The Beta-Delta-DELTA Sweet Spot*

David Laibson, Harvard University
Peter Maxted, Berkeley Haas

January 9, 2024

Abstract

When solving calibrated dynamic household finance models with present bias, backwards induction generates equilibria that are highly sensitive to various parameter choices, and hence non-robust. To address this problem, researchers have deployed a range of methodological compromises, such as assuming sufficiently little present bias or assuming substantial naivete. We show that non-robustness can instead be eliminated by using high frequency models (i.e., models with small time-steps). Specifically, robust behavior emerges in a "sweet spot" of time-steps that is consistent with empirical work studying the (psychological) present-bias wedge between *now* and *later*.

*We are grateful to seminar participants at SITE: Psychology and Economics and various institutions for helpful comments. Santiago Medina Pizarro and Kartik Vira provided excellent research assistance. The computations for this paper were run on the Research Computing Environment supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University. We acknowledge financial support from the Pershing Square Fund for Research on the Foundations of Human Behavior.

1 Introduction

Models with present-biased discounting have been used to explain a wide range of household financial choices, including consumption of nondurables, expenditure on durables, credit card and payday borrowing, mortgage refinancing, retirement saving, other forms of saving in illiquid assets, human capital investment, and program enrollment.[1] Despite these many applications, generic models with present bias are plagued by a methodological problem that is holding back applications: predictions are not robust to the way the "time-step" (i.e., period length) is modeled. To give an example, model predictions are materially affected by whether the modeled time-step is one week or one year, *ceteris paribus* (i.e., holding fixed all other variables, like the annualized interest rate and the annualized volatility of shocks).

In discrete time, present-biased preferences (often called β-δ preferences) are given by the discount function $\{1, \beta\delta, \beta\delta^2, \beta\delta^3, \ldots\}$. δ is the standard exponential discount factor and β is the short-run discount factor. Setting $\beta < 1$ creates an additional preference for utility "now" instead of "later." Note, however, that β-δ preferences alone do not specify the temporal division between "now" and "later." Instead, that division is determined implicitly by the modeler's choice of the time-step, which we henceforth denote by $\triangle$ (this is the "DELTA" in our title).

Dynamically inconsistent preferences like present bias are often modeled as an intrapersonal game (e.g., Strotz, 1955; Phelps and Pollak, 1968; Peleg and Yaari, 1973). When analytic solutions are not obtainable in models with present bias (e.g., by guessing and checking an equilibrium strategy), economists use backwards induction as a solution concept and solve for equilibrium numerically. However, backwards induction (with either partial or full sophistication) typically generates a sequence

[1] For some illustrative examples see Strotz (1955), Phelps and Pollak (1968), Laibson (1997), Angeletos et al. (2001), DellaVigna and Malmendier (2004), DellaVigna and Paserman (2005), Shapiro (2005), Amador et al. (2006), Ashraf et al. (2006), DellaVigna and Malmendier (2006), Carroll et al. (2009), Meier and Sprenger (2010), Ganong and Noel (2019), Lockwood (2020), Gerard and Naritomi (2021), Kuchler and Pagel (2021), Laibson et al. (2021), Allcott et al. (2022), Beshears et al. (2023), and Lee and Maxted (2023). For a literature review, see Cohen et al. (2020).

of policy functions that are highly sensitive to the time-step (as well as other model parameters). This sensitivity is caused by strategic feedback that exists because of the combination of (i) dynamically inconsistent preferences, and (ii) the ability of the present self to choose actions that materially control the distribution of the state variables of subsequent selves.[2] If policy functions are highly dependent on the (arbitrary) time-step that is used in a model, then model predictions are not methodologically robust.

In this paper we analyze the consumption functions of present-biased consumers and study the sensitivity of these policy functions to the choice of the time-step $\triangle$. In many papers in the household finance literature, the time-step is chosen either to match the frequency of an associated dataset or for technical convenience (e.g., common assumptions are quarterly or annual time-steps, but a rapidly growing literature is set in continuous time). In this paper, we first show that the time-step is a critical parameter in consumption-saving models with present-biased consumers.

To illustrate this sensitivity, Figure 1 below plots the solution to a single consumption-saving model with present-biased consumers ($\boldsymbol{\beta} = 0.5$), which has been solved with four different time-steps (holding all other parameters fixed on an annualized basis).[3] The flow of annualized consumption is reported on the vertical axis (i.e., $\frac{c}{\triangle}$) and the stock of cash on hand is reported on the horizontal axis. The consumption functions in Figure 1 are all annualized so that they can be directly compared across time-steps. The blue line reports the consumption function for the model with annual time-steps ($\triangle = 1$). The yellow line reports the annualized consumption function for the same model solved with semi-annual time-steps ($\triangle = \frac{1}{2}$). Similarly, the red line reports the annualized consumption function for the same model with approximately two-week time-steps ($\triangle = 1/25^{\text{th}}$ of a year), and the dashed line reports the annualized

[2]For discussion of policy function sensitivity and multiple equilibria in present-biased models, see Laibson (1994), Laibson et al. (1998), O'Donoghue and Rabin (1999), Harris and Laibson (2001, 2003), Krusell et al. (2002), Krusell and Smith Jr. (2003), and Cao and Werning (2018).

[3]For example, if a model with annual time-steps has a gross return of $\boldsymbol{R}$, then the same model with semi-annual time-steps has a gross return per time-step of $\boldsymbol{R}^{\frac{1}{2}}$.

consumption function for the same model posed in continuous time ($\triangle = dt$).[4]

This figure illustrates clear non-robustness associated with the time-step $\triangle$. Consumption in the annual model differs notably from the semi-annual model, both of which differ again from the two-week and continuous-time models.

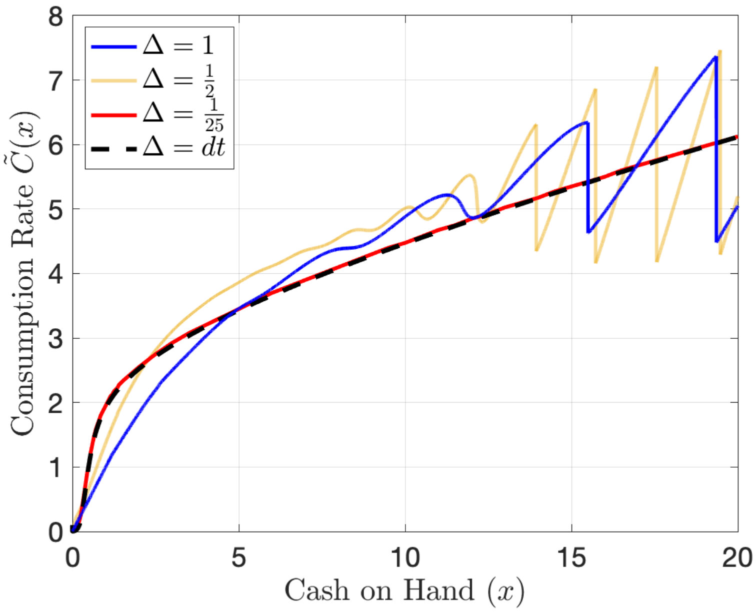


Figure 1: Consumption functions for a present-biased consumer ($\boldsymbol{\beta} = 0.5$) as $\triangle \to 0$. The model is presented in Sections 2 and 3 below. Section 4.2 gives calibration details.

The consumption functions in the annual and semi-annual model exhibit multiple regions of non-monotonicity and multiple downward discontinuities. Such properties are generally believed to be counterfactual, and as such are often referred to as "consumption pathologies" (e.g., Laibson et al., 1998; Harris and Laibson, 2003; Cao and Werning, 2018). For example, the $\triangle = 1$ model predicts that there exist regions of the state space where giving a consumer a one-penny windfall would cause their annual consumption to drop by roughly 40%.

As noted above, consumption pathologies like this arise because of the game-theoretic equilibrium concept that is used to analyze models with dynamically in-

[4]Our modeling of present bias in continuous time follows the Instantaneous Gratification (IG) specification of Harris and Laibson (2013). Details are provided in Section 3.

consistent preferences, where each "self" has preferences that are inconsistent with the preferences of their other temporally situated selves. When agents understand their own dynamically inconsistent preferences, strategic motives emerge in the intrapersonal game that is being played. Intuitively, if self t believes that self $t+1$ is prone to overconsumption, and that a little extra wealth at time $t+1$ will lead self $t+1$ to spend less, then self t might be willing to save more to push self $t+1$ to spend less. Effects like this are propagated and amplified by backwards induction, and produce the counterfactual consumption behavior that appears for $\triangle = 1$ and $\triangle = \frac{1}{2}$ in Figure 1.[5]

These kinds of counterfactual parameter sensitivities raise methodological questions about how to model the choices of present-biased consumers. Various methodological compromises have emerged in the literature: setting β close enough to one to make the pathologies vanish;[6] incorporating enough noise in the dynamic budget constraint to make the pathologies vanish;[7] making beliefs completely naive;[8] or working with two-period models (which prevent the pathologies from propagating in the first place).[9]

[5] These strategic properties can arise even if agents' understanding of their dynamic inconsistency is only partial, as in O'Donoghue and Rabin (2001). However, *completely* naive agents solve a maximization problem when deciding how to behave and consequently incorporate no strategic considerations of this kind.

[6] E.g., Laibson et al. (1998) set $\beta = 0.85$ to make their policy functions well-behaved. They write:

> "For the hyperbolic simulations, we would have preferred to have set β much lower – approximately equal to 0.6 – as Laibson has done in previous work on undersaving. Most of the experimental evidence suggests that the one-year discount rate is at least 40 percent. However, a value of 0.6 generates pathologies in discrete time simulations: strongly nonmonotonic and noncontinuous consumption functions. Such effects are commonplace in dynamic games such as the intrapersonal game that we consider. In our simulations, these pathologies vanish as β approaches unity. Specifically, we find that strong pathologies only arise for values of β below 0.8, which motivates our decision to adopt a value of 0.85."

Harris and Laibson (2003) prove uniqueness in the class of stationary equilibria for β in a neighborhood of unity.

[7] Again see Harris and Laibson (2003).

[8] This is the approach taken e.g. in Laibson et al. (2023). See also the discussion in Section 5.1 of DellaVigna (2018).

[9] See Harris and Laibson (2003) for an explanation of why anomalies first start to appear when the backward induction is iterated twice from period T.

On the other hand, Figure 1 also highlights that the problem of policy-function non-robustness may in fact be illusory once (many) models with present bias are appropriately calibrated. In contrast to the non-robust predictions that are produced for large time-steps, Figure 1 shows that robust behavior emerges for shorter time-steps of $\triangle = \frac{1}{25}$ or $\triangle = dt$. Indeed, the current paper demonstrates that there exists a time-step "sweet spot" interval that ranges from zero (i.e., dt in continuous-time notation) to roughly two weeks. We call this interval the "sweet spot" for five reasons.

First, time-steps in the sweet spot all generate essentially identical policy functions, so the choice of the time-step within the sweet spot will not affect the model's predictions. That is, the non-robust consumption behavior that exists for large time-steps fades away inside the sweet spot.

Second, the empirical intertemporal choice literature finds that present bias generates sharp discounting (e.g., $\beta < 0.9$) over a horizon of *minutes to days* (see e.g., McClure et al., 2007; Augenblick, 2018; Augenblick and Rabin, 2019). We call this the psychologically relevant range of time-steps for present bias, because this represents the duration of the psychological "present." Using computational methods, we show that the psychologically relevant range of time-steps is a strict subset of the sweet spot. By implication, any time-step in the sweet spot will produce policy functions that approximately match the policy functions that emerge from psychologically well-calibrated models.[10]

Third, the sweet spot includes the continuous-time limit, namely the Instantaneous Gratification (IG) specification of Harris and Laibson (2013) and Maxted (2022). This is a particularly useful boundary for the sweet spot because it is possible to theoretically characterize properties of this continuous-time model, including equilibrium uniqueness and consumption-function continuity. Quantitative homogeneity

[10] Our use of short time-steps only applies to "psychological" present bias, and may not apply in other settings where it is less natural to model present bias at a high frequency, such as time-inconsistent governmental preferences. Harstad (2020) provides a rich analysis of different mechanisms that can cause time inconsistency to arise, including time preferences, the rotation of political power (e.g., elections), preference aggregation, and generalizations of intergenerational altruism.

across the sweet spot implies that these desirable properties of the continuous-time case approximately characterize the entire sweet spot, which subsumes the psychologically relevant range.

Fourth, researchers wishing to characterize behavior in the psychologically relevant range of time-steps can use the continuous-time (IG) boundary of the sweet spot for quantitative modeling. The IG model is particularly computationally tractable, because it reduces to the solution of a differential equation that has desirable regularity properties.[11] For a recent set of papers that leverage the limiting IG specification to tractably model the financial choices of present-biased agents, see e.g. Grenadier and Wang (2007), Laibson et al. (2021), Acharya et al. (2022), Beshears et al. (2022), Maxted (2022), Rivera (2022), Shigeta (2022), and Lee and Maxted (2023).

Fifth, researchers who wish to use discrete-time methods will want to use time-steps that are maximally coarse to reduce computational burden. For example, solving a lifecycle model with five-minute time-steps may not be (currently) computationally feasible, but solving a lifecycle model with two-week time-steps typically is. Accordingly, researchers who prefer to use discrete-time methods will generally want to go to the outer boundary of the sweet spot for computational efficiency, even if this boundary is outside of the psychologically relevant range of time-steps, because policy functions are effectively homogeneous over the entire range of the sweet spot.

This paper documents these claims by studying a dynamic consumption-saving model that can be calibrated with a flexible time-step $\triangle$. We emphasize that the emergence of a time-step sweet spot relies on a key feature of the model: period-by-period shocks. It is well known that strategic behavior of the sort exhibited in Figure 1 for $\triangle = 1$ and $\triangle = \frac{1}{2}$ can be curtailed by adding more noise to the model, which reduces the extent to which any self can precisely predict and therefore manipulate the choices of future selves (e.g., Harris and Laibson, 2003). However, one cannot

[11]Further computational details are provided in Section 4.1 below. See Achdou et al. (2022) for an extensive presentation of finite-difference methods for solving the sorts of Hamilton-Jacobi-Bellman (HJB) equations that arise in consumption-saving models, and Maxted (2022) for a discussion of how these numerical methods extend to the case of present bias.

arbitrarily add noise to a calibrated economic environment. Instead, this paper's insight is that large time-steps – which are psychologically inappropriate in the first place – spuriously reduce the true noise that consumers face by providing implicit diversification. By reducing $\triangle$ to a more appropriate length, we unwind this implicit diversification and reintroduce the (approximately Brownian) noise necessary for reestablishing robust policy functions.[12]

The paper is organized with the following structure. In Section 2 we describe a simple discrete-time consumption-saving model with a flexible time-step $\triangle$. Section 3 presents the limiting continuous-time model. Section 4 describes our calibration and numerical algorithm. Section 5 reports our main results, showing that the sweet spot tends to emerge when time-steps are about two weeks or less. Comparative statics are given in Section 6. Section 7 outlines extensions to our model, including a discussion of naivete, and Section 8 concludes.

2 Discrete-Time Consumption-Saving Model

We begin by describing a discrete-time consumption model with present bias. The model is set up so that it converges to a continuous-time consumption model with present bias as the time-step $\triangle \to 0$. We postpone the presentation of the continuous-time model until Section 3.

After presenting the discrete-time model, we discuss how the calibration of the model varies with the model's *time-step* — i.e., the temporal distance between each period of the discrete-time model. Throughout this paper we denote the time-step of the discrete-time model by parameter $\triangle$. In most papers in the consumption-saving literature, the choice of $\triangle$ is made either for expositional purposes or to match the frequency of an associated dataset (e.g., annual or quarterly survey data). Our paper studies the consequences of making $\triangle$ small.

[12] We demonstrate the importance of noise with illustrative counter-examples, and show that without such noise there exist non-robust equilibria for arbitrarily small time-steps.

In the limit as $\triangle \to 0$ we pass to a continuous-time model. The benefit of continuous time is that it allows us to analytically characterize features of the unique equilibrium consumption function. Though we switch to continuous time in Section 3, knowledge of continuous-time methods is not needed to understand many of this paper's findings. Indeed, one of our key conclusions is that the limiting continuous-time model generates predictions that are analogous to those of discrete-time models with psychologically appropriate time-steps, since these models all exist within the sweet spot.

We adopt the (arbitrary) timing convention that $\triangle = 1$ refers to a model with annual time-steps. This annual benchmark model will be henceforth referred to as the $\triangle = 1$ model, or the 1-model in short. When $\triangle \neq 1$ we will refer to this as the $\triangle$-model (in practice, we will only be studying cases in which $\triangle < 1$). We will refer to the limiting continuous-time model as the dt-model.

2.1 The Discrete-Time Model

We now present a discrete-time model with an arbitrary time-step of $\triangle$ years. We will characterize this model so that it is internally consistent for all values of $\triangle > 0$.

Time periods are indexed by integers t. We emphasize that these periods have a length that we vary. For example, if $\triangle = 1$ then period $t+1$ is one year from period t, while if $\triangle = \frac{1}{52}$ then period $t+1$ is one week from period t.

Dynamic Budget Constraint. Assets in period t are denoted x_t. Let R denote the gross interest rate. The dynamic budget constraint is:

$$x_{t+1} = R(x_t - c_t) + z_{t+1}. \tag{1}$$

Variable c_t denotes consumption, and z_t is a stochastic "balance-sheet process" (detailed in the next paragraph). We impose a liquidity constraint that restricts the agent from borrowing against uncertain future wealth: $c_t \leq x_t$. The model can be generalized to include borrowing, but we omit this generalization to simplify exposition and minimize notation.

We assume the following reduced-form data generating process for z_t:

$$z_t = \overline{z} + \sigma \varepsilon_t.$$

$\overline{z}$ represents a deterministic income flow, and ε_t is an i.i.d. balance-sheet shock. Such period-by-period balance-sheet shocks are a reduced-form modeling tool for capturing the wide-ranging set of shocks that characterize a household's true economic problem, including, for example, asset return shocks, income shocks, unexpected medical costs, broken appliances, surprising utility bills, and even taste shocks.[13]

As we demonstrate below, such period-by-period noise is critical for establishing a sweet spot as $\triangle \to 0$. We assume that $\varepsilon_t \sim \mathcal{N}(0, 1)$, but this can be generalized.[14] We could also make the balance-sheet process autocorrelated, but this comes at the cost of additional state variables and complicates exposition.

The restriction that $c_t \leq x_t$ means that assets cannot become negative under the control of the agent. However, if $\sigma > 0$ then a sufficiently negative shock can cause assets to fall below $x = 0$. In order to prevent arbitrarily negative consumption, our model allows the agent to endogenously declare bankruptcy. Bankruptcy is modeled as a stopping problem. Details of the endogenous bankruptcy choice are given below.

Utility and Value. The agent has sophisticated present-biased time preferences $(1, \beta\delta, \beta\delta^2, \beta\delta^3, ...)$. Present-biased preferences are time-inconsistent. We model consumption as a dynamic game played by different temporal selves of the consumer.

In the continuation region the agent accumulates utils through consumption, as is standard. Upon declaring bankruptcy, we assume that the agent earns a certainty

[13] See Strack and Taubinsky (2021).

[14] The important requirements are that ε_t is a random variable and that the sum of ε's converges to Brownian motion as the model's time-step is taken to zero.

equivalent utility flow of u_B in the current and all future periods.[15] Introducing notation that we use below, we let $V_B = \frac{u_B}{1-\delta}$ denote the exponentially discounted value of bankruptcy.

Equilibrium: T-Horizon Game. Time is finite, and the consumer makes a consumption/bankruptcy decision at each period $t \in \{1, 2, ..., T-1\}$. We study the behavior of the sequence of policy functions for large T. Backward induction is applied to look for strategies in each period t that comprise a Markov equilibrium in cash on hand x_t. Starting from a terminal value condition, V_T, an equilibrium is characterized by the following system of Bellman equations defined on $x \in (-\infty, \infty)$ and $t \in \{1, 2, ..., T-1\}$:

$$C_t(x_t) \in \underset{c \leq x_t}{\operatorname{argmax}} \;\; u(c) + \beta\delta \mathbb{E}_t V_{t+1}(x_{t+1}), \tag{2}$$

$$W_t(x_t) = \max \Big\{ u(C_t(x_t)) + \beta\delta \mathbb{E}_t V_{t+1}(x_{t+1}),\; u_B + \beta\delta V_B \Big\}, \tag{3}$$

$$V_t(x_t) = \begin{cases} u(C_t(x_t)) + \delta \mathbb{E}_t V_{t+1}(x_{t+1}) & \text{if } x_t \notin \mathcal{B}_t^* \\ V_B & \text{if } x_t \in \mathcal{B}_t^* \end{cases}, \tag{4}$$

$$\mathcal{B}_t^* = \Big\{ x_t \mid u_B + \beta\delta V_B \geq u(C_t(x_t)) + \beta\delta \mathbb{E}_t V_{t+1}(x_{t+1}) \Big\}. \tag{5}$$

$C_t(x_t)$ denotes an equilibrium consumption function and $\mathcal{B}_t^*$ denotes the set of points at which the agent chooses to declare bankruptcy.[16] The agent's continuation-value function is given by $V_t(x_t)$. $W_t(x_t)$ is the current-value function.

In equation (2), $C_t(x_t)$ is the level of consumption that occurs during a single period of length $\triangle$ years. For example, if $\triangle = \frac{1}{52}$ then C_t is the consumption expenditure over one week. This implies that C_t is not comparable as we vary $\triangle$. Accordingly,

[15] Below, our calibrated model sets u_B well below $u(\bar{z})$ in order to disincentive bankruptcy. Such a restriction is not necessary, though it is economically natural.

[16] The bankruptcy decision has some similarities with the sovereign debt and default model of Eaton and Gersovitz (1981). See also Alfaro and Kanczuk (2017) for an application of quasi-hyperbolic discounting to the Eaton and Gersovitz framework.

we introduce a new variable, $\widetilde{C}_t$, which is the annualized rate of consumption:

$$\widetilde{C}_t(x) = \frac{C_t(x)}{\triangle}.$$

Equation (3) shows that bankruptcy is chosen to maximize the current-value function W_t. We assume that bankruptcy is declared at points of indifference between bankruptcy and continuation, and hence a weak inequality is used in equation (5).[17]

Our equilibrium concept requires further discussion. In discrete time we study Markov equilibria of a finite T-horizon game, where T is large. Even when T is large, this is not necessarily equivalent to studying stationary Markov equilibria. If the Bellman operator of the present-biased consumer was a contraction mapping then these two equilibrium approaches would be equivalent. However, for arbitrary $\beta \in (0, 1)$ the Bellman operator of the present-biased consumer is not a contraction mapping (Harris and Laibson, 2001, 2003). Consequently, it is unknown whether or not the T-horizon game converges to a stationary equilibrium as $T \to \infty$ (and even if it does, that equilibrium may not be unique).[18] We choose to study a T-horizon game because backward induction from a terminal payoff is a standard solution technique.

This discussion highlights some of the difficulties that researchers can encounter when solving discrete-time models with present-biased consumers. These complexities again justify our $\triangle \to 0$ approach, as they contrast with both continuous time – where we prove that there exists a unique stationary Markov equilibrium (Section 3) – and also with the sweet spot in discrete time – where the model is quantitatively close to the unique stationary equilibrium of the continuous-time model.

[17] This assumption simplifies the analysis when we pass to the limiting continuous-time model. For details, see footnote 29 in Section 3.

[18] Krusell and Smith Jr. (2003) and Cao and Werning (2018) show equilibrium multiplicity in deterministic consumption models.

2.2 Calibration with a Variable Time-Step

In many dynamic household finance models, the period length is one year or one quarter. In this paper we let the period length vary. This means that we need to calibrate the model in a way that is consistent with any choice of $\triangle$.

We do this in two steps. First, we describe the calibration for the annual benchmark case of $\triangle = 1$. Given a calibration under the $\triangle = 1$ benchmark, we then show how to recalibrate the model for shorter (or longer) period lengths.

The crucial property of our time-step reduction method is that it simply involves a recalibration of the consumption model already described. Any researcher struggling with consumption pathologies need only recalibrate their model for an appropriate time-step $\triangle$, while keeping the model (and accompanying numerical methods) unchanged. This is a key observation for the applied researcher, as it implies that shortening the time-step to enter the sweet spot is easy to implement.

Calibration when $\triangle = 1$ (1-Model). The model described in Section 2.1 consists of the parameters $\bar{z}, \sigma, R, \delta, \beta$, and u_B, as well as utility function $u(c)$. We denote the 1-model calibration of these parameters by deploying **boldface font**. Recall our convention that $\triangle = 1$ refers to annual time-steps. Thus, $\boldsymbol{\bar{z}}$ is the average annual flow of income, $\boldsymbol{\sigma}$ is the annual standard deviation of balance-sheet shocks, $\boldsymbol{R}$ is the annual interest rate, $\boldsymbol{\delta}$ is the annual exponential discount factor, $\boldsymbol{\beta}$ is the short-run discount factor that discounts all utility experienced "later" (which here means from one year onward), $\boldsymbol{u_B}$ is the annual utility flow accrued in bankruptcy, and $\boldsymbol{u}(\cdot)$ is the utility function defined over an annualized consumption level.

Generalized Calibration for Arbitrary $\triangle$. We now specify how to recalibrate the model for any $\triangle$ so that it remains consistent with the benchmark 1-model:

$$\bar{z} = \triangle\, \bar{\boldsymbol{z}} \tag{6}$$

$$\sigma = \sqrt{\triangle}\, \boldsymbol{\sigma} \tag{7}$$

$$R = \boldsymbol{R}^{\triangle} \tag{8}$$

$$\delta = \boldsymbol{\delta}^{\triangle} \tag{9}$$

$$\beta = \boldsymbol{\beta} \tag{10}$$

$$u(c) = \triangle \boldsymbol{u}\left(\frac{c}{\triangle}\right) \tag{11}$$

$$u_B = \triangle \boldsymbol{u_B}. \tag{12}$$

In equation (6), per-period income $\bar{z}$ scales in proportion to $\triangle$ (e.g., when $\triangle = \frac{1}{52}$ then the consumer earns $\frac{1}{52}$ of their average annual salary each week). In (7), the per-period standard deviation of shocks scales in proportion to $\sqrt{\triangle}$. This ensures that the variance of shocks remains constant when aggregated to the benchmark $\triangle = 1$ frequency. Specifically, it ensures that when $\triangle \neq 1$, the variance of cumulated shocks over $1/\triangle$ periods equals the single-period variance of shocks in the benchmark 1-model.[19] Intuitively, the property that $\bar{z}$ scales with $\triangle$ while σ scales with $\sqrt{\triangle}$ means that period-by-period noise dominates period-by-period consumption as the time-step $\triangle$ shrinks. As discussed in Section 5, this property is key for our sweet-spot result.

Setting $R = \boldsymbol{R}^{\triangle}$ ensures that a stock of wealth saved for the same temporal length earns the same return regardless of $\triangle$. Similarly, setting $\delta = \boldsymbol{\delta}^{\triangle}$ ensures that the discounted value of time-dated utility flows is not altered by $\triangle$.[20]

[19] Since shocks are independent, the variance of cumulated shocks in the $\triangle$-model over $1/\triangle$ periods is $\sum_1^{\frac{1}{\triangle}} \triangle\sigma^2 = \sigma^2$ (assuming that $\frac{1}{\triangle}$ is an integer).

[20] For example, when $\triangle = 1$ then a util earned one year from now is discounted by $\boldsymbol{\beta\delta}$. When $\triangle = \frac{1}{52}$, a util earned one year from now (but 52 periods from now), is discounted by $\boldsymbol{\beta}\left(\boldsymbol{\delta}^{\frac{1}{52}}\right)^{52} = \boldsymbol{\beta\delta}$.

In the standard model of present bias, short-run discount factor β is independent of time-step $\triangle$. Parameter β governs how the agent discounts "the future" relative to "the present." $\triangle$ governs the temporal boundary between the present and the future. However, $\triangle$ does not affect how the agent discounts the future relative to the present. So, $\beta = \boldsymbol{\beta}$.

In equation (11), our $\triangle$-dependent utility function explicitly defines utility as being based on consumption *per unit of time*. This is why utility is defined on $c/\triangle$ (recall that c denotes a consumption level). Then, this measure of utility per unit of time is scaled by the time-step $\triangle$ over which it is realized. Accordingly, equation (12) scales the annual bankruptcy utility $\boldsymbol{u_B}$ by the time-step $\triangle$ over which it is realized.

Discrete-Time Model Definition. The consumption-saving model presented above can be condensed into the following definition.

Definition 1. *A discrete-time consumption model is defined by the following set of seven parameters and one utility function:*

$$\{\triangle, \overline{\boldsymbol{z}}, \boldsymbol{\sigma}, \boldsymbol{R}, \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{u_B}, \boldsymbol{u}(c)\}.$$

These seven parameters and utility function, along with equations (1)–(12)*, describe the discrete-time consumption model with flexible* $\triangle$*.*

The discrete-time model is defined based on the $\triangle = 1$ calibration of $\bar{z}$, σ, R, δ, β, u_B, and $u(c)$. As $\triangle$ varies, so do these seven inputs as defined by equations (6) – (12). Again, we emphasize that the model presented in Section 2.1 is independent of the time-step. As $\triangle$ varies, all that changes is the model's calibration.

The explanation of $R = \boldsymbol{R}^{\triangle}$ is similar.

2.3 Technical Assumptions

For the remainder of this paper we impose the following set of technical assumptions. First, the utility function $u(c)$ has constant relative risk aversion (CRRA):

$$u(c) = \begin{cases} \triangle \frac{(c/\triangle)^{1-\rho}-1}{1-\rho} & \text{if } \rho \neq 1 \text{ and } c \geq 0 \\ \triangle \ln(c/\triangle) & \text{if } \rho = 1 \text{ and } c \geq 0 \cdot \\ -\infty & \text{if } c < 0 \end{cases} \tag{13}$$

The assumption of CRRA utility is common in consumption-saving models. Additionally, the continuous-time model outlined below relies on CRRA utility.[21]

Definition 1 can now be rewritten with the following set of eight parameters:

$$\{\triangle, \overline{\boldsymbol{z}}, \boldsymbol{\sigma}, \boldsymbol{R}, \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{u_B}, \boldsymbol{\rho}\},$$

where arbitrary utility function $\boldsymbol{u}(c)$ is replaced by the CRRA parameter $\boldsymbol{\rho}$.

Along with CRRA utility, we impose the following parameter restrictions for the remainder of the paper:

$$1 > \boldsymbol{\delta}\boldsymbol{R}^{1-\boldsymbol{\rho}}, \tag{14}$$

$$\boldsymbol{\rho} > 1 - \boldsymbol{\beta}, \tag{15}$$

$$\boldsymbol{\sigma} > 0, \tag{16}$$

$$-\infty < \boldsymbol{u_B} \leq \boldsymbol{u}(c) \text{ for some } c > 0. \tag{17}$$

Condition (14) ensures that every possible path of consumption leads to a present value of utility that is well-defined. Equation (15) ensures that dynamic inconsis-

[21] This is not strictly true, and the appendix of Harris and Laibson (2013) discusses generalizations of the CRRA assumption. We do not believe that our computational results depend on the CRRA special case, but this should be verified. An analysis of non-CRRA utility is beyond the scope of the current paper.

tency is not "too large."[22] Condition (16) implies that the dynamic budget constraint features noise. As will be detailed below, our $\triangle \to 0$ method for establishing a sweet spot relies on noise. The inequalities in equation (17) imply that bankruptcy is not infinitely aversive, but is still sufficiently aversive that it is not optimal to declare bankruptcy at all levels of cash on hand.[23]

Implications for the Bankruptcy Decision. Under these technical assumptions the optimal bankruptcy decision will be a threshold rule. There will exist a value $\underline{x}_t^* \geq 0$ such that bankruptcy is declared for all $x \leq \underline{x}_t^*$ and continuation is chosen for $x > \underline{x}_t^*$. Equivalently, $\mathcal{B}_t^* = (-\infty, \underline{x}_t^*]$. The existence of a threshold rule follows from the fact that the bankruptcy payoff $u_B + \beta\delta V_B$ is independent of x_t while $W_t(x)$ is continuous and increasing in x.

2.4 A Discussion of Consumption Pathologies

A detailed analysis of the consumption pathologies that can arise in discrete-time models can be found in Harris and Laibson (2003).[24] For intuition, consider self t's consumption choice (equation (2)). As is standard, the intertemporal consumption-saving decision is characterized by the trade-off of either consuming more today in order to increase utility $u(c_t)$, or passing more wealth to future selves in order to increase expected continuation value $\mathbb{E}_t V_{t+1}(x_{t+1})$. But, this trade-off is complicated by time-inconsistency. Given that self t and self $t+1$ disagree about the optimal consumption of self $t+1$, the increase in continuation value that self t earns from additional savings depends on the consumption choice of self $t+1$. This creates an incentive for self t to strategically manipulate their level of savings in order to "jump" to parts of the state space where self $t+1$ is expected to overconsume by less. Self t's

[22] As in Harris and Laibson (2013), intuitively this condition ensures that the agent's desire to consume immediately $(1-\beta)$ is less than the desire to smooth consumption (ρ).

[23] The restriction $-\infty < \boldsymbol{u_B}$ is necessary to make equilibrium well-defined, since when $\sigma > 0$ there is always a chance that a negative shock pushes $x_{t+1} < 0$.

[24] See also Krusell and Smith Jr. (2003), Chatterjee and Eyigungor (2016), and Cao and Werning (2018).

choice of consumption is strategic, and is picked with the understanding that cutting consumption today may be optimal if it jumps future selves to cash on hand levels featuring higher future saving rates.

Building on this intuition, an important feature of this sort of strategic consumption behavior is that it also begets further strategic behavior by other selves. This is why, for example, the consumption functions for $\triangle = 1$ and $\triangle = \frac{1}{2}$ in Figure 1 feature smaller non-monotonic "waves" at first, and then larger downward discontinuities at higher levels of cash on hand as those initial waves propagate upward.

3 The *dt*-Model: Passing to the Continuous-Time Limit

Having outlined our discrete-time consumption model with a flexible $\triangle$, we now pass to the continuous-time limit (i.e., the *dt*-model). This section follows Harris and Laibson (2013, henceforth HL13), making changes where necessary but avoiding details where repetitious. As in HL13, we refer to the present-biased agent in the *dt*-model as an IG agent, where IG stands for *Instantaneous Gratification*. We adopt this terminology because in the $\triangle \to 0$ limit of the $\triangle$-model, the current instantaneous self discounts all future selves by factor $\boldsymbol{\beta}$.

By making the psychological "present" vanishingly short, IG preferences are a convenient mathematical construct rather than a realistic representation of intertemporal decision-making. Nonetheless, the experimental literature estimates sharp discounting over horizons of less than one week (e.g., McClure et al., 2007; Augenblick, 2018; Augenblick and Rabin, 2019).[25] Accordingly, a contribution of this paper relative to HL13 is to validate their IG specification by demonstrating that it generates policy functions that closely approximate those of discrete-time models in which the duration of the psychological "present" is empirically realistic. This is an important property, since a key benefit of IG preferences is that their tractability allows us to prove that the consumption function of the *dt*-model is unique, continuous, and differentiable.

[25] For related discussions, see DellaVigna (2018) and Gottlieb and Zhang (2021).

The remainder of this section formalizes these results, and can be skipped without loss of continuity by readers not familiar with continuous-time methods.

3.1 The Continuous-Time Model

We start from a discrete-time model characterized by parameters $\{\triangle, \overline{\boldsymbol{z}}, \boldsymbol{\sigma}, \boldsymbol{R}, \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{u_B}, \boldsymbol{\rho}\}$. This section studies the model that results after taking $\triangle \to 0$ and passing to the continuous-time limit.[26]

Dynamic Budget Constraint. Cash on hand evolves according to:

$$dx_t = (\mu x_t + \overline{\boldsymbol{z}} - \tilde{c}_t)dt + \boldsymbol{\sigma} db_t. \tag{18}$$

The interest rate is given by $\mu = \ln(\boldsymbol{R})$, $\tilde{c}_t$ is the consumption flow (a rate), $\overline{\boldsymbol{z}}$ is the deterministic part of the income flow, $\boldsymbol{\sigma}$ is a scaling term, and b_t is a standard Brownian motion. As in discrete time, the dynamic budget constraint in equation (18) features a stochastic component ($\boldsymbol{\sigma} db_t$) to encapsulate the wide-ranging set of shocks – such as income and asset return shocks – that households face.

An important feature of continuous time is that the liquidity constraint will never bind in the interior of the wealth space. When $x > 0$, any finite rate of consumption is attainable so long as it persists for a short enough period of time.

Utility and Value. In continuous time, modeling is implemented with consumption rates rather than consumption levels. Flow utility is defined as follows (recall that we use tilde notation for rates):

$$\tilde{u}(\tilde{c}) = \begin{cases} \frac{\tilde{c}^{1-\boldsymbol{\rho}}-1}{1-\boldsymbol{\rho}} & \text{if } \boldsymbol{\rho} \neq 1 \text{ and } \tilde{c} \geq 0 \\ \ln(\tilde{c}) & \text{if } \boldsymbol{\rho} = 1 \text{ and } \tilde{c} \geq 0 \\ -\infty & \text{if } \tilde{c} < 0 \end{cases} \cdot \tag{19}$$

[26] Appendix A.3 heuristically derives the limiting dt-model. See HL13 for a formal argument.

The IG agent accumulates utils through consumption. Self t's continuation-value function is defined as follows:

$$v_t = \mathbb{E}_t \int_t^H e^{-\gamma(s-t)} \tilde{u}(\tilde{c}_s) ds + e^{-\gamma(H-t)} v_B.$$

We use H to denote the time at which bankruptcy is declared (i.e., the first-hitting time). $\gamma = -\ln(\boldsymbol{\delta})$ is the discount rate, and $v_B = \frac{\boldsymbol{u_B}}{\gamma}$ is the exponentially discounted value of bankruptcy.

Self t's current-value function is given by:

$$w_t = \boldsymbol{\beta} v_t.$$

As in discrete time, the IG agent discounts all future selves by $\boldsymbol{\beta}$. Unlike discrete time, in continuous time the current self lives for a single instant. Consumption of the current instantaneous self has no measurable impact on the overall value function, and therefore $w_t = \boldsymbol{\beta} v_t$.[27]

Equilibrium. In continuous time we study stationary Markov-perfect equilibria directly. Cash on hand x is the single state variable. Note that this equilibrium concept differs from the discrete-time model. In general, we would like to study stationary Markov-perfect equilibria. But, as discussed in Section 2.1, practical issues prevent us from always identifying such equilibria in discrete time. Alternatively, in continuous time there exist methods that allow us to characterize and solve for the unique stationary Markov-perfect equilibrium.

We start by describing the bankruptcy decision. In the dt-model, the equilibrium bankruptcy region is $\boldsymbol{\mathcal{B}}^* = (-\infty, 0]$. Because the liquidity constraint never binds in

[27]Note that despite each self living for a single instant, the IG specification nonetheless preserves dynamic inconsistency. Specifically, equation (21) below implies that each self's consumption is determined by $\tilde{u}'(\tilde{c}) = \boldsymbol{\beta} v'$. However, this consumption choice does not maximize v, which would be maximized if future selves were to (counterfactually) set $\tilde{u}'(\tilde{c}) = v'$. See Harris and Laibson (2013) for details.

the interior of the wealth space, the IG agent will never declare bankruptcy when $x > 0$. Alternatively, the utility function in equation (19) implies that bankruptcy will be declared whenever $x < 0$. Thus, $v(x) = v_B = \frac{\boldsymbol{u_B}}{\gamma}$ for all $x < 0$. A consequence of Brownian motion is that value matching holds at $x = 0$, giving $v(0) = v_B$. As in discrete time, we assume that bankruptcy is declared at points of indifference.

We are now prepared to define an equilibrium. A stationary Markov-perfect equilibrium is characterized by the following Bellman equation for the IG agent. This consists of a differential equation and an optimality condition for $x > 0$ (the continuation region), as well as the terminal bankruptcy payoff for $x \leq 0$.[28] Suppressing v's reliance on x for notational simplicity:

$$\gamma v = \tilde{u}(\tilde{c}) + (\mu x + \overline{\boldsymbol{z}} - \tilde{c})v' + \frac{1}{2}\boldsymbol{\sigma}^2 v'', \tag{20}$$

$$\tilde{u}'(\tilde{c}) = \boldsymbol{\beta} v', \tag{21}$$

$$v = v_B \text{ for all } x \leq 0. \tag{22}$$

Equation (21) defines the IG agent's consumption. Intuitively, the IG agent chooses consumption to equate the marginal utility of consumption, $\tilde{u}'(\tilde{c})$, with the marginal value of current wealth, $w' = \boldsymbol{\beta} v'$. Importantly, the existence of the additional discount factor $\boldsymbol{\beta}$ in equation (21) is a marker of dynamic inconsistency, and it implies that consumption is not chosen to maximize v (see also footnote 27 above).

Continuous-Time Model Definition. We now define the limiting dt-model.

Definition 2. *Given a discrete-time model* $\{\triangle, \overline{\boldsymbol{z}}, \boldsymbol{\sigma}, \boldsymbol{R}, \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{u_B}, \boldsymbol{\rho}\}$*, the associated dt-model is defined by equations* (19) – (22).

[28] The bankruptcy boundary condition implies that v has a convex kink at $x = 0$. Hence, we look for a viscosity solution to v.

3.2 Uniqueness in the dt-Model

To prove equilibrium uniqueness, we follow HL13 by introducing a mathematically convenient exponential agent with a reverse-engineered utility function of $\hat{u}$. We refer to the agent with utility function $\hat{u}$ as the $\hat{u}$ agent. $\hat{u}$ is reverse-engineered such that the value function of the $\hat{u}$ agent, denoted $\hat{v}$, is identical to the value function v of the IG agent. The modified utility function is:

$$\hat{u}(\hat{c}) = \frac{\psi}{\boldsymbol{\beta}}\tilde{u}\left(\frac{1}{\psi}\hat{c}\right) + \frac{\psi - 1}{\boldsymbol{\beta}}, \quad \text{where} \quad \psi = \frac{\boldsymbol{\rho} - (1 - \boldsymbol{\beta})}{\boldsymbol{\rho}}. \tag{23}$$

Since $\hat{u}(\hat{c}) < \tilde{u}(\hat{c})$ for any $\hat{c} > 0$, the $\hat{u}$ utility function penalizes the consumption of the $\hat{u}$ agent relative to the IG agent. This is because the $\hat{u}$ agent is time-consistent and chooses $\hat{c}$ to maximize $\hat{v}$, while the IG agent is time-inconsistent and does not choose $\tilde{c}$ to maximize v. Thus, the $\hat{u}$ utility function distorts the utility flows of the $\hat{u}$ agent downward to ensure that $\hat{v}(x) = v(x)$ for all $x > 0$.[29]

The Bellman equation for the $\hat{u}$ agent is defined by the following differential equation and optimality condition for $x > 0$, as well as the terminal bankruptcy payoff for $x \leq 0$:

$$\gamma\hat{v} = \hat{u}(\hat{c}) + (\mu x + \bar{\boldsymbol{z}} - \hat{c})\hat{v}' + \frac{1}{2}\boldsymbol{\sigma}^2\hat{v}'', \tag{24}$$

$$\hat{u}'(\hat{c}) = \hat{v}', \tag{25}$$

$$\hat{v} = v_B \text{ for all } x \leq 0. \tag{26}$$

The key difference between the Bellman equation for the IG agent and the $\hat{u}$ agent is the consumption choice (equations (21) and (25)). Both agents equate the marginal utility of consumption with the marginal value of current wealth, but only the time-inconsistent IG agent discounts all future selves by $\boldsymbol{\beta}$.

[29] In HL13 the $\hat{u}$ utility function is defined separately at $x = 0$ and $x > 0$. Here, the $x = 0$ boundary is included in the stopping region. Thus, the dt-model does not require a wealth-dependent definition of $\hat{u}$. Instead, a Dirichlet boundary condition is imposed at $x = 0$.

Proposition 1 (Value Function Equivalence)**.** *v is the value function of the IG agent if and only if v is the value function of the $\hat{u}$ agent.*

Proof. See Appendix A.1. □

Proposition 2 (Uniqueness)**.** *The dt-model has a unique equilibrium.*

Proof. See HL13 for full details. The intuition is given here. Optimization problems have a unique value function. The $\hat{u}$ agent is an exponential discounter who chooses consumption optimally to maximize $\hat{v}$. Therefore $\hat{v}$ is unique. Brownian motion makes $\hat{v}$ twice continuously differentiable on $(0, \infty)$. By value function equivalence (Proposition 1), v must also be unique and twice continuously differentiable on $(0, \infty)$. Using equation (21), continuous differentiability of v implies that the consumption function of the IG agent will be unique for all $x > 0$. □

3.3 IG Consumption

Value function equivalence can be used to link the IG agent's consumption function to that of the $\hat{u}$ agent. This is formalized below.

Corollary 1 (IG Consumption)**.** *The IG agent's consumption function can be characterized relative to the $\hat{u}$ agent's consumption function as follows:*

$$\tilde{c}(x) = \frac{1}{\psi}\hat{c}(x). \tag{27}$$

Proof. This follows from equations (21) and (25), using Proposition 1 to set $v = \hat{v}$. □

Since $\psi < 1$ when $\boldsymbol{\beta} < 1$, equation (27) gives the intuitive property that the IG agent chooses a higher rate of consumption than the exponential $\hat{u}$ agent. Additionally, we can characterize certain properties of the IG agent's consumption function.

Proposition 3 (Consumption in the *dt*-Model)**.**

(i) $\tilde{c}$ is continuously differentiable on the interior of the wealth space.

(ii) If $\mu \leq \gamma$ then $\tilde{c}$ is increasing on $(0, \infty)$.

(iii) If $\mu > \gamma$ then $\tilde{c}$ may be non-monotonic. Specifically, $\tilde{c}$ may be decreasing immediately to the right of the origin, but will eventually increase and remain increasing forever thereafter.

Proof. See Appendix A.2. □

We do not formally characterize the combination of parameters that causes consumption to be non-monotonic when $\mu > \gamma$. However, we find numerically that non-monotonicity tends to occur when $\boldsymbol{u_B}$ is high. To understand this, note that $\tilde{c}(x)$ is decreasing if and only if $v''(x) > 0$. From value matching, $\lim_{x \to {}^{+}0} v(x) = v_B = \frac{\boldsymbol{u_B}}{\gamma}$. If $\boldsymbol{u_B}$ is low and the stopping region is aversive, then v will be concave because the possibility of reaching the stopping region pulls down v near $x = 0$. Alternatively, if the stopping region is attractive then v may be convex because the possibility of reaching the stopping region props up v near $x = 0$. Non-monotonic consumption can also arise in models with exponential discounting ($\boldsymbol{\beta} = 1$) as long as $\boldsymbol{u_B}$ is large enough.[30] Appendix D provides a numerical example of non-monotonic consumption.

4 Numerical Methods and Calibration

Having presented our consumption-saving model allowing for full flexibility of the time-step $\triangle$, we now demonstrate why this flexibility is important. We proceed to solve our model numerically to show that consumption can be poorly behaved and non-robust for large time-steps, but there exists a $\triangle$ sweet spot over which consumption functions become quantitatively comparable.

4.1 Numerical Methods

Discrete Time. Since the Bellman equation of the present-biased agent (equations (2) – (5)) is not a contraction mapping, iterative methods may not converge to a

[30] This follows immediately from the $\hat{u}$ construction and equation (27).

stationary equilibrium. If they do, this equilibrium may not be unique. These issues commonly arise in the literature on present-biased preferences (with at least partial sophistication). In response, a goal of our sweet-spot analysis is to show that there exists a range of time-steps over which the model is quantitatively robust to these issues.

To solve the discrete-time model, we use a multigrid implementation of value function iteration (Chow and Tsitsiklis, 1991; Caldara et al., 2012). Value function iteration (VFI) is a standard technique for solving Bellman equations. Multigrid allows us to solve our model on a dense grid with relatively shorter runtimes. This is important as $\triangle \to 0$, because VFI methods can be slow to converge when $\delta \approx 1$.

In the main text we focus on calibrations with $\delta R < 1$, as our algorithm usually converges to a stationary policy function in these cases.[31] For the discrete-time model with $\delta R > 1$, our numerical methods almost never converge nor even cycle. Indeed, for these $\delta R > 1$ cases the consumption function can change quite drastically from one iteration to the next. We find that this failure of convergence occurs when the equilibrium supports expected asset accumulation. Due to additional complexities regarding numerical methods, which are not the focus of this paper, calibrations with $\delta R > 1$ are discussed in Appendix D.[32] Nevertheless, Appendix D shows that even in these $\delta R > 1$ cases, a sweet spot still emerges as $\triangle \to 0$.

Continuous Time. In contrast to discrete time, the continuous-time equilibrium is unique and there exists a well-developed theory on the numerical methods for characterizing it. The continuous-time equilibrium is computed using finite-difference

[31] In the $\delta R < 1$ cases where our algorithm fails to converge to a stationary solution, VFI enters a phase in which the resulting policy functions fluctuate slightly at a small number of grid points. These small fluctuations are inconsequential quantitatively, so, once we enter this phase, we stop the backward induction and report an arbitrarily chosen policy function. To reflect this approximate convergence, the consumption function that we select will be denoted $C(x)$ rather than $C_t(x)$.

[32] Specifically, equilibria that support expected asset accumulation create issues at the upper boundary of a finite numerical grid. This can create a feedback effect in which numerical error at the top of the grid trickles down and affects the consumption function throughout the state space, especially in cases where consumption is non-robust in the first place.

methods (Candler, 2001; Achdou et al., 2022). Barles and Souganidis (1991) prove that finite-difference methods converge to the unique viscosity solution of an HJB equation when certain conditions are satisfied. However, these conditions are violated by the Bellman equation of the dynamically-inconsistent IG agent (Maxted, 2022). To compute the IG agent's equilibrium, we instead solve for the time-consistent $\hat{u}$ agent's equilibrium. The IG agent's consumption function then follows from equation (27). Appendix C provides further details on our numerical methods.

4.2 Calibration

Recall that the discrete-time consumption model is comprised of the eight exogenous parameters $\{\triangle, \overline{\boldsymbol{z}}, \boldsymbol{\sigma}, \boldsymbol{R}, \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{u_B}, \boldsymbol{\rho}\}$. For the remainder of this paper, our baseline calibration when $\triangle = 1$ is as follows:

$$\overline{\boldsymbol{z}} = 1,\ \boldsymbol{\sigma} = \frac{1}{3},\ \boldsymbol{R} = 1,\ \boldsymbol{\delta} = 0.95,\ \boldsymbol{u_B} = \boldsymbol{u}\left(\frac{1}{10}\overline{\boldsymbol{z}}\right),\ \boldsymbol{\rho} = 1.$$

Robustness to these parameter choices is explored in Section 6.

$\boldsymbol{\beta}$ does not have a baseline calibration since this paper studies a variety of $\boldsymbol{\beta}$ values. However, most of our results use $\boldsymbol{\beta} = 0.5$. We choose $\boldsymbol{\beta} = 0.5$ for two reasons. First, Laibson et al. (2023) estimate $\boldsymbol{\beta}$ to be approximately 0.5.[33] Second, $\boldsymbol{\beta} = 0.5$ consumption functions are poorly behaved for large $\triangle$ (see Figure 1), generating a good test case for our $\triangle \to 0$ technique.

$\boldsymbol{R} = 1$ in the baseline calibration for simplicity.[34] $\boldsymbol{\delta} = 0.95$ is consistent with typical values in the literature (e.g., Kaplan and Violante, 2014; Kaplan et al., 2018). We set $\boldsymbol{\sigma} = \frac{1}{3}\overline{\boldsymbol{z}}$ to conservatively capture annual balance-sheet shocks.[35]

[33] Laibson et al. (2023) assume naive present bias while this paper assumes sophistication. An extension to naivete is presented in Section 7.2.

[34] Setting $\boldsymbol{R} = 1$ circumvents minor issues regarding differences in the present value of income as $\triangle$ is varied.

[35] Looking at persistent income shocks alone, the commonly used wage process estimated in Floden and Lindé (2001) features annual persistent log-wage volatility of 0.21. Households also face many other types of shocks in addition, such as transitory income shocks, asset return shocks, unexpected medical bills, taste shocks, etc.

Bankruptcy utility $\boldsymbol{u_B}$ is set to equal the utility accrued from consuming one-tenth of average income. This illustrative calibration makes bankruptcy highly aversive.[36]

5 Results: Consumption and the $\triangle$ Sweet Spot

Figure 1 in the introduction above plots the $\boldsymbol{\beta} = 0.5$ consumption function for $\triangle = 1$, $\triangle = \frac{1}{2}$, $\triangle = \frac{1}{25}$, and $\triangle = dt$ (continuous time). For the larger time-steps of $\triangle = 1$ and $\triangle = \frac{1}{2}$, Figure 1 shows the non-robustness that can occur.[37] However, for the $\triangle = \frac{1}{25}$ and $\triangle = dt$ time-steps that are closer to the psychologically relevant range, Figure 1 also highlights the emergence of the time-step sweet spot.

We now quantify the sweet spot over which consumption functions closely replicate one another. We take the dt-model as our benchmark and study the numerical convergence of the $\triangle$-model to the dt-model as $\triangle \to 0$. This section shows that $\triangle = dt$ provides a tractable approximation of the $\triangle$-model with psychologically appropriate time-steps. This is a helpful property, because the continuous-time IG specification is an appealing choice both analytically and numerically.[38]

We use a mean squared error metric to evaluate the distance between the consumption function of the $\triangle$-model and the dt-model. Our goal from a practical standpoint is to rule out the large and rapid variations in the consumption function that can make present-biased preferences non-robust and hence difficult for applied researchers to use. Squared error disproportionately punishes large deviations from the well-behaved dt-model, which is precisely our goal.

[36] Our setup nests other assumptions such as a reflecting barrier at $x = 0$. Though such barrier choices affect the resulting distribution of agents, this paper focuses only on policy functions.

[37] For similar results, see e.g. Harris and Laibson (2001, 2003), Krusell and Smith Jr. (2003), Chatterjee and Eyigungor (2016), and Morris and Postlewaite (2020).

[38] Analytically, continuous time allows us to characterize properties of the consumption function (Propositions 2 and 3). See also Laibson et al. (2021) and Maxted (2022) for further applications that rely on the tractability of IG preferences. Numerically, finite-difference methods are typically fast, particularly in comparison to discrete-time models with short time-steps (Achdou et al., 2022).

Define:

$$\triangle\text{-Model Error (DME)} = \frac{\int_{\underline{x}^*}^{20} \left(\tilde{C}(x) - \tilde{c}(x)\right)^2 dx}{20 - \underline{x}^*}. \tag{28}$$

Recall that the tilde notation denotes rates. The DME is defined in terms of consumption rates so that it is comparable across different $\triangle$'s. $\tilde{C}$ denotes the rate of consumption in the discrete-time model, and $\tilde{c}$ denotes the rate of consumption in the continuous-time model. We restrict the analysis to a bounded interval $(\underline{x}^*, 20]$. Because the model features only additive noise we expect pathologies to exist at large enough wealth levels for any $\triangle > 0$.[39]

Table 1 shows the DME as a function of the time-step $\triangle$ for $\boldsymbol{\beta} \in \{1, 0.5, 0.3\}$. Table 1 exhibits substantial deviations in consumption for large $\triangle$ when $\boldsymbol{\beta} \in \{0.5, 0.3\}$. For example, when $\boldsymbol{\beta} = 0.3$ the DME reaches almost 8 times $\overline{\boldsymbol{z}}$. However, our DME metric decreases drastically for small $\triangle$. This indicates the time-step sweet spot over which consumption functions are quantitatively homogeneous. Table 1 also shows that the precise upper boundary of the sweet spot is calibration-dependent; e.g., the sweet spot roughly emerges for monthly time-steps when $\boldsymbol{\beta} = 0.5$, and weekly time-steps when $\boldsymbol{\beta} = 0.3$ (well below most empirical estimates).

DME	$\boldsymbol{\beta} = 1$	$\boldsymbol{\beta} = 0.5$	$\boldsymbol{\beta} = 0.3$
$\triangle = 1$	0.0075	0.3189	2.5223
$\triangle = 1/5$	0.0003	0.0502	7.7697
$\triangle = 1/10$	0.0001	0.0047	3.8756
$\triangle = 1/25$	0.0001	0.0007	0.8111
$\triangle = 1/50$	0.0001	0.0003	0.0399

Table 1: DME values for $\boldsymbol{\beta} \in \{1, 0.5, 0.3\}$ and $\triangle \in \{1, \frac{1}{5}, \frac{1}{10}, \frac{1}{25}, \frac{1}{50}\}$.

Another insight from Table 1 is that in a classical world of $\boldsymbol{\beta} = 1$ the discrete-

[39] As the agent gets wealthier the effect of additive noise decreases. This makes future consumption more predictable and counterfactual pathologies more likely. Shrinking $\triangle$ to zero expands the interval on which consumption is well-behaved. In the continuous-time limit, Proposition 3 proves that consumption is continuous on the entire domain $(0, \infty)$. The addition of stochastic asset returns, which are multiplicative risks that scale with wealth, should also diminish this effect.

time and continuous-time models closely approximate one another, even for annual time-steps. This is perhaps unsurprising, and is likely the reason that the choice of time-step has thus far received little attention in the household finance literature. However, this property breaks down under present bias, where large time-steps can lead to non-robust properties that do not reflect the behavior of models with time-steps that are in the psychologically relevant range.

Intuition: The Importance of Period-by-Period Shocks. We now discuss the intuition for why a sweet spot emerges when time-steps are brought closer to the psychologically relevant range. For non-robust consumption pathologies like counterfactual downward discontinuities to arise, the current self must be able to both predict and control the wealth of future selves. However, as $\triangle \to 0$ the ability to manipulate the wealth of future selves becomes overwhelmed by the high-frequency noise in the system.

To see this effect, recall the dynamic budget constraint given by equation (1):

$$x_{t+1} = R(x_t - c_t) + \overline{z} + \sigma\varepsilon_{t+1}.$$

Self t's ability to predict the wealth of self $t+1$ depends on σ. As $\triangle \to 0$ the noise in the model (σ) decreases in proportion to $\sqrt{\triangle}$. Self t controls the wealth of future selves through the choice of consumption level c_t. As $\triangle \to 0$ consumption decreases approximately in proportion to $\triangle$. Thus, as $\triangle \to 0$ the noise in the model dominates each self's ability to control the wealth of future selves.[40]

Though it is well known that pathologies can fade away when noise is high (Harris and Laibson, 2003), in a calibrated model one cannot arbitrarily increase noise. This paper's insight is the recognition that unrealistically large time-steps provide implicit

[40] To check this intuition, Appendix Figure 5 studies what would happen in our $\triangle$-model if we were to set noise parameter σ equal to $\triangle\boldsymbol{\sigma}$ rather than $\sqrt{\triangle}\boldsymbol{\sigma}$. This alternate scaling keeps the ratio $\overline{z}/\sigma$ constant as $\triangle \to 0$. Appendix Figure 5 illustrates that without the requisite noise, consumption behavior remains pathological even for shorter time-steps.

diversification. Large time-steps thus deflate the true level of high-frequency noise that consumers face, which spuriously allows for non-robust and highly strategic consumption behavior to emerge. Reducing $\triangle$ to a length that is more psychologically appropriate undoes this illegitimate diversification, and allows for robust predictions to emerge.[41]

This discussion highlights the importance of period-by-period noise for generating robust policy functions as $\triangle \to 0$. To formalize this point, in Appendix B we study a deterministic economy and construct (a continuum of) "sawtooth equilibria" that contain an arbitrary number of policy-function discontinuities as $\triangle \to 0$ over any finite interval of cash on hand. This deterministic counter-example shows that noise is a necessary assumption for our results.

6 Robustness and Comparative Statics

To show the robustness of our sweet-spot result, we fix $\boldsymbol{\beta} = 0.5$ and repeat the analysis in Table 1 under alternate calibrations. Results are provided in Table 2. For comparison, the first column of Table 2 lists the DME for $\boldsymbol{\beta} = 0.5$ under the baseline calibration. Table 2 illustrates that the upper boundary of the sweet spot is somewhat calibration-dependent. However, conditional on a calibration, our $\triangle \to 0$ sweet-spot result continues to apply.

First we consider $\boldsymbol{\rho} \in \{0.75, 2\}$. On the low end we set $\boldsymbol{\rho} = 0.75$ because this is approximately the limit of where we can push our method for $\triangle \geq \frac{1}{50}$.[42] On the high end we double $\boldsymbol{\rho}$ from its baseline calibration. The DME is decreasing in $\boldsymbol{\rho}$. As CRRA parameter $\boldsymbol{\rho}$ increases the agent becomes less willing to engage in strategic behavior because this forces deviations from a smooth path of consumption.

Next we study robustness to the volatility of balance-sheet shocks, both halving noise to $\boldsymbol{\sigma} = \frac{1}{6}$ and doubling it to $\boldsymbol{\sigma} = \frac{2}{3}$. As discussed in Section 5, pathologies are

41 See also Pagel (2018) for a related result that time diversification makes portfolio allocations sensitive to $\triangle$ when consumers have loss-averse utility over news.

42 For $\boldsymbol{\beta} = 0.5$, restriction (15) requires $\boldsymbol{\rho} > 0.5$.

reduced as $\boldsymbol{\sigma}$ increases. We also analyze robustness to the assumption that balance-sheet shocks are Gaussian. Table 2 lists the DME under the alternate calibration that such shocks are uniformly distributed, with $\varepsilon \sim \mathcal{U}(-\sqrt{3}\sigma, \sqrt{3}\sigma)$. The uniform distribution creates explicit cutoffs and rules out large shocks, both of which increase the ability for self t to predict and strategically manipulate the consumption of the next self. This increases the DME relative to the baseline.

Our baseline assumption that $\boldsymbol{u_B} = \boldsymbol{u}\left(\frac{1}{10}\overline{\boldsymbol{z}}\right)$ imposes a strong bankruptcy penalty. The last column makes bankruptcy more attractive by setting $\boldsymbol{u_B} = \boldsymbol{u}\left(\overline{\boldsymbol{z}}\right)$. This increases the agent's desired consumption near $\underline{x}^*$. One effect of large time-steps here is that they make the liquidity constraint bind over a larger region of the state space,[43] which then increases the DME for high values of $\boldsymbol{u_B}$ when $\triangle$ is large.

DME	Baseline	$\boldsymbol{\rho} = 0.75$	$\boldsymbol{\rho} = 2$	$\boldsymbol{\sigma} = \frac{1}{6}$	$\boldsymbol{\sigma} = \frac{2}{3}$	$\varepsilon \sim \mathcal{U}$	$\boldsymbol{u_B} = \boldsymbol{u}\left(\overline{\boldsymbol{z}}\right)$
$\triangle = 1$	0.3189	1.2289	0.0035	0.6619	0.0510	0.5045	3.5798
$\triangle = \frac{1}{10}$	0.0047	1.6543	0.0001	0.6367	0.0004	0.0125	0.5736
$\triangle = \frac{1}{25}$	0.0007	0.3063	0.0001	0.2341	0.0003	0.0017	0.0584
$\triangle = \frac{1}{50}$	0.0003	0.0097	0.0001	0.0014	0.0003	0.0007	0.0286

Table 2: DME values for $\boldsymbol{\beta} = 0.5$ and $\triangle \in \{1, \frac{1}{10}, \frac{1}{25}, \frac{1}{50}\}$ under alternate calibrations.

7 Extensions

7.1 The $\hat{u}$ Agent in Discrete Time

In the dt-model the IG agent's consumption is $\frac{1}{\psi}$ times the $\hat{u}$ agent's consumption (equation (27)). Appendix Figure 6 shows that this property also holds approximately in the discrete-time model for small $\triangle$. This approximation result has useful numerical applications. Because the $\hat{u}$ agent is an exponential discounter, their consumption will not be subject to the sorts of strategic behavior that arises when $\beta < 1$.

[43] For example, when $\triangle = 1$ the agent's consumption is bounded by cash on hand: $C(x_t) \leq x_t$. For $\triangle = \frac{1}{50}$ it is still the case that consumption level $C(x_t) \leq x_t$, but it is the consumption rate that is important. The liquidity constraint will only bind when consumption rate $\tilde{C}(x_t) = 50x_t$, a much weaker restriction.

Numerical methods that are more efficient than VFI, but rely on well-behaved policy functions, can be utilized to calculate the consumption of the $\hat{u}$ agent. For small $\triangle$, multiplying the consumption of the $\hat{u}$ agent by $\frac{1}{\psi}$ thus provides a computationally efficient approximation to the consumption of the present-biased agent.

7.2 Naivete

Up to this point we have assumed that the present-biased agent is sophisticated, meaning that they are perfectly aware of their time-inconsistency. Indeed, it is their attempts to strategically interact with future selves that create non-robust consumption behavior. An alternative to sophistication is to assume partial or complete naivete (Akerlof, 1991; O'Donoghue and Rabin, 1999, 2001). Under naivete, self t erroneously believes that all future selves will behave according to short run discount factor $\beta^E > \beta$. Partial naivete sets $\beta^E \in (\beta, 1)$ and complete naivete sets $\beta^E = 1$.

As discussed in the introduction, naivete is one method that the literature has adopted to circumvent non-robustness issues. Such issues typically do not arise under complete naivete because the agent acts as if all future selves are time consistent. However, for all but complete naivete – a strong simplification – strategic behavior can still emerge. This section extends the model to allow for naivete.

Naivete in the Discrete-Time Model. Let β denote the agent's true present bias. Let β^E denote their expected present bias in all future periods. Under naivete, equations (2) and (5) become:

$$C_t(x_t) \in \underset{c \leq x_t}{\operatorname{argmax}} \;\; u(c) + \beta\delta\mathbb{E}_t V_{t+1}^E(x_{t+1}), \tag{29}$$

$$\mathcal{B}_t^* = \{x_t | u_B + \beta\delta V_B \geq u(C_t(x_t)) + \beta\delta\mathbb{E}_t V_{t+1}^E(x_{t+1})\}. \tag{30}$$

V_t^E is the continuation-value function, characterized by equations (2) through (5), that would obtain if the agent was sophisticated with a true short-run discount factor of $\boldsymbol{\beta^E}$. Numerically, we first compute V_t^E by solving the discrete-time model for a sophisticate with short-run discount factor $\boldsymbol{\beta^E}$. Given V_t^E, one final iteration is then needed to solve for the naif's policy functions at time t (as specified by equations (29) and (30)).

Naivete in the Continuous-Time Model. As in Laibson et al. (2021) and Maxted (2022), the naive IG agent's consumption is defined by

$$\tilde{u}'(\tilde{c}) = \boldsymbol{\beta} \frac{\partial v^E}{\partial x}. \tag{31}$$

Similar to discrete time, v^E is the value function that would obtain if the IG agent had short-run discount factor $\boldsymbol{\beta^E}$ instead of $\boldsymbol{\beta}$. To numerically solve for the consumption of the naive IG agent, we first solve for v^E. The naive IG agent's consumption is then defined implicitly by equation (31).

Consumption Functions under Naivete. Table 3 examines the DME metric under different levels of naivete. Two results are apparent. First, in all columns the DME is decreasing as $\triangle \to 0$, signaling again the emergence of a sweet spot. Second, for $\triangle = 1$ the DME is large regardless of $\boldsymbol{\beta^E}$. Though the consumption function of naive agents may not feature counterfactual pathologies, when written with large time-steps it still provides a poor approximation to models with smaller time-steps.

DME	Baseline	$\boldsymbol{\beta^E} = 0.6$	$\boldsymbol{\beta^E} = 0.7$	$\boldsymbol{\beta^E} = 0.8$	$\boldsymbol{\beta^E} = 0.9$	$\boldsymbol{\beta^E} = 1$
$\triangle = 1$	0.3189	0.1248	0.2220	0.2997	0.3315	0.3389
$\triangle = \frac{1}{10}$	0.0047	0.0003	0.0014	0.0023	0.0027	0.0028
$\triangle = \frac{1}{25}$	0.0007	0.0001	0.0002	0.0003	0.0004	0.0004
$\triangle = \frac{1}{50}$	0.0003	0.0001	0.0001	0.0002	0.0002	0.0002

Table 3: DME values for $\boldsymbol{\beta} = 0.5$ and $\triangle \in \{1, \frac{1}{10}, \frac{1}{25}, \frac{1}{50}\}$ across varying naivete.

8 Conclusion

This paper identifies a "sweet spot" of time-steps over which dynamic household finance models with present-biased agents produce robust, quantitatively homogeneous, policy functions. The sweet spot spans from zero (i.e., continuous time) to roughly two weeks, thus capturing the psychologically relevant range of present-bias horizons.

For researchers studying present bias in discrete time, our results highlight the importance of the time-step, and establish a range of time-steps over which model predictions are robust. Our results also imply that researchers can leverage the tractability of the continuous-time IG specification of present bias in order to closely approximate psychologically well-calibrated models. While our paper focuses specifically on present bias, one interesting pathway for future exploration is the extent to which our sweet-spot result continues to apply across the broader class of present-focused preferences (Ericson and Laibson, 2019), including true hyperbolic discounting.

References

Acharya, Subas, David Jimenez Gomez, Dmitrii Rachinskii, and Alejandro Rivera, "Present-Bias and the Value of Sophistication," *SSRN 3722608*, 2022.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 2022, *89* (1), 45–86.

Akerlof, George A., "Procrastination and Obedience," *The American Economic Review*, 1991, *81* (2), 1–19.

Alfaro, Laura and Fabio Kanczuk, "Fiscal Rules and Sovereign Default," *NBER w23370*, 2017.

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman, "Are High-Interest Loans Predatory? Theory and Evidence from Payday Lending," *The Review of Economic Studies*, 2022, *89* (3), 1041–1084.

Amador, Manuel, Iván Werning, and George-Marios Angeletos, "Commitment vs. Flexibility," *Econometrica*, 2006, *74* (2), 365–396.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg, "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 2001, *15* (3), 47–68.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *The Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Augenblick, Ned, "Short-Term Time Discounting of Unpleasant Tasks," *Mimeo*, 2018.

_ **and Matthew Rabin**, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *The Review of Economic Studies*, 2019, *86* (3), 941–975.

Barles, Guy and Panagiotis E. Souganidis, "Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations," *Asymptotic Analysis*, 1991, *4* (3), 271–283.

Beshears, John, James Choi, David Laibson, and Peter Maxted, "Present Bias Causes and Then Dissipates Auto-Enrollment Savings Effects," *AEA Papers and Proceedings*, 2022, *112*, 136–41.

_ **, James J. Choi, Christopher Clayton, Christopher Harris, David Laibson, and Brigitte C. Madrian**, "Optimal Illiquidity," *Mimeo*, 2023.

Caldara, Dario, Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, and Wen Yao, "Computing DSGE Models with Recursive Preferences and Stochastic Volatility," *Review of Economic Dynamics*, 2012, *15* (2), 188–206.

Candler, Graham V., "Finite-Difference Methods for Continuous-Time Dynamic Programming," *Computational Methods for the Study of Dynamic Economies*, 2001.

Cao, Dan and Iván Werning, "Saving and Dissaving with Hyperbolic Discounting," *Econometrica*, 2018, *86* (3), 805–857.

Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick, "Optimal Defaults and Active Decisions," *The Quarterly Journal of Economics*, 2009, *124* (4), 1639–1674.

Chatterjee, Satyajit and Burcu Eyigungor, "Continuous Markov Equilibria with Quasi-Geometric Discounting," *Journal of Economic Theory*, 2016, *163*, 467–494.

Chow, Chee-Seng and John N. Tsitsiklis, "An Optimal One-Way Multigrid Algorithm for Discrete-Time Stochastic Control," *IEEE Transactions on Automatic Control*, 1991, *36* (8), 898–914.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White, "Measuring Time Preferences," *Journal of Economic Literature*, 2020, *58* (2), 299–347.

DellaVigna, Stefano, "Structural Behavioral Economics," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 1, Elsevier, 2018, pp. 613–723.

_ **and M. Daniele Paserman**, "Job Search and Impatience," *Journal of Labor Economics*, 2005, *23* (3), 527–588.

_ **and Ulrike Malmendier**, "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

__ **and** __ , "Paying Not to Go to the Gym," *American Economic Review*, 2006, *96* (3), 694–719.

Eaton, Jonathan and Mark Gersovitz, "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *The Review of Economic Studies*, 1981, *48* (2), 289–309.

Ericson, Keith Marzilli and David Laibson, "Intertemporal Choice," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 1–67.

Floden, Martin and Jesper Lindé, "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?," *Review of Economic Dynamics*, 2001, *4* (2), 406–437.

Ganong, Peter and Pascal Noel, "Consumer Spending During Unemployment: Positive and Normative Implications," *American Economic Review*, 2019, *109* (7), 2383–2424.

Gerard, François and Joana Naritomi, "Job Displacement Insurance and (the Lack of) Consumption-Smoothing," *American Economic Review*, 2021, *111* (3), 899–942.

Gottlieb, Daniel and Xingtan Zhang, "Long-Term Contracting With Time-Inconsistent Agents," *Econometrica*, 2021, *89* (2), 793–824.

Grenadier, Steven R. and Neng Wang, "Investment Under Uncertainty and Time-Inconsistent Preferences," *Journal of Financial Economics*, 2007, *84* (1), 2–39.

Harris, Christopher and David Laibson, "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 2001, *69* (4), 935–957.

_ **and** _ , "Hyberbolic Discounting and Consumption," *Econometric Society Monographs*, 2003, *35*, 258–297.

_ **and** _ , "Instantaneous Gratification," *The Quarterly Journal of Economics*, 2013, *128* (1), 205.

Harstad, Bård, "Technology and Time Inconsistency," *Journal of Political Economy*, 2020, *128* (7), 2653–2689.

Kaplan, Greg and Giovanni L. Violante, "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 2014, *82* (4), 1199–1239.

_ **, Benjamin Moll, and Giovanni L. Violante**, "Monetary Policy According to HANK," *American Economic Review*, 2018, *108* (3), 697–743.

Krusell, Per and Anthony A. Smith Jr., "Consumption-Savings Decisions with Quasi-Geometric Discounting," *Econometrica*, 2003, *71* (1), 365–375.

_ **, Burhanettin Kuruşçu, and Anthony A. Smith Jr.**, "Equilibrium Welfare and Government Policy with Quasi-geometric Discounting," *Journal of Economic Theory*, 2002, *105* (1), 42–72.

Kuchler, Theresa and Michaela Pagel, "Sticking to Your Plan: The Role of Present Bias for Credit Card Paydown," *Journal of Financial Economics*, 2021, *139* (2), 359–388.

Laibson, David, "Self-Control and Saving," *Mimeo*, 1994.

_ , "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

Laibson, David I., Andrea Repetto, and Jeremy Tobacman, "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity*, 1998, *1998* (1), 91–196.

Laibson, David, Peter Maxted, and Benjamin Moll, "Present Bias Amplifies the Household Balance-Sheet Channels of Macroeconomic Policy," *NBER w29094*, 2021.

_ **, Sean Chanwook Lee, Peter Maxted, Andrea Repetto, and Jeremy Tobacman**, "Estimating Discount Functions with Consumption Choices Over the Lifecycle," *NBER w13314*, 2023.

Lee, Sean Chanwook and Peter Maxted, "Credit Card Borrowing in Heterogeneous-Agent Models: Reconciling Theory and Data," *SSRN 4389878*, 2023.

Lockwood, Benjamin B., "Optimal Income Taxation with Present Bias," *American Economic Journal: Economic Policy*, 2020, *12* (4), 298–327.

Maxted, Peter, "Present Bias Unconstrained: Consumption, Welfare, and the Present-Bias Dilemma," *Mimeo*, 2022.

McClure, Samuel M., Keith M. Ericson, David I. Laibson, George Loewenstein, and Jonathan D. Cohen, "Time Discounting for Primary Rewards," *Journal of Neuroscience*, 2007, *27* (21), 5796–5804.

Meier, Stephan and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

Morris, Stephen and Andrew Postlewaite, "Observational Implications of Non-Exponential Discounting," *Revue Économique*, 2020, *71* (2), 313–322.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

_ **and** _ , "Choice and Procrastination," *The Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

Pagel, Michaela, "A News-Utility Theory for Inattention and Delegation in Portfolio Choice," *Econometrica*, 2018, *86* (2), 491–522.

Peleg, Bezalel and Menahem E. Yaari, "On the Existence of a Consistent Course of Action When Tastes are Changing," *The Review of Economic Studies*, 1973, *40* (3), 391–401.

Phelps, Edmund S. and Robert A. Pollak, "On Second-Best National Saving and Game-Equilibrium Growth," *The Review of Economic Studies*, 1968, *35* (2), 185–199.

Rivera, Alejandro, "Contracting with a Present-Biased Agent: Sannikov Meets Laibson," *SSRN 4009513*, 2022.

Shapiro, Jesse M., "Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle," *Journal of Public Economics*, 2005, *89* (2-3), 303–325.

Shigeta, Yuki, "Quasi-Hyperbolic Discounting under Recursive Utility and Consumption-Investment Decisions," *Journal of Economic Theory*, 2022, p. 105518.

Strack, Philipp and Dmitry Taubinsky, "Dynamic Preference "Reversals" and Time Inconsistency," *NBER w28961*, 2021.

Strotz, Robert H., "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 1955, *23* (3), 165–180.

ONLINE APPENDIX

A The dt-Model

A.1 Proof of Proposition 1: Value Function Equivalence

This proof is similar to Theorem 2 of Harris and Laibson (2013), and is included for complete detail.[44] Let $f(\alpha)$ be the unique value of $\tilde{c}$ satisfying $\tilde{u}'(\tilde{c}) = \alpha$. Let $h(\alpha) = \tilde{u}(f(\boldsymbol{\beta}\alpha)) - \alpha f(\boldsymbol{\beta}\alpha)$. The IG consumer sets $\tilde{u}'(\tilde{c}) = \boldsymbol{\beta}v'$, and therefore $h(v') = \tilde{u}(f(\boldsymbol{\beta}v')) - v'f(\boldsymbol{\beta}v') = \tilde{u}(\tilde{c}) - v'\tilde{c}$. We can use $h(v')$ in order to rewrite the Bellman equation of the IG agent (equations (20) – (22)) in a reduced way:

$$\gamma v = (\mu x + \bar{\boldsymbol{z}})v' + \frac{1}{2}\boldsymbol{\sigma}^2 v'' + h(v'),$$
$$v = v_B \text{ for all } x \leq 0.$$

For the $\hat{u}$ agent, let $\hat{f}(\alpha)$ be the unique value of $\hat{c}$ satisfying $\hat{u}'(\hat{c}) = \alpha$. Let $\hat{h}(\alpha) = \hat{u}(\hat{f}(\alpha)) - \alpha\hat{f}(\alpha)$. The $\hat{u}$ agent sets $\hat{u}'(\hat{c}) = \hat{v}'$, and therefore $\hat{h}(\hat{v}') = \hat{u}(\hat{f}(\hat{v}')) - \hat{v}'\hat{f}(\hat{v}') = \hat{u}(\hat{c}) - \hat{v}'\hat{c}$. We can use $\hat{h}(\hat{v}')$ to rewrite the Bellman equation of the $\hat{u}$ agent (equations (24) – (26)) in a reduced way:

$$\gamma \hat{v} = (\mu x + \bar{\boldsymbol{z}})\hat{v}' + \frac{1}{2}\boldsymbol{\sigma}^2 \hat{v}'' + \hat{h}(\hat{v}'),$$
$$\hat{v} = v_B \text{ for all } x \leq 0.$$

The reduced Bellman of the IG agent and the reduced Bellman of the $\hat{u}$ agent will have identical solutions if and only if h is the same as $\hat{h}$. One can show directly that this is the case.

[44] A key difference is that Harris and Laibson (2013) use a wealth-dependent $\hat{u}$ function to account for binding constraints. This is not necessary here because bankruptcy occurs at the $x = 0$ boundary.

A.2 Proof of Proposition 3: Consumption in the dt-Model

Starting with clause (i), Brownian motion makes the value function twice continuously differentiable. Since $\tilde{u}'(\tilde{c}) = \boldsymbol{\beta} v'$ and $\tilde{u}''(\tilde{c})\frac{\partial \tilde{c}}{\partial x} = \boldsymbol{\beta} v''$, twice continuous differentiability of v gives continuous differentiability of $\tilde{c}$ for $x > 0$.

Clauses (ii) and (iii) are proven in Harris and Laibson (2013, Online Appendix G). See specifically Sections G.4 and G.5, and the results therein. The "once convex, always strictly convex" and "once concave, always strictly concave" results both hold on the interior of the wealth state space in the dt-model. The conclusion, which also holds in our model, is as follows:

1. If $\mu \leq \gamma$ then v is concave on $(0, \infty)$. Therefore $\tilde{c}$ is increasing on $(0, \infty)$.

2. If $\mu > \gamma$ then v may be convex immediately to the right of $x = 0$. But, v will eventually become concave and remain concave thereafter. Therefore $\tilde{c}$ may be decreasing immediately to the right of $x = 0$, but will eventually be increasing and remain increasing thereafter.

The HL13 online appendix can be found here: `https://scholar.harvard.edu/files/laibson/files/instantgrat_web_appendix.pdf`

A.3 Heuristic Derivation of the dt-Model from Discrete Time

Consider the discrete-time model of Section 2, characterized by parameters $\{\triangle, \boldsymbol{\overline{z}}, \boldsymbol{\sigma}, \boldsymbol{R}, \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{u_B}, \boldsymbol{\rho}\}$. To build intuition for the continuous-time model, we present a heuristic derivation by passing to the limit as $\triangle \to 0$. See Achdou et al. (2022, Online Appendix) for a similar exercise.

The model presented here makes four simplifying assumptions. First, we ignore the endogenous bankruptcy decision and assume that bankruptcy is declared for all $x \leq 0$.[45] Second, we ignore the liquidity constraint because it does not pass to continuous time. Third, we impose stationary Markov equilibrium for the discrete-time model, as this is the equilibrium selection used in the dt-model.[46] Fourth, we cast the model in terms of consumption rates (rather than consumption levels).

The dynamic budget constraint can be rewritten as:

$$x_{t+1} = R(x_t - \triangle \tilde{c}_t) + \overline{z} + \sigma \varepsilon_{t+1}, \tag{32}$$

where $\tilde{c}_t$ denotes the rate of consumption at time t. The Bellman equation of the present-biased consumer is:

$$\tilde{C}(x_t) = \operatorname*{argmax}_{\tilde{c}} \triangle \tilde{u}(\tilde{c}) + \beta\delta \mathbb{E}_t V(x_{t+1}) \tag{33}$$

$$W(x_t) = \triangle \tilde{u}(\tilde{C}(x_t)) + \beta\delta \mathbb{E}_t V(x_{t+1}) \tag{34}$$

$$V(x_t) = \triangle \tilde{u}(\tilde{C}(x_t)) + \delta \mathbb{E}_t V(x_{t+1}). \tag{35}$$

Note in the above expressions that $\tilde{C}(x_t)$ is a consumption rate.

[45] In discrete time this assumption creates a discontinuity in $W(x)$ at $x = 0$. It is irrelevant in continuous time because the borrowing constraint never binds for $x > 0$, and value matching holds at $x = 0$.

[46] This ignores issues of equilibrium existence and uniqueness.

Recall that $\delta = \boldsymbol{\delta}^{\triangle}$ and $R = \boldsymbol{R}^{\triangle}$. We can approximate δ and R for small $\triangle$ as:

$$\delta = \exp(\triangle \ln(\boldsymbol{\delta})) \;\approx 1 - \triangle\gamma,$$
$$R = \exp(\triangle \ln(\boldsymbol{R})) \approx 1 + \triangle\mu,$$

where $\gamma = -\ln(\boldsymbol{\delta})$ and $\mu = \ln(\boldsymbol{R})$. Using the approximation for R along with equation (6), budget constraint (32) becomes

$$x_{t+1} = x_t + \triangle(\mu x_t + \overline{\boldsymbol{z}} - \tilde{c}_t) + \sigma\varepsilon_{t+1} - \mu\triangle^2\tilde{c}_t.$$

Dropping higher-order terms, in the limit as $\triangle \to 0$ we obtain:

$$dx_t = (\mu x_t + \overline{\boldsymbol{z}} - \tilde{c}_t)dt + \boldsymbol{\sigma} db_t, \tag{36}$$

where b_t is a standard Brownian motion. This is the continuous-time dynamic budget constraint stated in equation (18).

Next, using the approximation for δ along with equation (35) yields:

$$V(x_t) = \triangle\tilde{u}(\tilde{c}_t) + (1 - \triangle\gamma)\mathbb{E}_t V(x_{t+1}).$$

Subtracting $(1 - \triangle\gamma)V(x_t)$ from both sides of this equation:

$$\triangle\gamma V(x_t) = \triangle\tilde{u}(\tilde{c}_t) + (1 - \triangle\gamma)\mathbb{E}_t\left[V(x_{t+1}) - V(x_t)\right].$$

Dividing both sides by $\triangle$ and taking the limit as $\triangle \to 0$ gives:

$$\lim_{\triangle\to 0} \gamma V(x_t) = \lim_{\triangle\to 0} \tilde{u}(\tilde{c}_t) + \lim_{\triangle\to 0}(1 - \triangle\gamma)\left[\frac{\mathbb{E}_t V(x_{t+1}) - V(x_t)}{\triangle}\right].$$

The term in brackets is $\frac{\mathbb{E}_t[dV(x_t)]}{dt}$. Using Itô's Lemma, this is given by $(\mu x_t + \overline{\boldsymbol{z}} - \tilde{c}_t)V' + \frac{1}{2}\boldsymbol{\sigma}^2 V''$. We've now recovered differential equation (20).

B Non-Robust Equilibria in a Deterministic Model

B.1 Deterministic Eat-the-Pie Model

Here we consider a deterministic Eat-the-Pie model of consumption. The agent begins with a stock of wealth x_0 and earns no further income ($\bar{z} = 0$). The agent has present-biased time preferences $(1, \beta\delta, \beta\delta^2, ...)$ and log utility:

$$u(c) = \triangle \ln(c/\triangle).$$

For this constructed equilibrium we set $\boldsymbol{\beta} = 0.4$, $\boldsymbol{\delta} = 0.95$, and $\boldsymbol{R} = 1$. The model is deterministic ($\boldsymbol{\sigma} = 0$) and bankruptcy is never declared by the agent in equilibrium ($\boldsymbol{u_B} = -\infty$).[47]

B.2 Phelps-Pollak Consumption Function

This Eat-the-Pie model features a linear equilibrium which we call the Phelps-Pollak equilibrium (Phelps and Pollak, 1968). Laibson (1994) shows that the Phelps-Pollak equilibrium is the unique equilibrium that is selected in the limit of a finite T-horizon game as $T \to \infty$.

The Phelps-Pollak equilibrium is characterized by the consumption rule $C(x) = \lambda x$, where λ is given by:

$$\lambda = \frac{1-\delta}{1-\delta(1-\beta)}.$$

Parameter λ varies with $\triangle$ since $\delta = \boldsymbol{\delta}^{\triangle}$.

[47] The assumption that $\boldsymbol{u_B} = -\infty$ is problematic when cash on hand is stochastic (see footnote 23). When the model is deterministic, the agent can always choose an equilibrium path that avoids bankruptcy.

B.3 The Constructed Sawtooth Consumption Function

Following the logic of Krusell and Smith Jr. (2003), we now construct a *continuum* of equilibria that feature a "sawtooth" consumption function. This consumption function is shown graphically in Figure 2 below. The sawtooth consumption function $C(x)$ is characterized by two (dotted/dashed) rays, an upper vector φx and a lower vector αx. Both φ and α depend on the time-step $\triangle$. At all except the countably infinite number of points where consumption jumps from the φ vector down to the α vector, the consumption function features an MPC of 1. We assume that $C(x) = \alpha x$ at points of discontinuity. This means that $C(x)$ is right-continuous (with an MPC of 1) at all wealth levels $x \in (0, \infty)$. Figure 2 also plots the savings function $x - C(x)$. Since the MPC is 1 at all non-jump points, savings is characterized by a step function that increases discretely whenever consumption jumps from the φ vector down to the α vector. We solve for α and φ as a function of $\triangle$ below.

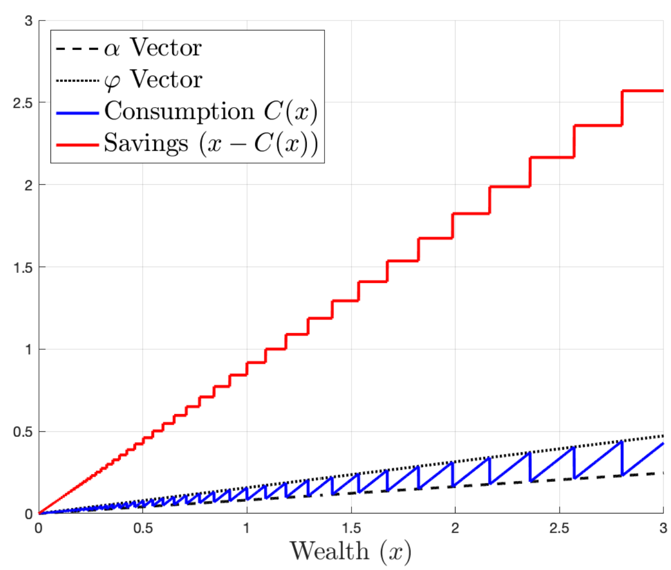


Figure 2: Constructed sawtooth equilibrium for $\triangle = 1$.

Let $\mathcal{X}$ denote the set of asset levels at which the sawtooth consumption function jumps down to the α vector. The sawtooth consumption function is constructed such that the agent moves down from one point in $\mathcal{X}$ to the next, always getting closer to the origin. When the agent consumes along the α vector, they consume α% of their assets and pass $(1-\alpha)$% on to the next self. From any point $x_\alpha \in \mathcal{X}$, this path of consumption yields a value of:

$$V(x_\alpha) = \sum_{s=t}^{\infty} \delta^{s-t} \triangle \ln\left(\frac{\alpha x_\alpha (1-\alpha)^{s-t}}{\triangle}\right) = \triangle\left[\frac{\ln(\alpha/\triangle)}{1-\delta} + \frac{\ln(x_\alpha)}{1-\delta} + \delta\frac{\ln(1-\alpha)}{(\delta-1)^2}\right]. \tag{37}$$

One implication of setting the MPC equal to 1 between the α and the φ vector is that $(1-\varphi) = (1-\alpha)^2$.[48] So, our constructed equilibrium has the property that

$$\varphi = \alpha(2-\alpha). \tag{38}$$

For this constructed equilibrium to hold, we impose that the current self is indifferent between consuming φx_α versus αx_α for all $x_\alpha \in \mathcal{X}$. At each $x_\alpha \in \mathcal{X}$ our indifference assumption yields the condition that

$$u(\alpha x_\alpha) + \beta\delta V((1-\alpha)x_\alpha) = u(\varphi x_\alpha) + \beta\delta V((1-\varphi)x_\alpha),$$

where $V(\cdot)$ is given by equation (37). Since $\varphi > \alpha$, $u(\alpha x_\alpha) < u(\varphi x_\alpha)$ but this is offset by $V((1-\alpha)x_\alpha) > V((1-\varphi)x_\alpha)$. Imposing equation (38), the above indifference condition can be restated as

$$u(\alpha x_\alpha) + \beta\delta V((1-\alpha)x_\alpha) = u(\alpha(2-\alpha)x_\alpha) + \beta\delta V((1-\alpha)^2 x_\alpha). \tag{39}$$

[48]Consider a wealth level $x_\alpha \in \mathcal{X}$. If the agent consumes φx_α then they jump to a point in $\mathcal{X}$ that is one point lower than the point they jump to by consuming αx_α. The agent needs to consume at rate α for two periods in order to get to wealth level $(1-\varphi)x_\alpha$. Consuming at rate α twice results in a wealth level of $(1-\alpha)^2 x_\alpha$ after two periods.

From equation (39) we can derive the following formula for α under log utility:

$$2-\alpha=(1-\alpha)^{-\frac{\beta\delta}{1-\delta}}. \tag{40}$$

Equation (40) implicitly defines α for any time-step $\triangle$. Given α, φ is given by equation (38).

B.4 Equilibrium

Now that we have outlined the construction of the sawtooth consumption function shown in Figure 2, we argue that this sawtooth consumption function is an equilibrium of the deterministic Eat-the-Pie model. Intuitively, the agent does not want to increase consumption because the sawtooth consumption function is designed to punish such increases. In particular, a small increase in consumption by self t discretely increases the consumption of self $t+1$ from the α vector to the φ vector, and pushes all selves $t+2$ onward to one kink-point lower than they would otherwise be. Similarly, present bias implies that the agent does not want to cut consumption at time t in order to increase the consumption of self $t+1$ (who has a right-MPC of 1).

In detail, let $g(c,x)=\max\{x_\alpha \in \mathcal{X} | x_\alpha \leq x-c\}$. In words, function $g(c,x)$ returns the closest point in $\mathcal{X}$ that is weakly below $x-c$. From any level of cash on hand, x_t, the current-value function for consumption choice c is:

$$W(c,x_t)=u(c)+\beta\delta u(x_t-c-(1-\alpha)g(c,x_t))+\beta\delta^2 V((1-\alpha)g(c,x_t)), \tag{41}$$

where $V(\cdot)$ is given by equation (37). The interpretation of $W(c,x_t)$ is as follows. The current self consumes c and earns utility $u(c)$. Next, the sawtooth consumption function implies that self $t+1$ will consume such that self $t+2$ has wealth $(1-\alpha)g(c,x_t)$. Self $t+1$ therefore consumes $(x_t-c)-(1-\alpha)g(c,x_t)$. From self $t+2$ onwards the agent returns to consuming along the α vector. The continuation-value function from $t+2$ onwards is given by $V\left((1-\alpha)g(c,x_t)\right)$.

Step 1: Equilibrium at Points in $\mathcal{X}$. Consider a point $x_\alpha \in \mathcal{X}$. Without loss of generality, we set $x_\alpha = 1$. The equilibrium is constructed so that $W(\alpha, 1) = W(\varphi, 1)$. We argue here that it is optimal for the self with wealth $x_\alpha = 1$ to choose either $C(1) = \alpha$ or $C(1) = \varphi$.

First, we show that $W(\alpha, 1) > W(c, 1)$ for all $c \in (\alpha, \varphi)$. To start, note that $W(c, 1)$ is continuous for $c \in (\alpha, \varphi]$ and $W(c, 1)$ is differentiable for $c \in (\alpha, \varphi)$. For $c \in (\alpha, \varphi)$, function $g(c, 1) = 1 - \varphi$ and therefore $\frac{\partial W(c,1)}{\partial c} = \frac{1}{c} - \beta\delta\frac{1}{1-c-(1-\alpha)(1-\varphi)}$. This derivative reaches a minimum as $c \to \varphi$. In the limit, $\lim\limits_{c \to^- \varphi} \frac{\partial W(c,1)}{\partial c} = \frac{1}{\varphi} - \beta\delta\frac{1}{\alpha(1-\varphi)}$. For our calibration, one can show that $\lim\limits_{c \to^- \varphi} \frac{\partial W(c,1)}{\partial c} > 0$ for all $\triangle \in (0, 1]$.[49] Since $\frac{\partial W(c,1)}{\partial c} > 0$ for all $c \in (\alpha, \varphi)$, $W(\varphi, 1) > W(c, 1)$ for all $c \in (\alpha, \varphi)$. Since $W(\alpha, 1) = W(\varphi, 1)$ by construction, it is also the case that $W(\alpha, 1) > W(c, 1)$ for all $c \in (\alpha, \varphi)$.

Second, we show that $W(\alpha, 1) > W(c, 1)$ for all $c \in (0, \alpha)$. Again, $W(c, 1)$ is continuous for $c \in (0, \alpha]$ and $W(c, 1)$ is differentiable for $c \in (0, \alpha)$. For $c \in (0, \alpha)$, function $g(c, 1) = 1 - \alpha$ and therefore $\frac{\partial W(c,1)}{\partial c} = \frac{1}{c} - \beta\delta\frac{1}{1-c-(1-\alpha)^2}$. This derivative reaches a minimum as $c \to \alpha$. In the limit, $\lim\limits_{c \to^- \alpha} \frac{\partial W(c,1)}{\partial c} = \frac{1}{\alpha} - \beta\delta\frac{1}{\alpha(1-\alpha)}$. For our calibration, one can show that $\lim\limits_{c \to^- \alpha} \frac{\partial W(\alpha,1)}{\partial c} > 0$ for all $\triangle \in (0, 1]$. Since $\frac{\partial W(c,1)}{\partial c} > 0$ for all $c \in (0, \alpha)$, $W(\alpha, 1) > W(c, 1)$ for all $c \in (0, \alpha)$.

Third, we argue that $W(\alpha, 1) > W(c, 1)$ for all $c \in (\varphi, 1)$. Since the current self is willing to cut consumption from φ to α in order to increase the consumption of future selves, the current self will not want to increase consumption above φ. Any consumption above φ decreases the consumption of selves $t + 2$ onward by at least one more kink-point, effectively reproducing the arguments above but with even less incentive to consume at time t and more incentive to pass liquidity to the future.

Step 2: Equilibrium at Points Outside $\mathcal{X}$. We've argued that our sawtooth equilibrium holds for all points in $\mathcal{X}$. Now we argue that the sawtooth consumption

[49]The important condition is $\beta < \frac{1}{2}$. Setting $c = \varphi$, one can show that $\frac{1}{\varphi} > \frac{\beta\delta}{1-\varphi-(1-\alpha)(1-\varphi)}$ reduces to the condition $(1-\alpha)^2 > \beta\delta(2-\alpha)$. Since $\alpha \to 0$ and $\delta \to 1$ as $\triangle \to 0$, $\beta < \frac{1}{2}$ is a necessary condition for this to hold.

function is also optimal at all points $x \notin \mathcal{X}$.

First, we rule out any deviations from the constructed equilibrium in which the current self increases consumption. Without loss of generality, continue to assume that $1 \in \mathcal{X}$ and consider a point $x \in (1, \frac{1}{1-\alpha})$.[50] The benefit of increasing consumption will be lower for all $x \in (1, \frac{1}{1-\alpha})$ than it is at $x = 1$. In Step 1 we showed that $C(1) = \alpha$ is (weakly) optimal. Thus, it will not be optimal to increase consumption for any $x \in (1, \frac{1}{1-\alpha})$.

Second, we rule out any deviations from the constructed equilibrium in which the current self decreases consumption. Here the argument is similar. Without loss of generality, continue to assume that $1 \in \mathcal{X}$ and consider a point $x \in (1-\alpha, 1)$.[51] The benefit to decreasing consumption will be lower for all $x \in (1-\alpha, 1)$ than it is in the limit as $x \to^{-} 1$. In Step 1 we showed that $C(1) = \varphi$ is (weakly) optimal. Thus, it will not be optimal to decrease consumption for any $x \in (\alpha, 1)$.

B.5 Equilibria as $\triangle \to 0$

For our chosen calibration of $\boldsymbol{\beta} = 0.4$ and $\boldsymbol{\delta} = 0.95$ we use equations (38) and (40) to characterize the sawtooth consumption function. We also present the linear Phelps-Pollak equilibrium for reference. Results are given in Table 4.

$\triangle$	α	φ	λ (Phelps-Pollak)
$\triangle = 1$	0.0821	0.1575	0.1163
$\triangle = 1/4$	0.0218	0.0431	0.0313
$\triangle = 1/16$	0.0055	0.0110	0.0080
$\triangle = 1/100$	8.8811 e-4	1.7754 e-3	1.2810 e-3
$\triangle = 1/1000$	8.8877 e-5	1.7775 e-4	1.2822 e-4
$\triangle = 1/10000$	8.8884 e-6	1.7777 e-5	1.2823 e-5

Table 4: Equilibrium values of α, φ, and λ. α, φ, and λ decrease approximately in proportion with time-step $\triangle$. Additionally, λ always lies between α and φ.

[50] If there exists a discontinuity at 1, the next highest discontinuity occurs at $\frac{1}{1-\alpha}$.

[51] If there exists a discontinuity at 1, the next lowest discontinuity occurs at $1-\alpha$.

B.6 Comment on Discrete vs. Continuous Time

The above analysis shows that the sawtooth consumption function characterized by α and φ contains an arbitrary number of downward discontinuities over any finite interval of the state space as $\triangle \to 0$. In contrast, Harris and Laibson (2013) prove that in continuous time with vanishingly small Brownian noise for asset returns, this Eat-the-Pie problem has a unique consumption function characterized by consumption rate $\lambda_{dt} = \frac{-\ln(\boldsymbol{\delta})}{\boldsymbol{\beta}}$. Note that $\lambda_{dt} = \lim_{\triangle \to 0} \frac{\lambda}{\triangle}$, where λ is the discrete-time Phelps-Pollak equilibrium defined above. Thus, the unique consumption function in the continuous-time Instantaneous Gratification case is linear in wealth.

What this comparison highlights is that the sawtooth equilibrium exists for every $\triangle > 0$, but it does not pass to the continuous-time Instantaneous Gratification case. Intuitively, the sawtooth equilibrium can support lower consumption than the Phelps-Pollak equilibrium because consumption deviations by self t will be punished by overconsumption by self $t+1$. In order for this punishment to be possible, there must exist values of x at which the sawtooth consumption function $C(x) > \alpha x$. However, in the continuous-time Instantaneous Gratification model (with vanishing Brownian motion), such punishments do not arise.

C Numerical Methods

C.1 Discrete Time

Here we outline the algorithm used to solve our discrete-time $\triangle$-model. We opt for simplicity whenever possible. As discussed in Section 2.1, the numerical methods remain the same for all $\triangle > 0$. All that changes is the model's calibration.

We use VFI to solve a T-horizon game for large T, and then check if the policy functions have converged. Our VFI has two state variables: assets x and period t. We construct a uniform grid for state variable x with a step-size denoted by $xjump$ and maximum grid value of $xmax$.[52] Our grid $\mathcal{G}$ consists of the set $\{0,\ xjump,\ 2 \times xjump,\ ...,\ xmax\}$. At each point in the grid and for each $t \in \{1, 2, ..., T-1\}$, the agent solves the following consumption problem:

$$C_t(x_t) \in \underset{c}{\operatorname{argmax}} \;\; u(c) + \beta\delta\mathbb{E}_t V_{t+1}(R(x_t - c) + z_{t+1}),$$

$$W_t(x_t) = \max\left\{u(C_t(x_t)) + \beta\delta\mathbb{E}_t V_{t+1}(R(x_t - C_t(x_t)) + z_{t+1}),\ u_B + \beta\delta V_B\right\}$$

$$V_t(x_t) = \begin{cases} u(C_t(x_t)) + \delta\mathbb{E}_t V_{t+1}(R(x_t - C_t(x_t)) + z_{t+1}) & \text{if } x_t \notin \mathcal{B}_t^* \\ V_B & \text{if } x_t \in \mathcal{B}_t^* \end{cases},$$

$$\mathcal{B}_t^* = \{x_t | u_B + \beta\delta V_B \geq u(C_t(x_t)) + \beta\delta\mathbb{E}_t V_{t+1}(R(x_t - C_t(x_t)) + z_{t+1})\},$$

subject to the restrictions that:

1. $V_T(x_T) = \begin{cases} \frac{u(\overline{z})}{1-\delta} & \text{if } x_T \geq 0 \\ V_B & \text{if } x_T < 0 \end{cases}$

[52] We use a uniform grid for three reasons. First, the general suggestion in the numerical literature is to place more gridpoints where the consumption function has the most curvature. In addition to placing gridpoints near the borrowing constraint, one would also ideally place grid points in areas where the consumption function exhibits large and rapid fluctuations. As our simulations show, this non-robust behavior is more likely to occur for large wealth values. Balancing the desire for gridpoints near the origin and near the top of the state space, we simply choose a uniform grid. Second, uniform grids are easy to understand. Third, with only one state variable our algorithm is fast enough to allow for many gridpoints, even as $\triangle \to 0$. Thus, we can ignore issues of grid point efficiency and simply use the brute-force approach of many grid points.

2. $x_t - C_t(x_t) \in \left\{0,\ \frac{xjump}{R},\ \frac{2\times xjump}{R},\ ...,\ \frac{xmax}{R}\right\}$

Restriction 1 serves as an initial "guess" to start our Value Function Iteration.[53] Restriction 2 is imposed to ensure that $R(x_t - C_t(x_t))$ lies on the grid in period $t+1$.[54] The practical implication of Restriction 2 is that the agent chooses consumption from a discrete choice set.

Restriction 2 also means that our discrete-time numerical methods become less accurate as $\triangle$ shrinks. This follows from the discretization of the state space into grid $\mathcal{G}$, which we keep constant regardless of $\triangle$. Restriction 2 implies that consumption level $C_t(x_t)$ is restricted to the set $\{x,\ x - \frac{xjump}{R},\ x - \frac{2\times xjump}{R},\ ...,\ x - \frac{xmax}{R}\}$. In this paper we care about consumption rate $\tilde{C}_t(x_t)$. By the same logic, $\tilde{C}_t(x_t)$ is restricted to the set $\{\frac{1}{\triangle}x,\ \frac{1}{\triangle}\left(x - \frac{xjump}{R}\right),\ \frac{1}{\triangle}\left(x - \frac{2\times xjump}{R}\right),\ ...,\ \frac{1}{\triangle}\left(x - \frac{xmax}{R}\right)\}$. As $\triangle \to 0$ the step size between each choice for consumption rate $\tilde{C}_t(x_t)$ increases. The benefit of our multigrid algorithm is that it allows us to efficiently use a dense grid.

Stopping Criteria. We do not always achieve complete policy function convergence from our value function iteration. In order for the VFI to terminate we must therefore impose a stopping criteria. In the case where $\delta R < 1$ we stop the VFI if either (i) the consumption function converges, or (ii) there are fewer than 40 unique consumption functions out of the last 50 iterations. Case (i) accounts for convergence, and case (ii) accounts for cycles.[55] In the case where criteria (ii) is met, we plot the final iteration of the consumption function. Every $\delta R < 1$ calibration of the discrete-time model in this paper is terminated due to either (i) or (ii) (i.e., the consumption

[53] In defining V_T we implicitly assume that the agent is infinitely lived and has a deterministic utility flow for all periods $t \geq T$. This can be generalized. Because the Bellman operator is not a contraction mapping when $\beta < 1$, this initialization can affect the equilibrium to which we converge, particularly in the cases where the equilibrium is non-robust.

[54] The restriction that $x \leq xmax$ means that any large positive shock that pushes $x_{t+1} > xmax$ will be wasted. So, the consumer will endogenously avoid holding assets near the top of the grid. When $\delta R < 1$ we find that this has little effect on the resulting equilibrium, as long as $xmax$ is relatively large. However, our algorithm can be updated to allow the agent to accumulate off the grid by extrapolating the payoffs for above-grid asset levels.

[55] These cycles are commonly found in the literature. For example, a discussion of VFI cycles can be found in Krusell et al. (2002).

function either converges or enters a cycle).

When $\delta R > 1$ we are no longer able to establish clean stopping criteria. Whenever the VFI does not terminate due to (i) or (ii), we simply terminate the VFI after a large and predefined number of runs. Results for $\delta R > 1$ cases are presented in Appendix D.

C.2 Continuous Time

Achdou et al. (2022) provide an excellent set of resources on continuous-time finite difference methods. We refer interested readers to their paper.

To solve for the equilibrium of the IG consumer, one needs to solve the HJB equation of the $\hat{u}$ agent and then back out the IG agent's consumption from the $\hat{u}$ equilibrium. The Bellman equation of the IG agent fails to meet a monotonicity condition that is required for finite difference methods to converge (Barles and Souganidis, 1991). This failure means that one cannot directly tackle the Bellman of the IG agent with a finite difference scheme. Since $\beta < 1$ is a necessary condition for the failure of monotonicity, the exponential $\hat{u}$ agent provides a solution. Further details are presented in Maxted (2022).

D $\triangle$-Model when $\delta R > 1$

Our value function iteration rarely converges when $\delta R > 1$. In particular, the consumption functions that we solve for become particularly sensitive whenever the agent accumulates assets (in expectation) near the top of the grid. In these accumulating equilibria, consumption function $C_t(x_t)$ depends on the consumption function at higher levels of wealth, which can be poorly behaved, and the strategic behavior that exists at high levels of wealth then "trickles down" to lower levels of wealth. Nevertheless, our $\triangle \to 0$ method still moderates issues with equilibrium non-robustness. However, these cases are relegated to the appendix because they require a more careful treatment from a numerical standpoint.

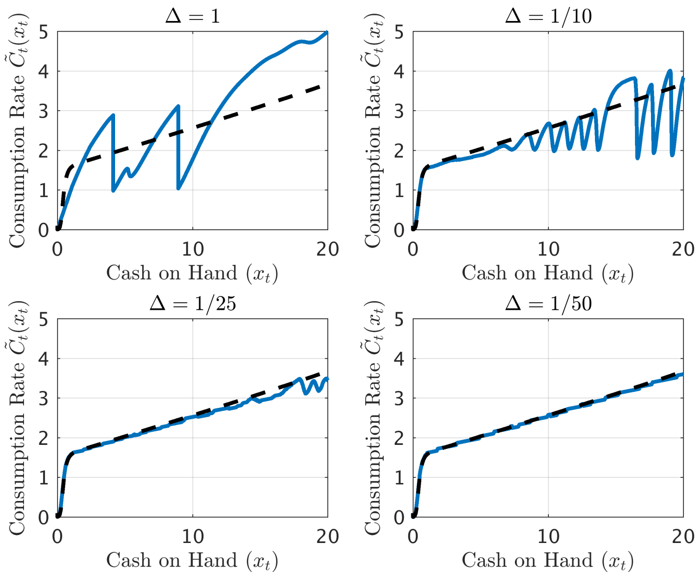
D.1 Numerical Methods

We explore two calibrations with $\delta R > 1$. First, we set $\boldsymbol{R} = 1.07$ and keep all other parameters of the baseline calibration constant. Second, we set $\boldsymbol{R} = 1.07$ and also increase $\boldsymbol{u_B}$ to $\boldsymbol{u}\left(\frac{9}{10}\overline{\boldsymbol{z}}\right)$. This high-$u_B$ calibration is included in order to generate non-monotonic consumption in the continuous-time model. These calibrations are illustrated in Figure 3 below.

In discrete time, the policy functions that we compute do not converge over the entire wealth space. Because we need our algorithm to terminate, we simulate a finite T-horizon game with a large and predefined T. Convergence is outlined in Figure 4 below. An important property of Figure 4 is that the consumption function converges (approximately) on an expanding wealth interval as $\triangle \to 0$, thus producing a growing interval of robust behavior as the time-step shrinks. In continuous time, there exists a stationary policy function that is easily characterized by our finite difference methods.

D.2 Policy Functions

(a) $\boldsymbol{\beta} = 0.5$ and $\boldsymbol{R} = 1.07$



(b) $\boldsymbol{\beta} = 0.5$, $\boldsymbol{R} = 1.07$, and $\boldsymbol{u_B} = \boldsymbol{u}\left(\frac{9}{10}\overline{\boldsymbol{z}}\right)$

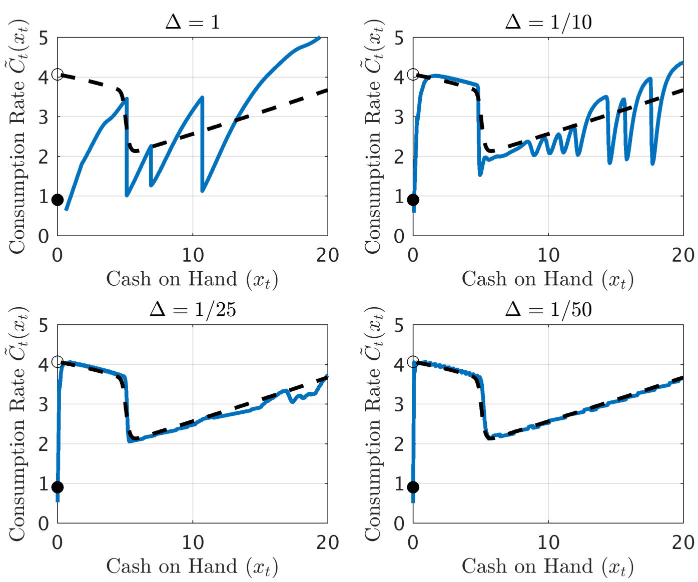
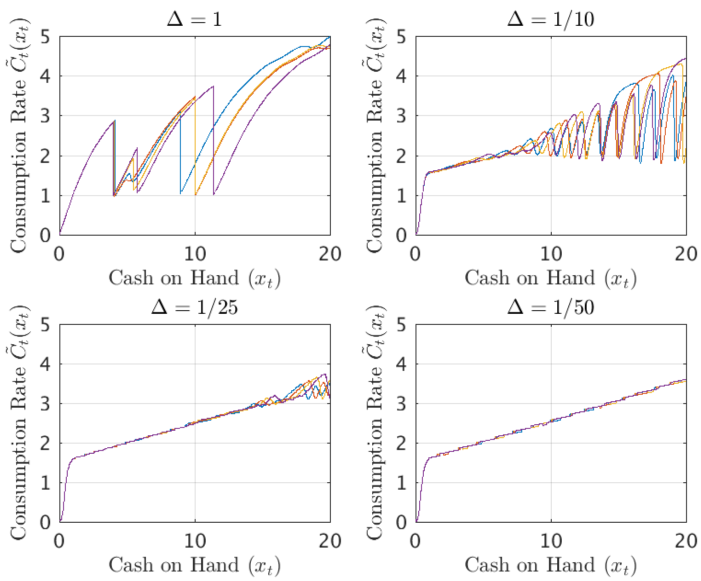


Figure 3: Consumption functions for $\triangle \in \{1, \frac{1}{10}, \frac{1}{25}, \frac{1}{50}\}$. The blue line plots the final iteration of the $\triangle$-model consumption function. The dashed black line plots the dt-model consumption function. Since consumption functions do not converge in the discrete-time model, the blue line is the (arbitrary) final iteration of the VFI.

D.3 Policy Function Convergence

(a) $\boldsymbol{\beta} = 0.5$ and $\boldsymbol{R} = 1.07$



(b) $\boldsymbol{\beta} = 0.5$, $\boldsymbol{R} = 1.07$, and $\boldsymbol{u_B} = \boldsymbol{u}\left(\frac{9}{10}\overline{\boldsymbol{z}}\right)$

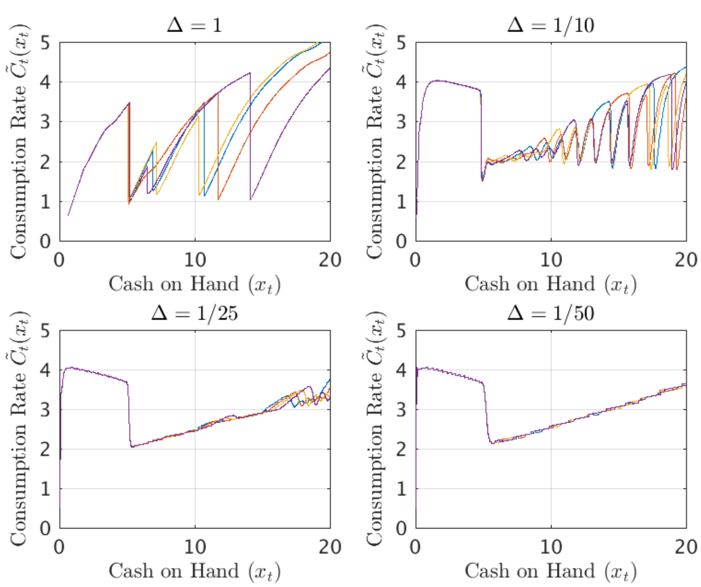


Figure 4: Consumption functions for $\triangle \in \{1, \frac{1}{10}, \frac{1}{25}, \frac{1}{50}\}$. Each panel plots the $t = 1$, $t = 26$, $t = 51$, and $t = 76$ consumption function. Consumption functions approximately converge on an interval around $x = 0$ that expands as $\triangle \to 0$.

E Additional Figures

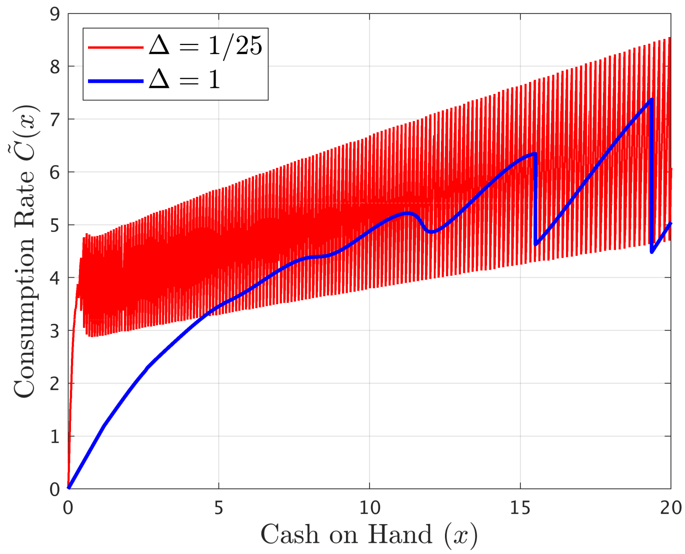


Figure 5: Consumption functions holding the $\frac{\bar{z}}{\sigma}$ ratio constant ($\boldsymbol{\beta} = 0.5$).

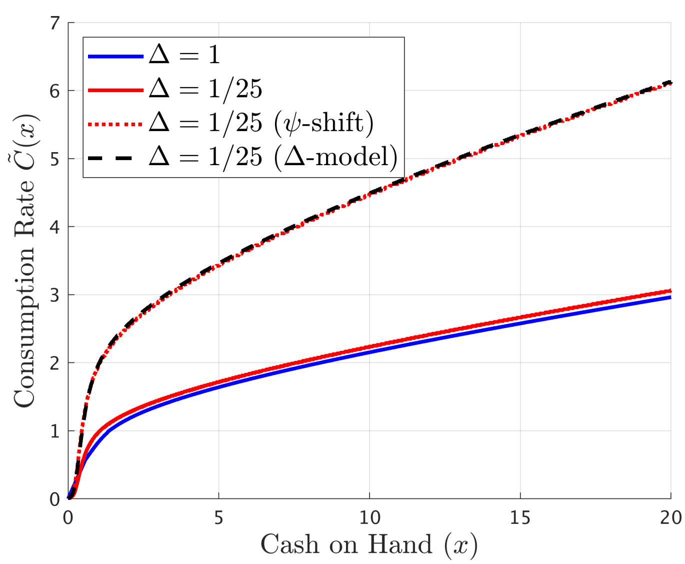


Figure 6: The $\hat{u}$ agent in discrete time ($\boldsymbol{\beta} = 0.5$). The bottom two curves plot the $\hat{u}$ agent's consumption for $\triangle \in \{1, \frac{1}{25}\}$. The dotted red line is $\frac{1}{\psi}$ times the consumption of the $\hat{u}$ agent for $\triangle = \frac{1}{25}$. The dashed black line is the consumption of the present-biased agent ($\boldsymbol{\beta} = 0.5$) for $\triangle = \frac{1}{25}$.