Naïve stochastic present bias and credit

A. Citanna
New York University - Abu Dhabi

P. Siconolfi
Columbia Business School

November 11, 2022

Abstract

Consumers with false beliefs and uncertain intrapersonal variations in present bias are shown to display heterogeneous behavior in the face of exploitative credit contracts. Their degree of optimism, the intensity of their precautionary savings motives, and their 'satiation' level may be responsible for under or over borrowing with respect to sophisticated consumers. Arising inefficiencies may not asymptotically vanish, even with nonsatiation and in the absence of additional contractual frictions such as market power or adverse selection. (JEL D14, D81, D86, G51)

Keywords: Present bias, optimism, borrowing, long-run welfare.

1 Introduction

Much progress has been made in the last twenty-five years in the study of the consequences of time inconsistency, and present bias, on economic decisions.[1] Particularly, contractual consequences of naïveté and present bias have been explored for credit markets (see, e.g., Heidhues and Köszegi (2010) and, more recently, Gottlieb and Zhang (2021)).

[1]See, e.g., Frederick, Loewenstein and O'Donoghue (2002), Laibson (2006) for earlier, and O' Donoghue and Rabin (2015) for more recent, general assessments of the present bias literature.

Most of the studies assume that individual preferences are represented by a single procrastination parameter, while variations may occur across individuals, at the population level. Evidence from the psychology and neuroscientific literature instead highlights that even for a single individual behavior, and thus the parameters representing it, may be subject to variations over time. These intrapersonal variations stem from influences of the environment on the neuro-cerebral system. In other words, while the chemistry-physics of the wiring and functioning of this system is surely constrained by genetics, and may lead an individual to be prone to procrastination as a by-product of the interaction of its limbic system with the frontal cortex,[2] epigenetic phenomena shape the brain and affect behavior, even past the fetal or childhood portion of life. In fact, sensory inputs produced by stressful external situations may alter individual behavior, for instance in the ability to perform tasks at the 'appropriate' or planned time –e.g., procrastinating. Personal events, such as divorce, the birth of a child, or unemployment, as well as macroeconomic events such as a recession —as they may produce stress— may affect the intensity of procrastinating behavior.[3] Personality psychology has recognized that individual traits are subject to intrapersonal variations due to internal or external 'states' or situations.[4] Thus, an individual should be more properly characterized by a distribution or process of states affecting their procrastination parameter, as well as by their perception of, or beliefs over, these states.

It is an easy further step to consider that firms offering credit to individuals can estimate the potential situations an individual may face through observables such as age, gender, profession, marital status, and other socio-economic variables. In other words, the firm may have quite accurate statistical information on the probability of future variation in the intensity of procrastination, whereas the individual may entertain beliefs which disagree with such probabilistic assessment. Does this more complex picture of naïve present bias have economic consequences? In this paper, we take this approach and study the case of borrowing behavior.

Positing present bias and the consumers' wrong estimation of their tendency to pro-

[2]For a dual-self theory of present-bias see, e.g., McClure, Laibson, Loewenstein, and Cohen, 2004.

[3]For an accessible, divulgative but authoritative summary of this viewpoint, see Sapolsky (2017).

[4]See, e.g., the survey by Ness, Foley, and Heggestad (2021), and van Eerde (2003), that support the possibility that the intensity of procrastination depend on external situations or states, which are inherently random.

crastinate, we take a closer look at other independent features of consumers' preferences, reflecting ancillary behavioral features such as risk aversion and prudence. We then find that these additional features combine with present bias and naïveté to deliver a wider range of behaviors. The ensuing heterogeneity of consumers' behavior may open the door to more successful identification strategies to test the proposed theories of present-biased-driven borrowing behavior. As in the data prudence and risk aversion are often correlated with the consumers' socioeconomic status, the latter may be responsible for consumers' over or underborrowing, and for their resulting long-run welfare.

In our model individuals display time inconsistency of the β-δ type, and seek to solve a consumption-savings problem. They face idiosyncratic uncertainty over time, in particular regarding the future value of the procrastination parameter β. They have beliefs μ over β, and incorrectly perceive their procrastination tendencies. Firms are fully rational and aware of an individual's misperception of their present bias. They have correct beliefs π about a consumer's present bias realizations and future income. They take advantage of individuals' misperceptions by offering flexible contracts, i.e., menu contracts with multiple repayment options to be exercised by the consumer's future selves. Our consumers can borrow today and at future dates as long as they have lifetime income to pay back.

Within this setup, we study the contract that a naïve present-biased consumer would receive at the equilibrium of a competitive market for contracts. We show that two features of individual preferences and beliefs play a key role: the intensity of precautionary savings motives, or prudence; and the level of naïveté.

It turns out that prudence is the main feature of preferences linked to the naïve consumers' borrowing behavior. This is most starkly seen when the firms' assessment of the individual intrapersonal variations of procrastination intensity corresponds to a probability distribution with full support. Then, when the prudence-to-risk aversion ratio is greater than two, present-biased and partially naïve consumers display underborrowing. The prudence-to-risk aversion ratio is $v'''v'/(v'')^2$, where v', v'' and v''' are the utility first, second and third derivatives, and $-v'''/v''$ measures precautionary savings motives (see, e.g., Kimball 1990). A log utility displays a ratio of two. Underlying the effect of prudence on borrowing behavior is consumption variability in the optimal credit contract, which in turn depends on the consumer's beliefs being pessimistic or optimistic. We define pessimism, and show that pessimistic consumers are immune from manipulative distortions. Indeed, their optimal contract prescribes no variation in future consumption, or full insurance, just as the committed sophisticated consumer likes. However, as long as the consumer is risk averse and not pessimistic about her own tendency to procrastinate, incentive constraints distort the sophisticated contract requiring variability in future consumption across self state realizations. As a result, with sufficient prudence a precautionary savings motive translates this future consumption variation into higher savings at time zero, or less borrowing than the sophisticated consumer would sign up for.

We should stress that mixed patterns of borrowing behavior can obtain while we fix preferences and income parameters (with over- or under-borrowing depending on individual perceptions of their procrastination tendencies, or on their subjective beliefs).

As for welfare consequences, we prove that the inefficiency due to naïve present bias distortions does not vanish asymptotically. Thus, under naïve uncertain present bias markets fail from a long-run welfare perspective even if there is no adverse selection –even if firms know the consumer's type before the contract offer. Unlike in the special case of a deterministic individual procrastination parameter studied in Citanna, Gottlieb, Siconolfi and Zhang (2022b), persistent inefficiencies arise even with long-term contracting no matter what the level of expected lifetime income is. In Citanna et al. (2022b) the lifetime income level mattered because of 'imaginary satiation': selves which do not realize under π can achieve their consumption upper bound in the baseline contract option when they are rich enough. Here either the true distribution displays full support, and hence there are no imaginary selves, or we impose a nonsatiation condition. Yet, consumers suffer from the negative consequences of their naïveté and bias no matter the contractual time horizon. The reason is the induced volatility of future consumption in the optimal contract, due to optimism and uncertain present bias.

Related literature Following Strotz's 1956 pioneering analysis and work by, e.g., O' Donoghue and Rabin (1999, 2001), the literature has examined the combined effects of present bias with the individual's optimistic assessment of their bias, or naïveté. This literature has focused on the case where the individual's true procrastination parameter β is certain. An individual is present biased if $\beta < 1$. The individual is deemed (fully) naïve if they believe future selves are biased instead according to a value $\hat{\beta} = 1$, and is partially naive if $\hat{\beta} \in (\beta, 1)$. We embed all of these cases in our framework, which generalizes also the notion of perception-perfect equilibrium (as introduced for the special case of no uncertainty in O' Donoghue and Rabin (1999, 2001)).

Specifically for borrowing behavior, the closest study to our own is Heidhues and Köszegi (2010) (see also Köszegi (2014)) and Sulka (2022)). Focusing on the effects of naïveté as a determinant of borrowing behavior, and thus comparing naïve and sophisticated consumers, Heidhues and Köszegi (2010) conclude that naïve present-biased agents overborrow in credit markets, reflecting the 'exploitative nature' of financial contracts.[5] We show that the effect on borrowing depends primarily on the prudence-to-risk aversion ratio. Heidhues and Köszegi's overborrowing result followed from having imposed that consumers can only borrow today.[6] With uncertain present bias, overborrowing is obtained, without limits to future borrowing, when the consumers' prudence-to-risk aversion ratio is low, namely, less than or equal to 2.

By using time-consistent sophisticated 'long-run preferences' as a welfare measure for present-biased naïves, recently Gottlieb and Zhang (2021) argued that naïve present-biased consumers do not suffer much from their bias unless there are other imperfections –e.g., market power, or adverse selection. Dropping their assumption that the consumption upper bound is not attainable, the analysis of this special case where the consumers' present bias parameter is certain (i.e., where $S = \{\beta, \hat{\beta}\}$ and $\pi(\beta) = 1 = \mu(\hat{\beta})$) in Citanna et al. (2022b) already unveils that the inefficiency due to naïve present bias may not dissipate even when the contracting horizon is large. This is shown to be a function of the individual's lifetime expected income. Our analysis here of the case when the consumers' present bias parameter is uncertain shows that the inefficiencies extend more broadly in these environments, and do not depend on expected lifetime income. Thus, even if there is no adverse selection problem, markets do not perform well.

Elasticity of substitution, prudence are also typically correlated with wealth: the

[5] Contracts are exploitative when their primary aim is to take advantage of agents' mistakes. See Köszegi (2014).

[6] Heidhues and Koszegi (2010) and Koszegi (2014) correspond, in our model, to the special case where the state space is $S = \{\beta, \hat{\beta}\}$, the time horizon is $T = 2$ and beliefs are $\pi(\beta) = 1$ while $\mu(\hat{\beta}) = 1$. We show in Citanna et a. (2022b) that Köszegi's (2014) example is reversed if the unrealistic one-time borrowing restriction is removed.

micro data also shows that many wealthier households display larger prudence and elasticity coefficients greater than one.[7] Empirical evidence in support of overborrowing because of naïve present bias seems hard to come by (see Zinman, 2014).[8] Our results suggest that one way around the empirical difficulties is to take into account consumers' preferences heterogeneity, and use wealth differences as a proxy for heterogeneous prudence and intertemporal substitution motives.

Our granular approach to borrowing by present-biased naïves is inspired by Strotz's' (1956) own analysis and O' Donoghue and Rabin's (1999) comments in a footnote,[9] conjecturing that consumer behavior may depend on ancillary features of preferences (e.g., concavity of utility). We show that their conjectured conclusions are almost correct, because even utilities more concave than log may display underborrowing if beliefs are sufficiently optimistic.

From a methodological viewpoint, we cast the contracting problem under naïve present bias as one of dynamic hidden information (of the kind studied, e.g., by Thomas and Worrall (1988)). As a result, a key inverse marginal utility inequality must hold. When the inverse marginal utility function is concave and under nonsatiation, the inverse marginal utility inequality is strict via Jensen's inequality, yielding underborrowing.

Section 2 introduces the environments. Section 3 defines the allocation problem, and efficiency. Section 4 studies the case where there are no imaginary selves, i.e., where the true selves distribution has full support. Section 4.2 defines pessimism, and characterizes borrowing behavior and welfare for naïve individuals. Section discusses the case where the full support condition fails. Appendix A contains some proofs. Appendix B adds ancillary results: it discusses pessimism, and a closed economy.

[7]Havranek's (2015) metastudy underscores a large variance in the elasticity of intertemporal substitution estimates across socio-economic groups. Guvenen (2006) provides evidence of the correlation between wealth heterogeneity and the elasticity. Positive correlation between prudence and wealth has been documented in experimental studies, most recently, e.g., by Noussair, Trautmann and van de Kuilen (2014).

[8]Meier and Sprengen (2010), e.g., find, in a lab experiment, that there is overborrowing with respect to the time consistent sophisticates, but their study is moot with respect to the role of naîvete per se.

[9]"That sophistication can hurt you is, however, implicit in Pollak (1968). In the process of demonstrating a mathematical result, Pollak shows that sophisticates and naifs behave the same for logarithmic utility. From this, it is straightforward to show that for utility functions more concave than the log utility function, sophisticates save more than naifs (i.e., sophistication mitigates self-control problems), whereas for less concave utility functions, sophisticates save less than naifs (i.e., sophistication exacerbates self-control problems)." From O' Donoghue and Rabin (1999, f.note 24, p. 119).

2 The model

Time and uncertainty are represented by a tree structure. There are $T < +\infty$ dates. Individual uncertainty is generated by a time-homogeneous Markov transition $\pi : S \to \Delta(S)$ where S is a finite set of personal states. We let $s^t = (s_0, ..., s_t)$, with $s_t \in S$, be a history of length t from initial state s_0, S^t be the collection of histories s^t, and $\{S^t\}_{t \in T}$ is the corresponding tree. All time-indexed variables are functions (processes) on S^T adapted to the tree. Hereafter, we write $\pi(s^t)$ for the probability of history s^t. For any process $y_t, t \geq 0$ we let $y^{t+1,\tau} = (y_{t+1}, ..., y_\tau)$, for $\tau > t$, denote the continuation path from $t+1$ to τ.

There is a continuum of ex-ante identical individuals (or consumers). Individuals have one self at each $t \in T$, 'self t', with possibly multiple realized types under π. Individuals have *subjective beliefs* $\mu : S \to \Delta(S)$ on the Markov states, and are discounted expected utility maximizers.

There is one physical good. The consumption process is $x_t, t \geq 0$, with $x_t(s^t) \geq 0$, all $s^t \in S^t$ and $t \geq 0$. Time inconsistent preferences are of the (β, δ)-form, representing present bias.[10] Self 0's utility is

$$u(x_0) + \mathbb{E}_{0,\mu} \sum_{t>0} \delta^t v(x_t)$$

and self t's utilities are

$$v(x_t) + \beta_t \mathbb{E}_{t,\mu} \sum_{\tau>t} \delta^t v(x_\tau),$$

where $\beta_t, t > 0$ is self t's discount factor (process) between period t and periods $\tau > t$, and $\delta \leq 1$, $\mathbb{E}_{t,\mu}$ is the expectation operator under subjective probability μ conditional on the information available at t, and u and v are Bernoulli indexes $u, v : \mathbb{R}^*_+ \to \mathbb{R}^*$. Distinguishing between future utility v and present utility u allows to incorporate time inconsistency also for self 0 via a factor $\beta_0 \leq 1$ and by setting $u = \frac{v}{\beta_0}$, the case $\beta_0 = 1$ corresponding to no procrastination at $t = 0$. Observe that the consumption space is the closed interval $[0, \infty]$: the upper bound ∞ is attainable. This matters when π does not have full support. Attainability is discussed in Citanna et al. (2022b).

[10] See, e.g., Phelps and Pollak (1968), Pollak (1968), or Laibson (1994, 1997).

We introduce standard assumptions on discount factors and utilities.

A1 u and v are continuously differentiable on $\mathbb{R}_{++}$, strictly increasing and concave on $\mathbb{R}_+$, with at least one strictly concave, and $v(0) > -\infty$.

$\beta_t(s^t) = \beta(s_t)$, all $t > 0$, with $\beta(s_t)$ increasing in s_t, and $\max_{s \in S} \beta(s) < 1$.

As a result, $v(x) = \infty$ only if $x = \infty$ (but not necessarily), and similarly for u. To include log utility or CRRA functions more concave than log, the consumption lower bound needs to be rescaled up to a positive number. The time independence of β_t and our assumptions on π and μ imply that the economy is Markovian. Restricting attention to Markovian environments is common to much of the time-consistency literature. The monotonicity assumption on $\beta(s)$ is without loss of generality since it always holds true up to a relabeling. We will comment below in Section 4 about the upper bound on the betas.

Throughout, we assume that firms offering contracts to individuals know the true distribution π and also the utilities u and v, thus β_0. When π is nondegenerate, firms face uncertainty about the individuals' procrastination tendencies. Over time individuals get to know their tendency to procrastinate while firms do not, in what we see as a form of hidden information. Uncertainty about utilities, or β_0, create a much different adverse selection problem.

Aggregate resource consumption at date t depends on the distribution of selves, via π. With the usual identification of ex-post resources with expectations,[11] $\mathbb{E}_\pi(x_t)$ are the aggregate resources consumed by contract x at t. Total lifetime consumption induced by process x is $\sum_{t \leq T} \rho^t \mathbb{E}_\pi(x_t)$, with $\rho > 0$ a market discount rate. Individuals are endowed with material resources over time. Total lifetime resources are $r > 0$.[12]

Our focus on lifetime quantities and the use of a market discount rate reflect the possibility of borrowing and saving with a 'deep-pocketed' outside agent, in the tradition of Green (1987).[13]

---

[11] This identification can be justified using a law of large numbers as in Uhlig (1996), or another aggregation technology reflecting the pooling of individual risks.

[12] Making the dependence of r on T, ρ explicit does not alter any conclusion of the paper, while it makes notation and some arguments cumbersome.

[13] Closed economies with no time aggregation can be accommodated by adding a time consistent

Lifetime resources r result from the aggregation of endowments, which can fluctuate over time and be stochastic. As long as endowment uncertainty is known by the individual, it plays no role. Thus, we omit the specification of the endowment distribution throughout.

3 Efficiency

We define the allocation problem and efficiency in these environments. An *allocation* is a contract (process) $x \equiv (x_t, t \geq 0)$, and G is the set of all such allocations, with $G = \times_{t,s^t}[0,\infty]$. An allocation x is *feasible* if

$$\sum_{t \leq T} \rho^t \mathbb{E}_\pi(x_t) - r \leq 0. \tag{F}$$

The distinctive feature of the environments is that firms offer contractual options to consumers, which here are represented by s^t-contingent net trades. Self 0 selects a contract x at time 0, i.e., a menu of options, and then at any $t > 0$ self t selects one from the available options, acting on the contract according to their own preferences and beliefs. A caveat is in order when π does not have full support. As selves' choices at $\tau > 0$ affect utility of selves at t preceding τ, who hold beliefs μ, behavior must be specified even for selves (histories) that may not realize under π but realize under μ. Implicitly, we are assuming that individual expectations over histories s^τ are rational at any time, and common across selves t that precede τ. They are sensitive to information revealed by state realizations, via transition μ. In other words, we are going to apply the requirements of perfection-perfect equilibrium (see O'Donoghue and Rabin (1999, 2001); also see Gottlieb and Zhang (2021, Appendix E, p. 10)).

This is equivalent to letting selves t change any history s^T into $s^{T\prime}$ while constrained by the choices of past selves at histories s^{t-1}, thus inducing manipulated contingent consumptions $x_t(s^{T\prime})$ much like in a dynamic hidden information economy.

For any $t > 0$, we define *self-t choice* as a $\hat{\phi}_t \in S$, a history of choices up to time t is $\hat{\phi}^t = (\hat{\phi}_1, ..., \hat{\phi}_t)$. Given s^T, consumption process x and history $\hat{\phi}^{t-1} \in S^{t-1}$ of

agent. The same borrowing behavior of the naïve present biased individual is then obtained while fixing the utility level of the time-consistent agent.

past menu choices, self t's choice $\hat{\phi}_t$ picks an element in the *time-t menu*, $x_t(\hat{\phi}^{t-1}, s)_{s\in S}$, and modifies menus available to subsequent selves. Selves choices are described as (pure) strategies. A (pure) strategy for self t is σ_t, a map from histories $\hat{\phi}^{t-1}$ of choice outcomes and s^t into a choice $\hat{\phi}_t$ at t. A strategy process is $\sigma \equiv (\sigma_t, t > 0)$.

As any self makes contractual choices anticipating that future selves will choose among the available options according to their own preferences, selves' strategies satisfy a sequential rationality, or subgame perfection, requirement. As customary (see, e.g., Luttmer and Mariotti (2003) and references therein), and given that the economy is Markovian, we impose a further behavioral restriction and use *Markov Perfect Equilibrium* (MPE) as a solution concept to characterize the selves' interaction.[14,15] We let $\Sigma^M(x)$ denote the set of MPE strategy profiles of game $x_t, t \geq 0$, i.e., of subgame perfect equilibrium profiles where strategies are Markovian. In fact, in these environments when studying MPE allocations we can further restrict attention at no loss of generality to incentive compatible strategies, i.e., strategies which involve truthtelling at every history. A strategy process σ is said to be *truthtelling* if $\sigma_t(\hat{\phi}^{t-1}, s^t) = s_t$, all $\hat{\phi}^{t-1}$ and s^t, all $t > 0$, denoted as σ^{id}. We say that x is *incentive compatible* (IC) if $\sigma^{id} \in \Sigma^M(x)$. An allocation x^{IC} is *incentive compatible efficient* (IC-efficient) if it solves

$$\begin{array}{l} \max_{x\in G} u_0(id, x; \mu) \quad s.to \\ \sigma^{id} \in \Sigma^M(x) \text{ and F.} \end{array} \qquad (IC\text{-problem})$$

An IC-efficient allocation takes into account material balance as well as Markov perfect individual behavior, thus it defines a second best in these economies. Equilibrium allocations of some competitive mechanism (e.g., a perfectly competitive market à la Debreu, as in Citanna and Siconolfi (2020); or a Bertrand competitive game among credit firms, as in, e.g., Heidhues and Köszegi (2010)) are IC-efficient allocations.

[14] More complex SPE strategies also give a lot of coordination power to the individuals' multiple selves –generally allowing for commitment behavior to emerge– against the spirit of present bias and the empirical evidence supporting it. A result on borrowing that depended on complicated strategic behavior in the intrapersonal game would be less compelling.

[15] Recall that a strategy profile σ is *Markovian* if σ_t depends on s^t only via its payoff-relevant component. As $(\beta_t, \mu_t)(s^t) = (\beta(s_t), \mu(s_{t+1}|s_t))$, the payoff-relevant component is s_t, i.e., $\sigma_t(\hat{\phi}^{t-1}, s^t) = \sigma_t(\hat{\phi}^{t-1}, s_t)$, all $\hat{\phi}^{t-1}$, $t > 0$ and s^t. Obviously, as the time horizon is finite, strategies are time dependent, and not time homogeneous, as underscored by the subscript t.

Finally, the following terminology is used throughout. We say that an individual is *sophisticated* if her beliefs coincide with π. Otherwise, the individual is *unsophisticated.* Given concavity of the utility functions, a full insurance consumption process is always incentive compatible and it is weakly preferred by a sophisticate to any other process, strictly if v is strictly concave. Thus, for a sophisticated individual, the constraint $\sigma^{id} \in \Sigma^M(x)$ in the IC-problem is never binding. We refer to the sophisticated problem as the FB-problem, for 'First Best' (a term explained by the full insurance nature of the contract).

In the literature on welfare comparisons, a further distinction in made by using the 'long-run preferences' of the individual (see, e.g., Gottlieb and Zhang, 2021). Then, a time-consistent sophisticate is compared with a time-inconsistent unsophisticated individual. We label the time-consistent sophisticated consumer *long-run sophisticate*, and the allocation is *asymptotically efficient* if it approximates

$$\max_{x \in G} \mathbb{E}_{0,\pi} \sum_{t=0}^{T} \delta^t v(x_t) \text{ s. to } F$$

as the time horizon T grows arbitrarily large.

Full vs. partial naïveté. We say that consumers are *fully unsophisticated* if $S = \{1, 2\}$, and both π and μ put probability one on one state opposite to each other, i.e., $\pi(s'|s) = 1 - \mu(s'|s) \in \{0, 1\}$, all $s, s' \in S$. In fact, consumers are *fully naïve* or *optimistic* if $\pi(1|s) = 1 = \mu(2|s)$ for $s = 1, 2$; they are *fully pessimistic* if $\pi(2|s) = 1 = \mu(1|s)$. Thus, fully naïve consumers' subjective beliefs are completely false, and repeatedly optimistic. Fully unsophisticated consumers are otherwise repeatedly pessimistic. Outside these cases, we say that consumers are simply *(partially) unsophisticated.* Thus, if $S = \{1, 2\}$ but π has full support, consumers are partially unsophisticated. We will discuss below how to distinguish naïves from pessimists in these less extreme settings.

Fully naïve individuals are called partially naïve in O' Donoghue and Rabin (1999, 2001)) since $\beta(2) < 1$; they are fully naïve in these authors' definition when $\beta(2) = 1$ –a case we have excluded; see also Gottlieb and Zhang (2021). The important point to note here is that these authors' definition of partial naïveté is embedded as a special case into our setup. Fully naïve behavior is studied in Citanna et al. (2022b).

3.1 An equivalent formulation

As often done in dynamic hidden information environments, we cast the IC-problem in the payoff space, where it is easier to characterize. Given a consumption process x, we let $u_0 \equiv u(x_0)$ and $v_t(s^t) \equiv v(x_t(s^t))$ be the corresponding utiles, i.e., u_0 is a utility level at $t = 0$, and $v_t(s^t)$ is a utility level at history s^t. Vice versa, given a utile process $u_0, v_t, t > 0$, strict monotonicity and continuity of the utility functions u and v yield a unique corresponding consumption process x with $x_0 = u^{-1}(u_0)$ and $x_t(s^t) = v^{-1}(v_t(s^t))$ provided that

$$u_0 \in [u(0), u(\infty)] \text{ and } v_t(s^t) \in [v(0), v(\infty)], \text{ all } s^t \text{ and } t > 0. \tag{B}$$

Given $u_0, v_t, t > 0$, let $v_\tau(s,t), \tau > t$ be the derived utile continuation process from s^t when $\phi_t(s^t) = s$, with entries $v_\tau(s,t)(s^\tau) = v_\tau(s^{t-1}, s, s^{t,\tau})$ for every history $s^\tau = (s^t, s^{t,\tau})$. Then, process $v_t, t > 0$ is incentive compatible if, for all $s^t, t > 0$,

$$v_t(s^t) + \beta(s_t)\mathbb{E}_{s^t,\mu} \sum_{\tau > t} \delta^{\tau - t} v_\tau \geq v_t(s^{t-1}, s) + \beta(s_t)\mathbb{E}_{s^t,\mu} \sum_{\tau > t} \delta^{\tau - t} v_\tau(s,t), \tag{IC}$$

all $s \in S$. It is feasible if

$$u^{-1}(u_0) + \mathbb{E}_\pi \sum_{t>0} \rho^t v^{-1}(v_t) \leq r. \tag{F}$$

Then, the IC-problem can be written as

$$\begin{gathered} \max_{u_0, v_t, t>0} u_0 + \mathbb{E}_\mu \textstyle\sum_{t>0} \delta^t v_t \\ \text{s. to IC, F and B.} \end{gathered}$$

For any given s^{T-1}, the incentive constraints at T read

$$v_T(s^{T-1}, s) \geq v_T(s^{T-1}, s'), \text{ all } s, s',$$

thereby implying s_T-invariance of the random variable v_T. Hereafter, we drop the incentive constraints associated to self T and we identify process $u_0, v_t, t > 0$ with $u_0, (v_t(s^t))_{t<T}, v_T(s^{T-1})$. We further write $u_0^{FB}, v_t^{FB}, t > 0$ for a solution to the sophisticated problem, and $u_0^{IC}, v_t^{IC}, t > 0$ for a solution to the IC-problem.

With this formulation, we study welfare implications of naïve present bias and compare the unsophisticated and sophisticated individuals' borrowing behavior. Recall that an individual is sophisticated if $\mu = \pi$. We say that the unsophisticated individual *underborrows* with respect to her sophisticated self if $u_0^{IC} \leq u_0^{FB}$. Otherwise, the individual *overborrows.* The individual *strictly underborrows* if $u_0^{IC} < u_0^{FB}$. Finally we say that a utile process $u_0, v_t, t > 0$ is *interior* if both $u_0 > u(0)$ and $u_0 < u(r)$, that is, if period 0 consumption is positive but less than the total resources available. It is *strongly interior* if, further, $v_t > v(0)$, all $t > 0$.

4 The analysis

We consider first the case where π has full support, as this eliminates rules out the existence of purely imaginary selves, i.e., of $s \in S$ with $\pi(s|s') = 0$ all $s' \in S$ and $\mu(s|s'') > 0$, some $s'' \in S$, simplifying existence arguments and allowing us to focus on the main feature of the problem:

A2.1 (π has full support) $\pi(\cdot|s) \gg 0$, all $s \in S$.

Assumption A2.1 leads to $v_t(s^t) < v(\infty)$ for all s^t, via feasibility, and to the existence of a finite (in consumption space) solution to both the sophisticated and unsophisticated problems, no matter the lifetime income level r.

4.1 A simple setting

It is convenient to start from the simple case where $T = 2$ and $S = \{1, 2\}$. Further, we let $\delta = \rho = 1$, and take $u = v$ (i.e., $\beta_0 = 1$), normalizing $v(0) = 0$. We finally assume that $\lim_{x \to 0} v'(x) = \infty$. As a result, all solutions are strongly interior, and are unique.

By strict concavity of v, $v_t^{FB}(s^t)$ is s^t-invariant, $t = 1, 2$, and by incentive compatibility, $v_2^{IC}(s^2) = v_2^{IC}(s^1)$. Thus, since s_0 is given, hereafter we drop any reference to it and write s for history $s^1 = (s_0, s)$ ——and thus use $\pi(s), \mu(s)$ for the corresponding probabilities, and $v_t(s)$ for the utiles.

Thus, the IC-problem takes the form

$$
\begin{aligned}
&\max\nolimits_{v_t, t=0,1,2} v_0 + \textstyle\sum_s \mu(s) \sum_{t=1,2} v_t(s) \text{ s. to B and} \\
&v_1(1) + \beta(1) v_2(1) \geq v_1(2) + \beta(1) v_2(2), \qquad (IC-1) \\
&v_1(2) + \beta(2) v_2(2) \geq v_1(1) + \beta(2) v_2(1), \qquad (IC-2) \\
&v^{-1}(v_0) + \textstyle\sum_s \pi(s) \sum_{t=1,2} v^{-1}(v_t(s)) \leq r. \qquad (F)
\end{aligned}
$$

The FB-problem is identical to the IC-problem except for not including the IC constraints, and for having $\mu = \pi$. First order conditions for the FB-problem are

$$v'(x_0^{FB}) = v'(\frac{r - x_0^{FB}}{2}). \qquad \text{(FOC-FB)}$$

Crucially, the solution to the IC-problem satisfies the 'inverse marginal utility' equation

$$\frac{1}{v'(v^{-1}(u_0^{IC}))} = \mathbb{E}_\pi \frac{1}{v'(v^{-1}(v_t^{IC}))}, \text{ all } t > 0.$$

Indeed, consider the following variational argument:[16] change v_0 by one (sufficiently small) utile, while reverse changing by one utile $v_t(s)$, all s, first for $t = 1$ and then for $t = 2$. Such changes are clearly incentive compatible, and viable as the solution is interior, while they do not change lifetime utility. They generate a change in consumption of resources equal to $\frac{1}{v_0'} - \mathbb{E}_\pi \frac{1}{v_t'}$. Optimality requires that such resource changes be equal to zero, that is, that the inverse marginal utility equation holds true.

Then, the effect of unsophistication on borrowing depends on whether subjective beliefs are pessimistic or optimistic, and in the latter case on the curvature of the inverse marginal utility function. Beliefs are *optimistic,* or *naïve,* if $\mu(2) > \pi(2)$, while they are *pessimistic* if $\mu(1) > \pi(1)$.

Indeed, if the unsophisticated solution displays variability in future consumption across self types, i.e., $x_t^{IC}(1) \neq x_t^{IC}(2)$ for some $t > 0$, and if the function $1/v'$ is concave, there is strict underborrowing, $v_0^{FB} > v_0^{IC}$**;** while if $1/v'$ is convex, there is

[16] This is reminiscent of arguments in the hidden information literature (see, e.g., Golosov, Kocherlakota and Tzivinski, 2003), where a similar inverse Euler equation is a standard property.

overborrowing, and $v_0^{FB} < v_0^{IC}$. Why?

By feasibility it is $\mathbb{E}_\pi x_t = \frac{r-x_0}{2}$. Then, when $1/v'$ is concave, the inverse marginal utility equation implies, via Jensen's inequality,

$$\frac{1}{v'(x_0^{IC})} < \frac{1}{v'(\mathbb{E}_\pi x_t^{IC})},$$

which by comparison with FB-FOC delivers strict underborrowing, or saving more.

Consumption variability, though, is entirely determined by beliefs: it is present if beliefs are naïve, while with pessimistic beliefs variability is absent. Let's see why.

For a pessimist, it must be that $\sum_{t=1,2} v_t^{IC}(s)$ is weakly decreasing in s. For if not, since $\mu(1) - \pi(1) = \pi(2) - \mu(2) > 0$, the reverse implies $\mathbb{E}_\mu \sum_{t=1,2} v_t^{IC} < \mathbb{E}_\pi \sum_{t=1,2} v_t^{IC}$. Thus,

$$\begin{aligned} v_0^{IC} + \mathbb{E}_\mu \sum_{t=1,2} v_t^{IC} &< v_0^{IC} + \mathbb{E}_\pi \sum_{t=1,2} v_t^{IC} \leq v_0^{FB} + \mathbb{E}_\pi \sum_{t=1,2} v_t^{FB} \\ &= v_0^{FB} + \mathbb{E}_\mu \sum_{t=1,2} v_t^{FB} \end{aligned}$$

where the weak inequality comes from $v_t^{IC}, t \geq 0$ feasible in the FB-problem, and the equality holds because v_t^{FB} is s-invariant. As $v_t^{FB}, t \geq 0$ satisfies all the constraints in the IC-problem, this is a contradiction to the optimality of $v_t^{IC}, t \geq 0$. Hence, combining $\sum_{t=1,2} v_t^{IC}(s)$ weakly decreasing with the two incentive constraints, and $1 > \beta(2)$, the only solution is pooling, or $v_t^{IC}(s)$ is s-invariant, all $t > 0$.[17] By strict concavity of v and since $\rho = 1$, $v_t^{IC}(s)$ is also t-invariant, and it thus coincides with the FB-solution.[18]

The naïve solution instead can never be pooling, thus it cannot be the sophisticated one. Indeed, consider the following contract change: decrease v_0 by one utile, $\Delta v_0 = -1$,

[17] Since $1 > \beta(2)$, summing the constraint $v_1(1) + v_2(1) \geq v_1(2) + v_2(2)$ with incentive constraint IC-2, delivers the inequality $v_2(1) \geq v_2(2)$; while summing incentive constraints IC-1 and IC-2 delivers $v_2(2) \geq v_2(1)$.

[18] To see the role of the assumption $\beta(s) < 1$, all s, we observe that the IC-problem with $\beta(s)$ is equivalent to the IC-problem with coefficients $1/\beta(s)$ (and where dates $t = 1, 2$ are switched). If $\beta(2) > 1 > \beta(1)$, then $1/\beta(1) > 1 > 1/\beta(2)$: a pessimist in the original problem would become optimist in the equivalent problem. Thus, pessimists would also be affected by their beliefs, obtaining underborrowing for any false belief. If instead $\beta(2) = 1 > \beta(1)$, then $1/\beta(1) > 1 = 1/\beta(2)$, and the pessimistic consumer would behave as a time-consistent individual.

while decreasing $v_1(2)$ by κ utiles and increasing $v_2(2)$ by $\frac{\kappa}{\beta(1)}$ utiles. Pick κ so that the overall change in self 0's utility is zero, that is, κ satisfies the equation

$$\Delta v_0 + \mu(2)(\Delta v_1(2) + \Delta v_2(2)) = 0,$$

that is

$$-1 + \mu(2)(-1 + \frac{1}{\beta(1)})\kappa = 0 \text{ or } \kappa = \frac{\beta(1)}{(1-\beta(1))\mu(2)}.$$

Since $1 > \beta(2) > \beta(1)$ and $(v_t, t \geq 0) \gg 0$, these changes in allocations are incentive compatible, viable and do not change overall utility. However, they generate a change in resource consumption equal to

$$\begin{aligned} & -\frac{1}{v_0'} + \pi(2)[\frac{1}{v_1'(2)}\Delta v_1(2) + \frac{1}{v_2'(2)}\Delta v_2(2)] \\ = \quad & -\frac{1}{v_0'} + \pi(2)[\frac{1}{v_1'(2)} - \frac{1}{\beta(1)}\frac{1}{v_2'(2)}]\frac{\beta(1)}{(\beta(1)-1)\mu(2)} \end{aligned}$$

If the naïve solution were pooling, it would coincide with the FB-solution, and then $v_1'(2) = v_2'(2) = v_0'$. Then $\pi(2) < \mu(2)$ would imply a saving in resource consumption, contradicting optimality. Equivalently, this establishes the variability in future consumption at the naïve solution. In other words, the variability in consumption is induced by the interaction of incentive constraints and naïve beliefs.

Observe that with naïveté $\sum_{t=1,2} v_t(s)$ is strictly increasing in s, while the reverse holds for pessimists. In other words, the unsophisticated individual tries to consume more in what she believes are the relatively more likely events, though her ability to do so is limited by incentive constraints.

Thus, we reach the following conclusions:

A naïve consumer (strictly) underborrows when the inverse marginal utility $1/v'$ *is (strictly) concave, overborrows when it is strictly convex. A pessimistic consumer behaves as a sophisticated consumer.*

4.2 The general full support case

We now characterize the borrowing behavior of fully naïve individuals for more general environments. Throughout, we assume that $v(0) = 0$.

4.2.1 Borrowing behavior

As noted in the simple setting, the key feature of the solution to the IC-problem is the 'inverse marginal utility inequality'[19] linking time-zero and future marginal utilities at any naïve optimal contract,

$$\frac{1}{u_0'} \leq \left(\frac{\rho}{\delta}\right)^t \mathbb{E}_\pi \frac{1}{v_t'}, \text{ for all } t > 0. \tag{IMU}$$

The intuition for why this is the case, and a short argument, was given in the simple setting, where the solution to the unsophisticated problem was strongly interior. If the naïve solution is not interior, the equality turns into the expression (IMU). Concavity of $1/v'$ then allows for the application of Jensen's inequality while preserving the inequality, and yields underborrowing by comparison with the optimality conditions of the FB-problem. To establish borrowing behavior, we add just an ancillary simplifying assumption, guaranteeing that at a solution $u_0 > u(0)$, that is:

A2.2 $u'(0) > v'(\frac{r}{\sum_{t>0}\rho^t}) \max_{t>0}(\frac{\delta}{\rho})^t$.

The borrowing behavior of unsophisticated consumers is then stated in the next proposition.

Proposition 1 (No overborrowing) *Under A1, A2.1 and A2.2, if $1/v'$ is concave, the unsophisticated individual does not overborrow.*

Proposition 1 rules out overborrowing no matter the consumer's lifetime income r —fully naïve borrowers instead may overborrow for high enough r, as in Citanna et al. (2022b). It does not rule out that the unsophisticated consumer mimics the sophisticate. It is also moot on consumers' behavior when the inverse marginal utility is

[19] An inequality is obtained because, while the upper bound $v(\infty)$ is not binding, the lower bound $v(0)$ may be –e.g., when v is linear and $\rho \neq 1$.

not concave. We then move on to further exploring borrowing behavior with the goal of finding conditions under which *strict* underborrowing occurs. In the process, and in line with what already highlighted in the simple setting, we discover the role that naïveté plays in determining consumers' behavior.

Pessimism, and strict underborrowing. Naïveté, or optimism, is clearly more than just having wrong beliefs. In the simple setting, where $T = S = 2$, it was identified with First Order Stochastic Dominance (FOSD) of μ over π.[20] Note that optimism gave rise to the property $\sum_t v_t^{IC}(2) > \sum_t v_t^{IC}(1)$, and as a consequence $\mathbb{E}_\mu \sum_{t>0} v_t^{IC} > \mathbb{E}_\pi \sum_{t>0} v_t^{IC}$: the consumer believes she can make a positive (utile) payoff while betting at market odds. This is because we can interpret $\mathbb{E}_\pi \sum_{t>0} v_t$ as the break-even cost of betting against a sophisticated, risk-neutral individual, or the market. It turns out that this arbitrage property is the key feature determining borrowing behavior. However, due to the strategic interaction among selves, when $S > 2$ defining optimism on the sole basis of the stochastic order properties of beliefs does not generally lead to the arbitrage property. Thus, hereafter we identify optimism directly with the belief that one can make (unbounded amounts of) money while betting at market odds.

To make the presentation smoother, we focus again hereafter on environments where the sophisticated allocation is strongly interior. This is implied by an Inada condition,

A2.3 $\lim_{x\to 0} v'(x) = \infty$.

Note that under A2.3 v is strictly concave. Naïveté is defined in opposition to pessimism. Consider the problem

$$\max_{m_t, t>0} \mathbb{E}_\mu \sum_{t>0} \delta^t m_t - \mathbb{E}_\pi \sum_{t>0} \delta^t m_t \text{ s. to } m_t \text{ satisfies IC.} \tag{P}$$

We say that the unsophisticated consumer is *pessimistic* if problem (P) has a solution. Otherwise, we say that the consumer is *naïve* or *optimistic*.

[20] Recall that, for $\hat{\pi}, \hat{\mu} \in \Delta(S)$, $\hat{\pi}$ first order stochastic dominates $\hat{\mu}$ if

$$\sum_{s' \leq s} \hat{\pi}(s') \leq \sum_{s' \leq s} \hat{\mu}(s'), \text{ all } s,$$

with at least one strict inequality.

In problem (P), a (risk neutral) consumer is betting on incentive compatible money profiles, $m_t, t > 0$. Her (believed) expected revenue is $\mathbb{E}_\mu \sum_{t>0} \delta^t m_t$. Instead, the break-even cost of betting against a sophisticated individual, or the market, is $\mathbb{E}_\pi \sum_{t>0} \delta^t m_t$. The objective in (P) is then the consumer's expected profit. A solution to (P) exists if and only if the expected profit is nonpositive for all incentive compatible, admissible bets $m_t, t > 0$, and then its value is zero. Thus, the unsophisticated consumer is a pessimist if she thinks she cannot make money betting against the market, or against a sophisticated consumer.

Often, but not always, pessimism ensues when the consumer thinks that she is more likely to be less patient than what she actually will turn out to be. In other words, existence of a solution to problem (P) is often related to the ranking of the transitions π and μ in terms of first order stochastic dominance. For example, in the simple setting a consumer is pessimistic if and only if $\pi(s_0, \cdot)$ FOSD $\mu(s_0, \cdot)$. However, beliefs comparison via first order stochastic dominance will not be enough to determine pessimism in general. In the Appendix, we explore related notions of pessimism which are only based on beliefs comparison.

The next lemma characterizes pessimism and generalizes the conclusions on pessimism reached in the simpler economy with $T = S = 2$.

Lemma 1 *Under A1 and A2, an unsophisticated individual behaves as a sophisticate if and only if she is a pessimist.*

Thus, when the consumer is not pessimistic, the sophisticated allocation cannot be IC-optimal. False beliefs have an effect on the consumer. The proof of Lemma 1 requires a preliminary closer study of the optimal sophisticated and unsophisticated allocations, and the associated first order conditions, a task which we defer to the Appendix.

Armed with a characterization of pessimism, we can state the effects of naïveté on borrowing behavior.

Proposition 2 *Under A1 and A2:*

i) if $1/v'$ is strictly concave, the naïve individual strictly underborrows;

ii) if $1/v'$ is strictly convex, the naïve individual overborrows.[21]

[21] It follows that in the class of constant relative risk aversion utilities, $v = \ln$ separates under- from

In the absence of risk aversion –that is, when v is linear– A2.3 fails. An optimistic risk neutral consumer then may behave as a sophisticate; much depends on the details of π, μ, and δ, ρ, and we cannot come to sharper conclusions than those obtained in Proposition 1.

4.2.2 Welfare

We now explore the welfare consequences in terms of long-run preferences. Throughout, we maintain A2.1, so that π has full support, and A2.3 –or an Inada condition.

Recall that here $u_0 = v_0/\beta_0$. We are comparing

$$W^C(r,T) = \max \mathbb{E}_\pi \sum_{t \le T} \delta^t v_t \text{ s. to } \mathbb{E}_\pi \sum_{t \le T} \rho^t v^{-1}(v_t) \le r$$

with $W^{IC}(r,T) = \mathbb{E}_\pi \sum_{t \le T} \delta^t v_t^{IC}(r,T)$. Again, we observe that $W^C(r,T) - W^{IC}(r,T) \ge 0$ for all r, T, as the $v_t^{IC}(r,T), t \ge 0$ solution is in the constraint set of the time consistent sophisticate.

We now establish that the inefficiency due to naïve present bias does not vanish asymptotically. To this end, we say that an individual is T^**-naïve* if she is naïve for some $T^* > 1$.

Proposition 3 *Let v be bounded, and $1 > \rho \ge \delta$ and $\beta_0 = \beta(s)$ for some $s \in S$. Under A1 and A2, for any T^*-naïve consumer* $\lim_T W^C(r,T) - W^{IC}(r,T) > 0$*, all r.*

Citanna et al. (2022b) already established, in fully naïve environments, that the limit $W^C(r,T) - W^{IC}(r,T)$ as T grows arbitrarily large stays positive for r large enough. Thus, we already knew that naïve present bias can have negative welfare consequences for consumers even when markets otherwise function well –i.e., in the absence of market power or of adverse selection problems. The additional conclusion here is that uncertainty over the procrastination parameter β induces permanent welfare losses no matter the consumer's expected lifetime income r. This is in stark contrast to what found by

over-borrowing. When $v = \ln$, partially naïve individuals and sophisticates choose identical saving (recall that for these utilities the consumption space lower bound must be raised to $\underline{x} > 0$ to have $v(\underline{x}) > -\infty$). Consumers with more concave utility than ln ($1/v'$ convex) will overborrow.

Gottlieb and Zhang (2021)[22], i.e., the asymptotically vanishing inefficiency due to naïve present bias.

4.3 The case with imaginary selves

Without A2.1 there are imaginary selves –histories s^t with $\mu(s^t|s_0) > 0$ but $\pi(s^t|s_0) = 0$. In general, $v_t(s^t)$ is no longer guaranteed to be bounded away from $v(\infty)$ –particularly when v is bounded, i.e., $v(\infty) < \infty$. When $v_t(s^t) = v(\infty)$ for some s^t, we say there is 'imaginary satiation'. If $v_t(s^t) < v(\infty)$ for all s^t, we say that there is 'nonsatiation for the imaginary self'. With risk aversion, the interplay between satiation over imaginary selves and the prudence-to-risk aversion ratio (the shape of $1/v'$) determines borrowing behavior. As long as the upper bound $v(\infty)$ is never binding, IMU holds and the conclusions of Propositions 1 and 2.i hold true. A sufficient condition that guarantees that the utility upper bound is never binding, even when the π-probability is zero for some history s^t, is as follows. Let $S = \{1, ...\bar{s}\}$.

A2.1'

$$(\frac{1 + \beta_{\bar{s}} \sum_{t>0} \delta^t}{\beta_1^{T-1}})^{T-1} V_T^{**} < v(\infty)$$

where V_T^{**} is the maximum future utility achievable over the feasible set at some contingency of positive π-probability. We then obtain the following.

Proposition 4 *Under A1, A2.1' and A2.2 hold. i) If* $1/v'$ *is concave, the unsophisticated individual does not overborrow, and the naïve individual strictly underborrows. ii) If* v *is bounded, and* $1 > \rho \geq \delta$ *and* $\beta_0 = \beta(s)$ *for some* $s \in S$*, for any* T^**-naïve consumer* $\lim_T W^C(r,T) - W^{IC}(r,T) > 0$*, all* r *consistent with A2.1'.*

Without A2.1, though, Proposition 2.ii may not extend. Indeed, when the function $1/v'$ is convex, the consumer's borrowing behavior now depends on her level of naïveté.

[22] Gottlieb and Zhang (2021, Theorem 1) claimed a vanishing inefficiency for all r provided utility v is bounded. This statement is corrected in Citanna, Gottlieb, Siconolfi and Zhang (2022a), where it is shown to apply only for small values of r.

When her subjective assessment μ and π are sufficiently close to full naïveté, under nonsatiation and following Citanna et al. (2022b) she will not overborrow.[23]

Thus, under nonsatiation, concavity of $1/v'$ leads to borrowing properties more robust in beliefs. Instead, under satiation, or in environments where only $v_t \leq v(\infty)$ can be binding, the inequality in IMU gets reversed, and the conclusions become symmetric: from

$$\frac{1}{u_0'} \geq \left(\frac{\rho}{\delta}\right)^t \mathbb{E}_\pi \frac{1}{v_t'}$$

and convexity of $1/v'$ we obtain overborrowing, strict under optimism. Thus, under satiation, it is convexity of $1/v'$ that leads to more robust borrowing properties.

References

[1] **Citanna, Alex, and Paolo Siconolfi**. 2020. "Constrained Efficient Markets for Manipulation Economies." International Economics Review 61, 1531–1567.

[2] **Citanna, Alex, Daniel Gottlieb, Paolo Siconolfi, and Xingtan Zhang.** 2022a. "Corrigendum: Long-Term Contracting with Time-Inconsistent Agents." Mimeo.

[3] **Citanna, Alex, Daniel Gottlieb, Paolo Siconolfi, and Xingtan Zhang.** 2022b. "Naïve Time-Inconsistent Borrowers with Satiation or Borrowing Limits." Mimeo.

[4] **Frederick, Shane, George Loewenstein, and Ted O'Donoghue**. 2002. "Time Discounting and Time Preference: A Critical Review." Journal of Economic Literature 40, 351–401.

[5] **Golosov, Michael, Narayana Kocherlakota, and Alec Tsyvinski**. 2003. "Optimal Indirect and Capital Taxation." Review of Economic Studies 70, 569–587.

[23] For example, let $S = T = 2$, $\delta = \rho = 1$, $u(x) = x/c$ and $v(x) = a - a/(1+x)$, where $0 < a$, $c > 1$, and $\beta(1) < ac < 1$. Straightforward computations show that, when r is sufficiently small so that A-0 in Citanna et al. (2022b) holds, underborrowing is strict for a fully naïve consumer.

[6] **Gottlieb, Daniel, and Xingtan Zhang**. 2021. "Long-Term Contracting with Time-Inconsistent Agents." Econometrica 89, 793–824.

[7] **Green, Ed. J.**. 1987. "Lending and the Smoothing of Uninsurable Income." In E. C. Prescott and N. Wallace (eds.), Contractual Arrangements for Intertemporal Trade, Minneapolis: University of Minnesota Press, 3–25.

[8] **Guvenen, Fatih**. 2006. "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective." Journal of Monetary Economics 53, 1451–1472.

[9] **Havranek, Tomas**. 2015. "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." Journal of the European Economic Association 13, 1180–1204.

[10] **Heidhues, Paul, and Botond Köszegi**. 2010. "Exploiting Naïveté About Self-Control in the Credit Market." American Economic Review 100, 2279–2303.

[11] **Kimball, Miles S.**. 1990. "Precautionary Saving in the Small and in the Large." Econometrica 58, 53–73.

[12] **Köszegi, Botond**. 2014. "Behavioral Contract Theory." Journal of Economic Perspectives 52, 1075–1118.

[13] **Laibson, David**. 1994. "Self-Control and Saving." Mimeo, Cambridge, MA, Harvard University.

[14] **Laibson, David**. 1997. "Golden Eggs and Hyperbolic Discounting." Quarterly Journal of Economics 112, 443–478.

[15] **Laibson, David**. 2006. "Intertemporal Decision Making." Encyclopedia of Cognitive Science, Wiley & Sons Ltd. online.

[16] **Luttmer, Erzo, and Thomas Mariotti**. 2003. "Subjective Discounting in an Exchange Economy." Journal of Political Economy 111, 959–989.

[17] **McClure, Samuel, David Laibson, George Loewenstein, and Jonathan Cohen**. 2004. "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." Science 306, 503–507.

[18] **Meier, Stephan, and Charles Sprengen**. 2010. "Present-Biased Preferences and Credit Card Borrowing." American Economic Journal: Applied Economics 2, 193–210.

[19] **Ness, Alisha, Kira Foley, Eric Heggestad.** 2021. "Intra-Individual Variability in Personality: A Methodological Review." In: "Measuring and Modeling Persons and Situations," Wood, Dustin, Stephen Read, Peter Harms, Andrew Slaughter (Ed.s). Academic Press, 313-353.

[20] **Noussair, Charles, Stefan Traumann, and Gijs van de Kuilen**. 2014. "Higher Order Risk Attitudes, Demographics, and Financial Decisions." Review of Economic Studies 81, 325–355.

[21] **O'Donoghue, Ted, and Matthew Rabin**. 1999. "Doing it Now or Later." American Economic Review 89, 103–124.

[22] **O'Donoghue, Ted, and Matthew Rabin**. 2001. "Choice and Procrastination." Quarterly Journal of Economics 116, 121–60.

[23] **O' Donoghue, Ted, and Matthew Rabin**. 2015. "Present Bias: Lessons Learned and to Be Learned," American Economic Review (Papers and Proceedings) 105, 273–279.

[24] **Phelps, Edmund, and Robert Pollak**. 1968. "On Second-Best National Saving and Game-Equilibrium Growth." Review of Economic Studies 35, 185–199.

[25] **Pollak, Robert**. 1968. "Consistent Planning." Review of Economic Studies 35, 201–208.

[26] **Sapolsky, R.**. 2017. "Behave: The Biology of Humans at Our Best and Worst." London: Penguin Press.

[27] **Strotz, R.H.**. 1956. “Myopia and Inconsistency in Dynamic Utility Maximization.” Review of Economic Studies 23, 165–80.

[28] **Sulka, T.** 2022. “Exploitative Contracting in a Life-Cycle Savings Model.” Mimeo, DICE, University of Düsseldorf.

[29] **Thaler, Richard, and Cass Sunstein**. 2008. “Nudge: Improving Decisions About Health, Wealth, and Happiness.” New Haven: Yale University Press.

[30] **Thomas, Jonathan, and Timothy Worrall**. 1988. “Self-Enforcing Wage Contracts.” Review of Economic Studies 55, 541–554.

[31] **Uhlig, Harald.** 1996. “A Law of Large Numbers for a Large Economy.” Economic Theory 8, 41–50.

[32] **van Eerde, Wendelien.** 2003. “A Meta-Analytically Derived Nomological Network of Procrastination.” Personality and Individual Differences 35, 1401-1418.

[33] **Zinman, Jonathan**. 2014. “Consumer Credit: Too Much or Too Little (or just right)?” Journal of Legal Studies 43, S209–S237.

5 Appendix A

Throughout the appendix, we use the nimbler notation β_s for $\beta(s)$ and $\hat{\rho}$ for ρ/δ. Also, we let $\mu(s^t)$ be the probability of history s^t under the transition μ.

We prove Proposition 1 and Lemma 1 in steps. We develop first order conditions for the sophisticated problem and the IC-problem (in the utility space). Under A1 and A2.1, standard arguments yield existence of a solution to the sophisticated and unsophisticated problems, where feasibility holds as equality. Indeed, under A2.1 feasibility creates an upper bound on utiles v_t which remain strictly below $v(\infty)$, and compactness of the constraint set ensues. Further, under A2.2, $u(x_0^{FB}) > u(0)$, and

$$u'(x_0^{FB})\hat{\rho}^t \geq v'(x^{FB}(s^t)), \text{ all } s^t, \tag{FOC*-FB}$$

with equality if $v(s^t) > 0$. If v is strictly concave, then the solutions are unique, and $v_t^{FB}(s^t) = v_t^{FB}$, all s^t, all $t > 0$.

Next, we characterize the first order conditions for the IC-problem. The IC-problem in the utility space is convex, as the incentive constraints are now linear in utiles. They can be expressed as $M\ \overrightarrow{v}\ \geq 0$, for an appropriate matrix M having as many rows as the number of incentive constraints, and as many columns as histories s^t, properly arranged in ascending order while respecting the time order, and $\overrightarrow{v}$ is the correspondingly arranged utile vector. The KT conditions are necessary and sufficient if a Slater constraint qualification condition holds —that this is the case it is shown separately below, in Appendix B.

As A2.1 implies $v_t < v(\infty)$, all $t > 0$, at any IC-efficient allocation, then $u_0, v_t, t > 0$ is a solution to the IC-problem with $u_0 > u(0)$ if and only if there is a vector $\alpha > 0$ of multipliers such that, for all s^t and all $t > 0$,

$$\alpha M(s^t) \leq \frac{\rho^t u_0'}{v'(s^t)}\pi(s^t) - \delta^t \mu(s^t), \tag{$FOC(s^t)$}$$

with equality if $v(s^t) > 0$. Next, for an arbitrary $t > 0$ consider the pooling $v_{t'}(t)$, $t' > 0$, defined as: $v_{t'}(t) = 1$, if $t' = t$; while $v_{t'}(t) = 0$ otherwise. As $M\ \overrightarrow{v}\ = 0$ for all pooling processes $v_t, t > 0$, it is $\sum_{s^t} M(s^t) = 0$. From FOC(s^t) and using the fact

$\sum_{s^t} M(s^t) = 0$, we immediately derive the key consequence.

Auxiliary Lemma 1 *Under A1 and A2.1, if at a solution to the IC-problem $u_0 > u(0)$, then an inverse marginal utility inequality holds,*

$$\frac{1}{u_0'} \leq \hat{\rho}^t \mathbb{E}_\pi \frac{1}{v_t'}, \text{ all } t > 0. \tag{IMU}$$

We are now in a position to prove Proposition 1.

Proof of Proposition 1: By contradiction, suppose that $u_0^{IC} > u_0^{FB}$, and then $u_0^{FB} < u(r)$. By concavity of u,

$$u'(x_0^{FB}) \geq u'(x_0^{IC}), \tag{*}$$

while from feasibility there exists $t > 0$ such that

$$\mathbb{E}_\pi(x_t^{FB}) > \mathbb{E}_\pi(x_t^{IC}) \geq 0. \tag{**}$$

In either case, as $u_0^{IC} > u(0)$, by IMU and Jensen's inequality applied to $1/v'$, at such t we get the inequality

$$\frac{1}{u'(x_0^{IC})} \leq \hat{\rho}^t \mathbb{E}_\pi(\frac{1}{v'(x_t^{IC})}) \leq \hat{\rho}^t \frac{1}{v'(\mathbb{E}_\pi(x_t^{IC}))} \Leftrightarrow \hat{\rho}^t u'(x_0^{IC}) \geq v'(\mathbb{E}_\pi(x_t^{IC})).$$

Combining this inequality with (**) and concavity of v, and using (*), we obtain

$$\hat{\rho}^t u'(x_0^{IC}) \geq v'(\mathbb{E}_\pi(x_t^{IC})) \geq v'(\mathbb{E}_\pi(x_t^{FB})) = \hat{\rho}^t u'(x_0^{FB}). \tag{***}$$

where the last equality comes from $u(x_0^{FB}) > u(0)$ and from FOC*-FB.

If v is strictly concave, it is $\mathbb{E}_\pi(x_t^{FB}) = x_t^{FB}$ and $v'(\mathbb{E}_\pi(x_t^{FB})) = v_t'^{FB}$. Then, by inequality (**), $v'(\mathbb{E}_\pi(x_t^{IC})) > v_t'^{FB}$. If u is strictly concave, (*) holds as a strict inequality. In either case, (*) and (***) combined lead to $u'(x_0^{FB}) > u'(x_0^{FB})$, an absurd yielding a contradiction. ■

FOC*-FB and FOC(s^t) can also be used to prove Lemma 1 and, in turn, Proposition 2.

Proof of Lemma 1: Under A2.1 and A2.3, and v strictly concave, both the IC-efficient and the sophisticated allocations are unique and strongly interior. If the sophisticated allocation is the solution to the IC-problem, under the Slater condition the corresponding linearized problem

$$\begin{gathered} \max_{\Delta u_0, \Delta v_t, t>0} \Delta u_0 + \mathbb{E}_\mu \sum_{t>0} \delta^t \Delta v_t \quad \text{s. to} \\ M\ \overrightarrow{\Delta v} \geq 0, \\ \frac{1}{u_0'^{FB}} \Delta u_0 + \mathbb{E}_\pi \sum_{t>0} \frac{\rho^t}{v_t'^{FB}} \Delta v_t \leq 0, \end{gathered}$$

has (zero as) a solution. Since feasibility holds as equality, and taking into account that $1/u_0'^{FB} = \hat{\rho}^t / v_t'^{FB}(s^t)$, all $s^t > 0$, after substituting out for Δu_0, the problem equivalently becomes problem (P).

Conversely, if linear problem (P) has a solution, then there cannot be any incentive compatible vector $\overrightarrow{v}$ such that $\mathbb{E}_\mu \sum_{t>0} \delta^t v_t - \mathbb{E}_\pi \sum_{t>0} \delta^t v_t > 0$.

By Farkas' Lemma, there exists a vector $\alpha > 0$ such that

$$\alpha M(s^t) = \delta^t(\pi(s^t) - \mu(s^t)), \tag{*}$$

all s^t. Now recall that $v_t^{FB}(s^t) = v_t^{FB} > 0$, and $\frac{u_0'^{FB}}{v_t'^{FB}} \rho^t = \delta^t$, while FOC($s^t$) reads $\alpha M(s^t) = \frac{u_0'}{v'(s^t)} \rho^t \pi(s^t) - \delta^t \mu(s^t)$. Thus, (*) is FOC($s^t$) computed at the sophisticated allocation. As the solution to the convex IC-problem is unique, the FB-allocation is the solution to the IC-problem, as wanted.■

We are now ready to prove Proposition 2.

Proof of Proposition 2: (i) We argue by contradiction so that by Proposition 1 we assume that $u_0^{IC} = u_0^{FB} = u_0$. If v is strictly concave IC and FB solutions are unique and by A2.3 they are strongly interior. Therefore by IMU and $\sum M(s^t) = 0$ the IC solution satisfies

$$\frac{1}{u_0'} = \mathbb{E}_\pi \frac{\hat{\rho}^t}{(v_t^{IC})'} \geq \frac{\hat{\rho}^t}{v_t'(\mathbb{E}_\pi x_t^{IC})}, \text{ all } t > 0, \tag{$\sharp$}$$

with strict inequality if the r.v. x_t^{IC} displays variability.

By strict concavity of v, it is $\mathbb{E}_\pi(x_t^{FB}) = x_t^{FB}$, so that by ($\sharp$) it is $x_t^{FB} = \mathbb{E}_\pi x_t^{IC}$ if x_t^{IC} is degenerate, while $x_t^{FB} > \mathbb{E}_\pi x_t^{IC}$ if x_t^{IC} displays variability. The latter together with Lemma 1 contradicts feasibility.

The proof of (ii) follows a very similar argument, and is left to the reader.■

Proof of Proposition 3: Fix r and drop reference to it throughout. Let $v_t^\kappa(T), t \leq T$ be a solution to the T-horizon problem $\kappa =$ C, IC, and $v_t^\kappa(\infty), t \geq 0$, a solution process when $T = \infty$. Let $V^\kappa(T)$ be the value of problem κ at T, and let $W^\kappa(T) = \mathbb{E}_\pi \sum_{t\geq 0} \delta^t v_t^\kappa(r,T)$. Finally, for each κ and $T > 1$ define $\phi^\kappa(T) \in \mathbb{R}^\infty$ as $\phi_t^\kappa(T) \equiv v_t^\kappa(T)$ if $t \leq T$ and $\phi_t^\kappa(T) \equiv 0$ if $t > T$. We consider sequences $\phi^\kappa(T), T > 1$.

Since π has full support and v is strictly concave, $\nu^\kappa(T)$ exists, is unique and $v_t^C(s^t,T) = v_t^C(T)$ for all s^t, $t' \leq T$ and $t < t'$. Thus, the following properties are readily established (see, e.g., Citanna et al. (2022b) for (i) and (ii)).

Auxiliary Lemma 2 *For all $T > 1$: i) $v_t^C(T) \geq v_{t'}^C(T)$, $t' \geq t$ for all $t < T$; ii) $\phi_t^C(T) \in [0, v(r)]$ all $t \geq 0$, with $v_0^C(T) \in [v((1-\rho)r), v(r)]$; for $\kappa =$ C, IC, iii) $W^\kappa(T) \in [0, \frac{v(r)}{1-\delta}]$; iv) $\phi^\kappa(T) \to v^\kappa(\infty)$ pointwise, and both $W^\kappa(T) \to W^\kappa(\infty)$ and $V^\kappa(T) \to V^\kappa(\infty)$.*

By Auxiliary Lemma 2.iv, $\phi^C(T) \to v^C(\infty)$ pointwise, $v^C(\infty)$ solves the C-problem for $T = \infty$ and $W^C(\infty) = \lim_{T\to\infty} W^C(T) = \sum_{t\geq 0} \delta^t v_t^C(\infty)$. Similarly since $\phi_t^{IC}(s^t,T) \in [0,1]$, and since $W^{IC}(T) < \frac{v(r)}{1-\delta}$, $W^{IC}(\infty) = \lim_{T\to\infty} W^{IC}(T) = \mathbb{E}_\pi \sum_{t\geq 0} \delta^t v_{\infty,t}^{IC}$ for any subsequence $\{\phi^{IC}(T)\}_{T>1}$ (without loss of generality, the sequence itself) such that $\phi^{IC}(T) \to v_\infty^{IC}$ pointwise.

To avoid the possibility that $\lim_{T\to\infty} v^{IC}(s^t,T) = 1$, pick $\lambda \in (0,1)$ and consider the process $\phi^\lambda(T) = \lambda\phi^C(T) + (1-\lambda)\phi^{IC}(T)$. For all $T > 1$, $W^{IC}(T) \leq W^\lambda(T) \leq W^C(T)$ by linearity, and $\sum_{t\geq 0} \rho^t \mathbb{E}_\pi v^{-1}(\phi_t^\lambda(T)) \leq r$ by the convexity of v^{-1}, and for all $t \geq 0$, $\phi_t^\lambda(T) \in [0, \bar{v}_\lambda]$, where $\bar{v}_\lambda = \lambda v(r) + 1 - \lambda < 1$ since $v_t^C(T) \in [0, v(r)]$. As $\phi^\lambda(T)$ is bounded, uniformly in T, at no loss of generality $\phi^\lambda(T) \to v_\infty^\lambda$ pointwise, and by linearity $v_\infty^\lambda = \lambda v^C(\infty) + (1-\lambda)v_\infty^{IC}$. Since $v^{-1}(\phi_t^\lambda(T)) \in [0, v^{-1}(\lambda v(r) + 1 - \lambda)]$, and since both $\sum_{t\geq 0} \delta^t \mathbb{E}_\pi v_t$ and $\sum_{t\geq 0} \rho^t \mathbb{E}_\pi v^{-1}(v_t)$ are continuous function on $\times_t \times_{s^t} [0, \bar{v}_\lambda]$, we get

$$\begin{aligned}
\lim_{T\to\infty} W^\lambda(T) &= \mathbb{E}_\pi \sum_{t\geq 0} \delta^t v_{\infty,t}^\lambda = W^\lambda(\infty), \text{ and} \\
\lim_{T\to\infty} \mathbb{E}_\pi \sum_{t\geq 0} \rho^t v^{-1}(\phi_t^\lambda(T)) &= \sum_{t\geq 0} \rho^t \mathbb{E}_\pi v^{-1}(v_{\infty,t}^\lambda) \leq r.
\end{aligned}$$

Therefore, v^λ_∞ belongs to the constraint set of the C-problem at $T = \infty$. Then, if $v^\lambda_\infty \neq v^C(\infty)$, by the strict convexity of v^{-1} it must be that $W^\lambda(\infty) < W^C(\infty)$ (and therefore, the thesis follows). Thus, to conclude the argument we need to show that $v^\lambda_\infty \neq v^C(\infty)$. Argue by contradiction by assuming that $v^\lambda_\infty = v^C(\infty)$, and therefore that $v^{IC}_\infty = v^C(\infty)$ and $\phi^{IC}(T) \to v^C(\infty)$, pointwise.

Consider the set of processes $\Phi(T^*) = \{\phi_t, t \geq 0 : \phi_t = v^C_t(\infty),\ t > T^*\}$, and the problem

$$\max_{(\phi_t, t\geq 0)\in\Phi(T^*)} \frac{\phi_0}{\beta_0} + \sum_{t>0} \delta^t \mathbb{E}_\mu \phi_t \text{ s. to B, F, IC.} \tag{T*}$$

Since $(\phi_t, t \geq 0) \in \Phi(T^*)$, the IC-constraints in the problem (T*) are the T^*-horizon IC constraints, and its FOC are the FOC of the IC problem with horizon T^* and resources $r - \sum_{t>T^*} \rho^t \mathbb{E}_\pi v^{-1}(v^C_t(\infty))$. Further, $v^C(\infty)$ is in the constraint set of the T*-problem, but by T^*-naïveté and Lemma 1, it cannot be an optimal solution to the T*-problem. Denoting by $\hat{v}_\infty$ a solution to the T*-problem, by strict concavity $\hat{v}_\infty \neq v^C(\infty) = v^{IC}_\infty$. Let V denote the value of the T^*-problem, and let $V^{IC} = \frac{v^{IC}_{\infty,0}}{\beta_0} + \sum_{t>0} \delta^t \mathbb{E}_\mu v^{IC}_{\infty,t}$. Then T^*-naïveté implies

$$V^{IC} = \frac{v^{IC}_{\infty,0}}{\beta_0} + \sum_{t>0} \delta^t \mathbb{E}_\mu v^{IC}_{\infty,t} < V = \frac{\hat{v}_{\infty,0}}{\beta_0} + \sum_{t>0} \delta^t \mathbb{E}_\mu \hat{v}_{\infty,t}.$$

Since $\phi^{IC}(T) \to v^{IC}_\infty = v^C(\infty)$ and since $\phi^C(T)$ is uniformly bounded, it is $V^C = V^{IC}$. Therefore, for T high enough

$$\frac{\phi^{IC}_0(T)}{\beta_0} + \sum_{t>0} \delta^t \mathbb{E}_\mu \phi^{IC}_t(T) < \frac{\hat{v}_{\infty,0}}{\beta_0} + \sum_{t=1}^{T} \delta^t \mathbb{E}_\mu \hat{v}_{\infty,t}.$$

By definition $\phi^{IC}_t(T) = v^{IC}_t(T), t \leq T$ is an optimal solution to the IC problem at T. On the other hand, $\hat{v}_{\infty,t}, t \leq T$, belongs to constraint set of the IC problem at T thereby turning the inequality above into a contradiction.■

Proof of Proposition 4: It suffices to show that, for every $T > 1$ and under A2.1',

$$v_t \leq \left(\frac{1 + \beta_{\bar{s}} \sum_{t>0} \delta^t}{(\beta_1)^{T-1}}\right)^{T-1} V^{**}_T < v(\infty), \text{ all } t > 0, \tag{BB}$$

for any process $v_t, t > 0$ in the constraint set of the IC-problem. Then, since the constraint set is then nonempty and compact in the standard Euclidean topology, a finite solution to the IC-problem exists via Weierstrass Theorem. If so, it is $v_t < v(\infty)$, all $t > 0$, and it is an easy step to see that IMU holds, and the conclusions of Proposition 1 hold. We now show that a solution exists.

The argument to show that BB holds is broken in two steps. Let $u_0, v_t,\ t > 0$ be in the constraint set of the IC-problem, and let $W(s^t|s) = \mathbb{E}_{s,\mu} \sum_{\tau > t} \delta^{\tau - t} v_\tau \mathbf{1}_{\{s^\tau \succ s^t\}}$, all $s \in S$.

Step 1: $v(s^t) = h \in [0, v(\infty)]$ *implies*

$$W(s^\tau|s) \geq (\beta_1)^{t-\tau-1} h, \text{ all } s \in S, \text{ all } \tau < t.$$

Proof: Let $\tau = t - 1$. By the incentive constraints IC, and $v \geq 0$ while $\beta_{\bar{s}} < 1$, it is

$$\begin{aligned} v(s^{t-1}, s) + W(s^{t-1}, s|s) &\geq v(s^{t-1}, s) + \beta_s W(s^{t-1}, s|s) \geq \\ v(s^t) + \beta_s W(s^t|s) &\geq h, \text{ for all } s. \end{aligned} \tag{1}$$

This implies

$$\begin{aligned} W(s^{t-1}|s') &= \sum_s \mu(s|s')[v(s^{t-1}, s) + W(s^{t-1}, s|s)] \\ &= \mathbb{E}_{s',\mu}[v(s^{t-1}, s) + W(s^{t-1}, s|s)] \geq h, \end{aligned}$$

all $s' \in S$, i.e., the statement at $\tau = t - 1$. Assume now by induction that the statement holds for $\tau < t - 1$, and consider $(s^{\tau-1}, s)$, with $s \neq s_\tau$. Then,

$$\begin{aligned} v(s^{\tau-1}, s) + W(s^{\tau-1}, s|s) &\geq v(s^{\tau-1}, s) + \beta_s W(s^{\tau-1}, s|s) \\ &\geq v(s^\tau) + \beta_s W(s^\tau|s) \geq \beta_s {\beta_1}^{t-\tau-1} h \geq \beta_1^{t-\tau} h. \end{aligned}$$

Therefore,

$$W(s^{\tau-1}|s') = \mathbb{E}_{s',\mu}[v(s^{\tau-1}, s) + W(s^{\tau-1}, s|s)] \geq \beta_1^{t-\tau} h,$$

all $s' \in S$, i.e., the statement is true at $\tau - 1$, as wanted. $\square$

Let $K = \frac{1 + \beta_{\bar{s}} \sum_{t>0} \delta^t}{(\beta_1)^{T-1}}$. It is $K \geq 1$.

Step 2: *It is* $v_t \leq K^{T-1} V_T^{**}$ *for all* $(u_0, v_t, t > 0)$ *elements of the constraint set in the IC-problem.*

Proof: Let $(u_0, v_t, t > 0)$ be an element of the constraint set in the IC-problem. We show that for each $t < T$ and history s^t with $\pi(s^t) > 0$, $v(s^\tau) \leq K^{T-(t+1)} V_T^{**}$, for all $s^\tau \succeq s^t$. Therefore, for $t = 0$, $v(s^t) \leq K^{T-1} V_T^{**}$, all $s^t, t > 0$.

The argument is by backward induction. For $t = T - 1$, the incentive constraints imply that $v(s^{T-1}, s) = v_T(s^{T-1})$, for all s. Hence, as $\pi(s^{T-1}) > 0$ implies $\pi(s^{T-1}, s) > 0$ for some s, it is $v(s^{T-1}, s) \leq V_T^{**}$, and then

$$v(s^{T-1}, s) = v_T(s^{T-1}) \leq V_T^{**} \text{ and } v(s^{T-1}) \leq V_T^{**}, \text{ all } s^{T-1} \text{ with } \pi(s^{T-1}) > 0,$$

proving the first step of the induction argument (with $K^{T-T} = 1$).

Suppose now the statement is true for $t < T - 1$ and all $s^\tau \succeq s^t$ with $\pi(s^t) > 0$. We show that is true at $t - 1$. As $\pi(s^{t-1}) > 0$, feasibility implies $v(s^{t-1}) \leq V_T^{**}$. By the induction assumption $v(s^\tau) \leq K^{T-(t+1)} V_T^{**}$, for all $s^\tau \succeq s^t$ if $\pi(s^t) > 0$. Therefore, for all s_t with $\pi(s^{t-1}, s_t) > 0$, it is

$$v(s^t) + \beta_{s_t} W(s^t | s_t) \leq (1 + \beta_{s_t} \sum_{\tau > t} \delta^{\tau - t}) K^{T-(t+1)} V_T^{**}.$$

Consider now (s^{t-1}, s) with $\pi(s^{t-1}, s) = 0$. Let $h(s^{t-1}, s) = \max_{s^\tau \succeq (s^{t-1}, s)} v(s^\tau) = v(s^{\tau^*})$ some $\tau^* \geq t$. From Step 1, and $v \geq 0$,

$$v(s^{t-1}, s) + \beta_{s_t} W(s^{t-1}, s | s_t) \geq (\beta_1)^{T-t} h(s^{t-1}, s)].$$

Therefore, by the incentive constraint associated to s^t and such (s^{t-1}, s), it is

$$v(s^t) + \beta_{s_t} W(s^t | s_t) \geq v(s^{t-1}, s) + \beta_{s_t} W(s^{t-1}, s | s_t) \geq (\beta_1)^{T-t} h(s^{t-1}, s),$$

thereby implying, for all (s^{t-1}, s) with $\pi(s^{t-1}, s) = 0$,

$$\begin{aligned} h(s^{t-1}, s) &\leq \frac{1 + \beta_{s_t} \sum_{\tau > t} \delta^{\tau - t}}{(\beta_1)^{T-t}} K^{T-(t+1)} V_T^{**} \\ &\leq \frac{1 + \beta_{\bar{s}} \sum_{t>0} \delta^t}{(\beta_1)^{T-1}} K^{T-(t+1)} V_T^{**} = K^{T-t} V_T^{**}. \end{aligned}$$

Thus, $v(s^\tau) \leq K^{T-t} V_T^{**}$ for all $s^\tau \succeq s^{t-1}$ and $\pi(s^{t-1}) > 0$, as wanted, concluding the argument. ■

6 Appendix B

6.1 Pessimism: Further analysis

A more detailed use of the first order conditions for the IC-problem allows to discuss a couple of notions of pessimism only based on belief comparison.

Let $\pi_t(s) = \sum_{s^{t-1}} \pi(s|s_{t-1})\pi(s^{t-1})$ be the t-th iterate of the initial distribution over selves through the transition π, and similarly we define μ_t. We say that the unsophisticated consumer is *mean pessimistic* if π_t FOSD μ_t for every $t > 0$. [24]

The consumer is *strongly pessimistic* if $\pi(\cdot|s)$ FOSD $\mu(\cdot|s)$ for all $s \in S$. Strong pessimism is the statement that every self thinks she is going to be less patient than what she actually is. When selves are independently and identically distributed under the subjective transition μ, i.e., $\mu(\cdot|s) = \mu(\cdot)$ all $s \in S$, we say that selves are *subjectively i.i.d.*.

Proposition 5 *Under A1 and A2, when v is strictly concave:*

i) an unsophisticated individual behaves as a sophisticate only if she is a mean pessimist;

ii) when selves are subjectively i.i.d., pessimism and strong pessimism coincide.

By (ii), when subjective beliefs are common across all selves, pessimism can be equivalently be expressed only in terms of conditionals.

[24] If the initial distributions $\pi(\cdot|s_0), \mu(\cdot|s_0)$ coincide with the invariant probability vectors of the transitions then $\pi(\cdot|s_0)$ FOSD $\mu(\cdot|s_0)$ implies mean pessimism.

Proof: We use the notation β_s, as introduced at the beginning of Appendix A, and we set $\delta = 1$ at no loss, to simplify the notation. Let $S = \{1, ...\bar{s}\}$ where $\bar{s} \geq s$, all $s \in S$.

(i) We simplify the IC-problem by eliminating the global constraints for each s^{t-1} and $t > 0$. Letting

$$\phi_s = \begin{cases} 2 & \text{if } s = 1 \\ s \pm 1 & \text{if } 1 < s < \bar{s} \\ \bar{s} - 1 & \text{if } s = \bar{s}. \end{cases}$$

the local IC constraints are written as

$$v_t(s^{t-1}, s) + \beta_s \mathbb{E}_{s,\mu} \sum_{\tau > t} v_\tau \geq v_t(s^{t-1}, \phi_s) + \beta_s \mathbb{E}_{s,\mu} \sum_{\tau > t} v_\tau(\phi_s, t).$$

Thus, using matrix M introduced earlier (just before Auxiliary Lemma 1), we have replaced the IC constraints $M \overrightarrow{v} \geq 0$ with the IC constraints $M' \overrightarrow{v} \geq 0$, where M' is the matrix obtained from M by deleting the rows indexed by (s^t, s), $s \neq \phi_s$. Using β_s increasing in s, it is clear that the local constraints imply the global ones. Thus, the simplified problem is equivalent to the IC-problem. Expression FOC(s^t) above shows that necessary and sufficient conditions for optimality in the simplified problem, computed at a strongly interior FB-efficient allocation, read exactly as

$$\alpha' M(s^t) = \pi(s^t) - \mu(s^t), \tag{GP'}$$

for all s^t and every $t > 0$, where α' is a vector of multipliers. Hence, by uniqueness of a solution, FOC(s^t) has a solution $\alpha \geq 0$ if and only if GP' has a solution $\alpha' \geq 0$. Hereafter, we explicitly compute the first order conditions of the simplified problem.

For a given pair process $(\alpha^+, \alpha^-)_t$, $t > 0$, and for any function $x = (x(s))_{s \in S}$ and any positive $t < T$ and $s \in S$, let

$$\begin{aligned} A(\alpha, x, s^t) &= [\alpha^+(s^t) + \alpha^-(s^t)]x(s_t) \\ &\quad - \alpha^+(s^{t-1}, s_t - 1)x(s_t - 1) - \alpha^-(s^{t-1}, s_t + 1)x(s_t + 1), \end{aligned}$$

where $\alpha^-(s^{t-1}, 1) \equiv \alpha^-(s^{t-1}, \bar{s}+1) \equiv 0$, and $\alpha^+(s^{t-1}, 0) \equiv \alpha^+(s^{t-1}, \bar{s}) \equiv 0$. When $x(s) = \beta_s$ we write $A(\alpha, \beta, s^t)$. When $x(s) = \beta_s \mu(s^{t+1,t'}|s)$, for $t < t'$, we write $A(\alpha, \beta\mu^{t'}, s^t)$;

and when we want to stress element $s_{t'}$ of $s^{t+1,t'}$ we write $A(\alpha, \beta\mu^{t',s_{t'}}, s^t)$. Finally, when $x(s) = 1$, we simply write $A(\alpha, 1, s^t) \equiv A(\alpha, s^t)$.

Equation $\alpha' M'(s^{t-1}, s) = \pi(s^{t-1}, s) - \mu(s^{t-1}, s)$ now reads

$$A(\alpha, s^t) + \sum_{\tau<t} A(\alpha, \beta\mu^{t,s_t}, s^\tau) = \pi(s^t) - \mu(s^t). \tag{2}$$

Using a recursive argument, it is seen that this system of equations has a unique solution given by

$$\begin{aligned} \hat{\alpha}^-(s^{t-1}, s) &= \gamma(\hat{\alpha}; s^{t-1}, s)\frac{1-\beta_{s-1}}{\beta_s - \beta_{s-1}} \\ \hat{\alpha}^+(s^{t-1}, s-1) &= \gamma(\hat{\alpha}; s^{t-1}, s)\frac{1-\beta_s}{\beta_s - \beta_{s-1}} \end{aligned}$$

for all $s > 1$, all s^{t-1} and all positive t, with the usual conventions for $s = 1, \bar{s}$, where function $\gamma_t(\alpha)$ is defined recursively via

$$\gamma(\alpha; 0, s) = \sum_{s' \geq s} [\pi(s') - \mu(s')],$$

and for $t > 1$, any history s^{t-1} and any s,

$$\gamma(\alpha; s^{t-1}, s) = \sum_{s' \geq s} [\pi(s^{t-1}, s') - \mu(s^{t-1}, s') - \sum_{\tau<t} A(\alpha, \beta\mu^{t,s'}, s^\tau)].$$

Thus, by the definition of $\hat{\alpha}$, $\hat{\alpha} \geq 0$ if and only if $\gamma(\hat{\alpha}; s^{t-1}, s) \geq 0$. Finally, notice that

$$\sum_{s_\tau} A(\alpha, \beta\mu^{t,s'}, s^\tau) = 0 \tag{‡}$$

which, since $\tau < t$, implies

$$\sum_{s^{t-1}} A_\tau(\alpha, \beta\mu^{t,s'}, s^\tau) = 0, \text{ all } \tau < t.$$

Thus, summing $\gamma(\hat{\alpha}; s^{t-1}, s)$ across s^{t-1} and using A2.3, i.e., strong interiority, the definition of γ and $\hat{\alpha}$, and of FOSD, we obtain the claim.□

(ii) As $\mu(\cdot|s) = \mu(\cdot)$ all $s \in S$, $\mu(s^{\tau+1,t}|s) = \mu(s^{\tau+1,t})$, all $s \in S$, and $A(\alpha, \beta\mu^t, s^\tau) = \mu(s^{\tau+1,t})A(\alpha, \beta, s^\tau)$. We are going to show that, for any process α and for all histories s^t and positive $t < T$,

$$A(\alpha, \beta, s^t) = \pi(s^{t-1})[\pi(s_t|s^{t-1}) - \mu(s_t)]. \tag{3}$$

First, writing (2) at $t+1$ and summing over s_{t+1}, and using (‡), it is

$$A(\alpha, \beta, s^t) = \pi(s^t) - \mu(s^t) - \sum_{\tau<t} A(\alpha, \beta\mu^t, s^\tau). \tag{B(t)}$$

Equation B(t) immediately implies (3) at $t = 1$. Suppose (3) is true at $t - 1 > 0$. Rearranging B($t-1$), we obtain

$$A(\alpha, \beta, s^{t-1}) + \sum_{\tau<t-1} \mu(s^{\tau+1,t-1})A(\alpha, \beta, s^\tau) = \pi(s^{t-1}) - \mu(s^{t-1}).$$

Thus, using independence,

$$\begin{aligned} A(\alpha, \beta, s^t) &= \pi(s^t) - \mu(s^t) - \mu(s_t)[A(\alpha, \beta, s^{t-1}) + \sum_{\tau<t-1} \mu(s^{\tau+1,t-1})A(\alpha, \beta, s^\tau)] \\ &= \pi(s^t) - \pi(s^{t-1})\mu(s_t), \end{aligned}$$

i.e., (3) at t, as wanted. Thus, again using B(t),

$$\sum_{\tau<t} \mu(s^{\tau+1,t})A(\alpha, \beta, s^\tau) = \pi(s^{t-1})\mu(s_t) - \mu(s^t)$$

and plugging this expression into γ_t, it is

$$\begin{aligned} \gamma(\alpha; s^{t-1}, s) &= \sum_{s' \geq s}[\pi(s^{t-1}, s') - \mu(s^{t-1}, s') - (\pi(s^{t-1})\mu(s') - \mu(s^{t-1}, s')] \\ &= \pi(s^{t-1}) \sum_{s' \geq s}[\pi(s'|s^{t-1}) - \mu(s')]. \end{aligned}$$

The conclusion now follows from the definition of $\hat{\alpha}$.■

We now show that the uniqueness of $\hat{\alpha}$ statement in the above proposition implies that a Slater constraint qualification condition holds for the IC-problem, and the KT conditions are thus both necessary and sufficient for a solution, as assumed in the paper. Recall the definition of matrix M.

Auxiliary Lemma (Slater) *There exists a $v_t, t > 0$ such that constraints F, B and IC hold with strict inequality.*

Proof: By translating by a strictly positive constant and multiplying by sufficiently small positive scalar, it suffices to show that there is a $\bar{v}$ such that $M \overrightarrow{\bar{v}} \gg 0$. As we argued in the previous proposition, $\alpha M = 0$ has the unique solution $\hat{\alpha} = 0$. Thus, for every s^t and $t > 0$ there is no solution $(\alpha_{\backslash s^t}, 1)$ with $\alpha_{\backslash s^t} \geq 0$ to $\alpha_{\backslash s^t} M_{\backslash s^t} + M_{s^t} = 0$, where $M_{\backslash s^t}$ is matrix M' without row s^t, and M_{s^t} is its s^t-th row. This implies via Farkas' Lemma the existence of a vector v^{s^t} such that $M_{\backslash s^t} v^{s^t} \geq 0$ and $M_{s^t} v^{s^t} > 0$. Then, $\bar{v} = \sum_{t,s^t} v^{s^t}$ gives the desired vector.■

6.2 A closed economy

It is straightforward to twist the baseline model to limit all agents' ability to borrow, including firms –as in a closed economy. We add (a continuum of) time-consistent individuals to the economy, with intertemporal discount rate equal to $\delta_2 \leq 1$. For the sake of simplicity we assume that time consistent and time inconsistent individuals have identical population size. Utilities, endowments and net trades are indexed by superscript $i = 1, 2$. Beyond A1-A2, we assume that also u^2, v^2 are differentiable, strictly increasing and concave, and $y_t^i(s^t) > 0$, all s^t, and $t \geq 0$, all i. Market clearing at each date t now is

$$\sum_i \mathbb{E}_\pi (x_t^i - y_t^i) \leq 0, \text{ all } t \geq 0,$$

Efficiency notions now include both (observable) types of individuals, and are defined in a straightforward manner. A first best $u_0^{i,FB}, v_t^{i,FB}, t > 0$, $i = 1, 2$ is the solution to the sophisticated consumer problem,

$$\begin{array}{c} \max_{u_0^i, v_t^i, t>0} u_0^1 + \mathbb{E}_\mu \sum_{t>0} v_t^1 \qquad \text{s. to} \\ u_0^2 + \mathbb{E}_\pi \sum_{t>0} \delta^t v_t^2 \geq \bar{v}^2 \\ \sum_i [(u^i)^{-1}(u_0^i) - y_0^i] \leq 0 \\ \sum_i \mathbb{E}_\pi [(v^i)^{-1}(v_t^i) - \sum_i y_t^i] \leq 0, t > 0 \\ u_0^i \geq 0,\ v_t^i \geq 0, \text{ all } t > 0, \text{ all } i. \end{array}$$

where $\bar{v}^2$ is a utility level parametrizing efficient allocations. An IC-efficient allocation that favors naïve present-biased consumers must solve the same problem but with in addition the incentive constraints for $i = 1$. Although now market clearing introduces an upper limit on net trades, it is evident that even in this variation the borrowing behavior results go through essentially unaltered. In particular, under concavity of $1/v^{1\prime}$, and for any fixed utility level for $i = 2$, the naïve individuals underborrow.

Proposition 6 *Under A1 and A2, let* $1/v^{1\prime}$ *be concave. Then, for any type* 2 *individuals' utility level, the naïve individual does not overborrow.*

It remains an open question to see how naïve consumers with different subjective preferences or probabilities μ^i behave relative to a population of sophisticates.