarily welfare improving.

We then generalize the vanishing inefficiency result for settings with one-sided commitment. In this case, the equilibrium converges to the path that maximizes the consumer's long-term preferences subject to renegotiation proofness constraints. Since, with commitment, the equilibrium maximizes long-term preferences without these additional constraints, removing commitment power does not help when the contracting horizon is long enough.

In the third part of the paper, we turn to the special case of deterministic endowments. This setting shuts down the effect of risk, allowing us to isolate the effect of smoothing consumption over time. We show that 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 consumers with the same "average impatience," the time-consistent consumer discounts periods further in the future by more than a time-inconsistent consumer. As a result, the time-inconsistent consumer is less hurt by front-loading payments from the periods sufficiently far in the future, which helps supporting long-term contracts.

Our last result concerns the effect of limiting the fee that firms can charge. This type of policy has been popular among regulators as a way to protect consumers.[4] We show that with one-sided commitment, limiting the fee that firms can charge is never welfare improving. The intuition is that when interest rates are low so customers would like to borrow, one-sided commitment prevents them from obtaining a long-term contract. Then, 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, imposing a maximum fee reduces savings, moving the equilibrium further away from the optimum.

The main message of our paper is that contract length is a key variable for the inefficiency of markets with dynamically inconsistent consumers. There are two possible interpretations for our main result. If one takes the models as currently formulated as good approximations of reality, then our results suggest that there is no role for regulation that corrects for present bias as long as contractual relationships are long enough. If, instead, one believes that inefficiency is a prevalent feature of actual markets with dynamically inconsistent consumers, then our results highlight that something must be missing from how these markets are typically modeled.

## Related Literature

Our paper fits into a recent literature on contracting with behavioral agents, summarized in Kőszegi (2014) and Grubb (2015). We build on the credit card model of Heidhues and Kőszegi (2010) by considering more than two consumption periods, allowing for uncertainty, and considering both two- and one-sided commitment.

[4]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, state-level nonforfeiture laws specify minimum payments that must be made to customers who surrender their permanent life insurance policies or annuities, with each state following a slight variation of the general guidelines from the National Association of Insurance Commissioners (NAIC). 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.

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 have commitment power but, because they face an unverifiable taste shock, they value the flexibility to adjust to different taste shocks. In Section 3, we consider a different incentive aspect: the agent's incentive to lapse and re-contract with other firms. Moreover, these papers study sophisticated agents, whereas our main focus is on partially naive agents.[5]

Finally, our paper is related to a literature on dynamic risk-sharing with one-sided commitment. Several papers show that front-loaded payment schedules help mitigate a consumer's lack of commitment power.[6] For example, Hendel and Lizzeri (2003) theoretically and empirically examine how life insurers mitigate reclassification risk by offering front-loaded policies. Similarly, several researchers show that mortgages are front loaded to mitigate prepayment risk.[7] 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. The main difference between these models and ours is that we assume that consumers are dynamically inconsistent.

The paper proceeds as follows. We first consider a general model with arbitrary income paths. In Section 2, we present the model with commitment. In Section 3, we introduce one-sided commitment and discuss the effect of removing commitment power on welfare. Then, in Section 4, we move to the special case of a constant income. Section 5 concludes. Extensions for the case of monopolistic firms and general (not quasi-hyperbolic) discounting, as well as all proofs are presented in the appendix.

# 2 Model with Commitment

There is one consumer (agent) and at least two firms (principals). Time is discrete and finite. To allow for arbitrary non-stationary settings, we model the stochastic environment as follows. There is a finite state space $\mathbb{S}_t$ for each $t \in \mathbb{N}$. The agent earns income $w(s_t)$ at state $s_t$. Let $p(s_t|s_\tau)$ denote the probability of reaching state $s_t$ conditional on state $s_\tau$. We say that state $s_t$ follows state $s_\tau$ if $p(s_t|s_\tau) > 0$. A state specifies all previously realized uncertainty, so a state cannot follow two different states. We consider the $T$-period truncation of this setting; that is, an environment with state spaces $\mathbb{S}_t$ and conditional probabilities $p(\cdot|\cdot)$ up to period $T$, at which point the game ends.

Without loss of generality, we assume that no uncertainty is realized before the initial period: $\mathbb{S}_1 = \{\emptyset\}$. Let $E[\cdot|s_t]$ denote the expectation operator conditional on state $s_t$ and let $E[\cdot]$ denote the unconditional (time-1) expectation. By taking degenerate distributions, our framework allows for deterministic income paths. Also, since the probabilities of reaching future states may depend on

[5] Our paper is also related to Bisin et al. (2015), who study the interaction between government policy and private commitments by present-biased voters and to Harris and Laibson (2001) and Cao and Werning (forthcoming), who study the Markov equilibria in infinite-horizon problems with sophisticated consumers and show there can be multiple non-smooth equilibria. Multiplicity and non-smoothness do not arise in our setting because our model has a finite (albeit arbitrary) horizon.

[6] This 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).

[7] See, e.g., Brueckner (1994) and Makarov and Plantin (2013)

the current state, our framework also allows for persistent shocks, which is important to encompass environments with reclassification risk.

Firms are risk neutral and can freely save or borrow at the interest rate $R \geq 1$, so that each firm maximizes its expected discounted profits. The expected profits at state $s_\tau$ of a firm who collects state-dependent payments $\{\pi(s_t)\}_{t\geq\tau}$ are

$$E\left[\sum_{t\geq\tau}\frac{\pi(s_t)}{R^{t-\tau}}\;\middle|\;s_\tau\right].$$

The agent has quasi-hyperbolic discounting and needs a firm to transfer consumption across states.[8] At state $s_\tau$, the agent evaluates the state-dependent consumption $\{c(s_t)\}_{t\geq\tau}$ according to

$$u(c(s_\tau)) + \beta E\left[\sum_{t>\tau}\delta^{t-s}u(c(s_t))\;\middle|\;s_\tau\right], \tag{1}$$

where $\beta \in (0,1)$ and $u : \mathbb{R}_+ \to \mathbb{R}$ is strictly increasing, strictly concave, and twice continuously differentiable. We are interested in time-inconsistent consumers who underestimate their bias – i.e., they are *partially naive* as defined by O'Donoghue and Rabin (1999). Such a consumer believes that, in all future periods, he will behave like someone with time-consistency parameter $\hat{\beta} \in (\beta, 1]$. For brevity, we will refer to a partially naive time-inconsistent consumer simply as a *time-inconsistent consumer*.[9] As a benchmark, we also consider the case of time-consistent consumers ($\hat{\beta} = \beta = 1$). Following most of the literature, we take the agent's long-run preferences as the relevant ones in our welfare calculations.[10] So, consumers maximize welfare in the time-consistent benchmark but not when they are time-inconsistent.

For simplicity, we will assume that firms know the consumer's preferences. This assumption is relaxed in Appendix B, where we allow firms not to know the consumer's naiveté parameter $\hat{\beta}$. For now, we also assume that the consumer has all bargaining power and that both parties are able to commit to long-term contracts, so the consumer offers a take-it-or-leave-it contract in the first period, which is honored until the game ends.

Whenever the consumer is not time consistent, his ranking of consumption streams depends on when the streams are evaluated. As usual, we model the behavior of such an agent by treating his decision in each period as if it was decided by a different "self." Because the consumer is naive, each self may mispredict how his future selves will choose. We are interested in Subgame Perfect Nash Equilibria (SPNE) of this game.[11]

[8]Equivalently, the income process $w(\cdot)$ can be interpreted as the smoothest consumption that the consumer can obtain without interacting with the firms in the model. In this interpretation, $w(s_t)$ includes the amount that the individual can borrow or save from other sources at state $s_t$.

[9]Subsection 4.3 considers sophisticates, who fully understand their time-inconsistency ($\hat{\beta} = \beta$).

[10]See, e.g., DellaVigna and Malmendier (2004); O'Donoghue and Rabin (1999, 2001).

[11]Our game-theoretical equilibrium concept coincides with the non-strategic competitive equilibrium of Heidhues and Kőszegi (2010). We formulate the model as a game because it can be more straightforwardly generalized to settings with one-sided commitment, as we do in Section 3.

## 2.1 Time-Consistent Consumers

As a benchmark, we first consider a time-consistent consumer. Because parties can commit to long-term contracts, the equilibrium consumption maximizes the agent's utility in period 1,

$$E\left[\sum_{t=1}^{T}\delta^{t-1}u\left(c\left(s_{t}\right)\right)\right], \tag{2}$$

subject to the zero profits constraint,

$$\sum_{t=1}^{T}E\left[\frac{w\left(s_{t}\right)-c\left(s_{t}\right)}{R^{t-1}}\right]=0. \tag{3}$$

Indeed, no firm would accept a contract with negative expected profits. If profits were positive, the agent would benefit by offering a contract with a slightly higher consumption. Because the objective function in (2) is strictly concave and (3) is a linear constraint, there is a unique solution. So, any SPNE of the game provides the same consumption, which solves the program above. Let $W^{C}$ denote the equilibrium welfare of the time-consistent consumer, which evaluates the objective (2) at the equilibrium consumption.

## 2.2 Time-Inconsistent Consumers

We now turn to time-inconsistent consumers. Any contract that is accepted with positive probability must maximize the consumer's utility in period 1 subject to two types of constraints: zero profits, which is the same as before, and incentive constraints, which are due to consumer naiveté.

Because the consumer mispredicts his future preferences, he may disagree with the firm about the actions that his future selves will take. So we need to distinguish between what the consumer believes that he will choose and what firms believe that the consumer will choose (which we interpret as the correct beliefs). This disagreement gives rise to two sets of incentive constraints. Following Heidhues and Kőszegi (2010), we refer to them as *perceived choice constraints (PCC)* and *incentive compatibility constraints (IC)*.

PCC requires the consumer to believe that his future selves will choose the actions that maximize his perceived utility. IC requires firms to believe that the consumer's future selves will choose the actions that maximize the consumer's true utility. The option that the consumer thinks that his future selves will choose is called the *baseline* option (B). The option that firms think that the consumer's future selves will choose is called the *alternative* option (A). In principle, these options can coincide, in which case the consumer and the firms agree about which actions will be chosen. But we will show that these options are always different in equilibrium.

A time-$t$ option history $h^{t}$ is a list of options chosen by the consumer up to time $t$: $h^{1}=\emptyset$, $h^{2}\in\{A,B\}$, $h^{3}\in\{AA,AB,BA,BB\}$, etc. 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 option histories when there are four periods.

An *equilibrium consumption vector* is the vector of state-dependent consumption in all option

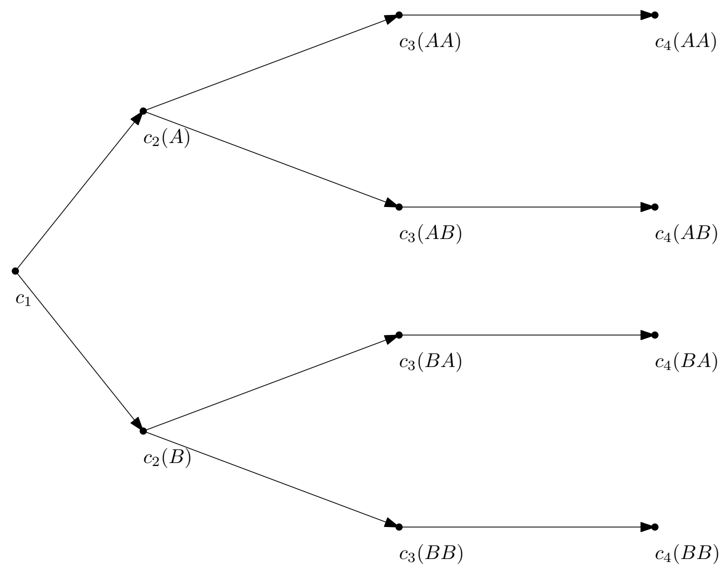


Figure 1: **Option histories**. The figure represents option histories when $T = 4$ and there is no uncertainty. The consumer thinks that he will choose the baseline option (B) in each node but ends up choosing the alternative option (A). So, the consumer initially believes that his consumption stream will be $(c_1, c_2(B), c_3(BB), c_4(BB))$. In period 2, he deviates to $c_2(A)$, 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, getting $c_3(AA)$ and $c_4(AA)$ in periods 3 and 4. With uncertainty, consumption also depends on the state of the world.

histories for all states that happen with positive probability:[12]

$$\mathbf{c} \equiv \{(c(s_1), c(s_2, h^2), c(s_3, h^3), \cdots, c(s_T, h^T)) : \ p(s_2|s_1)p(s_3|s_2)\cdots p(s_T|s_{T-1}) > 0\}.$$

A *consumption on the equilibrium path* is a vector of state-contingent consumption that happens with positive probability (using correct beliefs about the options that the consumer chooses):

$$\mathbf{c}^E \equiv \{(c(s_1), c(s_2, A), c(s_3, A, A), \cdots, c(s_T, A, ..., A)) : \ p(s_2|s_1)p(s_3|s_2)\cdots p(s_T|s_{T-1}) > 0\}.$$

Unlike the equilibrium consumption vector, the consumption on the equilibrium path only includes outcomes conditional on the consumer repeatedly picking option A.

The *equilibrium program* (P) is:

$$\max_{\{c(s_t,h^t)\}} u(c(s_1)) + \beta E\left[\sum_{t=2}^{T} \delta^{t-1} u(c(s_t, B, B, ..., B))\right],$$

subject to

$$\sum_{t=1}^{T} E\left[\frac{w(s_t) - c(s_t, A, A, ..., A)}{R^{t-1}}\right] = 0, \qquad \text{(Zero Profits)}$$

[12] Since a state of the world encodes all uncertainty realized up to that period, the distribution over future states conditional on $s_t$ can only have full support in the trivial case where no uncertainty was realized until state $s_t$.

$$u(c(s_\tau,(h_\tau,B))) + \hat{\beta} E\left[\left.\sum_{t>\tau} \delta^{t-\tau} u\left(c\left(s_t,(h_\tau,B,B,...,B)\right)\right)\right| s_\tau\right] \tag{PCC}$$

$$\geq u(c(s_\tau,(h_\tau,A))) + \hat{\beta} E\left[\left.\sum_{t>\tau} \delta^{t-\tau} u\left(c\left(s_t,(h_\tau,A,B,...,B)\right)\right)\right| s_\tau\right],$$

and

$$u(c(s_\tau,(h_\tau,A))) + \beta E\left[\left.\sum_{t>\tau} \delta^{t-\tau} u\left(c\left(s_t,(h_\tau,A,B,...,B)\right)\right)\right| s_\tau\right] \tag{IC}$$

$$\geq u(c(s_\tau,(h_\tau,B))) + \beta E\left[\left.\sum_{t>\tau} \delta^{t-\tau} u\left(c\left(s_t,(h_\tau,B,B,...,B)\right)\right)\right| s_\tau\right].$$

The following lemma establishes that the equilibrium program (P) characterizes the equilibrium consumption vector:

**Lemma 1.** ***c*** *is an equilibrium consumption vector if and only if it solves program (P).*

### 2.2.1 Auxiliary Program

Consider a dynamically consistent agent who differs from the one described in Subsection 2.1 in that he discounts consumption in the last period by an additional factor $\beta$. The equilibrium consumption for this agent solves the following *auxiliary program*:

$$\max_{\{c(s_t)\}} E\left[\sum_{t=1}^{T-1} \delta^{t-1} u\left(c(s_t)\right) + \beta\delta^{T-1} u\left(c(s_T)\right)\right], \tag{4}$$

subject to the zero profits constraint (3).

The following lemma establishes that the consumption on the equilibrium path for time-inconsistent agents coincides with the solution of the auxiliary program:

**Lemma 2.** *Suppose the consumer is time inconsistent. The consumption on the equilibrium path coincides with the solution of the auxiliary problem.*

The auxiliary program highlights that, in this model, underweighting consumption in the last period is the only distortion from time-inconsistency. To illustrate the lemma, consider the case of three periods and a constant income w. Since there is a single state of the world in each period, it can be omitted from the history. The equilibrium contract solves:

$$\max_{(c_1,c_2(A),c_2(B),c_3(A),c_3(B))} u(c_1) + \beta[\delta u(c_2(B)) + \delta^2 u(c_3(B))], \tag{5}$$

subject to

$$c_1 + \frac{c_2(A)}{R} + \frac{c_3(A)}{R^2} = w\left(1 + \frac{1}{R} + \frac{1}{R^2}\right), \tag{6}$$

$$u(c_2(B)) + \hat{\beta}\delta u(c_3(B)) \geq u(c_2(A)) + \hat{\beta}\delta u(c_3(A)), \tag{7}$$

$$u(c_2(A)) + \beta\delta u(c_3(A)) \geq u(c_2(B)) + \beta\delta u(c_3(B)), \tag{8}$$

where (6) is the zero profits constraint, (7) is the perceived choice constraint, and (8) is the incentive compatibility constraint.

Note first that incentive compatibility (8) must bind. Otherwise, we could increase $c_3(B)$ to achieve a higher utility without violating any constraint. Since (8) binds, we can rewrite the perceived choice constraint (7) as a monotonicity condition:

$$c_3(A) \leq c_3(B). \tag{9}$$

That is, because agents are more present-biased than they think they are, the contract that they actually pick ($A$) is more front-loaded than the contract that they think that they will pick ($B$). As usual, we ignore this monotonicity constraint for now and verify that it holds later.

Next, we contrast the marginal rate of substitution between $c_2(B)$ and $c_3(B)$ in the incentive compatibility constraint (8) and the objective function (5). The objective function evaluates consumption according to period-1 preferences, so the discount rate between periods 2 and 3 is $\delta$. The incentive compatibility constraint depends on the actual choice of the time-2 self, which discounts one period by $\beta\delta < \delta$. Therefore, front-loading the baseline consumption relaxes the incentive constraint. The solution of the program must then offer a maximally front-loaded baseline contract:

$$c_2(B) = 0. \tag{10}$$

Substituting (10) in the binding constraint (8) gives

$$u(c_2(A)) + \beta\delta u(c_3(A)) = u(0) + \beta\delta u(c_3(B)), \tag{11}$$

which can be rearranged as

$$\beta\delta\left[u(c_3(B)) - u(c_3(A))\right] = u(c_2(A)) - u(0) \geq 0.$$

Therefore, the monotonicity constraint (9) is automatically satisfied at the solution and can be ignored. Substituting (10) and (11) in the objective function, we obtain the objective of the auxiliary problem (up to a constant that can be omitted from the program):

$$u(c_1) + \beta[\delta u(c_2(B)) + \delta^2 u(c_3(B))] = u(c_1) + \delta u(c_2(A)) + \beta\delta^2 u(c_3(A)) - (1-\beta)\delta u(0). \tag{12}$$

Note that, in the proof above, we used equations (10) and (11) to eliminate the baseline consumption from the equilibrium program. Substituting the consumption that solves the auxiliary program back in these two equations, we can recover the baseline consumption. Since neither the auxiliary program nor equations (10) and (11) depend on the consumer's naiveté parameter $\hat{\beta}$, it follows that, in equilibrium, both the baseline consumption and the alternative consumption are not functions of $\hat{\beta}$.

**Corollary 1.** *There exists an equilibrium. Moreover, any equilibrium has the same equilibrium consumption vector, which is not a function of the consumer's naiveté $\hat{\beta}$ and is a continuous function of the agent's time-inconsistency parameter $\beta \in (0, 1]$.*

Since the equilibrium consumption vector is not a function of $\hat{\beta}$, the equilibrium obtained here would also be the equilibrium if we assumed that firms did not know the consumer's naiveté $\hat{\beta}$ (see Appendix B for a formal proof).

### 2.2.2 Vanishing Inefficiency

We now use Lemma 2 to obtain our main result. Let $W^I$ denote the equilibrium welfare of the time-inconsistent consumer, which evaluates the agent's consumption on the equilibrium path according to his long-term preferences (2) and recall that $W^C$ is the the welfare in the benchmark case of a time-consistent consumer. Since the time-consistent consumer maximizes welfare, $W^C - W^I \geq 0$ denotes the welfare loss from dynamic inconsistency.

**Theorem 1.** *Suppose $u$ is bounded and $\delta < 1$. Then, $lim_{T \nearrow +\infty} \left( W^C - W^N \right) = 0$.*

The theorem states that the welfare loss from dynamic inconsistency converges to zero as the contracting length grows. So, in any equilibrium, a time-inconsistent consumer gets approximately the maximum welfare possible if the number of periods is large. The assumption that $u$ is bounded and $\delta < 1$ ensure that the discounted welfare converges (otherwise, the discounted sum may not be well-defined). Note that this result is true for any fixed discount factors $\beta$ and $\delta$, so it is unrelated to the folk theorem literature from repeated games.

Recall that the time-inconsistency distorts consumption because it underweights the last period. Intuitively, because the effect of the last period vanishes as the number of periods grows, the solution of the auxiliary program converges to the equilibrium consumption with time-consistent consumers as $T \nearrow +\infty$. Therefore, even though the time-inconsistent consumer does not maximize his welfare function and has incorrect beliefs, the incentive constraints that keep forcing each future self to switch to the back-loaded repayment option imply that the equilibrium consumption along the equilibrium path converges to the welfare-maximizing consumption.

# 3 One-Sided Commitment

We now turn to the model in which consumers cannot commit to long-term contracts (one-sided commitment), keeping the assumption that the consumer has all bargaining power. As argued in the introduction, one-sided commitment is common in many markets. We model one-sided commitment as follows. The consumer offers a contract in each period. If a firm has accepted a contract, the consumer decides whether to keep it or replace it with a new one. If multiple firms accept a contract, the consumer picks each of them with some positive probability.

## 3.1 Benchmark: Time-Consistent Consumers

Consider first the benchmark case of a time-consistent consumer. With one-sided commitment, the consumer switches to a new contract whenever a firm is willing to provide him terms that are better than the terms of the original contract. For the purpose of characterizing the equilibrium consumption, there is no loss of generality in restricting attention to contracts in which the consumer never lapses ("renegotiation proofness"). To see this, consider an equilibrium in which the consumer lapses in some state of the world, replacing the original contract with a contract from another firm. Since the other firm cannot lose money by offering this new contract, the old firm could have accepted a contract that substituted the terms of the old contract from this period on with the terms of the new contract, and the consumer would have remained with the old firm. So, to characterize the consumption that can be supported in equilibrium, we can impose *non-lapsing constraints* which require contracts to be renegotiation proof. These constraints state the consumer's outside option at each state cannot exceed the value from keeping the current contract, where the outside option corresponds to the value from the best possible contract that other firms are willing to provide.[13]

Formally, the outside option at state $s_\tau$ is defined by the recursion:

$$V^C(s_\tau) \equiv \max_{\{c(s_t)\}} u\left(c(s_\tau)\right) + E\left[\sum_{t>\tau} \delta^{t-\tau} u\left(c\left(s_t\right)\right) \Bigg| \, s_\tau\right], \tag{13}$$

subject to

$$\sum_{t=\tau}^{T} E\left[\frac{w\left(s_t\right) - c\left(s_t\right)}{R^{t-\tau}} \Bigg| \, s_\tau\right] = 0, \tag{14}$$

and

$$u\left(c(s_{\tilde{\tau}})\right) + E\left[\sum_{t>\tilde{\tau}} \delta^{t-\tilde{\tau}} u\left(c\left(s_t\right)\right) \Bigg| \, s_{\tilde{\tau}}\right] \geq V^C(s_{\tilde{\tau}}), \;\; \forall s_{\tilde{\tau}} \text{ with } p(s_{\tilde{\tau}}|s_\tau) > 0. \tag{15}$$

Equation (14) is the zero profits condition, whereas (15) requires the new contract itself to be renegotiation proof. The equilibrium consumption with one-sided commitment solves this program at the initial period (i.e., at state $s_1 = \emptyset$).

While Program (13)-(15) characterizes the equilibrium consumption by a backward induction algorithm, an easier characterization can be obtained when the consumer is time-consistent. Consider, instead, the program that replaces the non-lapsing constraints by the requirement that, at each point in time, the expected future income cannot exceed the expected future consumption (FL):

$$\sum_{t\geq\tau} E\left[\frac{w\left(s_t\right) - c\left(s_t\right)}{R^{t-\tau}} \Bigg| \, s_\tau\right] \leq 0, \; \forall s_\tau. \tag{16}$$

Any contract that satisfies (14) and (16) is front loaded in the sense that, at each point in time, the accumulated profits cannot be negative. In a front-loaded contract, the consumer initially overpays to the firm and is repaid later. This overpayment discourages the consumer from switching contracts.

In general, (16) is a relaxation of the non-lapsing constraints: if the continuation contract gave positive expected profits at some state, a consumer would be able to increase his utility by replacing it with another contract that gives zero profits. When consumers are dynamically consistent, however, maximizing (13) subject to either (14) and (15) or (14) and (16) gives the same solutions. Suppose a solution to this latter program did not satisfy the non-lapsing constraints. Then, there would exist a continuation contract that gives zero profits while increasing the consumer's continuation utility. Substituting the original continuation contract by this new one would then increase the consumer's utility while giving non-negative profits at $t = 1$, contradicting the optimality of the original contract.

Following the same approach as in Corollary 1, we find that any SPNE must have the same equilibrium consumption.

[13]When a non-lapsing constraint binds, there are also equilibria in which the consumer lapses and recontracts with another firm. These equilibria are equivalent with the one with no lapsing in the sense that the consumer gets the same consumption and all firms make the same profits.

## 3.2 Time-Inconsistent Consumers

We now turn to the more interesting case of a time-inconsistent consumer. As with the time-consistent consumer, the equilibrium with one-sided commitment must satisfy non-lapsing constraints. Yet, because the consumer and the firms may disagree about the actions that will be chosen, we now need to distinguish between non-lapsing constraints according to the beliefs of the consumer and the beliefs of the firms. Equilibrium requires both of them to hold. To write down these constraints, we first define the outside options recursively.

The outside option at state $s_\tau$ given an option history $h^\tau$ according to the beliefs of firms is the highest utility that the consumer can actually obtain at that state. So, this "actual outside option" is the highest expected utility possible among contracts that are renegotiation proof and leave zero profits to the firm:

$$V(s_\tau, h^\tau) \equiv \max_{\{c(s_t,h^t)\}_{t\geq\tau}} u\left(c(s_\tau)\right) + \beta E\left[\sum_{t>\tau} \delta^{t-\tau} u\left(c(s_t, B, B, ..., B)\right)\middle| s_\tau\right],$$

subject to (PCC), (IC), the zero profits constraint

$$E\left[\sum_{t\geq\tau} \frac{w(s_t) - c(s_t, h^\tau, A, A, ..., A)}{R^{t-\tau}}\middle| s_\tau\right] = 0, \tag{17}$$

and the non-lapsing constraints

$$u\left(c_\tau(h^{\tilde{\tau}}, s_{\tilde{\tau}})\right) + \beta E\left[\sum_{t>\tilde{\tau}} \delta^{t-\tilde{\tau}} u\left(c(s_t, h^{\tilde{\tau}}, B, B, ..., B)\right)\middle| s_{\tilde{\tau}}\right] \geq V(h^{\tilde{\tau}}, s_{\tilde{\tau}}), \tag{18}$$

$$u\left(c_\tau(h^{\tilde{\tau}}, s_{\tilde{\tau}})\right) + \hat{\beta} E\left[\sum_{t>\tilde{\tau}} \delta^{t-\tilde{\tau}} u\left(c(s_t, h^{\tilde{\tau}}, B, B, ..., B)\right)\middle| s_{\tilde{\tau}}\right] \geq \hat{V}(h^{\tilde{\tau}}, s_{\tilde{\tau}}), \tag{19}$$

for all $s_{\tilde{\tau}}$ following $s_\tau$ and all $h^{\tilde{\tau}}$ that are continuation histories of $h^\tau$, where $\hat{V}$ is the "perceived outside option," which we define next.

The outside option at state $s_\tau$ given an option history $h^\tau$ according to the consumer's beliefs is the highest utility that the consumer believes that he would be able to obtain at that state. This perceived outside option is the highest perceived utility possible among contracts that are renegotiation proof and leave zero profits to the firm:

$$\hat{V}(h^\tau, s_\tau) \equiv \max_{\{c(h^t,s_t)\}_{t\geq\tau}} u\left(c_\tau(s_\tau)\right) + \hat{\beta} E\left[\sum_{t>\tau} \delta^{t-\tau} u\left(c(s_t, h^\tau, B, B, ..., B)\right)\middle| s_\tau\right],,$$

subject to (PCC), (IC), the zero profits constraint (17), and the non-lapsing constraints (18) and (19).

The *equilibrium program with one-sided commitment* (P1) adds the non-lapsing constraints to the program with two-sided commitment (P):

$$\max_{c(s_t,h^t)} u(c(s_1)) + \beta E\left[\sum_{t=1}^{T} \delta^{t-1} u(c(s_t, B, B, ..., B))\right],$$

subject to (Zero Profits), (PCC), (IC), and the non-lapsing constraints:

$$u\left(c(s_\tau, h^\tau)\right) + \beta E\left[\sum_{t>\tau} \delta^{t-\tau} u\left(c(s_t, h^\tau, B, B, ..., B)\right)\middle| s_\tau\right] \geq V(s_\tau, h^\tau) \;\; \forall (s_\tau, h^\tau), \tag{NL}$$

and

$$u\left(c(s_\tau, h^\tau)\right) + \hat{\beta} E\left[\sum_{t>\tau} \delta^{t-\tau} u\left(c(s_t, h^\tau, B, B, ..., B)\right)\middle| s_\tau\right] \geq \hat{V}(s_\tau, h^\tau) \;\; \forall (s_\tau, h^\tau). \tag{PNL}$$

Note that constraints (PNL) are associated with histories that are not reached in equilibrium. It may thus seem counter-intuitive why they would need to hold. But equilibrium requires the consumer to pick his optimal actions given how he thinks his future selves will behave. If (PNL) did not hold, the consumer would not expect his future selves to choose the baseline option, and so the consumption vector that solves the program would not correspond to an equilibrium consumption. Thus, both (NL) and (PNL) must hold in equilibrium.

**Lemma 3.** ***c*** *is an equilibrium consumption vector of the model with one-sided commitment if and only if it solves program (P1).*

### Auxiliary Program and Vanishing Inefficiency

As in the model with two-sided commitment, it is helpful to consider an auxiliary program corresponding to the equilibrium with a dynamically consistent agent who discounts the last period by an additional factor $\beta$. Since this agent is dynamically consistent, as shown in Subsection 2.1, we can replace the non-lapsing constraints by front-loading constraints. We refer to the maximization of (4) subject to the zero profits (3) and front-loading (16) constraints as the *auxiliary program with one-sided commitment*, which has a unique solution. The following lemma establishes that the solution of the auxiliary program coincides with the equilibrium with time-inconsistent agents:

**Lemma 4.** *Suppose the consumer is time inconsistent and has no commitment. The consumption on the equilibrium path coincides with the solution of the auxiliary problem with one-sided commitment.*

As in the model with commitment, the key distortion is that consumption in the last period is underweighted. To understand the illustrate the lemma, consider again the case of three periods and a constant income $w$, so we can omit the single state of the world from the history. The equilibrium contract solves:

$$\max_{(c_1, c_2(A), c_2(B), c_3(A), c_3(B))} u(c_1) + \beta[\delta u(c_2(B)) + \delta^2 u(c_3(B))], \tag{20}$$

subject to

$$c_1 + \frac{c_2(A)}{R} + \frac{c_3(A)}{R^2} = w\left(1 + \frac{1}{R} + \frac{1}{R^2}\right), \tag{21}$$

$$u(c_2(B)) + \hat{\beta}\delta u(c_3(B)) \geq u(c_2(A)) + \hat{\beta}\delta u(c_3(A)), \tag{22}$$

$$u(c_2(A)) + \beta\delta u(c_3(A)) \geq u(c_2(B)) + \beta\delta u(c_3(B)), \tag{23}$$

$$u(c_2(B)) + \hat{\beta}\delta u(c_3(B)) \geq \hat{V_2^N}, \tag{24}$$

$$c_3(B) \geq w, \tag{25}$$

$$u(c_2(A)) + \beta\delta u(c_3(A)) \geq V_2^N, \tag{26}$$

$$c_3(A) \geq w, \tag{27}$$

where $V_2^N$ is the actual outside option at time 2 and $\hat{V_2^N}$ is the perceived outside option at time 2. Equation (22) is the perceived choice constraint, (23) is the incentive compatibility constraint, and (24) - (27) are the non-lapsing constraints.

As in the model with two-sided commitment, the incentive compatibility constraint (23) must bind. Otherwise, we would be able to increase $c_3(B)$, achieving a higher utility without violating any constraint. Then, we can again rewrite the perceived choice constraint (22) as the monotonicity constraint (9), which will be ignored for now and verified later.

Next, we compare the marginal rate of substitution between the $c_2(B)$ and $c_3(B)$ in the incentive compatibility constraint (23), the non-lapsing constraint (24), and in the objective function. The incentive constraint depends on the actual choice of the time-2 self, which discounts one period by $\beta\delta$. The non-lapsing constraint depends on the agent's prediction of his future self's choice, so it discounts one period by $\hat{\beta}\delta > \beta\delta$. And the objective function is evaluated in period 1, so the discount rate between periods 2 and 3 is $\delta > \hat{\beta}\delta$. Therefore, front-loading the baseline consumption while preserving the incentive constraint weakens the non-lapsing constraint and increases the agent's utility, so the solution of the program must offer a maximally front-loaded baseline contract: $c_2(B) = 0$.

As before, we can substitute $c_2(B) = 0$ in the (binding) incentive compatibility constraint (23), obtaining

$$u(c_2(A)) + \beta\delta u(c_3(A)) = u(0) + \beta\delta u(c_3(B)). \tag{28}$$

Notice that we can rewrite (28) as

$$\beta\delta\left[u(c_3(B)) - u(c_3(A))\right] = u(c_2(A)) - u(0) \geq 0,$$

which shows that the monotonicity constraint holds. Substituting $c_2(B) = 0$ and (28) in the objective function, we obtain the objective of the auxiliary problem (up to a constant):

$$u(c_1) + \beta[\delta u(c_2(B)) + \delta^2 u(c_3(B))] = u(c_1) + \delta u(c_2(A)) + \beta\delta^2 u(c_3(A)) - (1-\beta)\delta u(0). \tag{29}$$

Next, we verify that the non-lapsing constraints of the baseline consumption (24) and (25) can be ignored. Intuitively, because the baseline consumption is more front-loaded than the actual consumption, whenever the actual consumption is front-loaded enough to prevent agents from lapsing, the baseline consumption will also satisfy the non-lapsing constraints. For equation (25), we have

$$c_3(B) \geq c_3(A) \geq w,$$

where the first inequality follows from the monotonicity condition, and the second inequality follows from (27). For (24), the result follows from the fact that $c_2(B) = 0$.

Therefore, we can simplify the original program (20)-(27) as:

$$\max_{(c_1, c_2(A), c_3(A))} u(c_1) + \delta u(c_2(A)) + \beta\delta^2 u(c_3(A)), \tag{30}$$

subject to

$$c_1 + \frac{c_2(A)}{R} + \frac{c_3(A)}{R^2} = w\left(1 + \frac{1}{R} + \frac{1}{R^2}\right), \tag{31}$$

$$u(c_2(A)) + \beta\delta u(c_3(A)) \geq V_2^N, \tag{32}$$

$$c_3(A) \geq w. \tag{33}$$

To conclude the proof, notice that this is the equilibrium program associated with a dynamically consistent agent who under-weights consumption in the last period by an additional term $\beta$. Then, as argued in Section 2.1, we can replace the non-lapsing constraint (32) by the constraint on the continuation profits at $t = 2$:

$$c_2(A) + \frac{c_3(A)}{R} \geq w\left(1 + \frac{1}{R}\right). \tag{34}$$

As in the model with two-sided commitment, we can recover the baseline consumption using $c_2(B) = 0$ and (28). Then, since neither the auxiliary program nor these equations depend on $\hat{\beta}$, it follows that, in equilibrium, both the baseline consumption and the alternative consumption are not functions of the consumer's naiveté.

**Corollary 2.** *Consider the model with one-sided commitment. There exists an SPNE. Moreover, any SPNE has the same equilibrium consumption vector, which is not a function of the consumer's naiveté $\hat{\beta}$, and is a continuous function of the agent's time-inconsistency parameter $\beta \in (0, 1]$.*

As shown in Appendix B, Corollary 2 implies that the equilibrium contracts obtained here coincide with the equilibrium contracts when firms do not know the consumer's naiveté $\hat{\beta}$.

As in the case of two-sided commitment, we can now use Lemma 4 to show that the welfare loss from dynamic inconsistency vanishes as the contracting length grows. Let $W_1^C$ and $W_1^I$ denote the equilibrium welfare of the time-consistent consumer (Subsection 3.1) and time-inconsistent consumer (Subsection 3.2), respectively.

**Theorem 2.** *Suppose $u$ is bounded and $\delta < 1$. Then, $lim_{T \nearrow +\infty} \left(W_1^C - W_1^N\right) = 0$.*

## 3.3 Removing Commitment Power

We now turn to the welfare effect of removing commitment power. We first show that, for a fixed contract length, removing commitment power can make the consumer better off. Recall that, on the equilibrium path, a time-inconsistent agent gets the same consumption as a dynamically consistent consumer who under-weights the last period by an additional factor $\beta$. Commitment power 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 consumer is sufficiently time inconsistent ($\beta$ is low), in which case the last period consumption is close to zero. So, if consuming zero in the last period hurts the agent enough and $\beta$ 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:

$$\mathcal{V}_S \equiv \max_{\{c(s_t)\}} \sum_{t=1}^{T-1} E\left[\delta^{t-1} u\left(c(s_t)\right)\right] + \delta^{T-1} u(0),$$

subject to

$$\sum_{t=1}^{T-1} E\left[\frac{c(s_t)}{R^{t-1}}\right] \leq \sum_{t=1}^{T} E\left[\frac{w(s_t)}{R^{t-1}}\right]$$

Let $\mathcal{V}_{NS}$ denote the agent's welfare from consuming the endowment in each state:

$$\mathcal{V}_{NS} \equiv \sum_{t=1}^{T} E\left[\delta^{t-1} u\left(w(s_t)\right)\right].$$

**Proposition 1.** *Suppose agents are time inconsistent and $\mathcal{V}_{NS} > \mathcal{V}_S$. There exists $\bar{\beta} > 0$ such that if $\beta < \bar{\beta}$, the welfare with one-sided commitment is greater than the welfare with two-sided commitment.*

Notice that for generic endowment paths, the condition that $\mathcal{V}_{NS} > \mathcal{V}_S$ fails when $T$ is large enough. So, as the contracting length grows, it becomes increasingly hard to satisfy the conditions for the time-inconsistent consumer to obtain a higher welfare without commitment. In fact, by Theorems 1 and 2, if the contracting length is large enough, removing commitment power cannot increase welfare. To see this, recall that, with two-sided commitment, a dynamically consistent consumer maximizes welfare subject to zero profits. Removing commitment power is equivalent to introducing front-loading constraints, so the welfare with one-sided commitment cannot be higher. But, since the welfare of time-inconsistent consumers converges to the welfare of dynamically consistent consumers, the same must be true when the consumer is dynamically inconsistent.

# 4 Deterministic Income

In the general model considered so far, contracting played two distinct roles: it allowed the agent to shift consumption over time and to insure against risk. We now focus on the intertemporal aspect by shutting down the risk-sharing channel. This simplification allows us to derive clearer implications of one-sided commitment and contrast it more clearly with existing results, since this model with $T = 3$ is analogous to the one from Heidhues and Kőszegi (2010).[14]

Formally, we assume that the agent gets a constant income of $w > 0$ in each period. We say that the *market breaks down* if the agent gets the same consumption as the endowment along the equilibrium path: $c_t^E = w$ for all $t$. If the market does not break down, we say that the equilibrium features a *long-term contract*.

## 4.1 Equilibrium Contracts

We start with the benchmark case of a time-consistent agent:

**Lemma 5.** *Suppose the agent is time consistent and has no commitment power. Then, the market breaks down if $R \leq \frac{1}{\delta}$, and the equilibrium features a long-term contract if $R > \frac{1}{\delta}$.*

[14]While Heidhues and Kőszegi (2010) assume that there is no consumption in the first period. This assumption does not affect the equilibrium with one-sided commitment since any contract accepted in the initial period will renegotiated in the second period if it does not solve our equilibrium program.

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 save ($R > \frac{1}{\delta}$). In this case, agents make an irreversible up-front payment that prevents them from lapsing in future periods. If the consumer would prefer to borrow at the prevailing interest rate ($R \leq \frac{1}{\delta}$), the market breaks down since they cannot commit not to drop any contract.

We now turn to time-inconsistent agents. Our next result shows that the conditions for long-term contracting with time-consistent and time-inconsistent agents coincide:

**Lemma 6.** *Consider a time-inconsistent agent in the one-sided commitment environment. The market breaks down if and only if $R \leq \frac{1}{\delta}$.*

Recall that the time-inconsistent agent behaves like a time-consistent agent except for the last period. Therefore, if $R > \frac{1}{\delta}$, then a long-term contract is provided and, except for the last period the equilibrium consumption is increasing over time. 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. However, if $R > \frac{1}{\beta\delta}$ even the last-period self strictly prefers not to lapse.

Because the condition for long-term contracts to be provided is the same for time-consistent and time-inconsistent consumers, one might be tempted to conclude that dynamic inconsistency neither helps nor hurts the ability to provide long-term contracts. But this interpretation is not warranted, as holding $\delta$ fixed also makes time-inconsistent agents more impatient than time-consistent agents (because of the additional discount factor $\beta$). 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 $\alpha = (\alpha_1, \cdots, \alpha_T)$, where $\alpha_i > 0$ and $\sum \alpha_i = 1$. An agent's *$\alpha$-weighted measure of impatience* is:

$$\alpha_1 + \alpha_2\beta\delta + \cdots + \alpha_T\beta\delta^{T-1}. \tag{35}$$

That is, if a time-consistent agent has discount parameter $\delta_C$ and a time-inconsistent agent has discount parameters $(\delta_I, \beta)$, they have the same $\alpha$-weighted impatience if:

$$\alpha_1 + \alpha_2\delta_C + \cdots + \alpha_T\delta_C^{T-1} = \alpha_1 + \alpha_2\beta\delta_I + \cdots + \alpha_T\beta\delta_I^{T-1}. \tag{36}$$

Intuitively, both agents discount the stream of utils $\alpha$ in the same way. Of course, they still discount other streams differently.[15] Simple algebraic manipulations show that, for any fixed vector of weights, $\beta\delta_I < \delta_C$ and $\beta\delta_I^{T-1} > \delta_C^{T-1}$. 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. Since $\frac{1}{\delta_I} < \frac{1}{\delta_C}$, Lemmas 5 and 6 imply that it is easier to sustain long-term contracting with time-inconsistent than with time-consistent consumers.

**Proposition 2.** *Fix a vector of weights $\alpha$ and consider a time-consistent and a time-inconsistent consumer with the same $\alpha$-weighted impatience. If the market breaks down for the time-inconsistent consumer, it also breaks down for time-consistent consumer.*

[15]One example of $\alpha$-weighted impatience is the effective discount factor introduced by Chade et al. (2008), which corresponds to an $\alpha$-weighted impatience with uniform weights ($\alpha_i = \frac{1}{T}$ for all $i$).

The prediction from Proposition 2 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.[16] But notice that whether we control for impatience is key for the predictions of the model. Lemmas 5 and 6 show that time-consistent and time-inconsistent agents with the same "long-term discount factor" $\delta$ are equally likely to obtain long-term contracts. However, holding the long-term discount factor fixed conflates the effects of discounting and time inconsistency. Proposition 2 shows that, controlling for discounting, time-inconsistent agents are actually more likely to obtain long-term contracts.

## 4.2 Maximum Fees

We now consider the effects of imposing a maximum fee in each period. As before, 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 $\underline{c}$.[17]

Of course, if the agent is time consistent, imposing a consumption floor introduces additional constraints in the agent's welfare maximization program, which cannot in increase welfare. Suppose, instead, that the agent is time inconsistent. Recall that the equilibrium contract solves the auxiliary program, which coincides with the equilibrium program of a dynamically consistent agent that discounts the last period more heavily than a time-consistent agent. 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. 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 time-inconsistent agent already under-saves relative to the welfare-maximizing amount, this policy hurts them whenever it is binding:[18]

**Proposition 3.** *Suppose the agent is time inconsistent and has no commitment power. Then, mandating a minimum consumption weakly decreases welfare.*

## 4.3 Sophisticated Consumers

The main focus of our paper is on consumers who underestimate their present bias. However, as a benchmark, we now consider a present-biased consumer who is perfectly aware of his bias

[16] Note 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.

[17] Any contract would give negative profits if the consumption floor exceeded the agent's endowment (or, equivalently, if the maximum fee was negative). Then, no firm would offer any contract that the agent would pick, and the agent would consume his endowment in each period. Therefore, there is no loss of generality in considering consumption floors that do not exceed the agent's endowment, $\underline{c} \leq w$, or, equivalently, non-negative maximum fees.

[18] One-sided commitment and naiveté are important for this result. With two-sided commitment, imposing a maximum fee sometimes helps time-inconsistent consumers (Heidhues and Kőszegi, 2010). But, as described in footnote 4, settings with this type of policy often have one-sided commitment. Moreover, as we show in Appendix B, a maximum fee has an ambiguous effect on welfare when consumers are sophisticated.

($\beta = \hat{\beta} < 1$).

As in the case of time-consistent consumers, in the equilibrium with two-sided commitment, any contract that is accepted by a firm must maximize the utility of the period-1 self subject to the zero profits constraint. Recall that the period-1 self discounts all future periods by the additional term $\beta$. So introducing any small amount of naiveté discontinuously shifts the equilibrium consumption from the one in which an additional discount $\beta$ is applied to all future periods to the one that solves the auxiliary program, in which this additional discount only applies to the last period.[19] In particular, unlike with partially naive consumers, the consumption path of a sophisticated consumer does not converge to the one that maximizes welfare as the number of periods grows.

With one-sided commitment, the equilibrium consumption must also satisfy the non-lapsing constraints, defined recursively as in (13)-(15). But, with sophistication, the front loading constraints are not sufficient to ensure that the consumer will not lapse. The more time inconsistent the consumer, the higher the front load required to prevent lapsing. The next lemma characterizes the conditions for the market to break down with sophisticated consumers:

**Lemma 7.** *Consider a time-inconsistent sophisticated agent in the one-sided commitment environment. There exists $r_T(\beta, \delta) > \frac{1}{\delta}$ such that the market breaks down if and only if $R \leq r_T(\beta, \delta)$. Moreover:*

1. *$r_T(\beta, \delta)$ is decreasing in $\beta$ and in $T$, and*
2. $\lim_{\beta \nearrow 1} r_T(\beta, \delta) = \lim_{T \nearrow \infty} r_T(\beta, \delta) = \frac{1}{\delta}$.

To understand the existence of the cutoff $r_T(\beta, \delta)$, note that when the interest rate is low, the consumer would like to borrow. However, because he cannot commit to repay his debt, no principal would lend him money. So, in equilibrium, the agent consumes his endowment in each period. Claim 1 states that market breaks down "less often" when agents have a higher time-consistency parameter $\beta$ and when there are more periods $T$. As argued previously, the amount of front-loaded payments needed to support long-term contracting is decreasing in the agent's time-consistency parameter $\beta$. Moreover, with a longer horizon, principals have more instruments to avoid the market from breaking down. Claim 2 states that, as agents become 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 time-consistent agents.[20]

Comparing the conditions from Lemmas 6 and 7, we find that 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-loaded contracts than they actually will, making it easier to convince them to accept a contract and to keep them by shifting payments into the future.

We conclude by extending Proposition 2 to incorporate sophisticated time-inconsistent agents:

[19]The discontinuity of the equilibrium consumption in the agent's naiveté was previously shown by Heidhues and Kőszegi (2010).

[20]It is straightforward to show that the equilibrium contract of a sophisticated agent is continuous in $\beta$, so that "almost time-consistent" sophisticated agents get approximately the same contract as time-consistent agents do. However, although the cutoff for a sophisticated agent converges to the cutoff of time-consistent agents as the horizon grows, it is not true that sophisticated agents with "sufficiently long" horizons get approximately the same consumption as time-consistent agents.

**Proposition 4.** *Fix a vector of weights $\alpha$ and consider time-consistent, sophisticated, and naive time-inconsistent agents with the same $\alpha$-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.*

Proposition 4 shows 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.

# 5 Conclusion

In this paper, we study contractual relationships between time-inconsistent consumers and risk neutral firms. We show that the welfare loss from time-inconsistency vanishes as the number of periods grows. We also study the effect of removing commitment power from consumers. For each fixed contracting horizon, removing commitment power increases welfare when consumers are sufficiently time-inconsistent. However, removing commitment power does not help if the contracting horizon is long.

Our results suggest that enforcing long-term contracts may be enough to ensure efficiency with naive consumers. With sophisticated agents, the equilibrium consumption does not converge to the one that maximizes their long-term preferences. Therefore, when the contracting horizon is long enough, making individuals aware of their dynamic inconsistency hurts them. This finding contrasts with a general intuition that educating behavioral individuals about their biases would increase their welfare.

The equilibrium consumption does not depend on the degree of naiveté as long as the consumer is still partially naive ($\hat{\beta} > \beta$). However, it jumps discontinuously at the point at which the consumer becomes sophisticated ($\hat{\beta} = \beta$). This discontinuity in the perceived time inconsistency $\hat{\beta}$ contrasts with the continuity in the actual time inconsistency $\beta$. The equilibrium consumption is continuous in $\beta$ both for partially naive and for sophisticated consumers. In particular, "almost time-consistent" agents get approximately the same consumption as time-consistent agents. Yet, "almost sophisticated" agents get the same consumption as any other naive agent, which is bounded away from the consumption of a sophisticated consumer.

The paper focuses on one particular deviation from rationality – dynamic inconsistency – for two reasons. First, leaving a long-term contract is fundamentally an intertemporal decision, and dynamic inconsistency is the most well-studied bias in intertemporal decision-making. Second, there is evidence that this bias is important in credit markets where consumers are allowed to leave previous agreements, such as mortgages or credit cards. Still, policies that remove commitment power may also be important in settings with other biases. For example, many countries have regulations regarding cooling-off periods, during which firms must allow consumers to return products. Cooling-off periods may be an effective policy for consumers who suffer from projection bias – that is, they mispredict their future tastes, overestimating how much it will resemble their current tastes.

More generally, removing commitment power is a particularly weak type of paternalistic policy. Instead of having a regulator decide which policy to ban, this decision is made by one's future selves. The study of the regulation of commitment power in settings with other biases is left for future work.

# Appendix

## Appendix A. Proofs

**Proof of Lemma 1.** The proof is similar to the proof with one-sided commitment (Lemma 3) and is therefore omitted. □

**Proof of Lemmas 2 and 4.** Consider one-sided commitment case first (Lemma 2). We claim that all IC constraints are binding. To simplify the exposition, we focus on the case of $T = 4$ and no uncertainty here (the proof for general $T$ is essentially the same except for additional notation; the proof for stochastic income is presented in the online appendix). There are two ICs:

$$u(c_2(A)) + \beta[\delta u(c_3(AB)) + \delta^2 u(c_4(AB))] \geq u(c_2(B)) + \beta[\delta u(c_3(BB)) + \delta^2 u(c_4(BB))], \tag{A1}$$

$$u(c_3(AA)) + \beta\delta u(c_4(AA)) \geq u(c_3(AB)) + \beta\delta u(c_4(AB)). \tag{A2}$$

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, (A1) must be binding at an optimum (otherwise, we can raise $c_4(BB)$, giving the agent a higher utility). Substitute the binding (A1) 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, (A2) must bind (otherwise, we can raise $c_4(AB)$, increasing the agent's utility). Substituting the binding (A2) 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, $\hat{V}_2$. Suppose $\hat{V}_2 = u(\hat{c}_2) + \hat{\beta}(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4))$. We obtain

$$\begin{aligned} &u(c_2(B)) + \hat{\beta}(\delta u(c_3(BB)) + \delta^2 u(c_4(BB)) \\ &= u(0) + \frac{\hat{\beta}}{\beta}\beta(\delta u(c_3(BB)) + \delta^2 u(c_4(BB)) \\ &= u(0) + \frac{\hat{\beta}}{\beta}(u(c_2(A)) + \beta(\delta u(c_3(AB)) + \delta^2 u(c_4(AB))) - u(0)) \\ &= (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}(u(c_2(A)) + \beta(\delta u(c_3(AB)) + \delta^2 u(c_4(AB)))), \end{aligned}$$

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 $u(c_2(A))+\beta(\delta u(c_3(AB))+\delta^2 u(c_4(AB) \geq V_2$. Since $V_2$ 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

$$\begin{aligned}
&u(c_2(B)) + \hat{\beta}(\delta u(c_3(BB)) + \delta^2 u(c_4(BB)) \\
&\geq (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}V_2 \\
&\geq (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}\left[u(\hat{c}_2) + \beta(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4)\right] \\
&= (1 - \frac{\hat{\beta}}{\beta})u(0) + (\frac{\hat{\beta}}{\beta} - 1)u(\hat{c}_2) + \left[u(\hat{c}_2) + \hat{\beta}(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4)\right] \\
&\geq u(\hat{c}_2) + \hat{\beta}(\delta u(\hat{c}_3) + \delta^2 u(\hat{c}_4)) = \hat{V}_2,
\end{aligned}$$

where the first line follows from the non-lapsing constraint at time 2, the second uses the revealed preference, and the last line uses $\hat{c}_2 \geq 0$ and $\hat{\beta} \geq \beta$. This shows that the perceived non-lapsing constraints hold.

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

$$\begin{aligned}
u(c_3(AB)) + \hat{\beta}\delta u(c_4(AB)) &= u(0) + \hat{\beta}\delta u(c_4(AB)) \\
&= (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}\left(u(c_3(AA)) + \beta\delta u(c_4(AA))\right) \\
&\geq u(c_3(AA)) + \hat{\beta}\delta u(c_4(AA)),
\end{aligned} \tag{A3}$$

and

$$\begin{aligned}
&u(c_2(B)) + \hat{\beta}[\delta u(c_3(BB)) + \delta^2 u(c_4(BB))] \\
&= u(0) + \hat{\beta}[\delta u(c_3(BB)) + \delta^2 u(c_4(BB))] \\
&= (1 - \frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}\left[u(c_2(A)) + \beta[\delta u(c_3(AB)) + \delta^2 u(c_4(AB))]\right] \\
&\geq u(c_2(A)) + \hat{\beta}[\delta u(c_3(AB)) + \delta^2 u(c_4(AB))].
\end{aligned} \tag{A4}$$

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 $u(c_t^E) + \delta u(c_{t+1}^E) + \cdots + \beta\delta^{t-4}u(c_4^E) \geq V_t$.

Therefore, the original program reduces to the auxiliary program:

$$\max_{(c_1,\cdots,c_4)} u(c_1) + \delta u(c_2) + \delta^2 u(c_3) + \beta\delta^3 u(c_4), \tag{A5}$$

subject to

$$\sum_{t=1}^{4} \frac{c_t}{R^{t-1}} = \sum_{t=1}^{4} \frac{w}{R^{t-1}}, \tag{A6}$$

$$u(c_t) + \delta u(c_{t+1}) + \cdots + \beta\delta^{T-t}u(c_T) \geq V_t^A, \forall 2 \leq t \leq 4. \tag{A7}$$

With two-sided commitment, the same arguments given above go through except that we can omit the non-lapsing constraints (A7). In all, the consumption on the equilibrium path coincides with the solution of the auxiliary problems. □

**Proof of Theorems 1 and 2.** We consider the case with one-sided commitment (Theorem 2). We omit the proof for the two-sided commitment case (Theorem 1), which is similar. We need to show that $\lim_{T\nearrow+\infty}(W_1^C - W_1^N) = 0$.

For each parameter $\beta$, let $V^A(\beta)$ denote the maximum value attained by the solution of the auxiliary program with one-sided commitment (Lemma 2). Notice that the feasible set is independent of $\beta$. When $\beta = 1$, the auxiliary program becomes a time-consistent agent's program, so that $V^A(1) = W_1^C$. We have also $\lim_{T\nearrow\infty}(W_1^N - V^A(\beta)) = \lim_{T\nearrow\infty}(1-\beta)E\delta^{T-1}u(c(s_T)) = 0$. Since the objective function is linear in $\beta$, it follows from the Envelope Theorem that $\frac{\partial V^A(\beta)}{\partial\beta} = E\delta^{T-1}u\left(c(s_T)\right) \geq \delta^{T-1}u\left(0\right)$. Applying Lagrange's Mean Value Theorem gives

$$V^A(1) - V^A(\beta) = \frac{\partial V^A(\beta)}{\partial\beta}|_{\beta=\beta'}(1-\beta) \geq \delta^{T-1}u(0)(1-\beta),$$

$$V^A(\beta) - V^A(0) = \frac{\partial V^A(\beta)}{\partial\beta}|_{\beta=\beta''}\beta \geq \delta^{T-1}u(0)\beta,$$

where $\beta' \in (\beta, 1), \beta'' \in (0, \beta)$. Taking $T$ to infinity leads to

$$\lim_{T\nearrow\infty} V^A(1) \geq \lim_{T\nearrow\infty} V^A(\beta) \geq \lim_{T\nearrow\infty} V^A(0).$$

To obtain the theorem, 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.

Let $c_C^1$ denote 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

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

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

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

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

**Proof of Lemma 3.** We show that the Subgame Perfect Nash Equilibrium of the game is outcome-equivalent to the solution of the maximization programs (P) and (P1). We are interested in the SPNE of the game played by the firms and by the different selves of the consumer. Suppose the t-period self of the consumer offers a contract $\mathcal{C}'_t$. Specifically, a contract at time $t$, $\mathcal{C}'_t$, specifies consumption on each possible state in each future time $\tau \geq t$. Denote the set of possible states by $K_{t,\tau}$, in which the first subscript corresponds to the time in which the contract is offered and the second subscript corresponds to the decision-making time $\tau$. The contract specifies consumption for each different income states, so the contracting space must be greater than the space of income states. In addition, SPNE imposes no restrictions on $K_{t,\tau}$, i.e., $K_{t,\tau}$. To keep analysis tractable, we assume that $K_{t,\tau}$ has a product structure. Otherwise, we can always add more states that are never reached so that it has a product structure and the resulting equilibrium is outcome-equivalent to the original equilibrium. 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 $\tau$ from an earlier history $h_t \in H_t$ for $\tau > t$.

We next write down the agent's strategy profile given a contract. Consider an agent who makes a decision at time $\tau$. Suppose the income-irrelevant messages that has been reached is $h_{\tau-1}$, which is an element in $H_{\tau-1}$. At time $\tau$, the agent learns the income state, i.e., $s_\tau$ is realized. The agent needs to decide which income-irrelevant message/action $a_\tau \in \Delta(H_\tau(h_{\tau-1}))$ to send, where $\Delta(\cdot)$ represents the distribution. If there is one-sided commitment, the agent also needs to decide whether he will lapse or not, in which case, the strategy can be summarized by a pair $(d_\tau, a_\tau)$, where $d_\tau \in \Delta(\{0,1\})$. If $d_\tau = 1$ with probability 1, then the agent stays, otherwise the contract is lapsed with a positive probability.

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. 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").

The main claim is that for any competitive equilibrium, there exists a SPNE that gives the agent an exactly same actual consumption and perceived consumption stream. For any SPNE, there exists a competitive equilibrium that gives the agent exactly same actual consumption and perceived consumption.

For a fixed competitive equilibrium, consider a candidate SPNE in which the contracts offered are identical to the offered in the competitive equilibrium and where the agent always chooses the alternative options. Since the competitive equilibrium satisfies the non-lapsing, perceived choice, and incentive constraints, it is clear that the candidate equilibrium is an SPNE. In addition, the SPNE gives the agent exactly same actual consumption and perceived consumption as the competitive equilibrium.

We now turn to the opposite direction. We say that two SPNE are *equivalent* if all selves of the consumers have same actual and perceived consumption in both. We will establish the result through two separate claims:

**Claim 1.** *Fix an SPNE. There exists an equivalent SPNE in which the agent never lapses ($d_\tau = 1, \forall \tau$).*

*Proof.* Consider an SPNE in which the agent lapses in some period $d_\tau = 0$ with a positive probability, replacing it with a contract $\mathcal{C}''_\tau$ from another principal. Since the other principal cannot lose money by offering this new contract, the old principal could have accepted 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 an SPNE that is equivalent to the original SPNE. □

**Claim 2.** *Fix an SPNE. There is an equivalent SPNE that offers two options following any history: $\#|H_t(h_{t-1})| \leq 2$, for all $h_{t-1} \in H_{t-1},\ t \geq 2$.*

*Proof.* From the previous claim, we can restrict attention for SPNE in which the agent never lapses. Suppose $t_1 < t_2 < t_3$. Note that 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 – one that the agent will choose and another one that the agent thinks that he will 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. □

Given these two claims, a contract offered by self $t$, $\mathcal{C}'_t$, must maximize the agent's utility subject to the zero profits, incentive compatibility, perceived choice, and non-lapsing constraints. In other words, the contracts $\{\mathcal{C}'_t\}_{t=1,\cdots,T}$ must form a competitive equilibrium with one-sided commitment. In addition, this competitive equilibrium gives the agent exactly same actual consumption and perceived consumption, concluding the proof of Lemma 3. □

**Proof of Corollaries 1 and 2.** Consider the case of one-sided commitment (Corollary 2). The case with two-sided commitment (Corollary 1) is analogous. We can focus on the auxiliary program. Let $x(s_t) \equiv u(c(s_t))$ denote the agent's utility from the consumption he gets in state $s_t$. We study the dual program:

$$\max_{\{x(s_t)\}} \sum_{t=1}^{T} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \frac{w(s_t) - u^{-1}\left(x(s_t)\right)}{R^{t-1}}, \tag{A8}$$

subject to

$$\sum_{t=1}^{T-1} \sum_{s_t \in \mathbb{S}_t} \delta^{t-1} p(s_t|s_1) x(s_t) + \beta \sum_{s_T \in \mathbb{S}_T} \delta^{T-1} p(s_T|s_1) x(s_T) \geq \underline{u}, \tag{A9}$$

and

$$\sum_{t \geq \tilde{\tau}} \sum_{s_t \in \mathbb{S}_t} \delta^{t-\tilde{\tau}} p(s_t|s_{\tilde{\tau}}) x(s_t) + \beta \sum_{s_T \in \mathbb{S}_T} \delta^{T-\tilde{\tau}} p(s_T|s_{\tilde{\tau}}) x(s_T) \geq V^A(s_{\tilde{\tau}}) \ \ \forall s_{\tilde{\tau}} \in \mathbb{S}_{\tilde{\tau}}(s_\tau), \forall \tau, \tag{A10}$$

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 unique. Moreover, the consumption on the equilibrium path is continuous in $\beta \in (0, 1]$. Finally, the program does not involve $\hat{\beta}$, so the equilibrium consumption vector is not a function of the consumer's naiveté. □

**Proof of Proposition 1.** First, the welfare with two-sided commitment approaches to $\mathcal{V}_S$ as $\beta$ 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 $\beta$ 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

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

which is positive whenever

$$\frac{u'(0)}{R\delta u'(W)} > \beta. \tag{A11}$$

The perturbation has exactly the same effect on the objective function (scaled down by $\delta^t$ and multiplied by the probability of reaching state $s_{T-1}$). Thus, as long as $\beta$ satisfies (A11), 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 $T$:

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

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

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

for all $\tilde{\tau} = 2, ..., T$, 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 (A11) 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. □

**Proof of Lemma 5.** The equilibrium contract solves:

$$\max_{\{c_t\}} \sum_{t=1}^{T} \delta^{t-1} u(c_t), \tag{A12}$$

subject to

$$\sum_{t=1}^{T} \frac{c_t}{R^{t-1}} = \sum_{t=1}^{T} \frac{w}{R^{t-1}}, \tag{A13}$$

$$\sum_{t=\tau}^{T} \frac{c_t}{R^{t-\tau}} \geq \sum_{t=\tau}^{T} \frac{w}{R^{t-\tau}}, \forall 2 \leq \tau \leq T. \tag{A14}$$

Since the objective function is strictly concave and the set of feasible contracts is convex, the solution is unique.

The Lagrangian is

$$\mathcal{L} = \sum_{t=1}^{T} \delta^{t-1} u(c_t) - \sum_{\tau=1}^{T} \lambda_\tau \left( \sum_{t=\tau}^{T} \frac{c_t}{R^{t-1}} - \sum_{t=\tau}^{T} \frac{w}{R^{t-1}} \right), \tag{A15}$$

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

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

therefore, $c_1 \geq c_2 \geq \cdots \geq c_T$. 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 in the last period. So it must be the case that $c_T = w$. Now we have $c_1 \geq \cdots \geq c_{T-1} \geq w$. 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 $c_1 = \cdots = c_T = w$. 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 (A14) holds automatically. Solving the problem with only the zero-profit condition gives $c_1 < c_2 < \cdots < c_T$. Notice that $c_1 < w$ because otherwise, $w \le c_1 < c_2 < \cdots < c_T$, contradicted to the zero profit condition. Similarly we have $c_T > w$. Let $\xi$ be the smallest index such that

$$c_1 < \cdots < c_\xi < w \le c_{\xi+1} < \cdots < c_T.$$

It is clear that (A14) holds strictly for $\tau \ge \xi$+1. Now consider $\tau \le \xi$, we have

$$\begin{aligned}\sum_{t=\tau}^{T} \frac{c_t}{R^{t-1}} &= \sum_{t=1}^{T} \frac{c_t}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{c_t}{R^{t-1}} \\ &= \sum_{t=1}^{T} \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{c_t}{R^{t-1}} \\ &> \sum_{t=1}^{T} \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{w}{R^{t-1}} \\ &= \sum_{t=\tau}^{T} \frac{w}{R^{t-1}}. \end{aligned} \tag{A16}$$

So the resource constraint (A14) holds strictly for $\tau \le \xi$. In all, the equilibrium contract features growing consumption and the long-term contract is supported in this market. □

**Proof of Lemma 6.** We can rewrite the non-lapsing constraints in the auxiliary problem as

$$\sum_{t=\tau}^{T} \frac{c_t}{R^{t-1}} \ge \sum_{t=\tau}^{T} \frac{w}{R^{t-1}}, \forall 2 \le \tau \le T. \tag{A17}$$

Consider the following program.

$$\max_{c_t} u(c_1) + \delta u(c_2) + \cdots + \delta^{T-2} u(c_{T-1}) + \beta \delta^{T-1} u(c_T), \tag{A18}$$

subject to

$$\sum_{t=1}^{T} \frac{c_t}{R^{t-1}} = \sum_{t=1}^{T} \frac{w}{R^{t-1}}, \tag{A19}$$

$$\sum_{t=\tau}^{T} \frac{c_t}{R^{t-1}} \ge \sum_{t=\tau}^{T} \frac{w}{R^{t-1}}, \forall 2 \le \tau \le T. \tag{A20}$$

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 $\beta$, 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 2 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 $c_1 \geq \cdots \geq c_{T-1} \geq w$. 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 $c_1 = \cdots = c_T = w$.

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:

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

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 $u'(\tilde{c}_1) = \cdots = (\delta R)^{T-2} u'(\tilde{c}_{T-1}) \leq \beta(\delta R)^{T-1} u'(\tilde{c}_T)$. Since $\delta R > 1$, we have $\tilde{c}_1 < \tilde{c}_2 < \cdots < \tilde{c}_{T-1}$. 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 $\xi$ 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

$$\begin{aligned}
\sum_{t=\tau}^{T} \frac{\tilde{c}_t}{R^{t-1}} &= \sum_{t=1}^{T} \frac{\tilde{c}_t}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{\tilde{c}_t}{R^{t-1}} \\
&= \sum_{t=1}^{T} \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{\tilde{c}_t}{R^{t-1}} \\
&> \sum_{t=1}^{T} \frac{w}{R^{t-1}} - \sum_{t=1}^{\tau-1} \frac{w}{R^{t-1}} \\
&= \sum_{t=\tau}^{T} \frac{w}{R^{t-1}}. \qquad \text{(A21)}
\end{aligned}$$

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

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

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 $u'(\tilde{c}_1) = \cdots = (\delta R)^{T-2} u'(\tilde{c}_{T-1})$, $\sum_{t=1}^{T-1} \frac{\tilde{c}_t}{R^{t-1}} = \sum_{t=1}^{T-1} \frac{w}{R^{t-1}}$, and $\tilde{c}_T = w$.

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

**Proof of Proposition 2.** Presented in the text. □

**Proof of Proposition 3.** 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 $(c_1^1, c_2^1, \cdots, c_T^1)$ the maximizer to the program $V_1^A$, and $(c_1^2, c_2^2, \cdots, c_T^2)$ 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 $c_1^1 < c_2^1 < \cdots < c_{T-1}^1$ 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 $(c_{k+1}^1, \cdots, c_T^1)$ and $(c_{k+1}^2, \cdots, c_T^2)$ 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

$$\begin{aligned}
& u(c_1^1) + \delta u(c_2^1) + \cdots + \delta^{T-1} u(c_T^1) \\
&= u(c_1^1) + \delta u(c_2^1) + \cdots + \beta \delta^{T-1} u(c_T^1) + (1-\beta) \delta^{T-1} u(c_T^1) \\
&\geq u(c_1^2) + \delta u(c_2^2) + \cdots + \beta \delta^{T-1} u(c_T^2) + (1-\beta) \delta^{T-1} u(c_T^1) \\
&\geq u(c_1^2) + \delta u(c_2^2) + \cdots + \beta \delta^{T-1} u(c_T^2) + (1-\beta) \delta^{T-1} u(c_T^2) \\
&= u(c_1^2) + \delta u(c_2^2) + \cdots + \delta^{T-1} u(c_T^2), \qquad \text{(A22)}
\end{aligned}$$

where the first inequality is from the fact that $(c_1^1, c_2^1, \cdots, c_T^1)$ 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. □

**Proof of Lemma 7.** The existence and uniqueness of the equilibrium consumption follows from the same argument as in the time-consistent case (rewriting in terms of the utility of consumption, we obtain a program that corresponds to the maximization of a strictly concave function subject to linear constraints). Let $\lambda_1$ denote the Lagrangian multiplier associated with the zero-profit constraint, and let $\lambda_\tau$ denote the Lagrangian multiplier associated with the non-lapsing constraints. The Lagrangian optimality conditions imply:

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

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

We first examine the conditions for the market to break down: $c_1 = \cdots = c_T = w$. 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.

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

Inversing 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:

$$1 \geq \beta r + \beta(1-\beta)r^2 + \cdots + \beta(1-\beta)^{T-2}r^{T-1}. \tag{A23}$$

The right-hand-side is an increasing function of $r$. If $r = 0$, LHS>RHS. If $r \to \infty$, LHS<RHS. So there exists a unique $r_T(\beta)$ such that the (A23) becomes an equality. Since 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 $\beta$. Recall that $r_T(\beta,\delta) = \frac{r_T(\beta)}{\delta}$, where $r_T(\beta)$ solves the equation (A23). It is sufficient to show that $r_T(\beta)$ is decreasing in $\beta$. The right hand side of (A23) is a geometric series, implying

$$1 = \beta r\frac{1-(1-\beta)^{T-1}r^{T-1}}{1-(1-\beta)r}. \tag{A24}$$

Rearranging terms leads to

$$1 - r + \beta(1-\beta)^{T-1}r^T = 0. \tag{A25}$$

Taking derivative with respect to $\beta$, we obtain

$$\begin{aligned} r_T'(\beta) &= \frac{(1-T\beta)(1-\beta)^{T-2}r_T^T(\beta)}{1-\beta(1-\beta)^{T-1}r_T^{T-1}T} \\ &= \frac{(1-T\beta)(1-\beta)^{T-2}r_T^{T+1}(\beta)}{r_T-\beta(1-\beta)^{T-1}r_T^T(\beta)T} \\ &= \frac{(1-T\beta)(1-\beta)^{T-2}r_T(\beta)^{T+1}}{r_T(\beta)-(r_T(\beta)-1)T} \\ &= \frac{(1-T\beta)(1-\beta)^{T-2}r^{T+1}}{T-(T-1)r_T(\beta)}. \end{aligned} \tag{A26}$$

We can rewrite equation (A25) as

$$r^T = \frac{1}{\beta(1-\beta)^{T-1}}(r-1). \tag{A27}$$

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 (A27) is a line with a slope $\frac{1}{\beta(1-\beta)^{T-1}}$. We can verify that when $\beta = \frac{1}{T}$. The lefe-hand-side and the right-hand-side of equation (A27) 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 (A26), we know that $r'(\beta) < 0$. So $r_T(\beta,\delta)$ is a decreasing function of $\beta$.

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., $(w,w,\cdots,w)$ is the solution to the sophisticates' program. Consider the non-lapsing constraint at time 2. First, since $(w,\cdots,w)$ 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 the limiting results. As $\beta \to 1$, the right-hand-side of equation (A23) becomes $\beta r$. So, $\lim_{\beta\nearrow 1} r_T(\beta) = 1$. It then follows that $\lim_{\beta\nearrow 1} r_T(\beta,\delta) = \frac{\lim_{\beta\nearrow 1} r_T(\beta)}{\delta} = \frac{1}{\delta}$. As $T \to +\infty$, the right-hand-side of equation (A23) becomes $\frac{\beta r}{1-(1-\beta)r}$. Solving $\frac{\beta r}{1-(1-\beta)r} = 1$ gives $r = 1$. Thus, $\lim_{T\nearrow\infty} r_T(\beta,\delta) = \frac{\lim_{T\nearrow\infty} r_T(\beta)}{\delta} = \frac{1}{\delta}$. □

**Proof of Proposition 4.** We compare when the market breaks down for each type of agents while fixing an $\alpha$-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 next show that it is easier to sustain long-term contracting with sophisticates than with time-consistent agents. 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 $c_1 = c_2 = \ldots = c_T = w$, 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

$$\left(\beta\delta_I^{T-1}R^{T-1} - 1\right) u'(w)\epsilon. \tag{A28}$$

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

$$\beta\delta_I^{T-1}R^{T-1} \leq 1. \tag{A29}$$

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 (A29) implies $\delta_C R \leq 1$. That is, whenever the market breaks down for sophisticates, it also breaks down for time-consistent agents.

## Appendix B. Additional Results

### B.1 Market Power

Throughout the paper, we assumed that the consumer had all bargaining power. We now consider the case in which a firm has all the bargaining power. Since the firm can always commit to a contract, there is no loss of generality in assuming that the firm makes a take-it-or-leave it offer to the consumer, which happens at time 1. We assume that the interest rate $R$ is strictly greater than 1 so that the firm's profit is bounded above as the contracting horizon goes to infinity. Let $\underline{U}$ denote the consumer's outside option ("reservation utility").

As before, our main focus is on (partially naive) time-inconsistent consumers. As a benchmark, we also consider time-consistent consumers. Let $W_T^C(\underline{U})$ and $W_T^I(\underline{U})$ denote the welfare of time-consistent and time-inconsistent consumers, respectively. Let $V_T^C(\underline{U})$ and $V_T^I(\underline{U})$ denote the firm's profit when the consumer is time consistent and time inconsistent, respectively. We will omit the subscript $T$ for notational simplicity.

We now show that when the firm has the bargaining power, the inefficiency also vanishes as the horizon grows. However, unlike in the case where bargaining power is on the consumer's side, the equilibrium converges to a different point on the Pareto frontier.

**Proposition 5.** *Suppose $u$ is bounded, $\delta < 1$, and $R > 1$. Then,*

$$\lim_{T\nearrow\infty}(W^C(\underline{U}') - W^N(\underline{U})) = 0, \quad \lim_{T\nearrow\infty}(V^C(\underline{U}') - V^N(\underline{U})) = 0,$$

*where $\underline{U}' \equiv \underline{U} + (1-\beta)\frac{\delta}{1-\delta}u(0)$.*

*Proof.* For simplicity, we will only present the proof for the case with a constant deterministic income. The general proof follows the same steps as the proof of Theorem 1. The equilibrium profit with time-consistent agents solves:

$$V_T^C(\underline{U}) := \max_{\{c_t\}} \sum_{t=1}^{T} \frac{w - c_t}{R^{t-1}} \tag{B1}$$

subject to

$$\sum_{t=1}^{T} \delta^{t-1}u(c_t) \geq \underline{U} \tag{B2}$$

With time-inconsistent agents, the equilibrium profits are determined by:

$$V^N_{T,\beta,\hat{\beta}}(\underline{U}) := \max_{\{c_t\}} \sum_{t=1}^{T} \frac{w - c_t(A \cdots A)}{R^{t-1}} \tag{B3}$$

subject to (IC), (PCC), and

$$u(c_1) + \beta \sum_{t=2}^{T} \delta^{t-1} u(c_t(B \cdots B)) \geq \underline{U}. \tag{B4}$$

Consider the following auxiliary program:

$$V^A_{T,\beta}(\underline{U}') := \max_{\{c_t\}} \sum_{t=1}^{T} \frac{w - c_t}{R^{t-1}} \tag{B5}$$

subject to

$$\sum_{t=1}^{T-1} \delta^{t-1} u(c_t) + \beta \delta^{T-1} u(c_T) \geq \underline{U}' \tag{B6}$$

We show that the equilibrium consumption must solve the auxiliary program for

$$\underline{U}' = \underline{U} + (1-\beta)u(0)(\delta + \cdots + \delta^{T-2}).$$

For simplicity, we present the proof for $T = 3$. The proof for general $T$ is similar and is therefore omitted. The equilibrium consumption for time-inconsistent agents solves the following program:

$$\max_{(c_1, c_2(A), c_2(B), c_3(A), c_3(B))} w - c_1 + \frac{w - c_2(A)}{R} + \frac{w - c_3(A)}{R^2} \tag{B7}$$

subject to

$$u(c_1) + \beta[\delta u(c_2(B)) + \delta^2 u(c_3(B))] \geq \underline{U}, \tag{B8}$$

$$u(c_2(B)) + \hat{\beta}\delta u(c_3(B)) \geq u(c_2(A)) + \hat{\beta}\delta u(c_3(A)), \tag{B9}$$

$$u(c_2(A)) + \beta\delta u(c_3(A)) \geq u(c_2(B)) + \beta\delta u(c_3(B)). \tag{B10}$$

By the same argument as in the main text, it follows that the IC constraint (B10) must bind and $c_2(B) = 0$. Using (B10) to substitute for $c_3(B)$, the participation constraint (B8) becomes

$$u(c_1) + \beta\delta u(c_2(B)) + \beta\delta^2 u(c_3(B)) \tag{B11}$$

$$= u(c_1) + \beta\delta u(c_2(B)) + \delta(u(c_2(A) + \beta\delta u(c_3(A)) - u(c_2(B)) \tag{B12}$$

$$= u(c_1) + \delta u(c_2(A) + \beta\delta^2 u(c_3(A)) - \delta(1-\beta)u(c_2(B)) \tag{B13}$$

$$= u(c_1) + \delta u(c_2(A) + \beta\delta^2 u(c_3(A)) - \delta(1-\beta)u(0) \tag{B14}$$

We can verify that (B9) holds as long as the IC binds and $c_2(B) = 0$. Thus, the equilibrium consumption for time-inconsistent consumers solves:

$$\max_{(c_1, c_2, c_3)} w - c_1 + \frac{w - c_2}{R} + \frac{w - c_3}{R^2} \tag{B15}$$

subject to

$$u(c_1) + \delta u(c_2) + \beta\delta^2 u(c_3) \geq \underline{U} + \delta(1-\beta)u(0), \tag{B16}$$

This is the same program as the program with a dynamically consistent agent who discounts the last period by an extra $\beta$.

We now obtain the convergence result. Note that the participation constraints must be binding both in the auxiliary program and in the program for time-consistent consumers. So $W^C = \underline{U}$ and $\sum_{t=1}^{T-1} \delta^{t-1} u(c_t^A(\beta, T)) + \beta\delta^{T-1} u(c_T^A(\beta, T)) = \underline{U}$, where $c^A(\beta, T, \underline{U}) := (c_1^A(\beta, T, \underline{U}), \cdots, c_T^A(\beta, T, \underline{U}))$ denotes the equilibrium consumption in the auxiliary program. Omitting the dependence of $c^A$ on $\beta, T$, and $\underline{U}$ for notational simplicity, we have:

$$\begin{aligned}
W_{\beta,T}^A(\underline{U}) &= \sum_{t=1}^{T} \delta^{t-1} u(c_t^A) \\
&= \sum_{t=1}^{T-1} \delta^{t-1} u(c_t^A) + \beta\delta^{T-1} u(c_T^A) + (1-\beta)\delta^{T-1} u(c_T^A) \\
&= \underline{U} + (1-\beta)\delta^{T-1} u(c_T^A) \\
&= W^C + (1-\beta)\delta^{T-1} u(c_T^A)
\end{aligned}$$

So $\lim_{T\nearrow\infty} |W^C - W^A| = \lim_{T\nearrow\infty} (1-\beta)\delta^{T-1} u(c_T^A) = 0$ (since $u$ is bounded and $\delta < 1$).

We now turn to the firm's profit. Let $\lambda$ denote the Lagrangian multiplier with the constraint (B6). FOC gives

$$\lambda\delta^{t-1} u'(c_t^A) = \frac{1}{R^{t-1}}, \forall t = 1, \cdots, T-1$$

and

$$\lambda\beta\delta^{T-1} u'(c_T^A) = \frac{1}{R^{T-1}}$$

Differentiating w.r.t $\beta$ on the binding participation constraint gives

$$\sum_{t=1}^{T-1} \delta^{t-1} u'(c_t^A) \frac{\partial c_t^A}{\partial \beta} + \beta\delta^{T-1} u'(c_T^A) \frac{\partial c_T^A}{\partial \beta} + \delta^{T-1} u(c_T^A) = 0.$$

Now,

$$\begin{aligned}
\frac{\partial V^A(\beta)}{\partial \beta} &= \sum_{t=1}^{T} \frac{-\frac{\partial c_t^A}{\partial \beta}}{R^{t-1}} \\
&= -\sum_{t=1}^{T-1} \lambda\delta^{t-1} u'(c_t^A) \frac{\partial c_t^A}{\partial \beta} - \lambda\beta\delta^{T-1} u'(c_T^A) \frac{\partial c_T^A}{\partial \beta} \\
&= \lambda\delta^{T-1} u(c_T^A) \geq \lambda\delta^{T-1} u(0)
\end{aligned}$$

where the inequality comes from $c_T^A \geq 0$. Applying Lagrange's Mean Value Theorem gives

$$\begin{aligned}
V^A(1) - V^A(\beta) &= \frac{\partial V^A(\beta)}{\partial \beta}\Big|_{\beta=\beta'} (1-\beta) \geq \lambda\delta^{T-1} u(0)(1-\beta) \\
V^A(\beta) - V^A(0) &= \frac{\partial V^A(\beta)}{\partial \beta}\Big|_{\beta=\beta''} \beta \geq \lambda\delta^{T-1} u(0)\beta
\end{aligned}$$

where $\beta' \in (\beta, 1), \beta'' \in (0, \beta)$. Since $\lambda$ is bounded because $\lambda = \frac{1}{u'(c_1^A)} \leq \frac{1}{u'\left(w \sum_{t=1}^{T} \frac{1}{R^{t-1}}\right)} \leq \frac{1}{u'\left(\frac{w}{1-\frac{1}{R}}\right)}$, sending $T$ to infinity leads to

$$\lim_{T \nearrow \infty} V^A(1) \geq \lim_{T \nearrow \infty} V^A(\beta)) \geq \lim_{T \nearrow \infty} V^A(0).$$

In order to show that $\lim_{T \nearrow \infty}(V^A(1) - V^A(\beta)) = 0$, it is sufficient to show that $\lim_{T \nearrow \infty}(V^A(1) - V^A(0)) = 0$.

We write the program for $V^A(0)$:

$$V_T^A(0) := \max_{\{c_t\}} \sum_{t=1}^{T} \frac{w - c_t}{R^{t-1}} \tag{B17}$$

subject to

$$\sum_{t=1}^{T-1} \delta^{t-1} u(c_t) \geq \underline{U} \tag{B18}$$

It is immediate that $c^T$ must be equal to $0$. Then the program reduces to

$$V_T^A(0) := \max_{\{c_t\}} \sum_{t=1}^{T-1} \frac{w - c_t}{R^{t-1}} + \frac{w}{R^{T-1}} \tag{B19}$$

subject to

$$\sum_{t=1}^{T-1} \delta^{t-1} u(c_t) \geq \underline{U} \tag{B20}$$

Then $V_T^A(0) = V_{T-1}^A(1) + \frac{w}{R^{T-1}}$. It follows that $\lim_{T \nearrow \infty}(V_T^A(1) - V_T^A(0)) = \lim_{T \nearrow \infty}(V_T^A(1) - V_{T-1}^A(1) - \frac{w}{R^{T-1}}) = \lim_{T \nearrow \infty}(V_T^A(1) - V_{T-1}^A(1)) - \lim_{T \nearrow \infty} \frac{w}{R^{T-1}} = \lim_{T \nearrow \infty}(V_T^A(1) - V_{T-1}^A(1)) = 0$. $\square$

### B.2 Removing Commitment Power for Sophisticates

In this appendix, we show that, for a fixed contract length, removing commitment power can make the sophisticated consumer better off. The intuition for this result is that commitment power allows the consumer to borrow, and time-inconsistent consumers 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 consumer is very time inconsistent ($\beta$ is small) and saving is important for welfare ($\delta$ is large). In that case, removing commitment power increases welfare.

**Proposition 6.** *Suppose the consumer 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.*

*Proof.* Fix an equilibrium with two-sided commitment. 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. Let $x(s_t) = u(c(s_t))$. By the dual program, there exists some utility level $\underline{u}$ to the agent, for which $\{x(s_t)\}$ solves the program:

$$\max_{\{x(s_t)\}} \sum_{t=1}^{T} \sum_{s_t \in \mathbb{S}_t} p(s_t|s_1) \frac{w(s_t) - u^{-1}(x(s_t))}{R^{t-1}}, \tag{B21}$$

subject to

$$x(s_1) + \beta \sum_{t=2}^{T} \sum_{s_t \in \mathbb{S}_t} \delta^{t-1} p(s_t|s_1) x(s_t) \geq \underline{u}, \tag{B22}$$

and

$$x(s_{\tilde{\tau}}) + \beta \sum_{t>\tilde{\tau}} \sum_{s_t \in \mathbb{S}_t} \delta^{t-\tilde{\tau}} p(s_t|s_{\tilde{\tau}}) x(s_t) \geq V^S(s_{\tilde{\tau}}) \ \ \forall s_{\tilde{\tau} \in \mathbb{S}_\tau}, \forall \tau, \tag{B23}$$

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 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

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

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

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

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 6. $\square$

### B.3 Results with General Discounting Models

In the text, we assumed quasi-hyperbolic discounting, which is the canonical model of present bias. In this appendix, we generalize the analysis for arbitrary preferences with present bias. We establish a version of the equivalence between the equilibrium and the solution of an auxiliary program, and a weak version of the welfare result, which shows that a naive consumer saves more than a sophisticated consumer. For the ease of exposition, we assume that income is deterministic.

To allow for arbitrary time discounting, we assume that at time $\tau \in \{1, 2, ...\}$, the agent evaluates a consumption stream $\{c_t\}_{t\geq\tau}$ according to

$$u(c_\tau) + \sum_{t>\tau} D_{t-\tau} u(c_t), \tag{B24}$$

where $D_t \in (0, 1)$ is a decreasing discount factor. It is well known that these preferences are time consistent if and only if $D_t = D_1^t$. We assume, instead, that preferences are present biased:

$$D_{i+j} > D_i D_j$$

for all $i, j$. This inequality states that the individual becomes more impatient as a period approaches. For example, with quasi-hyperbolic discounting, we have

$$D_{i+j} = \beta\delta^{i+j} > \beta^2\delta^{i+j} = D_i D_j.$$

It is straightforward to verify that the inequality also holds under hyperbolic discounting, where $D_t = \frac{1}{1+kt}$.

The agent can be naive or sophisticated. A *naif* has true time-consistency parameter $D$ but believes that, in the future, he will behave like an agent with time-consistency parameter $\hat{D} = (\hat{D}_1, \cdots, \hat{D}_{T-2})$ where $\hat{D}_i > D_i$. A *sophisticate* perfectly knows his time-consistency parameter: $\hat{D}_i = D_i$. We assume that a naive agent overestimates the patience of his future selves:

$$\frac{\hat{D}_{i+1}}{\hat{D}_i} \geq \frac{D_{i+1}}{D_i}$$

with strict inequality for at least one $i$. With quasi-hyperbolic discounting, this condition becomes $\hat{\beta} \geq \beta$.

Consider the following auxiliary program

$$u(c_1) + \frac{D_{T-1}}{D_{T-2}} u(c_2) + \frac{D_{T-1}}{D_{T-3}} u(c_3) + \cdots + D_{T-1} u(c_T)$$

We claim that a naif's program is equivalent to solving the auxiliary program. For notational simplicity, we present the proof for the special case of $T = 4$. When $T = 4$, a naif solves the following program

$$\max_{(c)} u(c_1) + D_1 u(c_2(B)) + D_2 u(c_3(BB)) + D_3 u(c_4(BB)), \tag{B25}$$

subject to

$$
\begin{aligned}
&c_1 + \frac{c_2(A)}{R} + \frac{c_3(AA)}{R^2} + \frac{c_4(AA)}{R^3} = w(1 + \frac{1}{R} + \frac{1}{R^2} + \frac{1}{R^3}),\\
&u(c_2(B)) + \hat{D}_1 u(c_3(BB)) + \hat{D}_2 u(c_4(BB)) \geq u(c_2(A)) + \hat{D}_1 u(c_3(AB)) + \hat{D}_2 u(c_4(AB)),\\
&u(c_2(A)) + D_1 u(c_3(AB)) + D_2 u(c_4(AB)) \geq u(c_2(B)) + D_1 u(c_3(BB)) + D_2 u(c_4(BB)),\\
&u(c_3(AB)) + \hat{D}_1 u(c_4(AB)) \geq u(c_3(AA)) + \hat{D}_1 u(c_4(AA)),\\
&u(c_3(AA)) + D_1 u(c_4(AA)) \geq u(c_3(AB)) + D_1 u(c_4(AB)).
\end{aligned}
$$

Consider the following perturbation $c_2(B) - \epsilon$ and $c_4(BB) + \frac{1}{D_2}\epsilon$. This does not change the IC, but improves the objective function by $-D_1 + \frac{D_3}{D_2}$, which is positive because of $D_3 > D_2 D_1$.

Consider the following perturbation $c_3(BB) - \frac{1}{D_1}\epsilon$ and $c_4(BB) + \frac{1}{D_2}\epsilon$. This does not affect PPC because $\frac{\hat{D}_2}{D_2} \geq \frac{\hat{D}_1}{D_1}$. The perturbation does not change the IC, but improves the objective function by $-\frac{D_2}{D_1} + \frac{D_3}{D_2} > 0$, which holds because of $D_3 D_1 > D_2^2$.

Substituting $c_2(B) = c_3(BB) = 0$, we obtain the new objective function

$$
u(c_1) + \frac{D_3}{D_2} u(c_2(A)) + \frac{D_3 D_1}{D_2} u(c_3(AB)) + D_3 u(c_4(AB)).
$$

Consider the following perturbation $c_3(AB) - \epsilon$ and $c_4(AB) + \frac{1}{D_1}\epsilon$. This does not change the IC, but improves the objective function by $-\frac{D_3 D_1}{D_2} + \frac{D_3}{D_1} > 0$, which holds because of $D_2 > D_1 D_1$.

Substituting $c_3(AB) = 0$, we obtain the objective function in our auxiliary program (up to a constant):

$$
u(c_1) + \frac{D_3}{D_2} u(c_2(A)) + \frac{D_3}{D_1} u(c_3(AA)) + D_3 u(c_4(AA)).
$$

Note that, for arbitrary present-biased preferences, the welfare criterion is generally ambiguous. Instead of considering each potential welfare criterion, we focus on the effect of naiveté on savings. Recall that this property is at the heart of the market breakdown result as discussed in Section 4.

A naif's consumption stream is denoted by $c^N$, which is given by

$$
c^N \equiv (c_1^N, \cdots, c_T^N) = \arg\max u(c_1) + \frac{D_{T-1}}{D_{T-2}} u(c_2) + \cdots + D_{T-1} u(c_T).
$$

A sophisticate's consumption stream is denoted by $c^S$, which is given by

$$
c^S \equiv (c_1^S, \cdots, c_T^S) = \arg\max u(c_1) + D_1 u(c_2) + \cdots + D_{T-1} u(c_T).
$$

We are now ready to present the following result.

**Proposition 7.** *A naif saves more than a sophisticate in the first period, i.e.,* $c_1^N < c_1^S$.

*Proof.* Consider the equations

$$
u'(c_1) = x_2 u'(c_2) = \cdots = x_{T-1} u'(c_{T-1}) = D_{T-1} u'(c_T), \tag{B26}
$$

$$
\sum \frac{c_i - w}{R^{i-1}} = 0. \tag{B27}
$$

So a naif's problem corresponds to the solution of above equations associated with $x^N = (x_2, \cdots, x_{T-1}) = (\frac{D_{T-1}}{D_{T-2}}, \cdots, \frac{D_{T-1}}{D_1})$. A sophisticate's problem corresponds to $x^S = (x_2, \cdots, x_{T-1}) = (D_1, \cdots, D_{T-2})$. Since $D_{i+j} > D_i D_j$, it is clear that $x^N > x^S$.

Returning to the system of (B26) and (B27), we can calculate

$$\frac{\partial c_1}{\partial x_i} = \frac{\frac{u'(c_i)}{x_i u''(c_i)}}{1 + \sum_{j=2}^{T-1} \frac{u''(c_1)}{x_j u''(c_j)} + \frac{u''(c_1)}{D_{T-1} u''(c_T)}} < 0.$$

In other words, $c_1$ as a function of $x$ is decreasing in every coordinate of $x$. As a result, $c_1^N = c_1(x^N) < c_1(x^S) = c_1^S$, establishing our claim that a naif consumes less than a sophisticate. □

### B.4 Results with Unknown Naiveté Parameter

In this appendix, we show that our results remain when firms do not know the consumer's naiveté parameter $\hat{\beta}$. We first show that there is no loss of generality in assuming that all types of the consumers should choose the same contract, so that his actual choice of contracts will convey no information (Myerson (1983)). Essentially, the consumer should never need to communicate any information to the firm by his choice of mechanisms, because he can always build such communication into the mechanism itself.

Suppose two contracts $\mathcal{C}$ and $\mathcal{C}'$ are offered in equilibrium by different types of consumers, then we can consider a contract $\mathcal{C}''$ that gives the consumer an option to choose among $\mathcal{C}$ and $\mathcal{C}'$ right after the firm accepts the contract. Offering $\mathcal{C}''$ yields the exactly same outcomes. Therefore, there is no loss of generality in assuming that all types of the consumers should choose the same contract, conveying no information, in either one-sided commitment or two-sided commitment setting.

Because of the informed principal argument, we can assume that all-type agents offer the same contract and that is accepted by the Principal at time 1. Since as we shows in Corollary 1, the equilibrium contract does not depend on $\hat{\beta}$ when $\hat{\beta}$ is known. As a result, it follows that when $\hat{\beta}$ is unknown, there is a SPNE in which the equilibrium contract is given by the equilibrium contract when $\hat{\beta}$ is known.

Furthermore, the equilibrium is unique defined as outcome-equivalence. The proof is similar to Yilankaya (1999), who shows that if the ex-ante optimal mechanism is also the ex-post optimal mechanism, then the interim optimal mechanism is exactly given by the ex-post optimal mechanism. Formally, due to the principle of inscrutability, there is no loss of generality in assuming that all agent types will choose the same contract $\mathcal{C}_1$. All agent types will have payoff at least as large as their payoff from $\mathcal{C}_1$. Applying Lemma 1 implies that the ex-ante optimal contract for the agent is independent of perceived time-consistency parameter and is exactly $\mathcal{C}_1$. So, suppose that an agent type $\beta_i$ obtains a higher expected payoff than what he would obtain from $\mathcal{C}_1$. This would contradict the ex ante optimality of $\mathcal{C}_1$. All agent types will obtain payoff equal to what he would obtain from $\mathcal{C}_1$. Since the auxiliary program has unique consumption, the equilibrium is therefore unique in terms of outcome-equivalence.

### B.5 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 $T = 3, \beta = 0.9, R = 1.1, \delta = 1, w = 1, u(c) = \log(c + \epsilon)$, where $\epsilon$ is a very small positive number. 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 constraints hold:

$$c_1^1 + \frac{c_2^1}{R} + \frac{c_3^1}{R^2} = 2.7355 = 1 + \frac{1}{R} + \frac{1}{R^2}, \tag{B28}$$

$$u(c_2^1) + \beta u(c_3^1) = 0.0224 > 0 = V_2^S. \tag{B29}$$

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:

$$c_1^2 + \frac{c_2^2}{R} + \frac{c_3^2}{R^2} = 2.7355 = 1 + \frac{1}{R} + \frac{1}{R^2}, \tag{B30}$$

$$u(c_2^2) + \beta u(c_3^2) = 0.0241 > 0 = V_2^S. \tag{B31}$$

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:

$$\begin{aligned} u(c_1^1) + u(c_2^1) + u(c_3^1) &= \frac{1}{\beta}\left(u(c_1^1) + \beta(u(c_2^1) + u(c_3^1))\right) - \frac{1-\beta}{\beta}u(c_1^1) \\ &\geq \frac{1}{\beta}\left(u(c_1^2) + \beta(u(c_2^2) + u(c_3^2))\right) - \frac{1-\beta}{\beta}u(c_1^1) \\ &> \frac{1}{\beta}\left(u(c_1^2) + \beta(u(c_2^2) + u(c_3^2))\right) - \frac{1-\beta}{\beta}u(c_1^2) \\ &= u(c_1^2) + u(c_2^2) + u(c_3^2), \end{aligned} \tag{B32}$$

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.

Oren Bar-Gill. The law, economics and psychology of subprime mortgage contracts. *Cornell Law Review*, 94(5):1073–1152, 2009.

Alberto Bisin, Alessandro Lizzeri, and Leeat Yariv. Government policy with time inconsistent voters. *American Economic Review*, 105(6):1711–37, 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.

Dan Cao and Iván Werning. Saving and dissaving with hyperbolic discounting. *Econometrica*, forthcoming.

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.

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.

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. Overconfident consumers in the marketplace. *The Journal of Economic Perspectives*, 29(4):9–35, 2015.

Jonathan Gruber and Botond Koszegi. Is addiction "rational"? theory and evidence. *Quarterly Journal of Economics*, 116(4):1261–1303, 2001.

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.

Christopher Harris and David Laibson. Dynamic choices of hyperbolic consumers. *Econometrica*, 69(4):935–957, 2001.

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.

Igor Makarov and Guillaume Plantin. Equilibrium subprime lending. *Journal of Finance*, 68(3): 849–879, 2013.

Stephen Meier and Charles Sprenger. Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics*, 2(1):193–210, 2010.

Roger B Myerson. Mechanism design by an informed principal. *Econometrica: Journal of the Econometric Society*, pages 1767–1797, 1983.

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.

Ted O'Donoghue and Matthew Rabin. Studying optimal paternalism, illustrated by a model of sin taxes. *American Economic Review*, 93(2):186–191, May 2003.

Ted O'Donoghue and Matthew Rabin. Present bias: Lessons learned and to be learned. *American Economic Review*, 105(5):273–79, May 2015.

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.

Okan Yilankaya. A note on the seller's optimal mechanism in bilateral trade with two-sided incomplete information. *Journal of Economic Theory*, 87(1):267–271, 1999.

# Online Appendix - Not For Publication

**Proof of Lemma 2 for General Income Distributions.** In this appendix, we show that a naive agent's program is equivalent to the auxiliary program for general income distribution. We show the results for one-sided commitment. If there is two-sided commitment, we can simply ignore the non-lapsing constraints in the proof.

Recall that the naive agent's program is

$$\max_{c(s_t,h^t)} u(c(s_1)) + \beta E\left[\sum_{t=2}^{T} \delta^{t-1} u(c(s_t, B, B, ..., B))\right],$$

subject to

$$\sum_{t=1}^{T} E\left[\frac{w(s_t) - c(s_t, A, A, ..., A)}{R^{t-1}}\right] = 0, \tag{Zero Profits}$$

$$u(c(s_\tau, (h_\tau, B))) + \hat{\beta} E\left[\sum_{t>\tau} \delta^{t-\tau} u(c(s_t, (h_\tau, B, B, ..., B)))\middle| s_\tau\right] \tag{PCC}$$

$$\geq u(c(s_\tau, (h_\tau, A))) + \hat{\beta} E\left[\sum_{t>\tau} \delta^{t-\tau} u(c(s_t, (h_\tau, A, B, ..., B)))\middle| s_\tau\right],$$

and

$$u(c(s_\tau, (h_\tau, A))) + \beta E\left[\sum_{t>\tau} \delta^{t-\tau} u(c(s_t, (h_\tau, A, B, ..., B)))\middle| s_\tau\right] \tag{IC}$$

$$\geq u(c(s_\tau, (h_\tau, B))) + \beta E\left[\sum_{t>\tau} \delta^{t-\tau} u(c(s_t, (h_\tau, B, B, ..., B)))\middle| s_\tau\right],$$

and non-lapsing constraints:

$$u(c(s_\tau, h^\tau)) + \beta E\left[\sum_{t>\tau} \delta^{t-\tau} u(c(s_t, h^\tau, B, B, ..., B))\middle| s_\tau\right] \geq V(s_\tau, h^\tau) \ \forall (s_\tau, h^\tau), \tag{NL}$$

and

$$u(c(s_\tau, h^\tau)) + \hat{\beta} E\left[\sum_{t>\tau} \delta^{t-\tau} u(c(s_t, h^\tau, B, B, ..., B))\middle| s_\tau\right] \geq \tilde{V}(s_\tau, h^\tau) \ \forall (s_\tau, h^\tau). \tag{PNL}$$

We first note that the incentive compatibility constraints (IC) must be binding on the equilibrium path, because otherwise we can increase $c(s_T, h^\tau, B, B, ..., B)$ without affecting all other constraints while weakly increase the agent's perceived utility. Given incentive constraints are binding, we can simplify (PCC) as

$$u(c(s_\tau, h^\tau, B)) \leq u(c(s_\tau, h^\tau, A)). \tag{OA1}$$

Substituting the binding IC constraints in the objective gives

$$\sum_{t=1}^{T-1}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_1)\delta^{t-1}u(c(s_t,B,B,\cdots,B)) + \beta\sum_{s_T\in\mathbb{S}_T} p(s_T|s_1)\delta^{T-1}u(c(s_T,B,B,\cdots,B))$$
$$+ (\beta-1)\sum_{t=2}^{T-1}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_1)\delta^{t-1}u(c(s_t,B,B,\cdots,B,A))$$

Since $\beta < 1$, we want to choose $c(s_t, B, B, \cdots, B, A)$ as small as possible (subject to the constraints). We now show that under the optimal contract, $c(s_t, B, B, \cdots, B, A) = 0$. We need to verify that setting $c(s_t, B, B, \cdots, B, A) = 0$ would not violate all other constraints. First, (PCC) holds because (OA1) holds.

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

$$\hat{V}(s_\tau,h^\tau) = u(\hat{c}(s_\tau,B(h^\tau))) + \hat{\beta}\sum_{t>\tau}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_\tau)\delta^{t-\tau}u(\hat{c}(s_t,h^\tau,B,\cdots,B)). \quad \text{(OA2)}$$

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

$$u(c(s_\tau,h^\tau,B)) + \hat{\beta}\sum_{t>\tau}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_\tau)\delta^{t-\tau}u(c(s_t,h^\tau,B,\cdots,B))$$
$$= u(0) + \hat{\beta}\sum_{t>\tau}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_\tau)\delta^{t-\tau}u(c(s_t,h^\tau,B,\cdots,B)) \quad \text{(OA3)}$$
$$= (1-\frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}\left(u(c(s_\tau,h^\tau,A)) + \beta\sum_{t>\tau}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_\tau)\delta^{t-\tau}u(c(s_t,h^\tau,A,B,\cdots,B))\right) \quad \text{(OA4)}$$
$$\geq (1-\frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}V(s_\tau,h^\tau) \quad \text{(OA5)}$$
$$\geq (1-\frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}\left(u(\hat{c}(s_\tau,h^\tau,B)) + \beta\sum_{t>\tau}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_\tau)\delta^{t-\tau}u(\hat{c}(s_t,h^\tau,B,\cdots,B))\right) \quad \text{(OA6)}$$
$$= (1-\frac{\hat{\beta}}{\beta})u(0) + \frac{\hat{\beta}}{\beta}u(\hat{c}(s_\tau,h^\tau,B)) + \hat{\beta}\sum_{t>\tau}\sum_{s_t\in\mathbb{S}_t} p(s_t|s_\tau)\delta^{t-\tau}u(\hat{c}(s_t,h^\tau,B,\cdots,B)) \quad \text{(OA7)}$$
$$= (1-\frac{\hat{\beta}}{\beta})u(0) + (\frac{\hat{\beta}}{\beta}-1)u(\hat{c}(s_\tau,h^\tau,B)) + \hat{V}(s_\tau,h^\tau) \quad \text{(OA8)}$$
$$\geq \hat{V}(s_\tau,h^\tau), \quad \text{(OA9)}$$

where (OA3) is from $c(s_\tau, B(h^\tau)) = 0$, (OA4) is from (IC), (OA5) is from the actual non-lapsing constraints (NL), (OA6) is due to the revealed preference argument since $\hat{c}$ is also a candidate contract for the program $V(s_\tau, h^\tau)$, (OA7) is a simple algebra, (OA8) comes from the definition of $\hat{V}(s_\tau, h^\tau)$ is (OA2), and finally (OA9) is because of $\hat{c}(s_\tau, h^\tau, B) \geq 0$.

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.

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

subject to the zero-profit condition and the actual non-lapsing constraints (NL). Then, we can rewrite the non-lapsing constraints as the front-loading constraints (FL). So $c^{1E} = c^{1A}$. $\square$