How the Future Shapes Consumption with Time-Inconsistent Preferences*

James Feigenbaum† Sepideh Raei‡

October 20, 2023

Abstract

Time-inconsistent preferences, which are modeled by relative discount functions, are a common explanation for the empirical finding that lifecycle profiles of household consumption are typically hump-shaped rather than monotonic. More precisely, time-inconsistent preferences that are present-biased often generate a hump-shaped consumption profile over the lifecycle. We develop a general framework for understanding present bias in consumption through a future weighting factor that perturbs the discount factor of utility at future periods away from exponential discounting. Using our framework we derive necessary and sufficient conditions on the future weighting factors for the log consumption profile to be locally concave. We find that these conditions, which are necessary for the consumption profile to be hump-shaped, are stronger than just assuming a present bias. Furthermore, we explore the conditions under which the consumption profile determined in the first period of life Pareto dominates the realized consumption profile. Lastly, we explore the interconnections between these two sets of conditions, elucidating the linkages between the determinants of hump-shaped consumption profiles and the conditions necessary for the initial consumption path to achieve Pareto dominance.

JEL: D60, D90

Keywords: present bias, time-inconsistent preferences, consumption hump, commitment mechanisms, welfare comparison

---

*We would like to thank Frank Caliendo and Scott Findley for their input. We are grateful to our editor, Dr. Ronald Peeters, and the two anonymous reviewers for their excellent comments. Further thanks goes to participants at the SABE 2022 conference for their helpful discussion.

†Utah State University, Huntsman School of Business, Utah, United States; james.feigenbaum@usu.edu.

‡Utah State University, Huntsman School of Business, Utah, United States; sepideh.raei@usu.edu.

1 Introduction

The canonical life-cycle model predicts that consumption will grow smoothly for patient individuals and decay smoothly for impatient individuals. However, from an empirical standpoint, one of the most striking aspects of people's choices of consumption over the lifecycle is that this profile is generally hump-shaped. As was first documented by Thurow (1969), average consumption increases while consumers are young, peaks when they reach middle age, and decreases afterwards.[1] The literature has proposed two approaches to address this inconsistency. One strand of the literature refined the canonical theoretical framework by enhancing the Lifecycle/Permanent-Income Hypothesis with more complicated preferences and various technological frictions. Another set of solutions is focused on relaxing the rational paradigm. For example, when we allow for time-inconsistent preferences, the consumption hump can be attributed to the concept of "present bias", meaning that individuals have a tendency to place disproportionate weight on immediate rewards over future rewards.[2].

In this paper, we develop a general framework for understanding the concept of "present bias" in consumption through a "future weighting factor" that perturbs the discount factor of utility at future periods away from exponential discounting. This framework nests the well-known hyperbolic and quasihyperbolic discount functions. It also facilitates general statements about necessary conditions for presence of both present bias and a concave log consumption profile, a characteristic often associated with present bias.

Our framework allows us to determine the conditions under which the consumption profile is consistent with the empirical evidence. In a nutshell, our analysis reveals that the occurrence of a hump-shaped consumption profile will depend on the extent to which the discount function deviates from an exponential discount function.[3] We find that while present bias

[1]See Carroll and Summers (1991), Attanasio and Weber (1995), Attanasio et al. (1999), Browning and Crossley (2001), Gourinchas and Parker (2002), and Fernandez-Villaverde and Krueger (2011).

[2]Present bias, which is viewed as a form of misoptimization that accounts for a range of behavioral "mistakes," e.g. undersaving for retirement, has yielded a large literature that emphasizes the potential for policies like forced pensions or retirement saving subsidies to protect against or correct such mistakes (for a survey on present bias see O'Donoghue and Rabin (2015)).

[3]Note that in order for the lifecycle consumption profile to be hump-shaped, we need consumption to grow up to some age that will be the peak of the hump and decline thereafter. Thus in the vicinity of the peak, we need both that the slope of the consumption profile is decreasing and that the slope is positive before the peak and negative afterwards. Our focus in this paper will be the former condition rather than the latter because the dynamics of consumption growth are determined strictly by the discount function whereas the initial condition for consumption growth depends on the interest rate, which is exogenous in the present context. If we endogenize the interest rate, it will depend on both the preferences and technology, and the technology is beyond the scope of this paper. We will establish necessary and sufficient conditions for the log consumption profile to be strictly concave, so consumption growth is strictly decreasing. Then

is a necessary condition for a hump-shaped consumption profile, it is not a sufficient one. Subsequently, we utilize this general framework to investigate the conditions under which all of the different selves prefer the commitment path of consumption (the plan established in period zero) to the realized path (the actual consumption decisions). That is to say the commitment path Pareto dominates the realized path. Our work reveals that the welfare difference between the realized path and the commitment path for each self is a U-shaped function of the terminal future weighting factor. That means, if the terminal future weighting factor is large enough in magnitude, whether positive or negative, most selves will prefer the realized path to the commitment path.

In this paper, we propose a general representation of a discount function in the form of $D_t = D_1^t(1+\varepsilon_t)$ for $t = 0, \ldots, T$, where ε_t is the **extra weight** (compared to the exponential discounting case) that we put on the discount factor t periods in the future, and $T+1$ is the life span. If $\varepsilon_t \neq 0$ for some $t = 0, \ldots, T$, we will have a nonexponential discount function, in which case we must interpret the "time" that parameterizes the function as the delay or waiting time until we experience the consumption from the present moment.[4] In our discount function, we call ε_t the *future weighting factor*. All forms of discount function, including nonexponential ones, can be written as a specific case of this general function by finding the corresponding ε_t. An advantage of this novel approach is the opportunity it provides to understand the driving force behind the consumption hump.

Our method observes that an exponential discount function yields a linear log consumption profile. Deviations from this linearity can help identify present or future bias based on the sign and trend of the future weighting factors ε_t. A present bias comes from having all ε_t be positive and strictly increasing for $t > 1$, whereas a future bias comes from having all ε_t be negative and strictly decreasing for $t > 1$.[5] [6]

In a lifecycle model, we establish conditions for the future weighting factor that result

there will be a range of interest rates for which the initial consumption growth is positive and the terminal growth is negative.

[4]With the exponential discounting function popularized by Samuelson (1937), the concept of time can be assumed to be either absolute time, calendar time, or even waiting time, i.e. the time to consumption. As Strotz (1955a) showed, the equivalence of exponential discount functions under these three temporal measures is a consequence of the exponential function **not** exhibiting preference reversals. In contrast, for nonexponential discount functions, such as one that exhibits present bias, the 'time' parameter should be understood as the delay or waiting period before consumption occurs from the present time.

[5]For the case of a future bias we also need the additional requirement that $\varepsilon_t > -1$.

[6]An alternative way of expressing this concept is that positive future weights indicate that the discount function will exceed an exponential discount function. If this excess increases with the delay time, that signifies the presence of present bias everywhere. Conversely, negative future weights indicate the reverse relationship and are typically associated with future bias.

in local concavity of the log consumption profile, which is essential for a hump-shaped consumption curve. Specifically, the growth rate of the future weighting factor must exceed those at shorter delays for a given point in the lifecycle. This implies a slower decay of the discount function compared to an exponential one, It is, however, a stronger condition than assuming a present bias, indicating that present bias alone is insufficient to guarantee a hump-shaped consumption profile.

The quasihyperbolic discounting function is a canonical example used to demonstrate present and future bias. This is often also referred to as a β-δ discount function, where the parameter δ is the generalization of the common exponential discount factor and β is a multiplicative shifter of all the discounting at a positive delay. If $\beta < 1$, the discount function will be present-biased, and, if $\beta > 1$, the discount function will be future-biased. A present-biased quasihyperbolic discount function will also exhibit a strictly concave consumption profile, and, conversely, a future-biased quasihyperbolic discount function yields strictly convex consumption profiles. However, we can further generalize the quasihyperbolic discount function to obtain a β-δ-ω discount function, where ω is the multiplicative shifter specifically of the discounting at the longest delay. When $\beta = \omega$, this discount function reverts to the quasihyperbolic case, but there is a region of the parameter space where the discount function is present-biased and the corresponding consumption profiles are not strictly concave and, for some choices of the interest rate, will not be hump-shaped.

Another area in which present (and future) bias has garnered attention in the recent literature is welfare analysis. Since an individual with time-inconsistent preferences will choose a consumption profile that depends on the time of the choosing, it is not obvious which period of life should be the reference point for welfare comparison. The literature often defaults to using the preferences of the 'initial self' as a welfare benchmark, primarily because present bias is empirically prevalent. However, this method has drawn criticism. Gul and Pesendorfer (2004) describes it as 'odd,' and Dewatripont et al. (2004) argue that this approach lacks a solid normative foundation for equating welfare with initial-time preferences.

We address existing criticisms by using our proposed future weighting functional form to identify conditions on the future weighting factors under which the commitment path, determined in life's first period, will Pareto dominate subsequent realized paths. These conditions are linked to the strict concavity of the log consumption profile. If the profile is strictly concave, the terminal self will always benefit more on the commitment path. For earlier selves, welfare differences between the two paths depend on the growth rates of future weighting factors. These differences are U-shaped functions of the future weighting factor at the longest delay. If the discount function is decreasing in the delay, the minima of these Us will always be such that the log consumption profile is concave at the beginning of the lifecycle.

It is worth mentioning that in this paper we model the household's choices in discrete time. A companion paper, Feigenbaum and Raei (2023), addresses the same issues in continuous time. We obtain analogous results in the two papers, but the two approaches are complementary in the sense that particular results are often more easily discerned in one framework than the other. As a result, the two papers have advanced simultaneously. A proposition arising from progress in continuous time has equivalents in both continuous and discrete time, leading to further advancements in discrete time, and vice versa.

The obvious advantage of using discrete time is its alignment with the preferences of most economists, as economic data is typically collected at discrete intervals. Another benefit is the finite specification of future weighting factors in discrete time, simplifying analyses. For instance, in a four-period model with two future weighting factors, it is straightforward to graphically identify the parameter space where the commitment path dominates.

More generally, in both discrete and continuous time, the terminal future weighting factor is crucial. However, discrete time allows for more intuitive characterizations of how behavior and welfare depend on this terminal future weighting factor in terms of derivatives with respect to it. In contrast, continuous time complicates this with an infinite parameter set and required smoothness assumptions, making it harder to isolate the impact of the terminal future weighting factor. Thus, results cannot be expressed simply in terms of partial derivatives with respect to the terminal future weighting factor as they can be here.

In discrete time it is also natural to define the future weighting factors in terms of the deviation of the discount function from the geometric discount function defined by the discount function at a unit delay. In continuous time, there is no natural time scale, so we can define future weighting factors relative to any exponential discount function. Consequently, results in discrete time are often more economical than in continuous time. For example, present bias in discrete time simply means the future weighting factors are increasing in the delay.[7]

Related literature This paper contributes to the sizeable literature which has been developed to address the "lifecycle consumption puzzle",[8] which refers to the discrepancy

[7]On the flip side, results in continuous time are expressed in terms of integrals that are more easily manipulated than the corresponding sums in discrete time.

[8]See Deaton (1992) and Browning and Crossley (2001) for more recent overviews.

between the empirical observation of a hump-shaped consumption profile and the smooth consumption over the lifespan predicted by the Lifecycle/Permanent-Income Hypothesis of Friedman and Modigliani (Modigliani and Brumberg (1954), Friedman (2018)). One strand of literature develops a set of solutions to this inconsistency by adding elements that are directly observable such as family-size effects (Attanasio et al. (1999), Attanasio and Browning (1993), Browning et al. (1985)), consumption-leisure trade-offs (Heckman (1974), Bullard and Feigenbaum (2007)), wage income uncertainty and the precautionary saving motive (Nagatani (1972), Hubbard et al. (1994), Carroll (1994), Carroll (1997), Gourinchas and Parker (2002)), mortality risk (Feigenbaum (2008), Hansen and Imrohoroglu (2008)), and consumer durable (Fernandez-Villaverde and Krueger (2011)).

Another set of mechanisms that can explain the hump in the consumption profile relax the assumptions on preferences of the standard rational paradigm, epitomized by Samuelson (1937). One of the most popular of these is to allow for time-inconsistent preferences by generalizing the discount function from an exponential function. Strotz (1955b) was the first to explore such deviations from Samuelson's model. Phelps and Pollak (1968) later proposed the hyperbolic function as a specific alternative to the exponential function, and David Laibson's dissertation (Laibson (1994)) offered hyperbolic discounting as a solution to the consumption hump puzzle. Today, this strand of the literature generally attributes such consumption humps to the concept of "present bias".[9] We contribute to this strand of literature by showing that present bias is a necessary but not sufficient condition for the lifecycle consumption profile to be hump-shaped.

Another issue related to present and future bias that has been the focus of a relatively recent literature pertains to welfare analysis. For individuals with time-inconsistent preferences, it is unclear which consumption profile or preference period should serve as the reference point for welfare comparison due to their varying choices over time. A common solution to this problem in the literature is to use the preferences of the initial self to evaluate welfare (see for example Laibson (1996), Laibson (1997), Laibson (1998), Laibson et al. (1998), O'Donoghue and Rabin (1999), O'Donoghue and Rabin, O'Donoghue and Rabin (2001) among many others). In fact, Caliendo and Findley (2019) show that commitment to the time-zero consumption plan can improve the objective function for all selves if the num-

[9]See Harris and Laibson (2013), Grenadier and Wang (2007), Cao and Werning (2018) and Mu et al. (2016) Feldstein (1985), Caliendo and Aadland (2007), Griffin et al. (2012), Hong and Hanna (2014). There are also papers that approach this puzzle by combining behavioral and more traditional factors, such as Campbell and Mankiw (1989), who explain hump-shaped wages by adding rule-of-thumb consumers to the economy.

ber of selves exceeds a certain threshold which turned out to be quite small in their setting. We contribute to this strand of literature by specifying the conditions on future weighting factors under which the commitment path will Pareto dominate the realized path in discrete time.[10] In other words, we investigate the conditions under which all of the different selves prefer the commitment path of consumption (the plan established in period zero) to the realized path (the actual consumption decisions).

This paper is organized in the following way. Section 2 describes the model environment including the general format for the discount function. Section 3 develops the condition on the discount function for a concave or convex log consumption profile. Section 4 explores the condition on the discount function under which commitment to the initial plan would Pareto dominate the realized plan and investigate the relationship between the concavity condition and Pareto dominance condition. Section 5 generalizes the quasihyperbolic discount function so the terminal future weighting factor is a free parameter and uses this example to demonstrate quantitatively the results of the previous sections. Finally, section 6 concludes.

2 Model environment

We focus on a finite-horizon life-cycle model in which households live for $T+1$ periods. The household earns income $y_t \geq 0$ at age t for $t = 0, ..., T$, which can be consumed c_t or saved as k_{t+1} at a fixed gross interest rate $R \geq 0$. In what follows, we present the household optimization problem to introduce the consumption path under commitment and the realized consumption path.[11]

[10] We only compare the preferences of the households' various selves regarding the commitment path and the realized path. We do not make any claims regarding Pareto efficiency as in Richter (2020), i.e. we do not compare how the various selves value these two paths relative to other feasible consumption paths.

[11] It is worth mentioning that similar to Drouhin (2020), in this paper we use the "choice-based" methodology which compares the solutions of dynamic programs with different decision dates. It is the methodology used originally by Strotz (1956), and now standard in behavioral macroeconomics, since the pioneering work of Laibson (1994), Laibson (1997), O'Donoghue and Rabin (1999). It is "choice based" because it not only uses a utility function that represents the preference relation but also imposes the budgetary constraints that the decision maker faces.

2.1 Household optimization problem

At time t, a household with existing saving k_t maximizes

$$U_t = \sum_{s=t}^{T} D_{s-t} \ln c_{s|t}$$

subject to

$$c_{s|t} + k_{s+1|t} = y_s + Rk_{s|t}, \quad s = t, \dots, T,$$

where $D_t \geq 0$ is the discount function, and $c_{s|t}$ and $k_{s+1|t}$ are consumption and saving at period s as planned in period t.[12] We will normalize $D_0 = 1$ and will also assume that $D_1 > 0$. The latter condition ensures that it will never be optimal for the household to consume all of its remaining wealth in the present period, which would leave the future selves with utility of $-\infty$. Note that the household will solve this problem with $k_{t|t} = k_t$ and $k_{T+1|t} = 0$. To simplify notation, we will assume the household begins with $k_0 = 0$.[13]

We can define the present value of the income stream from period t onward as h_t in the following way:

$$h_t = \sum_{s=t}^{T} \frac{y_s}{R^{s-t}}. \tag{1}$$

Note that we can rewrite h_t as the sum of current period income and the present value of the income stream from period $t+1$ onward

$$h_t = y_t + \sum_{s=t+1}^{T} \frac{y_s}{R^{s-t}} = y_t + \frac{h_{t+1}}{R} \tag{2}$$

for $t < T$. We can combine the period budget constraints from t to T into a lifetime budget constraint as of t:

$$\sum_{s=t}^{T} \frac{c_{s|t} + k_{s+1|t}}{R^{s-t}} = \sum_{s=t}^{T} \frac{y_s + Rk_{s|t}}{R^{s-t}}.$$

[12]The results are not qualitatively different for other CRRA utility functions, but they are more complicated so we only consider the logarithmic case. In solving the model we will proceed as though the household is naive about its time-inconsistency and does not know it will revise its plans as its preferences change. We could alternatively assume that the household is sophisticated about its time-inconsistency. However, with logarithmic period utility, the realized path will be the same under both assumptions and so will the commitment path. Thus there is no loss of generality between naivete and sophistication in the results documented here. For more discussion see Marin-Solano and Navas (2009).

[13]Our results easily generalize if the household is endowed with savings or debt at birth.

Using (1) and (2), this simplifies to

$$\sum_{s=t}^{T} \frac{c_{s|t}}{R^{s-t}} = h_t + Rk_t, \tag{3}$$

which shows the lifetime budget constraint.

The Lagrangian of the household problem at t can then be written as

$$L_t = \sum_{s=t}^{T} \left[D_{s-t} \ln c_{s|t} - \frac{\lambda_t c_{s|t}}{R^{s-t}} \right] + \lambda_t [h_t + Rk_t]. \tag{4}$$

Therefore, the first order condition (FOC) with respect to consumption will be

$$\frac{\partial L_t}{\partial c_{s|t}} = \frac{D_{s-t}}{c_{s|t}} - \frac{\lambda_t}{R^{s-t}} = 0. \tag{5}$$

The initial consumption plan $c_{s|0}$ that is determined at $t = 0$, the first period of life, will be referred to hereafter as the **commitment path**. Note, however, that unless the discount function is exponential the household will only follow the initial plan at $t = 0$. Indeed, at each period t of life, the household will choose a new plan $c_{s|t}$, but only the choice of consumption at t, $c_t = c_{t|t}$, will adhere to this plan. As the household progresses from period to period, its preferences will unexpectedly change since we are assuming that the household is naive about the change in its future preferences. When it gets to $t + 1$, it will then have saving $k_{t+1} = k_{t+1|t}$, but it will solve (4) anew, updated to $t + 1$. The resulting consumption profile c_t, determined at each period t, will be referred to as the **realized path**.

While the FOC (5) governs the whole commitment path for consumption $c_{s|0}$ from $s = 0, \ldots, T$, along the realized path only the FOC with $s = t$ will actually matter. For $s = t$, (5) simplifies to

$$\frac{D_{t-t}}{c_{t|t}} - \frac{\lambda_t}{R^{t-t}} = 0. \tag{6}$$

Since $c_t = c_{t|t}$ and $D_0 = 1$, (6) reduces to

$$\lambda_t = \frac{1}{c_t}.$$

The future plan $c_{s|t}$ at t is only relevant to the extent that it determines the Lagrange multiplier λ_t. Using this, we can rearrange (5) to obtain

$$c_{s|t} = \frac{D_{s-t}R^{s-t}}{\lambda_t} = D_{s-t}R^{s-t}c_t.$$

Inserting these into the lifetime budget constraint (3), we get

$$\sum_{s=t}^{T} \frac{D_{s-t}R^{s-t}c_t}{R^{s-t}} = h_t + Rk_t,$$

which reduces to

$$c_t = \frac{h_t + Rk_t}{\sum_{s=t}^{T} D_{s-t}}. \tag{7}$$

Hence, on the realized path, the budget constraint at period t can be written as

$$k_{t+1} = k_{t+1|t} = y_t + Rk_t - c_t = y_t + Rk_t - \frac{h_t + Rk_t}{\sum_{s=t}^{T} D_{s-t}}. \tag{8}$$

We can use this to calculate the effective Euler equation realized by the household for a general discounting function D_t with log utility:[14]

$$c_{t+1} = R\frac{\sum_{s'=t+1}^{T} D_{s'-t}}{\sum_{s=t+1}^{T} D_{s-t-1}}c_t. \tag{9}$$

As mentioned above, since $D_1 > 0$, c_{t+1} will be strictly positive.

[14]Combining (2) and (8), we get,

$$\begin{aligned} h_{t+1} + Rk_{t+1} &= R\left(\frac{h_{t+1}}{R} + y_t + Rk_t - \frac{h_t + Rk_t}{\sum_{s=t}^{T} D_{s-t}}\right) \\ &= R\left(h_t + Rk_t - \frac{h_t + Rk_t}{\sum_{s=t}^{T} D_{s-t}}\right) \\ &= R\left(\frac{\sum_{s'=t+1}^{T} D_{s'-t}}{\sum_{s=t}^{T} D_{s-t}}\right)(h_t + Rk_t). \end{aligned}$$

Updating (7) to $t+1$, consumption at $t+1$ is

$$c_{t+1} = R\left(\frac{\sum_{s'=t+1}^{T} D_{s'-t}}{\sum_{s=t}^{T} D_{s-t}}\right)\frac{h_t + Rk_t}{\sum_{s=t+1}^{T} D_{s-t-1}} = R\frac{\sum_{s'=t+1}^{T} D_{s'-t}}{\sum_{s=t+1}^{T} D_{s-t-1}}c_t.$$

In the special case of an exponential discount function $D_t = \delta^t$, the ratio

$$\mathcal{D}_t = \frac{\sum_{s'=t+1}^{T} D_{s'-t}}{\sum_{s=t+1}^{T} D_{s-t-1}}$$

simplifies to the constant δ, and we get back the familiar Euler equation $c_{t+1} = \delta R c_t$. More generally, though, for a nonexponential discount function, the inverse ratio $\mathcal{D}_t^{-1}$ measures the gross rate of change in the sum of the discount functions relevant for the remaining lifespan as the household moves from t to $t+1$. That is to say the change from the sum $D_1 + \cdots + D_{T-t}$ applicable at t to the sum $1 + \cdots + D_{T-t-1}$ applicable at $t+1$. The richer consumption dynamics that can be obtained in equilibrium with nonexponential discounting functions stems entirely from the deviation of the $\mathcal{D}_t$ from a constant, which will depend on how the discount function D_t deviates from an exponential function.

2.2 Future Weighting Discount Function

This section presents the future weighting discount function that we have developed in this paper and examines its characteristics. Given a discount function $D_t \geq 0$ for $t = 0, ..., T$, we define the "future weighting factor" ε_t via

$$D_t = D_1^t (1 + \varepsilon_t), \tag{10}$$

where D_1 is the discount factor for one period ahead. This future weighting factor captures the extra (or diminished, if negative) weight that we put on the discounting t periods in the future. Since we normalize $D_0 = 1$, by definition we will have $\varepsilon_0 = \varepsilon_1 = 0$. This general form of discounting function that we developed here can in fact accommodate various forms of discounters such as the standard geometric discounter, for which $D_t = \delta^t$; immediate successor agents, for whom $D_1 = \delta$ and $D_2 = D_3 = \cdots = 0$ (see, Lane and Mitra (1981), Leininger (1986) and Bernheim and Ray (1987)); and quasihyperbolic agents, for whom $D_t = \beta\delta^t$ (Laibson (1997)). In the latter case, the discount function will be present-biased if $\beta < 1$ and future-biased if $\beta > 1$.

To be concrete, as an example, let us consider the future weighting factor for the quasi-hyperbolic case. Since

$$D_1 = \beta\delta,$$

ε_t can be calculated as

$$\frac{D_t}{D_1^t} = \frac{\beta\delta^t}{\beta^t\delta^t} = \beta^{1-t} = 1 + \varepsilon_t.$$

Hence

$$\varepsilon_t = \beta^{1-t} - 1. \tag{11}$$

As another common example, for a myopic discounting function that vanishes for $t \geq t^*$, we have $\varepsilon_t = -1$ for $t \geq t^*$.

Note that if $\varepsilon_t = 0$ for all t, the discount function reverts to an exponential form. Thus, the future weighting factor, ε_t, serves as a measure of how the discount function deviates from an exponential at delay t. A positive ε_t results in higher future discount factors relative to the exponential case, meaning utility from consumption t periods in the future will be weighted more heavily than it would be under an exponential discount function. On the commitment path, the household will allocate a higher consumption to a period t with positive ε_t relative to the consumption level under an exponential discount function. Conversely, with $\varepsilon_t < 0$ the weight on utility consumption t periods in the future will be lower relative to an exponential discount function.[15]

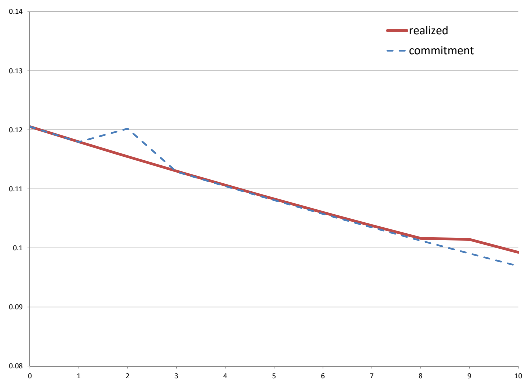
To have a better understanding of the role of ε_t in determining consumption behavior, figure 1 compares the consumption profile under the commitment path and the realized path for a model with $T = 10$. We consider two cases to demonstrate the role of an individual ε_t. First, we have a discount function for which ε_t is zero for all t except $t = 2$. Second, we have a discount function for which ε_t is zero for all t except $t = 8$.

In both plots, the blue dashed line shows the commitment path and the red solid line shows the realized path. In figure 1a, we see a spike in period two along the commitment path simply because $\varepsilon_2 > 0$ means that the household initially puts a higher weight on the utility from consuming two periods ahead compared to all other future periods which induced a higher consumption level for that period. Likewise, looking at figure 1b in which $\varepsilon_8 > 0$, the spike in the commitment path is at $t = 8$.

The effect of ε_t on the realized path is much more subtle than for the commitment path. With $\varepsilon_2 > 0$, shown in figure 1a, the household continually plans to have high consumption two periods ahead, as happens at $t = 2$ on the commitment path. However, with each new period, she reoptimizes and pushes forward when she intends to have high consumption. This trend continues until the household arrives at period nine of her lifetime, at which

[15] To be very precise, as we have defined the future weighting factor, we are talking about a departure from exponential discounting at the rate used between period 0 and 1. In discrete time, it is natural to think of the deviation of D_t from D_1^t, and this will yield some helpful simplifications.

Figure 1: consumption profile, commitment path and realized path dor a model with 10 periods



a . $\varepsilon_t = 0$ for all t except $t = 2$

b . $\varepsilon_t = 0$ for all t except $t = 8$

Note: on both graphs the horizontal axis is time (consumption period) and vertical access is the consumption level at each period.

point there no longer is a period two periods ahead. At this point, the household elevates its consumption for both the current and all future periods. Essentially, the additional consumption originally slated for two periods ahead is now dispersed across the remaining lifetime. Consequently, the realized consumption path is quite smooth, as it would be with exponential discounting, for $t < 9$. From this point, all future periods are discounted with the same rate. Consumption jumps up in these last two periods as she finally consumes the saving she accumulated to finance the planned extra consumption two periods ahead.

The same intuition applies to figure 1b in which $\varepsilon_8 > 0$. There, the future period with a higher discounting factor disappears after the second period. That is the reason why the realized consumption plan for $t \geq 3$ shifts upward. The high ε_8 disappears from her calculus once there no longer is a period eight periods ahead within her remaining time horizon. Consequently, she behaves like an exponential discounter thereafter, smoothing out over all the periods with $t \geq 3$ the extra consumption that she had previously intended, at $t = 2$, to save entirely for the last period.

Present (Future) bias in the context of future weighting discount function: As previously stated, the consumption-hump literature has conventionally described the impact of the discount function on the shape of the (log) consumption profile in relation to present (future) bias. By evaluating present and future bias using future weighting factors ε_t, we can gain fresh insights into the genesis of these ideas. Later, we will use this to demonstrate that possessing present bias alone does not necessarily lead to a consumption profile with a humped shape.

To commence, we will provide a brief overview of present bias and future bias, and then establish their implied conditions on future weighting factors, ε_t. A discount function exhibits present bias at $t > 0$ if it gives rise to the following type of preference reversal. Suppose for some allocation $\{c_t\}_{t=0}^T$, there exists $\xi_t > 0$ and $\xi_{t+1} \in (0, c_{t+1})$ such that the household would prefer at time 0 the original allocation over a forward-shifted allocation with c_t increased by ξ_t and c_{t+1} decreased by ξ_{t+1}. However, when the household gets to time t, it instead prefers the forward-shifted allocation over the original allocation. Thus the household would prefer not to shift consumption forward when the possibility of doing so is in the future, but it would opt to make that shift in the present. This is usually interpreted as the household putting an extra preference on consumption in the immediate present. On the other hand, future bias at $t > 0$ is defined similarly except the preference reversal goes the other way. The household would prefer the forward-shifted allocation over the original allocation when t is in the future, and prefers the original allocation when it reaches time t. We say a discount function is present-biased (future-biased) if it exhibits present (future) bias at all $t > 0$.

Assuming $D_s > 0$ for all s, we can define $m_s(t)$, as the perceived marginal rate of substitution between consumption at t and consumption at $t + 1$ as of time $s \leq t$ in the following way:[16]

$$m_s(t) = \frac{D_{t+1-s} u'(c_{t+1})}{D_{t-s} u'(c_t)}.$$

Now we can express the condition for preference reversals in terms of $m_s(t)$. Basically, the household will prefer the forward-shifted allocation at time 0 and the original allocation at t if

$$D_t u'(c_t)\xi_t - D_{t+1} u'(c_{t+1})\xi_{t+1} < 0 < u'(c_t)\xi_t - D_1 u'(c_{t+1})\xi_{t+1},$$

[16]While the concept of the marginal rate of substitution between consumption at t and $t+1$ is well-known (see, for example, FEI (2016)), our unique discount function necessitates a tailored definition to align with our future weighting approach.

which we can rearrange as

$$m_0(t) = \frac{D_1 u'(c_{t+1})}{u'(c_t)} < \frac{\xi_t}{\xi_{t+1}} < \frac{D_{t+1} u'(c_{t+1})}{D_t u'(c_t)} = m_t(t).$$

Therefore, the household will have a present bias at t if $m_0(t) < m_t(t)$. Since $\varepsilon_1 = 0$ by definition, by replacing D_t with $(1+\varepsilon_t)D_1$, we can rewrite $m_0(t) < m_t(t)$ as

$$1 < \frac{1+\varepsilon_{t+1}}{1+\varepsilon_t},$$

or equivalently

$$\varepsilon_t < \varepsilon_{t+1}.$$

We express this as the following lemma:[17]

Lemma 1. *A present-biased discount function will have* ***strictly increasing*** *and* ***positive*** *(for $t \geq 2$) future weighting factors.*

For our discussions throughout the rest of the paper, it will be helpful to define the future weighting growth factor

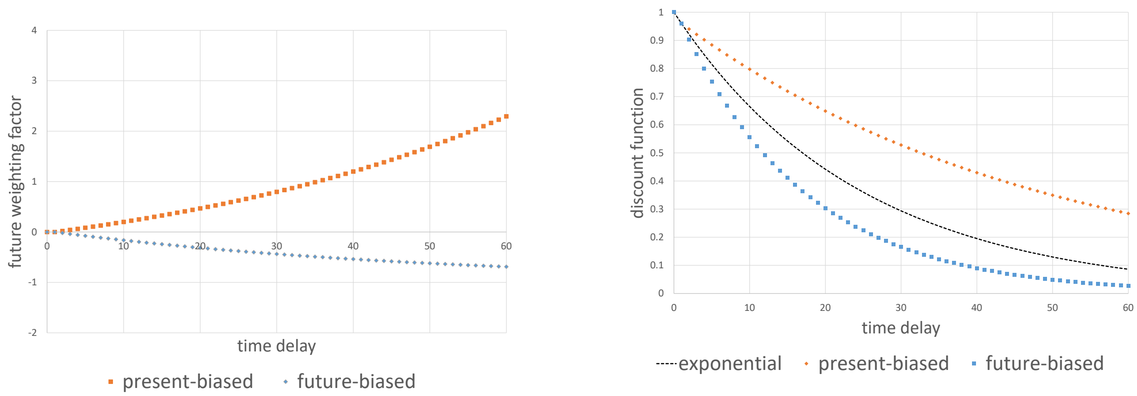
$$\phi_t = \frac{1+\varepsilon_{t+1}}{1+\varepsilon_t} \tag{12}$$

at t, assuming $\varepsilon_t > -1$, since many of our results depend on such ratios. Note that we have the theorem that $\phi_t \gtreqless 1$ if and only if $\varepsilon_{t+1} \gtreqless \varepsilon_t$. In the ϕ_t notation, a present-biased discount function will have $\phi_t > 1$ for $t > 0$. Conversely, a strictly positive and future-biased discount function will have strictly decreasing and negative (for $t \geq 2$) future weighting factors. To visualize this difference, we plot future weighting factors for both a present bias and a future-biased discount function in 2a. To put this in more graphical terms, a present-biased discount function will lie above the exponential function defined by the discounting between time delay 0 and time delay 1, and the divergence between the curves must increase with the time delay. A future-biased discount function will lie below the same exponential function, and the divergence between the curves must also increase. This is shown in 2a in which we compare a future-biased discount function along with a present-biased and an exponential one.[18]

[17]A related property of discount functions is increasing patience (Prelec (2004)). Since Prelec defines this concept in continuous time, we refer the reader to our companion paper in continuous time, Feigenbaum and Raei (2023), for an understanding of how it translates into a property of the future weighting factors.

[18]Note that a myopic discount function that is zero for t greater than equal to some $t^* > 1$ does not fit nicely into the categories of a present- or future-biased discount function because it does not satisfy the

Figure 2: Future-biased and present-biased discount function relative to an exponential discount function



a Each point in this graph represents ε_t for the relevant time delay.

b Each point in this graph represents the discount function for the relevant time delay defined as $D_1^t(1+\varepsilon_t)$

Note: These graphs help visualize the difference between a future-biased and present-biased discount factor relative to an exponential discount function, using our future weighting discount function framework.

3 Curvature of the log consumption profile

Empirically, lifecycle profiles of household consumption are hump-shaped, and time-inconsistency is often invoked as an explanation for this phenomenon. As we discussed in the previous section, ε_t is the parameter that controls the discounting weight of future periods. In this section, we explore how the value of the future weighting factor, ε_t, determines the curvature of the log consumption profile of the household. More precisely, we establish a necessary condition on ε_t under which the log consumption profile would be locally concave (convex) at age $T-t$. This in turn is a necessary condition for the consumption profile to have a local maximum at age $T-t$. [19]

caveat that the D_t are all positive, which is necessary for the marginal rate of substitution between c_t and c_{t+1} to be defined. There will be a future bias at t^*-1 since at time zero the household would prefer not to consume anything at t^*, but its (t^*-1)-utility is only defined if $c_{t^*}>0$. On the other hand, there will be a weak present bias at $t \geq t^*$ since at time zero the household will be indifferent between how it allocates consumption between t and $t+1$. However, at time t the household will prefer to have more consumption at t.

[19]If the consumption profile has a local maximum at t^*, it will, of course, also be necessary to have $\frac{c_{t^*}}{c_{t^*-1}} > 1 > \frac{c_{t^*+1}}{c_{t^*}}$. However, the main hurdle is constructing a model where the growth rate of consumption

As a first step, we will rewrite the Euler equation in terms of the future weighting factor. Replacing the general form of discounting function D_t in the household's Euler equation (9) with the form involving the future weighting discounting function (10) gives us

$$c_{t+1} = D_1 R \frac{\sum_{s'=t+1}^{T} D_1^{s'}(1+\varepsilon_{s'-t})}{\sum_{s=t+1}^{T} D_1^{s}(1+\varepsilon_{s-t-1})} c_t. \tag{13}$$

In this form, it is more apparent that the Euler equation reduces to the usual $c_{t+1} = D_1 R c_t$ when we have an exponential discounting function and $\varepsilon_2 = \varepsilon_3 = \cdots = \varepsilon_T = 0$. Alternatively, by setting $z = s - t$, we can rewrite this Euler equation (13) as

$$\frac{c_{t+1}}{c_t} = D_1 R \frac{\sum_{z'=1}^{T-t} D_1^{z'}(1+\varepsilon_{z'})}{\sum_{z=1}^{T-t} D_1^{z}(1+\varepsilon_{z-1})}. \tag{14}$$

We are interested in the log consumption profile, which will be concave if $\log(\frac{c_{t+1}}{c_t})$ decreases with t. We can take logs of both sides of the Euler equation (14) and difference it to obtain

$$\Delta \ln c_t = \ln(D_1 R) + \ln\left(\frac{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})}{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}\right). \tag{15}$$

Similarly, we can define the second-order difference

$$\Delta^2 \ln c_t = \ln\left(\frac{\sum_{z'=1}^{T-t-1} D_1^{z'}(1+\varepsilon_{z'})}{\sum_{z=1}^{T-t-1} D_1^{z}(1+\varepsilon_{z-1})}\right) - \ln\left(\frac{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})}{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}\right),$$

which simplifies to

$$\Delta^2 \ln c_t = \ln\left(\frac{\sum_{z'=1}^{T-t-1} D_1^{z'}(1+\varepsilon_{z'})}{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})} \frac{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}{\sum_{z=1}^{T-t-1} D_1^{z}(1+\varepsilon_{z-1})}\right). \tag{16}$$

The log consumption profile will be concave iff $\Delta^2 \ln c_t \leq 0$ for $t = 0, ..., T-2$. If $\Delta^2 \ln c_t < 0$ for all $t = 0, ..., T-2$, then the log consumption profile will be strictly concave. The reverse inequalities will yield convex and strictly convex profiles.[20]

changes. Adjusting the model so we quantitatively get growth rates both above and below 1 is a matter of calibration. In a partial-equilibrium environment where R is a free parameter, this is trivial. In a general-equilibrium environment, it is more challenging but still less of an issue than getting a concave profile in the first place.

[20]Unlike in continuous time, for the log consumption profile to be strictly concave (convex) at $t+1$ we must have $\Delta^2 \ln c_t$ be negative (positive). If the second difference vanishes, the profile must be locally linear.

Note that if we set the future weighting factors all to zero in (16), the argument of the logarithm is clearly one, so all of the surviving terms on the right-hand side are of first or higher order in the ε_t, corroborating again that the log consumption profile with an exponential discounting function is exactly linear. Any deviation from linearity is driven by the future weighting factors.

As we mentioned, the concavity of log consumption profile requires $\Delta^2 \ln c_t \leq 0$ for $t = 0, ..., T-2$, which implies that the log consumption is concave at $t+1$ if

$$\frac{\sum_{z'=1}^{T-t-1} D_1^{z'}(1+\varepsilon_{z'})}{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})} \frac{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}{\sum_{z=1}^{T-t-1} D_1^{z}(1+\varepsilon_{z-1})} \leq 1, \tag{17}$$

which can be simplified to[21]

$$\varepsilon_{T-t} \geq \frac{\sum_{s'=0}^{T-t-2} D_1^{s'}(1+\varepsilon_{s'+1})}{\sum_{s=0}^{T-t-2} D_1^{s}(1+\varepsilon_{s})}(1+\varepsilon_{T-t-1}) - 1. \tag{18}$$

We want to emphasize that all of the future weighting factors on the right-hand side of the above inequality (18) are at delays shorter than $T-t$. Thus the exact condition for concavity of the consumption profile at $t+1$ is a lower bound on ε_{T-t} that depends on future weighting factors at shorter delays.

If $\varepsilon_{T-t-1} = -1$, so $D_{T-t-1} = 0$, there are two possibilities in terms of the shape of the log consumption profile at $t+1$. These depend on ε_{T-t}. If $\varepsilon_{T-t} = -1$ too, then $\Delta \ln c_t = \Delta \ln c_{t+1}$, and the log consumption profile will be linear (and thus both weakly concave and weakly convex) in the vicinity of $t+1$. If, on the other hand, $\varepsilon_{T-t} > -1$, $\Delta \ln c_t > \Delta \ln c_{t+1}$, and the log consumption profile will be strictly concave in the vicinity of $t+1$.

If, on the other hand, $\varepsilon_0, \ldots, \varepsilon_t > -1$, we can define the average future weighting growth factor

$$\overline{\phi}_t = \frac{\sum_{s=0}^{t} D_s \phi_s}{\sum_{s'=0}^{t} D_{s'}}. \tag{19}$$

Then we can conveniently express the following result.

Proposition 2. *If $\varepsilon_s > -1$ for all $s = 0, \ldots, T-t-1$, the log consumption profile will be strictly concave locally at $t+1$ iff*

$$\phi_{T-t-1} > \overline{\phi}_{T-t-2} \tag{20}$$

[21] See appendix A for details on this calculation.

or equivalently that

$$\phi_{T-t-1} > \overline{\phi}_{T-t-1}. \tag{21}$$

The profile will be strictly convex locally if the inequalities are reversed.

It follows from (18) using (12) that strict concavity at $t+1$ requires $\phi_{T-t-1} > \overline{\phi}_{T-t-2}$. The second inequality then follows since $\overline{\phi}_t$ is a weighted average, so $\overline{\phi}_t \lesseqgtr \overline{\phi}_{t-1}$ iff $\phi_t \lesseqgtr \overline{\phi}_{t-1}$.

So the concavity condition at $t+1$ is that ϕ_{T-t-1} is bigger than a weighted average of the ϕ_s for $s = 0, ..., T-t-2$, where the weights are the D_s. That is to say, the log consumption profile will be concave when there are s periods remaining if and only if the future weighting growth factor at s is bigger than a weighted average of the future weighting growth factor at shorter delays.

As we showed in the previous section, a present-biased discount function will have $\phi_t > 1$ for all $t > 0$. Given the assumption of $\varepsilon_0 = \varepsilon_1 = 0$, we have $\phi_0 = 1$. Therefore, local concavity imposes a stronger condition on the shape of consumption profile compared to present bias.

Since a weighted average of a heterogeneous set must be less than the maximum in the set and greater than the minimum in the set, it follows immediately from Proposition 2 that if the ϕ_t are strictly increasing (decreasing) then the log consumption profile will be strictly concave (convex). Moreover, if the ϕ_t are increasing with $\phi_1 > 1$, the log consumption profile will be strictly concave. Likewise, if the ϕ_t are decreasing with $\phi_1 < 1$, the log consumption profile will be strictly convex.

Corollary 3. *For the entire log consumption profile to be strictly concave (convex), the $\Delta\varepsilon_s$ from $s = 1, \ldots, T-1$ must all be positive (negative) and the ϕ_s must all be greater (less) than the weighted average of previous ϕ_s, where the weight is the discount factors. This implies that present bias is a necessary and <u>not</u> sufficient condition for the log consumption profile to be strictly concave.*

This proposition can be proved by induction. Suppose the $\phi_i > 1$ for $s = 1, \ldots, s-1$. Then (20) implies $\phi_s > 1$, and $\varepsilon_{s+1} = \varepsilon_s + \Delta\varepsilon_s > \varepsilon_s > 0$. Note also that each successive iteration of (20) is the necessary condition for the log consumption profile to be concave one period earlier. Therefore, the condition that follows is necessary to ensure a strictly concave log consumption profile between $t = 0$ and $t = 2$.

$$\varepsilon_T > \frac{1}{D_1} \frac{\sum_{s'=1}^{T-1} D_1^{s'}(1+\varepsilon_{s'})}{\sum_{s=0}^{T-2} D_1^{s}(1+\varepsilon_s)}(1+\varepsilon_{T-1}) - 1.$$

Iterating forward in time, each log consumption growth ratio will depend on one more difference $\Delta\varepsilon_s$ than the ensuing log consumption growth ratio. Therefore, it is necessary to have $\Delta\varepsilon_s > 0$, or equivalently $\varepsilon_{s+1} > \varepsilon_s$, for the log consumption growth ratio to decrease over time. One interpretation of this outcome is that what is commonly referred to as present bias can be viewed as young households placing more weight on consumption in distant future. However, as the future approaches the present, consumption in these future periods gradually matters less to the household. As a result, the ε_t must increase with t because this implies that the extra weight associated with a specific age decreases as that age approaches and the delay time gets shorter.

4 Pareto dominance of the commitment path

In the preceding sections, we presented the household problem with a discounting function that is contingent upon the time between the consumption and the time at which the current self lives, rather than the calendar time at which consumption occurs. This gives rise to time-inconsistent preferences, where the marginal rate of substitution between consumption at different times is dependent on when the household evaluates the utility of these consumptions. Consequently, the household at different ages will value consumption plans differently. This multiplicity of selves can substantially complicate welfare analysis.

A common solution to tackle this complication in the literature is to use the preferences of the initial self to evaluate welfare. See, for example, Laibson (1997, 1996), Laibson et al. (1998), and O'Donoghue and Rabin, 1999, 2001). However, this method is not without criticism. For example, Dewatripont et al. (2004) states that there is "no normative foundation" for equating welfare with time-zero preferences.

A more recent literature explores conditions that can be imposed on the discount function under which committing to the initial plan of the time-zero self improves the welfare of all selves over the life cycle as compared to what they would actually obtain over the lifecycle. This approach provides a justification for singling out the preferences of the time-zero self. As an example, Caliendo and Findley (2019) show that with quasihyperbolic discounting commitment to the time-zero consumption plan can improve the objective function for all selves if the number of selves exceeds a certain threshold that turned out to be quite small in their setting.

In this section, we use our setup to explore conditions on our general formulation of the discount function under which committing to the initial plan will Pareto dominate the realized plan.

To provide a preview of our findings in this section, we show that the welfare difference between the realized path and the commitment path for each self is a U-shaped function of the terminal future weighting factor. If the terminal future weighting factor is large enough in magnitude, whether positive or negative, most of the selves will prefer the realized path to the commitment path. The last self, however, is an exception. If the log consumption profile is strictly concave, which will, as shown in the previous section, happen if all of the future weighting factors are sufficiently large relative to future weights at shorter delays the last self will prefer the commitment path.

These results suggest one does need to be careful when performing welfare analysis with time-inconsistent preferences. Absent other information, it will not be obvious that the consumption path chosen by the initial self will necessarily be the best choice to serve as the benchmark for the purpose of welfare exercises. On the contrary, for most of the parameter space of possible discount functions, neither the commitment path nor the realized path will Pareto dominate the other, which will leave us in the situation that Dewatripont et al. (2004), for example, complained about.

Deriving the expression for ΔU_τ: As a starting point for evaluating the welfare of different selves, we derive an expression for the difference between realized utility and commitment utility. The realized utility as of time τ is simply the realized value of the household's objective function at time τ, which we have already dealt with in previous sections:

$$U_\tau^* = \sum_{t=\tau}^{T} D_{t-\tau} \ln(c_t).$$

In contrast, the commitment utility at time τ is

$$U_\tau^c = \sum_{t=\tau}^{T} D_{t-\tau} \ln(c_{t|0}), \tag{22}$$

which is what you obtain if you insert the original consumption path as of time 0 into the objective function at time τ. What concerns us is ΔU_τ which is the difference in utility between the realized plan and the original plan at time τ:

$$\Delta U_\tau = U_\tau^* - U_\tau^c = \sum_{t=\tau}^{T} D_{t-\tau} \ln\left(\frac{c_t}{c_{t|0}}\right). \tag{23}$$

If $\Delta U_\tau > 0$, then following the realized consumption plan provides the household at age τ with a higher utility compared to the initial plan. Conversely, if $\Delta U_\tau < 0$, then committing to the initial plan is optimal for the household at age τ.

By definition, the commitment path must maximize lifetime utility at $t = 0$, so we must have $\Delta U_0 \leq 0$. We will say that the commitment path Pareto dominates the realized path if, for all $\tau = 0, \ldots, T$, $\Delta U_\tau \leq 0$ and if, for some $s \in 0, \ldots, T$, $\Delta U_s < 0$. This Pareto dominance provides a compelling justification for helping the household commit to the initial path without having the policy maker impose her norms about which selves matter more to the household. We will also use the term "almost Pareto dominates", which we define as follows. If $\Delta U_\tau \geq 0$ for all $\tau \in 1, \ldots, T$ then the realized path will almost Pareto dominate the commitment path. Note that the realized path can never Pareto dominate the commitment path.

First let us explore ΔU_τ for the exceptional case of a discount function that is not strictly positive, including the case of a myopic discount function. Let $t_* = \min\{t \in \{2, ..., T\} : D_t = 0\}$. Then the commitment consumption rule will be $c_{t_*} = 0$. This does not cause anything pathological for U_0 since $D_{t_*} \ln(c_{t_*}) = 0$. However, $D_{t_*-1} \ln(c_{t_*}) = -\infty$, so $U_\tau^c = -\infty$ for $\tau = 1, \ldots, t_*$. In contrast, since we have assumed $D_1 > 0$, the realized path of consumption will be positive for all t, and U_τ^* will be finite for all τ. Thus the commitment path cannot Pareto dominate the realized path. In the myopic case where $D_t = 0$ for all $t \geq t_*$, the realized path will almost Pareto dominate the commitment path, which implies using the commitment path as a reference point for welfare analysis is not appropriate.

For the remainder of this section, we will assume the discount function is strictly positive, so the $\varepsilon_t > -1$ for all t, and the ϕ_t are all defined.

We will begin by simplifying the expression for ΔU_τ. Note that for both paths, we have

$$c_t = c_0 \prod_{s=0}^{t-1} \frac{c_{s+1}}{c_s}, \tag{24}$$

and

$$c_{t|0} = c_0 \prod_{s=0}^{t-1} \frac{c_{s+1|0}}{c_{s|0}}. \tag{25}$$

Thus we can rewrite (23) as

$$\Delta U_\tau = \sum_{t=\tau}^{T} \sum_{s=0}^{t-1} D_{t-\tau} \left[\ln\left(\frac{c_{s+1}}{c_s}\right) - \ln\left(\frac{c_{s+1|0}}{c_{s|0}}\right) \right]. \tag{26}$$

This is convenient because we have previously specified the evolution of the realized path in terms of the effective Euler equation (14). The initial plan $c_{t|0}$, i.e. the consumption at period t as determined at period 0, can be obtained from (6):

$$c_{t|0} = D_t R^t c_0 = D_1^t (1 + \varepsilon_t) R^t c_0. \tag{27}$$

Thus consumption growth from t to $t+1$ along the commitment path simplifies to

$$\frac{c_{t+1|0}}{c_{t|0}} = D_1 R \frac{1 + \varepsilon_{t+1}}{1 + \varepsilon_t}. \tag{28}$$

Combining these expressions, (14) and (28), for consumption growth along the two paths, the difference in utility at age τ becomes

$$\Delta U_\tau = \sum_{t=\tau}^{T} \sum_{s=0}^{t-1} D_{t-\tau} \left[\ln \left(\frac{\sum_{z=1}^{T-s} D_1^z (1 + \varepsilon_z)}{\sum_{z'=1}^{T-s} D_1^{z'} (1 + \varepsilon_{z'-1})} \right) - \ln \left(\frac{1 + \varepsilon_{s+1}}{1 + \varepsilon_s} \right) \right], \tag{29}$$

which we can rewrite as

$$\Delta U_\tau = \sum_{t=\tau}^{T} \sum_{s=0}^{t-1} D_{t-\tau} \left[\ln \left(\frac{\sum_{z=0}^{T-s-1} D_1^z \phi_z}{\sum_{z'=0}^{T-s-1} D_{z'}} \right) - \ln \phi_s \right].$$

Using (19), this simplifies to

$$\Delta U_\tau = \sum_{t=\tau}^{T} \sum_{s=0}^{t-1} D_{t-\tau} \ln \left(\frac{\overline{\phi}_{T-s-1}}{\phi_s} \right). \tag{30}$$

Connecting concavity condition to the sign of ΔU_τ; Recall that the condition for strict concavity at t is that $\phi_{T-t} > \overline{\phi}_{T-t}$. However, if the log consumption profile is everywhere strictly concave (30) does not generally imply that $\Delta U_\tau < 0$ since the subscripts of ϕ and $\overline{\phi}$ are different. This difference arises from how the future weighting factors affect the two paths. The commitment path is obtained by iterating (28), so $c_{t|0}$ depends on $\phi_1, \dots, \phi_{t-1}$. In contrast, the realized consumption at t depends on the future weighting factors $\varepsilon_2, \dots, \varepsilon_{T-t}$ that still affect the household's problem at age t. The combined effect of the latter is conveyed by $\overline{\phi}_{T-t}$ instead of $\overline{\phi}_t$.

Nevertheless, for $\tau = T$, i.e. the terminal period,

$$\begin{aligned}\Delta U_T &= \sum_{s=0}^{T-1} D_0 \left[\ln \overline{\phi}_{T-s-1} - \ln \phi_s\right] \\ &= \sum_{s=0}^{T-1} \ln\left(\frac{\overline{\phi}_s}{\phi_s}\right).\end{aligned}$$

Thus strict concavity of the log consumption profile does imply that $\Delta U_T < 0$. Likewise, strict convexity implies that $\Delta U_T > 0$.

For $\tau < T$, strict concavity of the log consumption profile is not sufficient to unambiguously sign the whole sum in (30). As we describe in detail in Feigenbaum and Raei (2023), we can decompose (30) into terms that can be unambiguously signed if the log consumption profile is strictly concave, but we will have terms of both signs.

The role of the terminal future weighting factor; in this discrete-time context, we can characterize how the ΔU_τ depend on ε_T.[22] Subsequently, we can analyze the implications for the sign of ΔU_τ that arise from a strictly concave log consumption profile in relation to this characterization.

For $\tau \geq 1$, ΔU_τ only depends on ε_T through its dependence on ϕ_{T-1} and $\overline{\phi}_{T-1}$. We can rewrite (30) as

$$\Delta U_\tau = \sum_{t=\tau}^{T} \sum_{i=0}^{t-1} D_{t-\tau} \ln \overline{\phi}_{T-t+i} - \sum_{t=\tau}^{T} \sum_{i=0}^{t-1} D_{t-\tau} \ln \phi_i. \tag{31}$$

Note that the $D_{t-\tau}$ that appear in this equation will never depend on ε_T for $\tau > 0$, and we do not need to consider ΔU_0 since it must, by definition, be nonpositive.

To differentiate (31) with respect to ε_T, it will be helpful to compute the partial derivatives of the ϕ_t and $\overline{\phi}_t$. From (12), the former is

$$\frac{\partial \phi_t}{\partial \varepsilon_T} = \frac{\delta_{t,T-1}}{1 + \varepsilon_{T-1}}, \tag{32}$$

where δ_{ij} is the Kronecher delta, equaling 1 when i and j are the same and 0 otherwise. Likewise, the latter is, by (19),

$$\frac{\partial \overline{\phi}_t}{\partial \varepsilon_T} = \frac{1}{1 + \varepsilon_{T-1}} \frac{D_{T-1}\delta_{t,T-1}}{\sum_{s=0}^{T-1} D_s}. \tag{33}$$

[22]Note that in continuous time, we cannot simply isolate the effect of ε_T on the ΔU_τ.

Then, the corresponding derivatives of $\ln \phi_t$ and $\ln \overline{\phi}_t$ are

$$\frac{\partial \ln \phi_t}{\partial \varepsilon_T} = \frac{\delta_{t,T-1}}{1+\varepsilon_T} \tag{34}$$

and

$$\frac{\partial \ln \overline{\phi}_t}{\partial \varepsilon_T} = \frac{1}{1+\varepsilon_{T-1}} \frac{D_{T-1}\delta_{t,T-1}}{\sum_{s=0}^{T-1} D_s \phi_s}. \tag{35}$$

All four of these partial derivatives are nonnegative. Since both terms in (31) include contributions from ϕ_{T-1} and $\overline{\phi}_{T-1}$ that are strictly positive, this means that the first term, which accrues from the realized utility, is unambiguously positive while the second term, which accrues from the subtraction of the commitment utility, is unambiguously negative. This property that $\frac{\partial \phi_t}{\partial \varepsilon_s}$ and $\frac{\partial \overline{\phi_t}}{\partial \varepsilon_s}$ are nonnegative for all $t = 1, \ldots, T-1$ is unique to $s = T$, which elucidates the reason that ε_T is of special significance of all the future weighting factors.

An increase in ε_T will generate a spike in consumption at the end of life on the commitment path that will add to the commitment utility of all the household's selves. For the singular case of $\tau = T$, this spike will unambiguously decrease ΔU_T since only c_T matters for the welfare of the final self.[23] The initial self will save more to finance this spike in terminal consumption. However, the intermediate selves will end up diverting some of this additional saving to consumption at other ages. Thus an increase in ε_T will increase $c_{T|0}$ more than c_T, resulting in a net decrease of ΔU_T, but this is accomplished by decreasing the $c_{t|0}$ for $t < T$. This latter effect can make the ΔU_τ positive for $\tau < T$, rendering ΔU_τ nonmonotonic.

[23]This accounts for why strict concavity alone can guarantee that $\Delta U_T < 0$ since strict concavity imposes a strictly positive upper bound on ε_T that depends on the other future weighting factors, which must also be positive. This is a tighter bound on ε_T than what we need to get a negative ΔU_T.

Let us now focus on the case of $\tau < T$. Partially differentiating (31) with respect to ε_T,

$$\begin{aligned}
\frac{\partial \Delta U_\tau}{\partial \varepsilon_T} &= \frac{1}{1+\varepsilon_{T-1}} \sum_{t=\tau}^{T} \sum_{i=0}^{t-1} D_{t-\tau} \frac{D_{T-1}\delta_{T-t+i,T-1}}{\sum_{s=0}^{T-1} D_s \phi_s} - \frac{1}{1+\varepsilon_T} \sum_{t=\tau}^{T} \sum_{i=0}^{t-1} D_{t-\tau}\delta_{i,T-1} \\
&= \frac{1}{1+\varepsilon_{T-1}} \sum_{t=\tau}^{T} D_{t-\tau} \frac{D_{T-1}}{\sum_{s=0}^{T-1} D_s \phi_s} - \frac{D_{T-\tau}}{1+\varepsilon_T} \\
&= \sum_{t=\tau}^{T} D_{t-\tau} \frac{D_1^{T-1}}{\sum_{i=0}^{T-1} D_1^i (1+\varepsilon_{i+1})} - \frac{D_{T-\tau}}{1+\varepsilon_T} \\
&= D_1^{T-\tau} \left[\frac{\sum_{t=\tau}^{T} D_1^{t-1}(1+\varepsilon_{t-\tau})}{\sum_{i=0}^{T-1} D_1^i (1+\varepsilon_{i+1})} - \frac{1+\varepsilon_{T-\tau}}{1+\varepsilon_T} \right] \\
&= D_1^{T-\tau} \frac{(1+\varepsilon_T)\sum_{t=\tau}^{T} D_1^{t-1}(1+\varepsilon_{t-\tau}) - (1+\varepsilon_{T-\tau})\sum_{i=0}^{T-1} D_1^i (1+\varepsilon_{i+1})}{(1+\varepsilon_T)\sum_{i=0}^{T-1} D_1^i (1+\varepsilon_{i+1})}
\end{aligned}$$

and finally, after cancelling like terms in the numerator, we obtain

$$\frac{\partial \Delta U_\tau}{\partial \varepsilon_T} = D_1^{T-\tau} \frac{(1+\varepsilon_T)\sum_{t=\tau}^{T-1} D_1^{t-1}(1+\varepsilon_{t-\tau}) - (1+\varepsilon_{T-\tau})\sum_{i=0}^{T-2} D_1^i (1+\varepsilon_{i+1})}{(1+\varepsilon_T)\sum_{i=0}^{T-1} D_1^i (1+\varepsilon_{i+1})}. \tag{36}$$

Notice that the numerator is linear in ε_T. Moreover, given our assumptions that $D_1 > 0$ and the $\varepsilon_t > -1$, both the whole denominator and the coefficient of ε_T in the numerator are positive.[24]

Proposition 4. *For $\tau = 1, \ldots, T-1$, the difference ΔU_τ of lifetime welfare between the realized path and the commitment path is a U-shaped function of the future weighting factor at the longest delay, ε_T. The global minimum of ΔU_τ as a function of ε_T is*

$$\underline{\varepsilon}_T^\tau \equiv \frac{D_{T-\tau}}{D_{T-1}} \frac{\sum_{i=0}^{T-2} D_i}{\sum_{t=0}^{T-\tau-1} D_t} \bar{\phi}_{T-2}(1+\varepsilon_{T-1}) - 1. \tag{37}$$

Thus ΔU_τ is a strictly decreasing function of ε_T for $\varepsilon_T < \underline{\varepsilon}_T^\tau$ and ΔU_τ is strictly increasing for $\varepsilon_T > \underline{\varepsilon}_T^\tau$.

This follows immediately from (36). The location of the global minimum is obtained by solving for the ε_T that sets the numerator of (36) to zero. See appendix B for details.

This proposition will play a critical role in our objective of characterizing the region of

[24]For $\tau = T$, the numerator does not depend on ε_T and with those assumptions is strictly negative. Thus ΔU_T is strictly decreasing in ε_T.

the parameter space where the commitment path Pareto dominates the realized path, which requires the values of ΔU_τ to all be negative. Since, as a function of ε_T, ΔU_τ is U-shaped, the set of ε_T where $\Delta U_\tau \leq 0$ must be a neighborhood of the global minimum $\underline{\varepsilon}_T^\tau$.[25]

Notice that if $\varepsilon_2 = \cdots = \varepsilon_{T-1} = 0$, we will have $\phi_1 = \cdots = \phi_{T-2} = 1$, and consequently (37) simplifies to

$$\underline{\varepsilon}_T^\tau = \frac{D_{T-\tau}}{D_{T-1}} - 1 = D_1^{1-\tau} - 1.$$

In the canonical case where $D_1 < 1$, so the discount function is decreasing (except perhaps from D_{T-1} to D_T), this means that ΔU_τ will be decreasing with respect to ε_T at $\varepsilon_T = 0$ if $\tau > 1$, but ΔU_1 is minimized with respect to ε_T at $\varepsilon_T = 0$. In fact we can show that ΔU_1 is flat in all directions when the future weighting factors all vanish.

Lemma 5. *The gradient* $\nabla \Delta U_1 = 0$ *when* $\varepsilon_2 = \cdots = \varepsilon_T = 0$.

The proof is in appendix D. This is a consequence of the fact that the commitment and realized plans are the same at $t = 0$ so $c_0 = c_{0|0}$. We have

$$\begin{aligned} D_1 \Delta U_1 &= D_1 \sum_{t=1}^{T} D_{t-1} \ln\left(\frac{c_t}{c_{t|0}}\right) \\ &= \ln\left(\frac{c_0}{c_{0|0}}\right) + \sum_{t=1}^{T} D_t \ln\left(\frac{c_t}{c_{t|0}}\right) + \sum_{t=1}^{T} (D_1 D_{t-1} - D_t) \ln\left(\frac{c_t}{c_{t|0}}\right) \\ &= \Delta U_0 + \sum_{t=1}^{T} (D_1 D_{t-1} - D_t) \ln\left(\frac{c_t}{c_{t|0}}\right). \end{aligned}$$

When the future weighting factors all vanish, both factors, $(D_1 D_{t-1} - D_t)$ and $\ln\left(\frac{c_t}{c_{t|0}}\right)$, of the last term vanish at the origin so partial derivatives of this last term also vanish at the origin. Consequently, the gradient of ΔU_1 is proportional to the gradient of ΔU_0. Since the initial self must prefer the commitment path, ΔU_0 must be maximized at the origin. Therefore, its gradient must vanish, and the gradient of ΔU_1 must also vanish. This intuition does not extend to later τ because $\ln\left(\frac{c_\tau}{c_{\tau|0}}\right)$ for $\tau \geq 1$ only vanishes at the origin, so it has a nonzero gradient.

However, while the gradient of ΔU_1 must vanish at the origin, ΔU_1 differs from ΔU_0 in that the origin is a global maximum of ΔU_0 whereas it is not a global maximum of ΔU_1.

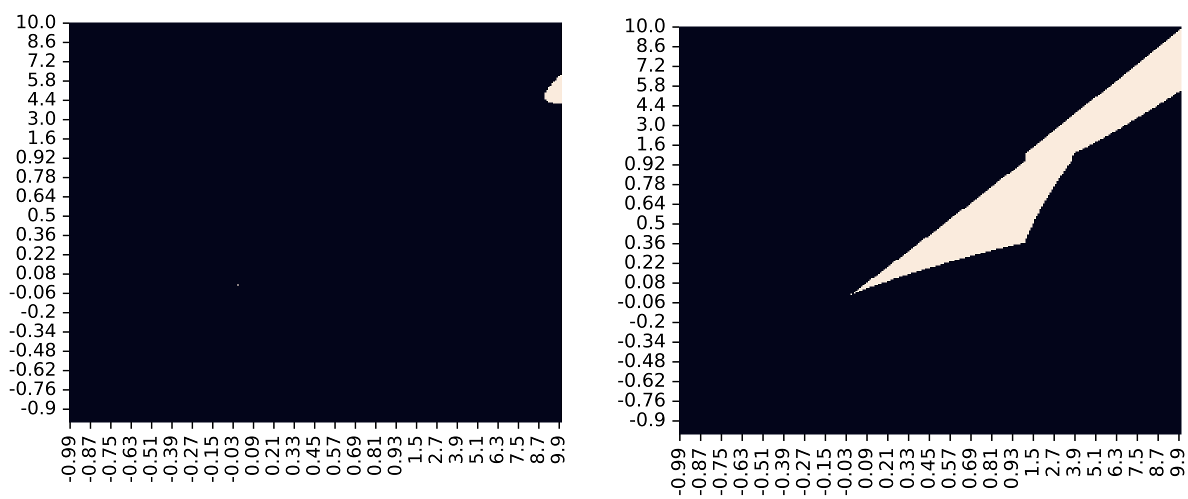
[25] Additional conditions will be necessary to guarantee that ΔU_τ is in fact negative in the neighborhood at this minimum.

On the contrary, we have already demonstrated that ΔU_1 is minimized with respect to ε_T at the origin. Thus the Hessians of ΔU_0 and ΔU_1 differ at the origin.

To demonstrate this, in appendix E we calculate the Hessian of ΔU_1 at the origin for $T = 3$. The diagonal elements are both positive, so the second self will prefer the realized path both if $\varepsilon_2(\neq 0)$ is small in magnitude while $\varepsilon_3 = 0$ and if $\varepsilon_3 \neq 0$ while $\varepsilon_2 = 0$. Nevertheless, the determinant of the Hessian is $D_1^3(1 - D_1^3)$. In the normal case where $D_1 < 1$, the Hessian of ΔU_1 will be positive definite, and the second self will prefer the realized path over the commitment path for any small deviation of one or both future weighting factors from zero. On the other hand, if $D_1 > 1$, ΔU_1 has a saddlepoint at the origin.

In figure 3, we show for two calibrations of D_1 graphs of a neighborhood of the origin in which white pixels show pairs $(\varepsilon_3, \varepsilon_2)$ for which the commitment path Pareto dominates the realized path whereas the black pixels correspond to pairs where at least one self prefers the realized path. In both cases we imagine a period is twenty years so a total life span is 80 years. In 3a, $D_1 = 0.44$, or 0.96 in annual terms. In 3b, it is 1.42, or 1.02 in annual terms.

Figure 3: Pixel plot of the combinations of ε_2 and ε_3 for which the commitment path Pareto dominate the realized path



a . for $D_1 = 0.96$

b . for $D_1 = 1.02$

Note: on both graphs, the bright area shows the region that Pareto condition holds. We have $\varepsilon_2 \in [-1, 10]$ on the y axis and $\varepsilon_3 \in [-1, 10]$ on the x axis.

As we can see in figure 3, Pareto dominance of the commitment path holds over a larger subset of the parameter space when D_1 is 1.02 as opposed to 0.96. In both calibrations the subset where the commitment path Pareto dominates the realized path lies entirely within the first quadrant, where both ε_2 and ε_3 are positive. However when D_1 is 1.02, the subset radiates from the origin. When D_1 is 0.96, Pareto dominance only occurs when ε_2 and ε_3 are both large and positive. Our result for the Hessian explains the difference between figures 3a and 3b. Note that in both cases of D_1, the sign of ΔU_1 is the determining factor whether the commitment path Pareto dominates the realized path for ε near the origin. In fact, for both $\tau = 2, 3$, the region where $\Delta U_\tau \leq 0$ in the vicinity of the origin, is described approximately by $\varepsilon_3 \geq m_\tau \varepsilon_2$ for some constant m_τ, independent of the value of D_1. But if $D_1 < 1$, ΔU_1 is nonnegative in a neighborhood of the origin. For $D_1 > 1$, there is a ray in the first quadrant where ΔU_1 is negative as in figure 3b.

For general T, we can understand fairly simply why ΔU_1 at the origin depends on ε_T so very differently from the ΔU_τ at later τ. If $\varepsilon_T \neq 0$ while the other future weighting factors vanish, the only time-inconsistency will be between the preferences of the initial self and the second self. Since ε_T only matters for the initial self, the third and later selves will all have the same preferences as the second self. To put it another way, the initial self weighs c_T differently from an exponential discounter, whereas the later selves are all exponential discounters. Thus, the optimal plan for the second self will be the realized path, as the third and later selves have no reason to alter the consumption path planned by the second self. The second self must prefer the realized path over the commitment path for the same reason that the initial self must prefer the commitment path over the realized path. This explains why we must have $\underline{\varepsilon}_T^1 = 0$ if $\varepsilon_2 = \cdots = \varepsilon_{T-1} = 0$. In this special case, whether the third and later selves will also prefer the realized path to the initial path, depends on how the two paths deviate.

Looking again at figure 3b, if we pick a fixed value for ε_2 then the region where the commitment path Pareto dominates the realized path appears to always be a finite interval with respect to ε_3. This is a general pattern that would also be apparent in figure 3a if we included a larger portion of the parameter space.

The economic rationale behind this result can be comprehended as follows. By decreasing ε_T from zero while keeping the other future weights fixed (not necessarily at zero in this case), the initial self will put diminishing weight on the terminal consumption, going to zero as $\varepsilon_T \to -1$ and $D_T \to 0$. Since the other selves will not put less weight on c_T, the initial path will become more and more objectionable as compared to the realized path. Thus ΔU_τ will get larger as $\varepsilon_T \to -1$.

To understand what happens in the opposite direction, let us focus again on the special case of $\varepsilon_2 = \cdots = \varepsilon_{T-1} = 0$. We have already discussed why ΔU_1 will increase from zero as we increase ε_T in this case. For the third and later selves, except for the terminal self, the intuition of what happens when we decrease ε_T will extend to the positive direction for small ε_T. The initial self puts more weight on $c_{T|0}$, which makes the initial path more preferable to these later selves as well. However, the increase in $c_{T|0}$ comes at the expense of reducing the initial allocation of consumption to earlier ages. That does not matter for the terminal self, but it does matter for $\tau = 2, \ldots, T-1$. This is a second-order effect so it is dominated by the first-order effect of high $c_{T|0}$ for small ε_T. But for large enough ε_T these selves will look upon the realized path more favorably. Indeed, for these τ we can show that ΔU_τ will eventually turn positive as we keep increasing ε_T. This leads us to the following lemma, which is true even when the $\varepsilon_2, \cdots, \varepsilon_{T-1}$ are not all zero.

Lemma 6. *For $\tau = 1, \ldots, T-1$, for any given choice of $\varepsilon_2, \ldots, \varepsilon_{T-1}$, we have*

$$\lim_{\varepsilon_T \to \infty} \Delta U_\tau = \lim_{\varepsilon_T \to -1} \Delta U_\tau = \infty. \tag{38}$$

See appendix C for the proof of the first equality. Note that with $\varepsilon_{T-1} \to -1$, we have $\phi_{T-1} = 0$ and $\overline{\phi}_{T-1} > 0$, while all of the other ϕ_t and $\overline{\phi}_t$ remain positive. Therefore, the second equality in (38) holds since all of the ΔU_τ for $\tau \geq 1$ depend on $\ln \phi_{T-1}$ with a positive coefficient in (31).

Combining Proposition 4 and Lemma 6 yields the main result of this section.

Proposition 7. *Fix $\varepsilon_2, \ldots, \varepsilon_{T-1}$. For $\tau = 1, \ldots, T-1$, there will exist $A^\tau \leq \underline{\varepsilon}_T^\tau$ and $B^\tau \geq \underline{\varepsilon}_T^\tau$ such that $\Delta U_\tau < 0$ iff $\varepsilon_T \in (A^\tau, B^\tau)$. There will also exist A^T such that $c_{T|0} < c_T$ iff $\varepsilon_T > A^T$. Then*

$$(A, B) \equiv (A^T, \infty) \cap \bigcap_{\tau=1}^{T-1} (A^\tau, B^\tau). \tag{39}$$

is the interior of the set of ε_T such that the commitment path Pareto dominates the realized path.[26]

Note that the bounds A^τ and B^τ in Proposition 7 depend on the other future weighting factors $\varepsilon_2, \cdots, \varepsilon_{T-1}$. We do not have a simple characterization of the A^τ and B^τ beyond what can be obtained by numerical solution of the equation $\Delta U_\tau = 0$. The set (A, B) could also be empty. A necessary and sufficient condition for $(A^\tau, B^\tau) \neq \emptyset$ is that $\Delta U_\tau < 0$ at $\varepsilon_T = \underline{\varepsilon}_T^\tau$.

[26] We do not characterize the boundary of the set because at least one of the ΔU_τ must be negative for the commitment path to Pareto dominate the realized path, so some but not all of the boundary will be included in the set.

We have already established that strict concavity of the log consumption profile is a sufficient condition for the terminal self to prefer the commitment path to the realized path, so $\varepsilon_T > A^T$. Does strict concavity buy us anything with regards to whether the earlier selves also prefer the commitment path?

In Section 3, we established that local concavity of the log consumption profile at age t depends on ε_{T-t+1}. Thus ε_T can only affect whether the log consumption profile is concave at the beginning of the lifespan. In order to get a better sense of how concavity later in the lifespan is connected to welfare, we would have to delve deeper into the future weighting factors at shorter delays. For now, we can establish the following lemma.

Lemma 8. *If the log consumption profile is locally concave at $t = 1$ and the commitment path Pareto dominates the realized path, then we must have $\varepsilon_T \in (\underline{\varepsilon}_T^1, B^T)$, and for $\tau = 2, \ldots, T-1$ we must have $\underline{\varepsilon}_T^\tau \in (\underline{\varepsilon}_T^1, B^\tau)$.*

From (37), we have that

$$\underline{\varepsilon}_T^1 = \overline{\phi}_{T-2}(1 + \varepsilon_{T-1}) - 1, \tag{40}$$

which implies that

$$\frac{1 + \underline{\varepsilon}_T^1}{1 + \varepsilon_{T-1}} = \overline{\phi}_{T-2}.$$

Because $\overline{\varepsilon}_{T-2}$ and $\overline{\varepsilon}_{T-1}$ are averages, the terminal future weighting factor $\varepsilon_T \lesseqgtr \underline{\varepsilon}_T^1$ if and only if $\phi_{T-1} \lesseqgtr \overline{\phi}_{T-1}$, which is the determining factor for whether the consumption profile $\ln c_t$ is strictly concave, convex, or linear for $t = 0, 1, 2$. If the log consumption profile is strictly concave *and* $\Delta U_1 < 0$, we must have $\varepsilon_T \in (\underline{\varepsilon}_T^1, B^1)$.

Note that for $\tau = 2, \ldots, T-1$, (37) gives

$$\underline{\varepsilon}_T^\tau \equiv \frac{D_{T-\tau}}{D_{T-1}} \frac{\sum_{i=0}^{T-2} D_i}{\sum_{t=0}^{T-\tau-1} D_t}(1 + \underline{\varepsilon}_T^1) - 1. \tag{41}$$

If we make the further assumption that D_t is decreasing in t for $t < T$, then this implies $\underline{\varepsilon}_T^\tau \geq \underline{\varepsilon}_T^1$. If there is any ε_T for which ΔU_τ is negative, the choice of ε_T for which it will be most negative, i.e. $\underline{\varepsilon}_T^\tau$, must be such that the log consumption profile is strictly concave on $t = 0, 1, 2$.

5 A Simple Example: Generalizing the Quasihyperbolic Discount Function

In order to elucidate the preceding results, let us consider a generalization of the quasihyperbolic discount function that informs much of this literature. Consider what we will call beta-delta-omega discounting with

$$D_t = \begin{cases} 1 & t = 0 \\ \beta\delta^t & 1 \leq t \leq T-1 \\ \omega\delta^T & t = T \end{cases} . \tag{42}$$

for $\beta, \delta, \omega > 0$. For the special case of $\beta = \omega$, this reduces to the standard quasihyperbolic function. The corresponding future weighting factors are

$$\varepsilon_t = \begin{cases} 0 & t = 0 \\ \beta^{1-t} - 1 & 1 \leq t \leq T-1 \\ \frac{\omega}{\beta^T} - 1 & t = T \end{cases} . \tag{43}$$

Note that the vector of future weighting factors is the same for all selves like in the quasihyperbolic special case. However, the terminal weighting factor, which differs in structure from ε_t for $t < T$, only matters to the initial self since the later selves do not include a point T periods in the future within their remaining time horizon. This innovation permits us to vary ε_T without affecting the quasihyperbolic structure at earlier delays (and after the initial decision point) simply by adjusting ω. Thus we can demonstrate the results of Section 4 by considering how the ΔU_τ vary with ω.

First let us consider how the shape of the log consumption profile depends on ω. The future weighting growth factors are

$$\phi_t = \begin{cases} 1 & t = 0 \\ \frac{1}{\beta} & 1 \leq t \leq T-2 \\ \frac{\omega}{\beta^2} & t = T-1 \end{cases} .$$

The discount function will be present-biased if $\phi_t > 1$ for all $t > 0$, which happens if $\beta < 1$ and $\omega > \beta^2$. The first condition is the familiar condition for present bias with quasihyperbolic discounting, and this will imply the second condition if $\omega = \beta$. More generally, beta-delta-omega discounting will be present biased for ω in a neighborhood of β that is unbounded to the right. That is to say if ω is sufficiently large that $\varepsilon_T > \varepsilon_{T-1}$.

To apply Proposition 2, we also need to compute weighted averages of the ϕ_t, which we called $\overline{\phi}_t$. For $1 \leq t < T-1$, the $\overline{\phi}_t$ will be the same as in the quasihyperbolic case:

$$\begin{aligned}
\overline{\phi}_t &= \frac{\sum_{s=0}^{t} D_s \phi_s}{\sum_{s'=0}^{t} D_{s'}} = \frac{1 + \frac{1}{\beta}\sum_{s=1}^{t} D_s}{\sum_{s'=0}^{t} D_{s'}} = \frac{1 - \frac{1}{\beta} + \frac{1}{\beta}\sum_{s=0}^{t} D_s}{\sum_{s'=0}^{t} D_{s'}} \\
&= \frac{1 - \frac{1}{\beta}}{\sum_{s'=0}^{t} D_{s'}} + \frac{1}{\beta} = \frac{1 - \frac{1}{\beta}}{\sum_{s'=0}^{t} D_{s'}} + \phi_t
\end{aligned}$$

Thus

$$\overline{\phi}_t - \phi_t = \frac{1 - \frac{1}{\beta}}{\sum_{s=0}^{t} D_s}.$$

If $\beta < 1$, $\phi_t > \overline{\phi}_t$ for $1 \leq t < T-1$ so the log consumption profile will be strictly concave except possibly at the beginning where $t = 0, 1, 2$. Likewise, if $\beta > 1$, the log consumption profile will be strictly convex except possibly at the beginning.

Note that only the shape of the consumption profile at the beginning will depend on ω under beta-delta-omega discounting. More precisely, the weighted average of the ϕ_t for $t = 0, \ldots, T-1$ is

$$\begin{aligned}
\overline{\phi}_{T-1} &= \frac{\sum_{s=0}^{T-1} D_s \phi_s}{\sum_{s'=0}^{T-1} D_{s'}} = \frac{1 + \frac{1}{\beta}\sum_{s=1}^{T-2} D_s + \frac{\omega}{\beta^2} D_{T-1}}{\sum_{s'=0}^{T-1} D_{s'}} \\
&= \frac{1 + \frac{1}{\beta}\sum_{s=0}^{T-1} D_s + \frac{\omega}{\beta^2}\beta\delta^{T-1} - \frac{1}{\beta} - \frac{1}{\beta}\omega\delta^{T-1}}{\sum_{s'=0}^{T-1} D_{s'}} \\
&= \frac{1}{\beta} + \frac{1 - \frac{1}{\beta}}{\sum_{s'=0}^{T-1} D_{s'}}
\end{aligned}$$

Thus the relevant difference is

$$\begin{aligned}
\overline{\phi}_{T-1} - \phi_{T-1} &= \frac{\beta-\omega}{\beta^2} + \frac{1-\frac{1}{\beta}}{\sum_{s'=0}^{T-1} D_{s'}} \qquad (44)\\
&= \frac{\beta-\omega}{\beta^2} + \frac{1-\frac{1}{\beta}}{1+\sum_{s'=1}^{T-1} \beta\delta^{s'}}\\
&= \frac{\beta-\omega}{\beta^2} + \frac{1-\frac{1}{\beta}}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}}\\
&= \frac{1}{\beta}\left[1-\frac{1}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}}\right] + \frac{1}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}} - \frac{\omega}{\beta^2}\\
&= \frac{\delta\frac{1-\delta^{T-1}}{1-\delta}}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}} + \frac{1}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}} - \frac{\omega}{\beta^2}.
\end{aligned}$$

That is to say,

$$\overline{\phi}_{T-1} - \phi_{T-1} = \frac{1+\delta\frac{1-\delta^{T-1}}{1-\delta}}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}} - \frac{\omega}{\beta^2}.$$

Thus the log consumption profile will be strictly concave for $t = 0, 1, 2$ iff

$$\omega > \underline{\omega} = \beta^2 \frac{1+\delta\frac{1-\delta^{T-1}}{1-\delta}}{1+\beta\delta\frac{1-\delta^{T-1}}{1-\delta}}.$$

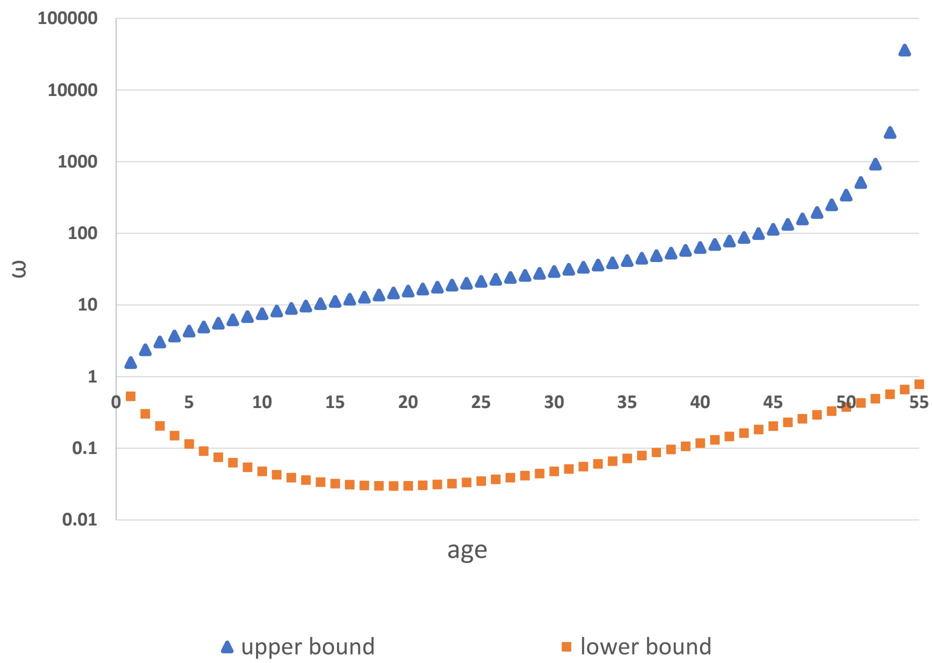
Note that if $\beta < 1$ we have $\beta^2 < \underline{\omega}$, so for $\omega \in (\beta^2, \underline{\omega})$ the discount function will be present-biased, yet the log consumption profile will be strictly convex for $t = 0, 1, 2$. From (44), we also have $\underline{\omega} < \beta$ since the whole log consumption profile is strictly concave if $\omega = \beta < 1$.

To demonstrate welfare results with beta-delta-omega discounting, it is necessary to calibrate the model. Since δ corresponds to the usual concept of a discount factor, it is common to calibrate $\delta = \frac{1}{R}$, so we adopt standard choices of $\delta = 0.96$ and $R = 1.0417$. The literature on quasihyperbolic discounting has not reached a consensus about how much present bias is observed in typical households. For simplicity, we assume also that $\beta = 0.96$.[27]

Note that ω is a linear transformation of ε_T that is easier to interpret since $\omega = \beta$ corresponds to the familiar beta-delta discounting function. Therefore, in the following we will discuss how remaining lifetime utility on the commitment path compares to the corresponding utility on the realized path in terms of ω. Thus, for $\tau = 1, \ldots, T-1$, the

[27]See for example Guo and Krause (2015).

Figure 4: The Pareto interval for beta-delta-omega discounting function



Note: this figure plots the interval $\{A^{\tau}_{\omega}, B^{\tau}_{\omega}\}$ for various levels of ω over the lifecycle of a household with a beta-delta-omega discount function. Within this interval ΔU_{τ} is less than zero.

interval where $\Delta U_\tau < 0$ is $\omega \in (A^\tau_\omega, B^\tau_\omega)$, where

$$A^\tau_\omega = \beta^T(1 + A^\tau) \tag{45}$$

and similarly for B^τ_ω. Likewise, analogous to (37),

$$\underline{\omega}_\tau = \beta^T(1 + \underline{\varepsilon}^\tau_T) \tag{46}$$

is the value of ω that minimizes ΔU_τ.

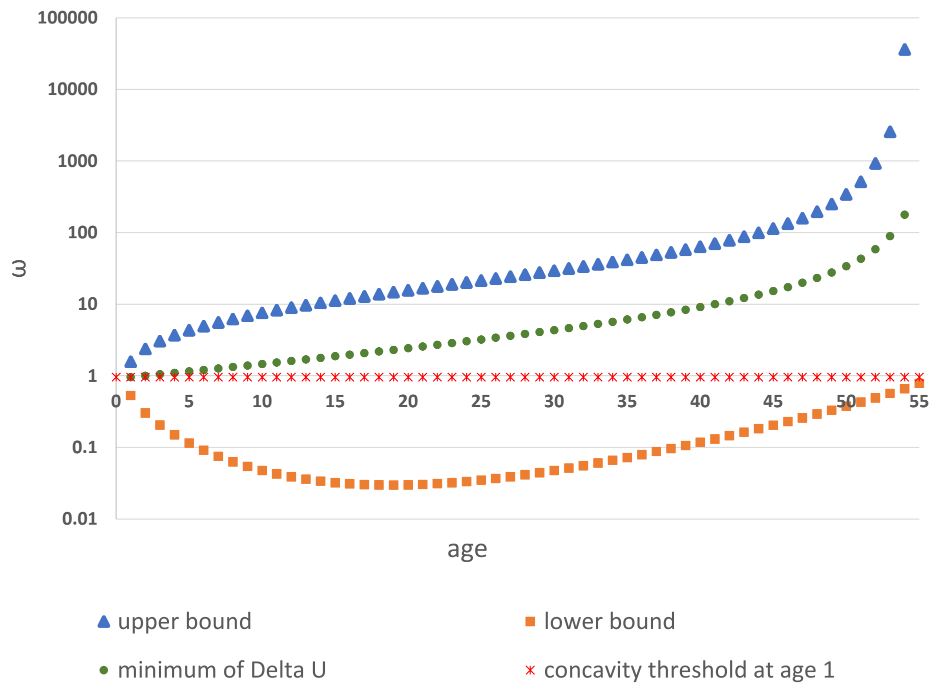
We provide figure 4 which shows the upper and lower bounds on ω within which following the commitment path with the beta-delta-omega discount function will Pareto dominate the realized path. More precisely, this figure plots the bounds of $(A^\tau_\omega, B^\tau_\omega)$ that we defined in section 4 as a consequence of Propositions 4 and lemma 6.

Note that the upper bound, B^τ_ω, of where $\Delta U_\tau < 0$ is strictly increasing with age. Thus it is the upper bound at age 1, $B^1_\omega = 1.57$ that is the binding upper bound on ω for when the commitment path Pareto dominates the realized path. In contrast, the lower bound, A^τ_ω has a more complicated, U-shaped profile. Since $A^1_\omega = 0.53 < A^T_\omega = 0.78$, it is the lower bound at age T that binds for Pareto dominance of the commitment path. Thus the commitment path Pareto dominates the realized path for $\omega \in (A^1_\omega, B^T_\omega) = (0.78, 1.57)$.

In figure 5, we add two curves to figure 4 that help to demonstrate Lemma 8, which encapsulates the relation between the concavity of the log consumption profile and the Pareto dominance of the commitment path of consumption. These curves are $\underline{\omega}^\tau$, the value of ω that minimizes ΔU_τ, and a dotted line at the value of $\underline{\omega}^1 = 0.96$. The latter is the threshold value such that the log consumption profile is strictly concave if $\omega > \underline{\omega}^1$. Since D_τ is strictly decreasing for $\tau < T$, For $\tau > 1$, Lemma 8 implies that the green curve, $\underline{\omega}^\tau$, is strictly above the dotted line, $\underline{\omega}^1$, for $\tau > 1$. That is to say ΔU_τ is minimized for $\tau > 1$ in the part of the parameter space where the log consumption profile is strictly concave. Consistent with this, the bulk of the interval (A^1_ω, B^T_ω) is in this subspace.

Note that the upper bound B^τ_ω is strictly increasing in figure 4 while the lower bound A^τ_ω is everywhere below the $\underline{\omega}^1$ line. This means that for $\omega \in (\underline{\omega}^1, B^1_\omega)$ the log consumption profile is both strictly concave and the commitment path Pareto dominates the realized path.

Figure 5: Visualization of Pareto interval along with the concave log consumption profile



Note: this figure plots the interval $\{A^{\tau}_{\omega}, B^{\tau}_{\omega}\}$ for various levels of ω over the lifecycle of a household with a beta-delta-omega discount function. Within this interval ΔU_{τ} is less than zero. Also it has $\underline{\omega}^{\tau}$, the value of ω that minimizes ΔU_{τ} (which is marked as "minimum of Delta U) and the value of $\underline{\omega}^{1}$ (which is marked as concavity threshold at age 1), which is the threshold value such that the log consumption profile is strictly concave if $\omega > \underline{\omega}^{1}$

6 Concluding remarks

In this paper we proposed a general representation of relative discounting functions that allows us to focus on how the discounting function deviates from an exponential discounting function, which does not exhibit time-inconsistency. We term the perturbation away from the exponential case a *future weighting factor* ε_t. This specific format of the discounting function provides a simple way to depict a future bias by having all ε_t be negative and decreasing for $t > 1$, and a present bias by having all ε_t be positive and increasing for $t > 1$.

We find that for the log consumption profile to be locally concave, which is necessary at the peak of a hump-shaped consumption profile, a future weighting growth factor must be bigger than the weighted average of future weighting growth factor at shorter delays, where the weights are the discount factor. This means that a present bias is a necessary but not sufficient condition for the entire log consumption profile to be strictly concave.

Also, using the proposed future weighting functional form, we explored the conditions on the future weighting factors under which the consumption profile that is determined in the first period of life will Pareto dominate the realized consumption profiles chosen at each period. This result is especially useful because Pareto dominance of the initial path is often used to motivate how one performs welfare analysis in these models with time-inconsistent preferences, where choosing a reference consumption plan for the analysis is a point of controversy in the literature. The results of our study suggest that one has to be cautious when analyzing welfare with time-inconsistent preferences. The consumption path chosen by one's initial self is not necessarily the best choice to serve as the benchmark for welfare purposes without additional information. As a matter of fact, neither the commitment path nor the realized path will dominate each other for most of the parameter space of possible discount functions.

References

Equivalent representations of non-exponential discounting models. Journal of Mathematical Economics, 66:58 – 71, 2016.

Orazio Attanasio and Martin Browning. Consumption over the life cycle and over the business cycle, 1993.

Orazio Attanasio and Guglielmo Weber. Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. Journal of Political Economy, 103(6):1121–57, 1995. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:103:y:1995:i:6:p:1121-57.

Orazio P. Attanasio, James Banks, Costas Meghir, and Guglielmo Weber. Humps and bumps in lifetime consumption. Journal of Business Economic Statistics, 17(1):22–35, 1999. ISSN 07350015. URL http://www.jstor.org/stable/1392236.

B. Douglas Bernheim and Debraj Ray. Economic Growth with Intergenerational Altruism. The Review of Economic Studies, 54(2):227–243, 04 1987. ISSN 0034-6527. doi: 10.2307/2297513. URL https://doi.org/10.2307/2297513.

Martin Browning and Thomas Crossley. Unemployment insurance benefit levels and consumption changes. Journal of Public Economics, 80(1):1–23, 2001. URL https://EconPapers.repec.org/RePEc:eee:pubeco:v:80:y:2001:i:1:p:1-23.

Martin Browning, Angus Deaton, and Margaret Irish. A profitable approach to labor supply and commodity demands over the life-cycle. Econometrica: journal of the econometric society, pages 503–543, 1985.

James Bullard and James Feigenbaum. A leisurely reading of the life-cycle consumption data. Journal of Monetary Economics, 54(8):2305–2320, 2007. URL https://EconPapers.repec.org/RePEc:eee:moneco:v:54:y:2007:i:8:p:2305-2320.

Frank Caliendo and David Aadland. Short-term planning and the life-cycle consumption puzzle. Journal of Economic Dynamics and Control, 31(4):1392–1415, 2007. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2006.05.002. URL https://www.sciencedirect.com/science/article/pii/S016518890600100X.

Frank Caliendo and T. Scott Findley. Commitment and welfare. Journal of Economic Behavior and Organization, 2019.

John Y Campbell and N Gregory Mankiw. Consumption, income, and interest rates: Reinterpreting the time series evidence. NBER macroeconomics annual, 4:185–216, 1989.

Dan Cao and Iván Werning. Saving and dissaving with hyperbolic discounting. Econometrica, 86(3):805–857, 2018. doi: https://doi.org/10.3982/ECTA15112. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA15112.

Christopher Carroll. Buffer-stock saving and the life cycle/permanent income hypothesis. The Quarterly Journal of Economics, 112(1):1–55, 1997. URL https://EconPapers.repec.org/RePEc:oup:qjecon:v:112:y:1997:i:1:p:1-55.

Christopher Carroll and Lawrence Summers. Consumption growth parallels income growth: Some new evidence. In National Saving and Economic Performance, pages 305–348. National Bureau of Economic Research, Inc, 1991. URL https://EconPapers.repec.org/RePEc:nbr:nberch:5995.

Christopher D Carroll. How does future income affect current consumption? The Quarterly Journal of Economics, 109(1):111–147, 1994.

Angus Deaton. Understanding Consumption. Oxford University Press, 1992. URL https://EconPapers.repec.org/RePEc:oxp:obooks:9780198288244.

Mathias Dewatripont, Isabelle Brocas, and Juan Carrillo. Commitment devices under self-control problems: an overview. Ulb institutional repository, ULB – Universite Libre de Bruxelles, 2004.

Nicolas Drouhin. Non-stationary additive utility and time consistency. Journal of Mathematical Economics, 86:1 – 14, 2020. ISSN 0304-4068. doi: https://doi.org/10.1016/j.jmateco.2019.10.005. URL http://www.sciencedirect.com/science/article/pii/S0304406819301077.

James Feigenbaum. Can mortality risk explain the consumption hump? Journal of Macroeconomics, 30(3):844–872, 2008.

James Feigenbaum and Sepideh Raei. Lifecycle consumption and welfare with nonexponential discounting in continuous time. forthcoming , Journal of Mathematical Economics, 2023. URL http://www.platonicadventures.com/research.html.

Martin Feldstein. The optimal level of social security benefits. The Quarterly Journal of Economics, 100(2):303–320, 1985.

Jesus Fernandez-Villaverde and Dirk Krueger. Consumption and saving over the life cycle: How important are consumer durables? Macroeconomic Dynamics, 15(5):725–770, 2011. URL https://EconPapers.repec.org/RePEc:cup:macdyn:v:15:y:2011:i:05:p:725-770_00.

Milton Friedman. Theory of the consumption function. Princeton university press, 2018.

Pierre-Olivier Gourinchas and Jonathan A. Parker. Consumption over the life cycle. Econometrica, 70(1):47–89, 2002. doi: https://doi.org/10.1111/1468-0262.00269. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00269.

Steven R Grenadier and Neng Wang. Investment under uncertainty and time-inconsistent preferences. Journal of Financial Economics, 84(1):2–39, 2007.

Barbara Griffin, Beryl Hesketh, and Vanessa Loh. The influence of subjective life expectancy on retirement transition and planning: a longitudinal study. Journal of Vocational Behavior, 81(2):129–137, 2012.

Faruk Gul and Wolfgang Pesendorfer. Self-control, revealed preference and consumption choice. Review of Economic Dynamics, 7(2):243 – 264, 2004.

Jang-Ting Guo and Alan Krause. Dynamic nonlinear income taxation with quasi-hyperbolic discounting and no commitment. Journal of Economic Behavior Organization, 109:101–119, 2015. ISSN 0167-2681. doi: https://doi.org/10.1016/j.jebo.2014.11.002. URL https://www.sciencedirect.com/science/article/pii/S0167268114002819.

Gary Hansen and Selahattin Imrohoroglu. Consumption over the Life Cycle: The Role of Annuities. Review of Economic Dynamics, 11(3):566–583, July 2008. doi: 10.1016/j.red.2007.12.004. URL https://ideas.repec.org/a/red/issued/06-155.html.

Christopher Harris and David Laibson. Instantaneous gratification. The Quarterly Journal of Economics, 128(1):205–248, 2013.

James Heckman. Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. American Economic Review, 64(1):188–94, 1974. URL https://EconPapers.repec.org/RePEc:aea:aecrev:v:64:y:1974:i:1:p:188-94.

Eunice Hong and Sherman D Hanna. Financial planning horizon: A measure of time preference or a situational factor? Journal of Financial Counseling and Planning, 25(2):184–196, 2014.

R Glenn Hubbard, Jonathan Skinner, and Stephen P Zeldes. The importance of precautionary motives in explaining individual and aggregate saving. In Carnegie-Rochester conference series on public policy, volume 40, pages 59–125. Elsevier, 1994.

David Laibson. Hyperbolic discounting and consumption. PhD diss. Massachusetts Institute of Technology, 1994.

David Laibson. Golden eggs and hyperbolic discounting. The Quarterly Journal of Economics, 112(2):443–478, 1997.

David Laibson. Life-cycle consumption and hyperbolic discount functions. European Economic Review, 42(3):861 – 871, 1998.

David I Laibson. Hyperbolic discount functions, undersaving, and savings policy. Working Paper 5635, National Bureau of Economic Research, June 1996.

David I. Laibson, Andrea Repetto, Jeremy Tobacman, Robert E. Hall, William G. Gale, and George A. Akerlof. Self-control and saving for retirement. Brookings Papers on Economic Activity, 1998(1):91–196, 1998.

John Lane and Tapan Mitra. On nash equilibrium programs of capital accumulation under altruistic preferences. International Economic Review, 22(2):309–331, 1981. ISSN 00206598, 14682354. URL `http://www.jstor.org/stable/2526279`.

Wolfgang Leininger. The Existence of Perfect Equilibria in a Model of Growth with Altruism between Generations. Review of Economic Studies, 53(3):349–367, 1986. URL `https://ideas.repec.org/a/oup/restud/v53y1986i3p349-367..html`.

Jesus Marin-Solano and Jorge Navas. Non-constant discounting in finite horizon: The free terminal time case. Journal of Economic Dynamics and Control, 33(3):666–675, 2009.

Franco Modigliani and Richard Brumberg. Utility analysis and the consumption function: An interpretation of cross-section data. Franco Modigliani, 1(1):388–436, 1954.

Congming Mu, Jinqiang Yang, et al. Optimal contract theory with time-inconsistent preferences. Economic Modelling, 52:519–530, 2016.

Keizo Nagatani. Life cycle saving: theory and fact. The American Economic Review, 62(3): 344–353, 1972.

Ted O'Donoghue and Matthew Rabin. The economics of immediate gratification. Journal of Behavioral Decision Making, 13(2):233–250.

Ted O'Donoghue and Matthew Rabin. Doing it now or later. American Economic Review, 89(1):103–124, March 1999.

Ted O'Donoghue and Matthew Rabin. Choice and procrastination. The Quarterly Journal of Economics, 116(1):121–160, 2001.

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

Edmund S Phelps and Robert A Pollak. On second-best national saving and game-equilibrium growth. The Review of Economic Studies, 35(2):185–199, 1968.

Drazen Prelec. Decreasing impatience: A criterion for non-stationary time preference and "hyperbolic" discounting. The Scandinavian Journal of Economics, 106(3):511–532, 2004. doi: https://doi.org/10.1111/j.0347-0520.2004.00375.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0347-0520.2004.00375.x.

Michael Richter. A time-inconsistent first welfare theorem:efficiency and the convexity of patience. working paper, 2020.

Paul A. Samuelson. A note on measurement of utility. The Review of Economic Studies, 4 (2):155–161, 1937.

R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. Review of Economic Studies, 23(3):165–180, 1955a.

Robert Henry Strotz. Myopia and inconsistency in dynamic utility maximization. The review of economic studies, 23(3):165–180, 1955b.

Lester C. Thurow. The optimum lifetime distribution of consumption expenditures. The American Economic Review, 59(3):324–330, 1969. ISSN 00028282. URL http://www.jstor.org/stable/1808961.

Appendices

A Simplifying the Concavity Condition

The log consumption profile is concave at $t+1$ iff we have

$$\frac{\sum_{z'=1}^{T-t-1} D_1^{z'}(1+\varepsilon_{z'})}{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})} \frac{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}{\sum_{z=1}^{T-t-1} D_1^{z}(1+\varepsilon_{z-1})} \leq 1.$$

We can rearrange this inequality as follows.

$$\frac{\sum_{z'=1}^{T-t-1} D_1^{z'}(1+\varepsilon_{z'})}{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})} \leq \frac{\sum_{z=1}^{T-t-1} D_1^{z}(1+\varepsilon_{z-1})}{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}$$

$$1 - \frac{D_1^{T-t}(1+\varepsilon_{T-t})}{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})} \leq 1 - \frac{D_1^{T-t}(1+\varepsilon_{T-t-1})}{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})}$$

$$\frac{1+\varepsilon_{T-t-1}}{\sum_{s=1}^{T-t} D_1^{s}(1+\varepsilon_{s-1})} \leq \frac{1+\varepsilon_{T-t}}{\sum_{s'=1}^{T-t} D_1^{s'}(1+\varepsilon_{s'})}$$

We wish to isolate ε_{T-t}, which appears in both the numerator and the denominator of the right-hand side.

$$\frac{1+\varepsilon_{T-t-1}}{\sum_{s=0}^{T-t-1} D_1^{s+1}(1+\varepsilon_{s})} \leq \frac{1+\varepsilon_{T-t}}{\sum_{s'=1}^{T-t-1} D_1^{s'}(1+\varepsilon_{s'}) + D_1^{T-t}(1+\varepsilon_{T-t})}$$

$$\frac{\sum_{s'=1}^{T-t-1} D_1^{s'}(1+\varepsilon_{s'}) + D_1^{T-t}(1+\varepsilon_{T-t})}{\sum_{s=0}^{T-t-1} D_1^{s+1}(1+\varepsilon_{s})}(1+\varepsilon_{T-t-1}) \leq 1+\varepsilon_{T-t}$$

$$\frac{\sum_{s'=1}^{T-t-1} D_1^{s'}(1+\varepsilon_{s'})}{\sum_{s=0}^{T-t-1} D_1^{s+1}(1+\varepsilon_{s})}(1+\varepsilon_{T-t-1}) \leq (1+\varepsilon_{T-t})\left[1 - \frac{D_1^{T-t}(1+\varepsilon_{T-t-1})}{\sum_{s=0}^{T-t-1} D_1^{s+1}(1+\varepsilon_{s})}\right]$$

$$\frac{\sum_{s'=1}^{T-t-1} D_1^{s'}(1+\varepsilon_{s'})}{\sum_{s=0}^{T-t-1} D_1^{s+1}(1+\varepsilon_{s})}(1+\varepsilon_{T-t-1}) \leq (1+\varepsilon_{T-t})\left[\frac{\sum_{z'=0}^{T-t-2} D_1^{z'+1}(1+\varepsilon_{z'})}{\sum_{z=0}^{T-t-1} D_1^{z+1}(1+\varepsilon_{z})}\right]$$

Thus we obtain the condition

$$\frac{\sum_{z=1}^{T-t-1} D_1^{z}(1+\varepsilon_{z})}{\sum_{z'=0}^{T-t-2} D_1^{z'+1}(1+\varepsilon_{z'})}(1+\varepsilon_{T-t-1}) \leq 1+\varepsilon_{T-t} \tag{47}$$

for concavity at $t+1$.

B Derivation of Eq. (37)

$$
\begin{aligned}
\underline{\varepsilon}_T^{\tau} &= \frac{\sum_{i=0}^{T-2} D_1^{i}(1+\varepsilon_{i+1})}{\sum_{t=\tau}^{T-1} D_1^{t-1}(1+\varepsilon_{t-\tau})}(1+\varepsilon_{T-\tau})-1 \\
&= \frac{\sum_{i=0}^{T-2} D_1^{i}(1+\varepsilon_{i+1})}{\sum_{t=0}^{T-\tau-1} D_1^{t+\tau-1}(1+\varepsilon_{t})}(1+\varepsilon_{T-\tau})-1 \\
&= D_1^{1-\tau}\frac{\sum_{i=0}^{T-2} D_1^{i}(1+\varepsilon_{i+1})}{\sum_{t=0}^{T-\tau-1} D_1^{t}(1+\varepsilon_{t})}(1+\varepsilon_{T-\tau})-1 \\
&= D_1^{1-\tau}\frac{\sum_{i=0}^{T-2} D_i\phi_i}{\sum_{t=0}^{T-\tau-1} D_t}(1+\varepsilon_{T-\tau})-1 \\
&= \frac{D_1^{T-\tau}(1+\varepsilon_{T-\tau})}{D_1^{T-1}(1+\varepsilon_{T-1})}\frac{\sum_{t=0}^{T-2} D_t}{\sum_{s=0}^{T-\tau-1} D_s}\frac{\sum_{i=0}^{T-2} D_i\phi_i}{\sum_{j=0}^{T-2} D_j}(1+\varepsilon_{T-1})-1 \\
&= \frac{D_{T-\tau}}{D_{T-1}}\frac{\sum_{t=0}^{T-2} D_t}{\sum_{s=0}^{T-\tau-1} D_s}\overline{\phi}_{T-2}(1+\varepsilon_{T-1})-1.
\end{aligned}
$$

C Derivation of limits of ΔU_τ

$$
\begin{aligned}
\lim_{\varepsilon_T\to\infty}\Delta U_\tau &= \lim_{\varepsilon_T\to\infty}\sum_{t=\tau}^{T}\sum_{i=0}^{t-1} D_{t-\tau}\ln\frac{\overline{\phi}_{T-t+i}}{\phi_i} \\
&= \sum_{t=\tau}^{T}\sum_{i=0}^{t-1} D_{t-\tau}\lim_{\varepsilon_T\to\infty}\ln\overline{\phi}_{T-t+i} - \sum_{t=\tau}^{T}\sum_{i=0}^{t-1} D_{t-\tau}\lim_{\varepsilon_T\to\infty}\ln\phi_i \\
&= \sum_{t=\tau}^{T}\sum_{i=0}^{t-2} D_{t-\tau}\lim_{\varepsilon_T\to\infty}\ln\overline{\phi}_{T-t+i} + \sum_{t=\tau}^{T} D_{t-\tau}\lim_{\varepsilon_T\to\infty}\ln\overline{\phi}_{T-t+(t-1)} - \sum_{t=\tau}^{T-1}\sum_{i=0}^{t-1} D_{t-\tau}\lim_{\varepsilon_T\to\infty}\ln\phi_i \\
&\quad -\sum_{i=0}^{T-1} D_{T-\tau}\lim_{\varepsilon_T\to\infty}\ln\phi_i \\
&= \sum_{t=\tau}^{T}\sum_{i=0}^{t-2} D_{t-\tau}\ln\overline{\phi}_{T-t+i} + \sum_{t=\tau}^{T} D_{t-\tau}\lim_{\varepsilon_T\to\infty}\ln\overline{\phi}_{T-1} - \sum_{t=\tau}^{T-1}\sum_{i=0}^{t-1} D_{t-\tau}\ln\phi_i - \sum_{i=0}^{T-2} D_{T-\tau}\ln\phi_i - D_{T-}
\end{aligned}
$$

$$
\lim_{\varepsilon_T\to\infty}\ln\phi_{T-1} = \lim_{\varepsilon_T\to\infty}\ln\varepsilon_T - \ln(1+\varepsilon_{T-1}) = \lim_{\varepsilon_T\to\infty}\ln\varepsilon_T
$$

$$\begin{aligned}
\lim_{\varepsilon_T\to\infty} \ln \overline{\phi}_{T-1} &= \lim_{\varepsilon_T\to\infty} \ln \left(\frac{\sum_{s=0}^{T-1} D_s \phi_s}{\sum_{s'=0}^{T-1} D_{s'}} \right) \\
&= \lim_{\varepsilon_T\to\infty} \ln \left(D_1^{T-1}(1+\varepsilon_T) \right) - \ln \left(\sum_{s'=0}^{T-1} D_{s'} \right) \\
&= \ln D_1^{T-1} - \ln \left(\sum_{s'=0}^{T-1} D_{s'} \right) + \lim_{\varepsilon_T\to\infty} \ln \varepsilon_T = \lim_{\varepsilon_T\to\infty} \ln \varepsilon_T
\end{aligned}$$

For $\tau < T$,

$$\begin{aligned}
\lim_{\varepsilon_T\to\infty} \Delta U_\tau &= \sum_{t=\tau}^{T} \sum_{i=0}^{t-2} D_{t-\tau} \ln \overline{\phi}_{T-t+i} - \sum_{t=\tau}^{T-1} \sum_{i=0}^{t-1} D_{t-\tau} \ln \phi_i - \sum_{i=0}^{T-2} D_{T-\tau} \ln \phi_i + \left(\sum_{t=\tau}^{T-1} D_{t-\tau} \right) \lim_{\varepsilon_T\to\infty} \ln \varepsilon_T \\
&= \left(\sum_{t=\tau}^{T-1} D_{t-\tau} \right) \lim_{\varepsilon_T\to\infty} \ln \varepsilon_T > 0.
\end{aligned}$$

D Gradient of ΔU_1 at the Origin

$$\Delta U_\tau = \sum_{t=\tau}^{T} \sum_{i=0}^{t-1} D_{t-\tau} \ln \frac{\overline{\phi}_{T-t+i}}{\phi_i}.$$

$$\Delta U_1 = \sum_{s=1}^{T} \sum_{i=0}^{s-1} D_{s-1} \ln \frac{\overline{\phi}_{T-s+i}}{\phi_i}$$

$$\frac{\partial \Delta U_1}{\partial \varepsilon_t} = \sum_{s=1}^{T} \sum_{i=0}^{s-1} \left[\frac{\partial D_{s-1}}{\partial \varepsilon_t} \ln \frac{\overline{\phi}_{T-s+i}}{\phi_i} + D_{s-1} \left(\frac{\partial \ln \overline{\phi}_{T-s+i}}{\partial \varepsilon_t} - \frac{1}{1+\varepsilon_{i+1}} \frac{\partial \varepsilon_{i+1}}{\partial \varepsilon_t} + \frac{1}{1+\varepsilon_i} \frac{\partial \varepsilon_i}{\partial \varepsilon_t} \right) \right]$$

$$\overline{\phi}_s = \frac{\sum_{i=0}^{s} D_i \phi_i}{\sum_{j=0}^{s} D_j}$$

$$\frac{\partial \ln \overline{\phi}_s}{\partial \varepsilon_t} = \frac{\sum_{i=0}^{s} \left(\frac{\partial D_i}{\partial \varepsilon_t} \phi_i + D_i \frac{\partial \phi_i}{\partial \varepsilon_t} \right)}{\sum_{i'=0}^{s} D_{i'} \phi_{i'}} - \frac{\sum_{j=0}^{s} \frac{\partial D_j}{\partial \varepsilon_t}}{\sum_{j'=0}^{s} D_{j'}}$$

$$\frac{\partial \phi_i}{\partial \varepsilon_t} = \frac{\partial}{\partial \varepsilon_t} \left(\frac{1+\varepsilon_{i+1}}{1+\varepsilon_i} \right) = \frac{\delta_{i+1,t}}{1+\varepsilon_i} - \frac{1+\varepsilon_{i+1}}{(1+\varepsilon_i)^2} \delta_{i,t}$$

$$\left. \frac{\partial \phi_i}{\partial \varepsilon_t} \right|_{\varepsilon=0} = \delta_{i+1,t} - \delta_{i,t}$$

$$\left.\frac{\partial \ln \overline{\phi}_s}{\partial \varepsilon_t}\right|_{\varepsilon=0} = \frac{\sum_{i=0}^{s}\left(\frac{\partial D_i}{\partial \varepsilon_t} + D_1^i(\delta_{i+1,t} - \delta_{i,t})\right)}{\sum_{i'=0}^{s} D_1^{i'}} - \frac{\sum_{j=0}^{s}\frac{\partial D_j}{\partial \varepsilon_t}}{\sum_{j'=0}^{s} D_1^{j'}} = \frac{\sum_{i=0}^{s} D_1^i(\delta_{i+1,t} - \delta_{i,t})}{\sum_{i'=0}^{s} D_1^{i'}}$$

$$\begin{aligned}\left.\frac{\partial \Delta U_1}{\partial \varepsilon_t}\right|_{\varepsilon=0} &= \sum_{s=1}^{T}\sum_{i=0}^{s-1} D_1^{s-1}\left(\frac{\partial \ln \overline{\phi}_{T-s+i}}{\partial \varepsilon_t} - \delta_{i+1,t} + \delta_{i,t}\right) \\ &= \sum_{s=1}^{T}\sum_{i=0}^{s-1} D_1^{s-1}\left(\frac{\sum_{j=0}^{T-s+i} D_1^j(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-s+i} D_1^{s'}} - \delta_{i+1,t} + \delta_{i,t}\right)\end{aligned}$$

Suppose $T = 2$.

$$\left.\frac{\partial \Delta U_1}{\partial \varepsilon_2}\right|_{\varepsilon=0} = \sum_{s=1}^{2}\sum_{i=0}^{s-1} D_1^{s-1}\left(\frac{\sum_{j=0}^{2-s+i} D_1^j(\delta_{j+1,2} - \delta_{j,2})}{\sum_{s'=0}^{2-s+i} D_1^{s'}} - \delta_{i+1,2} + \delta_{i,2}\right)$$

$$\begin{aligned}\left.\frac{\partial \Delta U_1}{\partial \varepsilon_2}\right|_{\varepsilon=0} &= \sum_{i=0}^{0} D_1^0\left(\frac{\sum_{j=0}^{1+i} D_1^j(\delta_{j+1,2} - \delta_{j,2})}{\sum_{s'=0}^{1+i} D_1^{s'}} - \delta_{i+1,2} + \delta_{i,2}\right) \\ &\quad + \sum_{i=0}^{1} D_1^1\left(\frac{\sum_{j=0}^{i} D_1^j(\delta_{j+1,2} - \delta_{j,2})}{\sum_{s'=0}^{i} D_1^{s'}} - \delta_{i+1,2} + \delta_{i,2}\right) \\ &= \frac{\sum_{j=0}^{1} D_1^j(\delta_{j+1,2} - \delta_{j,2})}{\sum_{s'=0}^{1} D_1^{s'}} \\ &\quad + D_1\frac{\sum_{j=0}^{0} D_1^j(\delta_{j+1,2} - \delta_{j,2})}{\sum_{s'=0}^{0} D_1^{s'}} + D_1\left(\frac{\sum_{j=0}^{1} D_1^j(\delta_{j+1,2} - \delta_{j,2})}{\sum_{s'=0}^{1} D_1^{s'}} - 1\right) \\ &= \frac{D_1}{1+D_1} + D_1\left(\frac{D_1}{1+D_1} - 1\right) = \frac{D_1 + D_1^2}{1+D_1} - D_1 = 0\end{aligned}$$

$$\left.\frac{\partial \Delta U_1}{\partial \varepsilon_t}\right|_{\varepsilon=0} = \sum_{s=1}^{T}\sum_{i=0}^{s-1} D_1^{s-1}\left(\frac{\sum_{j=0}^{T-s+i} D_1^j(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-s+i} D_1^{s'}} - \delta_{i+1,t} + \delta_{i,t}\right)$$

$$\begin{aligned} S &= \{(s,i) \in \mathbf{Z}^2 : 1 \le s \le T \wedge 0 \le i \le s-1\} \\ S' &= \{(s,i) \in \mathbf{Z}^2 : 0 \le i \le T-1 \wedge i+1 \le s \le T\}\end{aligned}$$

If $(s,i) \in S$, $1 \le s \le T \wedge 0 \le i \le s-1$. Thus $0 \le i \le s-1 \le T-1$, and $i+1 \le s \le T$, so $(s,i) \in S'$.

If $(s,i) \in S'$, $0 \le i \le T-1 \wedge i+1 \le s \le T$, $1 \le i+1 \le s \le T$ and $0 \le i \le s-1$.

$$\left.\frac{\partial \Delta U_1}{\partial \varepsilon_t}\right|_{\varepsilon=0} = \sum_{i=0}^{T-1} \sum_{s=i+1}^{T} D_1^{s-1} \left(\frac{\sum_{j=0}^{T-s+i} D_1^j (\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-s+i} D_1^{s'}} - \delta_{i+1,t} + \delta_{i,t} \right)$$

$$\begin{aligned} \sum_{i=0}^{T-1} \sum_{s=i+1}^{T} D_1^{s-1}(\delta_{i+1,t} - \delta_{i,t}) &= \sum_{s=t}^{T} D_1^{s-1} - (1-\delta_{tT}) \sum_{s=t+1}^{T} D_1^{s-1} \\ &= D_1^{t-1} + \delta_{tT} \sum_{s=t+1}^{T} D_1^{s-1} = D_1^{t-1} \end{aligned}$$

$$V_1 = \sum_{s=1}^{T} \sum_{i=0}^{s-1} \sum_{j=0}^{T-s+i} \frac{D_1^{s+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-s+i} D_1^{s'}}$$

Let $z = s - i$, so $i = s - z$

$$V_1 = \sum_{s=1}^{T} \sum_{z=1}^{s} \sum_{j=0}^{T-z} \frac{D_1^{s+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}}$$

$$\begin{aligned} S &= \{(z,j) \in \mathbf{Z}^2 : 1 \le z \le s \wedge 0 \le j \le T-z\} \\ S' &= \{(z,j) \in \mathbf{Z}^2 : 0 \le j \le T-1 \wedge 1 \le z \le \min\{s, T-j\}\} \end{aligned}$$

If $(z,j) \in S$, $1 \le z \le s \wedge 0 \le j \wedge j \le T-z$. Thus $0 \le j \le T-z \le T-1$. $1 \le z$, $z \le s$, and $z \le T-j$. Thus $1 \le z \le \min\{s, T-j\}$. So $(z,j) \in S'$.

If $(z,j) \in S'$, $0 \le j \le T-1 \wedge 1 \le z \le \min\{s, T-j\}$. Thus $1 \le z \le s$. Since $z \le T-j$, we have $j \le T - z$. Thus $0 \le j \le T - z$. Thus $(z, j) \in S$.

$$
\begin{aligned}
V_1 &= \sum_{s=1}^{T} \sum_{j=0}^{T-1} \sum_{z=1}^{\min\{s,T-j\}} \frac{D_1^{s+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}} \\
&= \sum_{j=0}^{T-1} \sum_{s=1}^{T} \sum_{z=1}^{\min\{s,T-j\}} \frac{D_1^{s+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}} \\
&= \sum_{s=1}^{T} \sum_{z=1}^{\min\{s,T-t+1\}} \frac{D_1^{s+t-2}}{\sum_{s'=0}^{T-z} D_1^{s'}} - (1 - \delta_{t,T}) \sum_{s=1}^{T} \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s+t-1}}{\sum_{s'=0}^{T-z} D_1^{s'}} \\
&= \sum_{s=1}^{T} \left[\sum_{z=1}^{\min\{s,T-t+1\}} \frac{D_1^{s+t-2}}{\sum_{s'=0}^{T-z} D_1^{s'}} - \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s+t-1}}{\sum_{s'=0}^{T-z} D_1^{s'}} + \delta_{t,T} \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s+t-1}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right] \\
&= \sum_{s=1}^{T} \left[\sum_{z=1}^{\min\{s,T-t+1\}} \frac{D_1^{s+t-2}}{\sum_{s'=0}^{T-z} D_1^{s'}} - \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s+t-1}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right]
\end{aligned}
$$

If $T = t = 2$,

$$
\begin{aligned}
V_1 &= \sum_{s=1}^{2} \left[\sum_{z=1}^{\min\{s,1\}} \frac{D_1^{s+2-2}}{\sum_{s'=0}^{2-z} D_1^{s'}} - \sum_{z=1}^{\min\{s,0\}} \frac{D_1^{s+2-1}}{\sum_{s'=0}^{2-z} D_1^{s'}} \right] \\
&= \sum_{s=1}^{2} \sum_{z=1}^{\min\{s,1\}} \frac{D_1^{s}}{\sum_{s'=0}^{2-z} D_1^{s'}} \\
&= \sum_{z=1}^{\min\{1,1\}} \frac{D_1^{1}}{\sum_{s'=0}^{2-z} D_1^{s'}} + \sum_{z=1}^{\min\{2,1\}} \frac{D_1^{2}}{\sum_{s'=0}^{2-z} D_1^{s'}} \\
&= \frac{D_1}{1 + D_1} + \frac{D_1^2}{1 + D_1} = D_1
\end{aligned}
$$

$$
\begin{aligned}
V_1 &= D_1^{t-1} \sum_{s=1}^{T} \left[\sum_{z=1}^{\min\{s,T-t+1\}} \frac{D_1^{s-1}}{\sum_{s'=0}^{T-z} D_1^{s'}} - \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right] \\
&= D_1^{t-1} \left[\sum_{s=1}^{T} \sum_{z=1}^{\min\{s,T-t+1\}} \frac{D_1^{s-1}}{\sum_{s'=0}^{T-z} D_1^{s'}} - \sum_{s=1}^{T} \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right] \\
&= D_1^{t-1} \left[\sum_{s=0}^{T-1} \sum_{z=1}^{\min\{s+1,T-t+1\}} \frac{D_1^{s}}{\sum_{s'=0}^{T-z} D_1^{s'}} - \sum_{s=1}^{T} \sum_{z=1}^{\min\{s,T-t\}} \frac{D_1^{s}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right] \\
&= D_1^{t-1} \left[\sum_{z=1}^{\min\{1,T-t+1\}} \frac{D_1^{0}}{\sum_{s'=0}^{T-z} D_1^{s'}} + \sum_{s=1}^{T-1} \frac{D_1^{s}}{\sum_{s'=0}^{T-\min\{s,T-t\}-1} D_1^{s'}} - \sum_{z=1}^{\min\{T,T-t\}} \frac{D_1^{T}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right] \\
&= D_1^{t-1} \left[\frac{1}{\sum_{s'=0}^{T-1} D_1^{s'}} + \sum_{s=1}^{T-1} \frac{D_1^{s}}{\sum_{s'=0}^{T-\min\{s,T-t\}-1} D_1^{s'}} - \sum_{z=1}^{T-t} \frac{D_1^{T}}{\sum_{s'=0}^{T-z} D_1^{s'}} \right]
\end{aligned}
$$

Suppose $T = 3$ and $t = 2$.

$$
V_1 = \sum_{s=1}^{3} \left[\sum_{z=1}^{\min\{s,2\}} \frac{D_1^{s}}{\sum_{s'=0}^{3-z} D_1^{s'}} - \sum_{z=1}^{\min\{s,1\}} \frac{D_1^{s+1}}{\sum_{s'=0}^{3-z} D_1^{s'}} \right]
$$

$$
\begin{aligned}
V_1 &= \sum_{z=1}^{\min\{1,2\}} \frac{D_1^1}{\sum_{s'=0}^{3-z} D_1^{s'}} - \sum_{z=1}^{\min\{1,1\}} \frac{D_1^2}{\sum_{s'=0}^{3-z} D_1^{s'}} \\
&\quad + \sum_{z=1}^{\min\{2,2\}} \frac{D_1^2}{\sum_{s'=0}^{3-z} D_1^{s'}} - \sum_{z=1}^{\min\{2,1\}} \frac{D_1^3}{\sum_{s'=0}^{3-z} D_1^{s'}} \\
&\quad + \sum_{z=1}^{\min\{3,2\}} \frac{D_1^3}{\sum_{s'=0}^{3-z} D_1^{s'}} - \sum_{z=1}^{\min\{3,1\}} \frac{D_1^4}{\sum_{s'=0}^{3-z} D_1^{s'}} \\
&= \frac{D_1}{1+D_1+D_1^2} - \frac{D_1^2}{1+D_1+D_1^2} \\
&\quad + \frac{D_1^2}{1+D_1+D_1^2} + \frac{D_1^2}{1+D_1} - \frac{D_1^3}{1+D_1+D_1^2} \\
&\quad + \frac{D_1^3}{1+D_1+D_1^2} + \frac{D_1^3}{1+D_1} - \frac{D_1^4}{1+D_1+D_1^2} \\
&= \frac{D_1 - D_1^4}{1+D_1+D_1^2} + D_1^2 \\
&= \frac{D_1 - D_1^4 + D_1^2 + D_1^3 + D_1^4}{1+D_1+D_1^2} = \frac{D_1 + D_1^2 + D_1^3}{1+D_1+D_1^2} = D_1
\end{aligned}
$$

If $T = t = 3$,

$$
\begin{aligned}
V_1 &= \sum_{s=1}^{3} \left[\sum_{z=1}^{\min\{s,1\}} \frac{D_1^{s+1}}{\sum_{s'=0}^{3-z} D_1^{s'}} - \sum_{z=1}^{\min\{s,0\}} \frac{D_1^{s+2}}{\sum_{s'=0}^{3-z} D_1^{s'}} \right] \\
&= \frac{D_1^2}{1+D_1+D_1^2} + \frac{D_1^3}{1+D_1+D_1^2} + \frac{D_1^4}{1+D_1+D_1^2} = D_1^2
\end{aligned}
$$

$$
V_1 = D_1^{t-1} \left[\frac{1}{\sum_{s'=0}^{T-1} D_1^{s'}} + \sum_{s=1}^{T-1} \frac{D_1^s}{\sum_{s'=0}^{T-\min\{s,T-t\}-1} D_1^{s'}} - \sum_{z=1}^{T-t} \frac{D_1^T}{\sum_{s'=0}^{T-z} D_1^{s'}} \right]
$$

If $T = 3$ and $t = 2$,

$$
\begin{aligned}
V_1 &= D_1 \left[\frac{1}{\sum_{s'=0}^{2} D_1^{s'}} + \sum_{s=1}^{2} \frac{D_1^s}{\sum_{s'=0}^{3-\min\{s,1\}-1} D_1^{s'}} - \sum_{z=1}^{1} \frac{D_1^3}{\sum_{s'=0}^{3-z} D_1^{s'}} \right] \\
&= D_1 \left[\frac{1}{1+D_1+D_1^2} + \frac{D_1}{1+D_1} + \frac{D_1^2}{1+D_1} - \frac{D_1^3}{1+D_1+D_1^2} \right] \\
&= D_1 \left[\frac{1-D_1^3}{1+D_1+D_1^2} + D_1 \right] = D_1 \frac{1 - D_1^3 + D_1 + D_1^2 + D_1^3}{1+D_1+D_1^2} = D_1
\end{aligned}
$$

$$V_1 = \sum_{s=1}^{T} \sum_{z=1}^{s} \sum_{j=0}^{T-z} \frac{D_1^{s+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}}$$

Let $S = \{(s,z) : 1 \leq s \leq T \wedge 1 \leq z \leq s\}$ and $S' = \{(s,z) : 1 \leq z \leq T \wedge z \leq s \leq T\}$. Let $(s,z) \in S$. Then $1 \leq s \leq T \wedge 1 \leq z \leq s$, so $1 \leq z \leq s \leq T$ and $z \leq s \leq T$, so $(s,z) \in S'$.

Let $(s,z) \in S'$. Then $1 \leq z \leq T \wedge z \leq s \leq T$, so $1 \leq z \leq s \leq T$ and $1 \leq z \leq s$.

$$\begin{aligned}
V_1 &= \sum_{z=1}^{T} \sum_{s=z}^{T} \sum_{j=0}^{T-z} \frac{D_1^{s+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}} \\
&= \sum_{z=1}^{T} \sum_{s=0}^{T-z} \sum_{j=0}^{T-z} \frac{D_1^{s+z+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}} \\
&= \sum_{z=1}^{T} \sum_{j=0}^{T-z} \sum_{s=0}^{T-z} \frac{D_1^{s+z+j-1}(\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-z} D_1^{s'}} \\
&= \sum_{z=1}^{T} \sum_{j=0}^{T-z} D_1^{z+j-1}(\delta_{j+1,t} - \delta_{j,t})
\end{aligned}$$

Let $T = t = 2$.

$$\begin{aligned}
V_1 &= \sum_{z=1}^{2} \sum_{j=0}^{2-z} D_1^{z+j-1}(\delta_{j+1,2} - \delta_{j,2}) \\
&= \sum_{j=0}^{1} D_1^{j}(\delta_{j+1,2} - \delta_{j,2}) + \sum_{j=0}^{0} D_1^{j+1}(\delta_{j+1,2} - \delta_{j,2}) = D_1
\end{aligned}$$

Let $T = t = 3$

$$\begin{aligned}
V_1 &= \sum_{z=1}^{3} \sum_{j=0}^{3-z} D_1^{z+j-1}(\delta_{j+1,3} - \delta_{j,3}) \\
&= \sum_{j=0}^{2} D_1^{j}(\delta_{j+1,3} - \delta_{j,3}) \\
&\quad + \sum_{j=0}^{1} D_1^{j+1}(\delta_{j+1,3} - \delta_{j,3}) \\
&\quad + \sum_{j=0}^{0} D_1^{j+2}(\delta_{j+1,3} - \delta_{j,3}) \\
&= D_1^2
\end{aligned}$$

Let $T = t = 2$

$$\begin{aligned} V_1 &= \sum_{z=1}^{3}\sum_{j=0}^{3-z} D_1^{z+j-1}(\delta_{j+1,2} - \delta_{j,2}) \\ &= \sum_{j=0}^{2} D_1^{j}(\delta_{j+1,2} - \delta_{j,2}) \\ &\quad + \sum_{j=0}^{1} D_1^{j+1}(\delta_{j+1,2} - \delta_{j,2}) \\ &\quad + \sum_{j=0}^{0} D_1^{j+2}(\delta_{j+1,2} - \delta_{j,2}) \\ &= D_1 - D_1^2 + D_1^2 = D_1 \end{aligned}$$

$$V_1 = \sum_{z=1}^{T}\sum_{j=0}^{T-z} D_1^{z+j-1}(\delta_{j+1,t} - \delta_{j,t})$$

$$\begin{aligned} S &= \{(z,j) : 1 \leq z \leq T \wedge 0 \leq j \leq T - z\} \\ S' &= \{(z,j) : 0 \leq j \leq T - 1 \wedge 1 \leq z \leq T - j\} \end{aligned}$$

Let $(z,j) \in S$. Then $1 \leq z \leq T \wedge 0 \leq j \leq T - z$. So $z \leq T - j$, and $1 \leq z \leq T - j$ while $0 \leq j \leq T - z \leq T - 1$. Thus $(z,j) \in S'$.

Let $(z,j) \in S'$. Then $0 \leq j \leq T - 1 \wedge 1 \leq z \leq T - j$. So $j \leq T - z$, so $0 \leq j \leq T - z$. $1 \leq z \leq T - j \leq T$. Thus $(z,j) \in S$.

$$\begin{aligned} V_1 &= \sum_{j=0}^{T-1}\sum_{z=1}^{T-j} D_1^{z+j-1}(\delta_{j+1,t} - \delta_{j,t}) \\ &= \sum_{z=1}^{T-(t-1)} D_1^{z+t-2} - (1 - \delta_{Tt})\sum_{z=1}^{T-t} D_1^{z+t-1} \end{aligned}$$

If $T = t$,

$$V_1 = \sum_{z=1}^{1} D_1^{z+t-2} = D_1^{T-1}$$

If $t < T$,

$$V_1 = \sum_{z=1}^{T-t+1} D_1^{z+t-2} - \sum_{z=1}^{T-t} D_1^{z+t-1}$$

Let $s = z - 1$, so $z = s + 1$.

$$\begin{aligned} V_1 &= \sum_{s=0}^{T-t} D_1^{s+1+t-2} - \sum_{z=1}^{T-t} D_1^{z+t-1} \\ &= \sum_{s=0}^{T-t} D_1^{s+t-1} - \sum_{z=1}^{T-t} D_1^{z+t-1} \\ &= D_1^{t-1} \end{aligned}$$

Thus

$$\left.\frac{\partial \Delta U_1}{\partial \varepsilon_t}\right|_{\varepsilon=0} = \sum_{i=0}^{T-1} \sum_{s=i+1}^{T} D_1^{s-1} \left(\frac{\sum_{j=0}^{T-s+i} D_1^j (\delta_{j+1,t} - \delta_{j,t})}{\sum_{s'=0}^{T-s+i} D_1^{s'}} - \delta_{i+1,t} + \delta_{i,t} \right) = D_1^{t-1} - D_1^{t-1} = 0$$

E Hessian of ΔU_1 at the Origin for $T = 3$

$$\begin{aligned} \Delta U_1 &= \sum_{s=1}^{T-1} \left[\ln \overline{\phi}_s \sum_{t=\max\{T-s-1,0\}}^{T-1} D_t - \ln \phi_s \sum_{t=\max\{0,s\}}^{T-1} D_t \right] \\ &= \sum_{s=1}^{T-1} \left[\ln \overline{\phi}_s \sum_{t=T-1-s}^{T-1} D_t - \ln \phi_s \sum_{t=s} D_t \right] \end{aligned}$$

Only ϕ_{T-1} and $\overline{\phi}_{T-1}$ will depend on ε_3, so

$$\Delta U_1 = \ln \overline{\phi}_{T-1} \sum_{t=0}^{T-1} D_t - D_{T-1} \ln \left(\frac{1 + \varepsilon_T}{1 + \varepsilon_{T-1}} \right)$$

$$\overline{\phi}_{T-1} = \frac{\sum_{z=0}^{T-1} D_z \phi_z}{\sum_{z=0}^{T-1} D_z}$$

$$\frac{\partial \ln \overline{\phi}_{T-1}}{\partial \varepsilon_T} = \frac{D_{T-1} \frac{1}{1+\varepsilon_{T-1}}}{\sum_{z=0}^{T-1} D_z \phi_z} = \frac{D_1^{T-1}}{\sum_{z=0}^{T-1} D_z \phi_z}$$

$$\frac{\partial \Delta U_1}{\partial \varepsilon_T} = \frac{D_1^{T-1}}{\sum_{z=0}^{T-1} D_z \phi_z} \sum_{t=0}^{T-1} D_t - \frac{D_{T-1}}{1 + \varepsilon_T}$$

If $\varepsilon_2 = \cdots = \varepsilon_{T-1} = 0$,

$$\begin{aligned}
\frac{\partial \Delta U_1}{\partial \varepsilon_T} &= D_1^{T-1} \frac{\sum_{t=0}^{T-1} D_1^t}{\sum_{z=0}^{T-2} D_1^z + D_1^{T-1}(1+\varepsilon_T)} - \frac{D_1^{T-1}}{1+\varepsilon_T} \\
&= D_1^{T-1} \frac{(1+\varepsilon_T)\sum_{t=0}^{T-1} D_1^t - \left[\sum_{z=0}^{T-2} D_1^z + D_1^{T-1}(1+\varepsilon_T)\right]}{\left(\sum_{z'=0}^{T-1} D_1^{z'} + D_1^{T-1}\varepsilon_T\right)(1+\varepsilon_T)} \\
&= D_1^{T-1} \frac{(1+\varepsilon_T)\sum_{t=0}^{T-2} D_1^t - \sum_{z=0}^{T-2} D_1^z}{\left(\sum_{z'=0}^{T-1} D_1^{z'} + D_1^{T-1}\varepsilon_T\right)(1+\varepsilon_T)} \\
&= D_1^{T-1} \frac{\sum_{t=0}^{T-2} D_1^t}{\left(\sum_{z'=0}^{T-1} D_1^{z'} + D_1^{T-1}\varepsilon_T\right)(1+\varepsilon_T)} \varepsilon_T
\end{aligned}$$

This is positive except when $\varepsilon_T = 0$.

$$\frac{\partial^2 \Delta U_1}{\partial \varepsilon_T^2} = D_1^{T-1} \sum_{t=0}^{T-2} D_1^t \frac{\left(\sum_{z=0}^{T-1} D_1^z + D_1^{T-1}\varepsilon_T\right)(1+\varepsilon_T) - \varepsilon_T\left[\sum_{z=0}^{T-1} D_1^z + D_1^{T-1}\varepsilon_T + D_1^{T-1}(1+\varepsilon_T)\right]}{\left[\left(\sum_{z'=0}^{T-1} D_1^{z'} + D_1^{T-1}\varepsilon_T\right)(1+\varepsilon_T)\right]^2}$$

$$\left.\frac{\partial^2 \Delta U_1}{\partial \varepsilon_T^2}\right|_{\varepsilon=0} = D_1^{T-1} \sum_{t=0}^{T-2} D_1^t \frac{\sum_{z=0}^{T-1} D_1^z}{\left[\sum_{z'=0}^{T-1} D_1^{z'}\right]^2} > 0$$

Thus if $\varepsilon_2 = \cdots = \varepsilon_{T-1} = 0$, $\Delta U_1 \geq 0$ with equality only if $\varepsilon_T = 0$.

$$\begin{aligned}
\Delta U_1 &= (D_1 + D_1^2(1+\varepsilon_2))\left(\ln\left(1 + \frac{D_1}{1+D_1}\varepsilon_2\right) - \ln(1+\varepsilon_2)\right) \\
&\quad + (1 + D_1 + D_1^2(1+\varepsilon_2))\ln\left(\frac{1 + D_1 + D_1^2 + D_1\varepsilon_2 + D_1^2\varepsilon_3}{1 + D_1 + D_1^2 + D_1^2\varepsilon_2}\right) - D_1^2(1+\varepsilon_2)\ln\left(\frac{1+\varepsilon_3}{1+\varepsilon_2}\right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \Delta U_1}{\partial \varepsilon_2} &= D_1^2\left(\ln\left(1 + \frac{D_1}{1+D_1}\varepsilon_2\right) - \ln(1+\varepsilon_2)\right) + (D_1 + D_1^2(1+\varepsilon_2))\left(\frac{\frac{D_1}{1+D_1}}{1 + \frac{D_1}{1+D_1}\varepsilon_2} - \frac{1}{1+\varepsilon_2}\right) \\
&\quad + D_1^2 \ln\left(\frac{1 + D_1 + D_1^2 + D_1\varepsilon_2 + D_1^2\varepsilon_3}{1 + D_1 + D_1^2 + D_1^2\varepsilon_2}\right) \\
&\quad + (1 + D_1 + D_1^2(1+\varepsilon_2))\left[\frac{D_1}{1 + D_1 + D_1^2 + D_1\varepsilon_2 + D_1^2\varepsilon_3} - \frac{D_1^2}{1 + D_1 + D_1^2 + D_1^2\varepsilon_2}\right] \\
&\quad - D_1^2 \ln\left(\frac{1+\varepsilon_3}{1+\varepsilon_2}\right) + D_1^2 \frac{1+\varepsilon_2}{1+\varepsilon_2}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \Delta U_1}{\partial \varepsilon_2} &= D_1^2\left(\ln\left(1+\frac{D_1}{1+D_1}\varepsilon_2\right)-\ln(1+\varepsilon_2)\right)+(D_1+D_1^2(1+\varepsilon_2))\left(\frac{D_1}{1+D_1+D_1\varepsilon_2}-\frac{1}{1+\varepsilon_2}\right)\\
&+D_1^2\ln\left(\frac{1+D_1+D_1^2+D_1\varepsilon_2+D_1^2\varepsilon_3}{1+D_1+D_1^2+D_1^2\varepsilon_2}\right)\\
&+(1+D_1+D_1^2(1+\varepsilon_2))\left[\frac{D_1}{1+D_1+D_1^2+D_1\varepsilon_2+D_1^2\varepsilon_3}-\frac{D_1^2}{1+D_1+D_1^2+D_1^2\varepsilon_2}\right]\\
&-D_1^2\ln\left(\frac{1+\varepsilon_3}{1+\varepsilon_2}\right)+D_1^2
\end{aligned}
$$

As a check,

$$
\begin{aligned}
\left.\frac{\partial \Delta U_1}{\partial \varepsilon_2}\right|_{\varepsilon_2=\varepsilon_3=0} &= (D_1+D_1^2)\left(\frac{D_1}{1+D_1}-1\right)\\
&+(1+D_1+D_1^2)\left[\frac{D_1-D_1^2}{1+D_1+D_1^2}\right]+D_1^2\\
&= -\frac{D_1+D_1^2}{1+D_1}+D_1-D_1^2+D_1^2=0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \Delta U_1}{\partial \varepsilon_2} &= (D_1+D_1^2(1+\varepsilon_2))\left(\frac{D_1}{1+D_1+D_1\varepsilon_2}-\frac{1}{1+\varepsilon_2}\right)+D_1^2\\
&+D_1^2\ln\left(\left(1+\frac{D_1}{1+D_1}\varepsilon_2\right)\frac{1+D_1+D_1^2+D_1\varepsilon_2+D_1^2\varepsilon_3}{(1+D_1+D_1^2+D_1^2\varepsilon_2)(1+\varepsilon_3)}\right)\\
&+(1+D_1+D_1^2(1+\varepsilon_2))\left[\frac{D_1}{1+D_1+D_1^2+D_1\varepsilon_2+D_1^2\varepsilon_3}-\frac{D_1^2}{1+D_1+D_1^2+D_1^2\varepsilon_2}\right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \Delta U_1}{\partial \varepsilon_2 \partial \varepsilon_3} &= D_1^2\left[\frac{D_1^2}{1+D_1+D_1^2+D_1\varepsilon_2+D_1^2\varepsilon_3}-\frac{1}{1+\varepsilon_3}\right]\\
&-D_1^3\frac{1+D_1+D_1^2(1+\varepsilon_2)}{(1+D_1+D_1^2+D_1\varepsilon_2+D_1^2\varepsilon_3)^2}
\end{aligned}
$$

$$
\begin{aligned}
\left.\frac{\partial^2 \Delta U_1}{\partial \varepsilon_2 \partial \varepsilon_3}\right|_{\varepsilon_2=\varepsilon_3=0} &= D_1^2\left[\frac{D_1^2}{1+D_1+D_1^2}-1\right]-D_1^3\frac{1+D_1+D_1^2}{(1+D_1+D_1^2)^2}\\
&= -D_1^2\frac{1+D_1}{1+D_1+D_1^2}-\frac{D_1^3}{1+D_1+D_1^2}
\end{aligned}
$$

$$
\left.\frac{\partial^2 \Delta U_1}{\partial \varepsilon_2 \partial \varepsilon_3}\right|_{\varepsilon_2=\varepsilon_3=0} = -\frac{D_1^2}{1+D_1+D_1^2}(1+2D_1)
$$

$$\left.\frac{\partial^2 \Delta U_1}{\partial \varepsilon_3^2}\right|_{\varepsilon_2=\varepsilon_3=0} = \frac{D_1^2}{1+D_1+D_1^2}(1+D_1)$$

Meanwhile,

$$\left.\frac{\partial^2 \Delta U_1}{\partial \varepsilon_2^2}\right|_{\varepsilon_2=\varepsilon_3=0} = \frac{D_1(1+D_1+4D_1^2+3D_1^3)}{(1+D_1)(1+D_1+D_1^2)}$$

$$\begin{aligned}
\Delta U_1 &= \frac{D_1^2+D_1^3}{2(1+D_1+D_1^2)}\varepsilon_3^2 - \frac{D_1^2+2D_1^3}{1+D_1+D_1^2}\varepsilon_2\varepsilon_3 + \frac{D_1+D_1^2+4D_1^3+3D_1^4}{2(1+D_1)(1+D_1+D_1^2)}\varepsilon_2^2 + O(\varepsilon^3) \\
&= \frac{1}{2}\frac{1}{1+D_1+D_1^2}\begin{bmatrix} \varepsilon_2 & \varepsilon_3 \end{bmatrix} \begin{bmatrix} D_1^2(1+D_1) & -D_1^2(1+2D_1) \\ -D_1^2(1+2D_1) & \frac{D_1+D_1^2+4D_1^3+3D_1^4}{(1+D_1)} \end{bmatrix} \begin{bmatrix} \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} + O(\varepsilon^3)
\end{aligned}$$

$$\begin{aligned}
& \begin{vmatrix} D_1^2(1+D_1) & -D_1^2(1+2D_1) \\ -D_1^2(1+2D_1) & \frac{D_1+D_1^2+4D_1^3+3D_1^4}{(1+D_1)} \end{vmatrix} \\
= \; & D_1^3+D_1^4+4D_1^5+3D_1^6 - D_1^4(1+2D_1)^2 \\
= \; & D_1^3+D_1^4+4D_1^5+3D_1^6 - D_1^4 - 4D_1^5 - 4D_1^6 \\
= \; & D_1^3 - D_1^6 \\
= \; & D_1^3(1-D_1^3)
\end{aligned}$$

If $D_1 > 1$, $\Delta U_1 < 0$ is possible. However, if $D_1 < 1$, the determinant is nonnegative. Thus in a deleted neighborhood of $(\varepsilon_2, \varepsilon_3) = (0,0)$, ΔU_1 must be strictly nonnegative.

F Sufficient Upper Bound on ε_T for Pareto Dominance of the Commitment Path

$$\Delta U_\tau = \sum_{t=\tau}^{T}\sum_{i=0}^{t-1} D_{t-\tau} \ln \frac{\overline{\phi}_{T-t+i}}{\phi_i}.$$

We can rewrite this as

$$\Delta U_\tau = \sum_{t=\tau}^{T}\sum_{j=T-t}^{T-1} D_{t-\tau} \ln \overline{\phi}_j - \sum_{t=\tau}^{T}\sum_{i=0}^{t-1} D_{t-\tau} \ln \phi_i,$$

where $j = T-t+i$. The first terms are all positive while the second terms are all negative.

$S = \{(t,i) : \tau \le t \le T \wedge 0 \le i \le t-1\}$. $S' = \{(t,i) : 0 \le i \le T-1 \wedge \max\{\tau, i+1\} \le t \le T\}$. Let $(t,i) \in S$, so $\tau \le t \le T \wedge 0 \le i \le t-1$. Then $0 \le i \le t-1 \le T-1$. We have both $\tau \le t$ and $i+1 \le t$, so $\max\{\tau, i+1\} \le t \le T$. Thus $(t,i) \in S'$.

Now let $(t,i) \in S'$, so $0 \le i \le T-1 \wedge \max\{\tau, i+1\} \le t \le T$. Then $\tau \le t \le T$. $0 \le i \le t-1$. Thus $(t,i) \in S$.

Let $S = \{(t,j) : \tau \le t \le T \wedge T-t \le j \le T-1\}$. Let $S' = \{(t,j) : 0 \le j \le T-1 \wedge \max\{\tau, T-j\} \le t \le T\}$. Let $(t,j) \in S$. Then $\tau \le t \le T \wedge T-t \le j \le T-1$. So $0 \le T-t \le j \le T-1$, and we have both $\tau \le t$ and $T-j \le t$, so $\max\{\tau, T-j\} \le t \le T$. Thus $(t,j) \in S'$. Let $(t,j) \in S'$. Then $0 \le j \le T-1 \wedge \max\{\tau, T-j\} \le t \le T$. $\tau \le t \le T$, and $T-t \le j \le T-1$. Thus $(t,j) \in S$.

$$\Delta U_\tau = \sum_{j=0}^{T-1} \sum_{t=\max\{\tau,T-j\}}^{T} D_{t-\tau} \ln \overline{\phi}_j - \sum_{i=0}^{T-1} \sum_{t=\max\{\tau,i+1\}}^{T} D_{t-\tau} \ln \phi_i.$$

The first terms are all positive and the second terms are all negative. Let us define

$$P_i^\tau = \sum_{t=\max\{\tau,T-i\}}^{T} D_{t-\tau} \tag{48}$$

and

$$Q_i^\tau = \sum_{t=\max\{\tau,i+1\}}^{T} D_{t-\tau} \tag{49}$$

Thus we have

$$\begin{aligned}
\Delta U_\tau &= \sum_{i=0}^{T-1} \left[P_i^\tau \ln \overline{\phi}_i - Q_i^\tau \ln \phi_i\right] \\
&= \sum_{i=0}^{T-1} \left[P_i^\tau \left(\ln \frac{\overline{\phi}_i}{\phi_i} + \ln \phi_i\right) - Q_i^\tau \ln \phi_i\right] \\
&= \sum_{i=0}^{T-1} \left[P_i^\tau \left(\ln \frac{\overline{\phi}_i}{\phi_i} + \sum_{j=i}^{T-1} \ln \phi_j - \sum_{j=i+1}^{T-1} \ln \phi_j\right) - Q_i^\tau \ln \phi_i\right] \\
&= \sum_{i=0}^{T-1} \left[P_i^\tau \left(\ln \frac{\overline{\phi}_i}{\phi_i} + \ln \frac{1+\varepsilon_T}{1+\varepsilon_i} - \sum_{j=i+1}^{T-1} \ln \phi_j\right) - Q_i^\tau \ln \phi_i\right]
\end{aligned}$$

$$\Delta U_\tau = \sum_{i=0}^{T-1} P_i^\tau \ln(1+\varepsilon_T) + \sum_{i=0}^{T-1} \left[P_i^\tau \left(\ln \frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{T-1} \ln \phi_j\right) - Q_i^\tau \ln \phi_i\right]$$

Suppose that $s < T-1$. Suppose that

$$\varepsilon_T \le B_s^\tau = \exp\left(\frac{\sum_{i=0}^{s}\left[P_i^\tau\left(\ln\frac{\phi_i}{\overline{\phi}_i} + \ln(1+\varepsilon_i) + \sum_{j=i+1}^{s}\ln\phi_j\right) + Q_i^\tau \ln\phi_i\right]}{\sum_{i'=0}^{T-1} P_{i'}^\tau}\right) - 1.$$

Then we will have

$$\sum_{i=0}^{T-1} P_i^\tau \ln(1+\varepsilon_T) + \sum_{i=0}^{s}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{s}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \le 0$$

since

$$\begin{aligned} 0 \;\;\ge\;\; & \sum_{i=0}^{s}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=s+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \\ & + \sum_{i=s+1}^{T-1}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \end{aligned}$$

we have

$$\begin{aligned} 0 \;\;\ge\;\; & \sum_{i=0}^{T-1} P_i^\tau \ln(1+\varepsilon_T) + \sum_{i=0}^{s}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{s}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \\ & + \sum_{i=0}^{s}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=s+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \\ & + \sum_{i=s+1}^{T-1}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \\ = \;\; & \sum_{i=0}^{T-1} P_i^\tau \ln(1+\varepsilon_T) + \sum_{i=0}^{s}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \\ & + \sum_{i=s+1}^{T-1}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] \\ = \;\; & \sum_{i=0}^{T-1} P_i^\tau \ln(1+\varepsilon_T) + \sum_{i=0}^{T-1}\left[P_i^\tau\left(\ln\frac{\overline{\phi}_i}{\phi_i} - \ln(1+\varepsilon_i) - \sum_{j=i+1}^{T-1}\ln\phi_j\right) - Q_i^\tau\ln\phi_i\right] = \Delta U_\tau. \end{aligned}$$