Time-Inconsistency, Commitment Device, and Allocation of Authority amongst Generations*

SangHo Yoon†

Work-in-Progress (July 30, 2010)

Abstract

The model of non-stationary (e.g., hyperbolic-like) future discounting suggests that the interests of subsequent generations may be in the conflict. If then, would it be always efficient for present generation to exhort its own will to future generations? The answer may not be always positive as devising and implementing the control strategy (i.e., commitment devise) is often very costly. I first identify the optimal condition for devising and implementing the control strategy (i.e., commitment device) by comparing naive and sophisticated dynasties. In particular, we investigates how a dynasty of subsequent generations may use the constitution as a vehicle to align the interests amongst generations. In doing so, we show why frequent revisions or adoptions of new constitutions may be costly to the society. The debate between Jefferson and Madison over the longevity of constitution exemplifies our argument.

JEL Classification: D11, D62, D91
Keywords: Hyperbolic and Exponential Discounting, Time-Inconsistency, Commitment Device, Authority over Time

1 Introduction

The model of non-stationary (e.g., hyperbolic-like) future discounting suggests that the interests of subsequent generations may be in the conflict. For example, consider the utility of generation i that is altruistic toward subsequent generations and, therefore, holds the preference in the following forward-looking and additively separable form.

$$U^i = \sum_{t=0}^{T-i} f(t)u(c_{i+t}) \tag{1}$$

*I thank. . . Any remaining errors are my responsibility.

†...

where $f(\tau)$ is a non-stationary discount function and $u(c_{i+\tau})$ is an instantaneous subutility function at i. Given a fixed consumption level $\bar{c}$ and a small incremental change x, generation i displays the 'preference reversal' such that

$$\frac{f(t)}{f(t-1)} < \frac{u(\bar{c})}{u(\bar{c}+x)} < \frac{f(t')}{f(t'-1)} \tag{2}$$

where $t < t'$; that is, generation i's preference for future generations' consumption relative to their succeeding generation's changes over time. Such a dynasty of generations with 'time-variant preference', $f(\tau)$ as depicted in eq. (2), faces the problem of intertemporal conflict of interest and in consequence, is subject to an inconsistent planning. Since the seminal work of Strotz (1956), the model of non-stationary future discounting has been framed either as a sequential game amongst multiple selves over time (Pollak, 1968; Peleg and Yaari, 1973; Goldman, 1980; Thaler, 1981; Schelling, 1984; Laibson, 1997; O'Donoghue and Rabin, 1999a; Palacios-Huerta, 2003) or though less frequently, as one amongst subsequent generations (Phelps and Pollak, 1968; Laibson, 1996; Barro, 1999).[1]

This paper extends the model of non-stationary future discounting. Its contribution lies on the application of non-stationary discounting model to the analysis on the intergenerational conflict of interests where an earlier generation is allowed to alter/bind the future generations' behavior by investing in the institutions that future generations inherit and become costly to deviate from.[2] That is, an earlier generation, if the dynasty is 'sophisticated' to understand the suboptimality of its optimal plan from the standpoint of future generations, does not only engage in "the strategy of consistent planning" of choosing the best plan that is actually followed by future generations (Strotz, 1956, p. 173), but also exerts its willpower upon future generations by altering their incentive structure or the domain of decision authority. In a sense, I consider an earlier generation to exercise both strategies of consistent planning and precommitment simultaneously, which are also designed voluntarily and endogenously.

1.1 A Brief Review

To make how this paper extends the model of intergenrational conflict crispier, let me step back a bit and consider a simple intergenrational sequential game in Phelps and Pollak (1968) that the dynasty of generations with 'time-variant' preference plays. The purpose of stepping back to a simple game here, however, is not to merely review the previous analysis on the problems of inconsistent planning and intergenerational conflict of interests, but is to identify the reason why an earlier generation may be motivated to alter/bind future generations.

To our knowledge, almost entire research on the non-stationary discounting of future regards the strategy of consistent planning as a sole endogenous option against the preference

[1] For the overview of literature on non-stationary future discounting, please see Loewenstein and Prelec (1992); Rabin (1998); Frederick, Loewenstein, and O'Donoghue (2002)

[2] For example, one may think of the U.S. constitution as a founding fathers' way not only to limit the scope of government over people but also the choice set of future generations.

reversal of future generations (or selves). Rather than finding/designing the technologies that enables its will to be exerted on future decisions, current generation of a hyperbolic dynasty is assumed to choose a suitable plan from which subsequent generation will not divert – in a game-theoretic term, this implies a subgame perfect equilibrium where the subsequent generations' options to deviate are eliminated. Furthermore, whether its participation is either voluntary or involuntary, the 'external' commitment technology that is designed, provided, and enforced by an outsider (e.g., social planner) dominates the welfare analysis of seeking the Pareto improvement.[3] To me, it is not clear why an earlier generation can exercise the strategy of consistent planning endogenously but has to resort to an external source for the 'commitment technology'. The first step of narrowing the asymmetric research programs in two strategies is to identify 'the value of precommitment' and examine the voluntary motive to implement it. Hopefully a simple extension of the analysis on the intergenerational game in Phelps and Pollak (1968) can illustrate why the voluntary motive for the precommitment strategy exists.

Given the characterization of generation i's preference in eq. (1) and its non-stationary discounting of future in eq. (2), generation i faces the problem of maximizing eq. (1) subject to

$$A_{i+t+1} = (1+r)(A_{i+t} - p_{i+t}c_{i+t}) \tag{3}$$

where A_{i+t} is the wealth generation $i+t$ has inherited from previous generation, r the interest rate, and p_τ the price of c_τ. To simplify the game even further, I assume no income where the problem is limited to consumption/saving decision of the initial endowment that maximizes the utility in eq. (1). Without the non-stationary discounting of future, the problem reduces to a standard intertemporal consumption/saving maximization that results in a usual intertemporal optimality condition of the equality between the marginal rate of substitution of succeeding consumptions and the ratio of discount and interest rates.

To derive any meaningful interpretation of the maximization problem with a time-variant preference, I adopt and apply the concept of naiveté and sophistication based on whether the dynasty of generations is aware of or unaware of subsequent recalculations of each generation's optimal plan: If the dynasty of generations is naive, a representative generation i falsely assumes that future generation $i+t$ will follow generation i's optimal plan, but if sophisticated, generation i correctly anticipates that future generation $i+t$'s optimal plan is different from generation i's and will not be followed.[4] Assuming that the price of consumption does not change over time $p = p_\tau$, I first note that the following Euler equation holds for the naive dynasty.

$$f(0)u'(c_i) = f(1)(1+r)u'(c_{i+1}) \tag{4}$$

[3] Harris and Laibson (2003, p. 287) provide a short description on the criteria of welfare analysis and commitment technologies.

[4] Though Strotz (1956, pp. 170-1) was not explicit about the distinction of naiveté and sophistication, his seminal essay noted the implication of periodical reevaluations and the "rational" response to the prospect of "intertemporal tussle". For further elaboration on naiveté and sophistication, see Harris and Laibson (2003, pp. 184–5) and the collaboration of O'Donoghue and Rabin (1999b,a, 2001).

of which the derivation can be found in Appendix A. Note that the planned sequence of consumption from eq. (4) resembles the property of "the optimal plan as seen today" in Strotz (1956, pp. 168–9) or the "first-best optimum" in Phelps and Pollak (1968, pp. 188–9); that is, letting $\mathbf{c}^*_{\mathbf{i,t}} = \{c^*_{i,i}, c^*_{i,i+1}, \ldots, c^*_{i,T-1}, c^*_{i,T}\}$ be the planned sequence of consumption that holds the first-order-condition (4) and assuming that future generation $i+t$ is effortlessly controlled to follow the generation i's plan, there is no other sequence of consumption $\mathbf{c}^{-*}_{\mathbf{i,t}} = \{c^{-*}_{i,i}, c^{-*}_{i,i+1}, \ldots, c^{-*}_{i,T-1}, c^{-*}_{i,T}\}$ such that $V^*_{i,t} = \sum_{t=0}^{T-i} f(t)u(c^*_{i,i+t}) \geq \sum_{t=0}^{T-i} f(t)u(c^{-*}_{i,i+t})$. In other words, the planned sequence of consumption $\mathbf{c}^*_{\mathbf{i,t}}$ leads to the maximum obtainable utility by a dynasty of 'time-variant' generations.

It would be, however, unrealistic to assume that "*even under conditions of certainty*, [subsequent generations $i+t$] will actually follow" generation i's optimal plan (Strotz, 1956, p. 170). Indeed, generation i's decision with the naiveté is based on a false assumption of the dynamic consistency of its optimal plan $\mathbf{c}^*_{\mathbf{i,t}}$. Actual consumptions of a naive dynasty follow a quite different path as subsequent generation $i+t$ recalculates own optimal plan. Given that the dynasty is naive such that the intergenerational tussle is unanticipated, $\mathbf{c}^*_{\mathbf{i,i}} = \{c^*_{i,i}, c^*_{i+1,i+1} \ldots, c^*_{T-1,T-1}, c^*_{T,T}\}$ is actual consumption path of a naive dynasty that holds the first-order-condition in eq. (4) only for the first period of each generation's optimal plans. I also let the utility from actual consumption path be $V^*_{i,i} = \sum_{t=0}^{T-i} f(t)u(c^*_{i+t,i+t})$.

This leads to a game-theoretic consideration of the strategic response against the anticipated behavior of future generation $i+t$, if the dynasty is sophisticated. Assuming that differentiable strategy profiles exist, a Markov-perfect subgame-perfect equilibrium for generation i in a sophisticated dynasty is described by the following Euler equation.

$$\begin{aligned} f(0)u'(c_i) = {} & f(1)(1+r)u'(c_{i+1}) \\ & + \sum_{t=2}^{T-i} \left(f(t) - \frac{f(1)}{f(0)} f(t-1) \right) \prod_{k=1}^{t-1} \left(1 - p\frac{\partial c_{i+k}}{\partial A_{i+k}} \right) p(1+r)^t \frac{\partial c_{i+t}}{\partial A_{i+t}} u'(c_{i+t}) \end{aligned} \tag{5}$$

of which derivation is again available in Appendix A.

Though the Euler equation for sophisticated dynasty eq. (5) looks complex and scary, its basic interpretation essentially remains unchanged as one for the naive dynasty. However, it represents a 'correct' first-order-condition that with a non-stationary discounting, a small permutation in generation i's consumption c_i requires a larger compensation in future generation $i+t$'s consumption c_{i+t} than what is found in an optimality condition either for standard intertemporal problem or for a naive dynasty (e.g., eq. (4)). Such an intuition is easily confirmed by the positive second term in the right-hand-side of eq. (5), which is immediately available from $\frac{f(t)}{f(t-1)} - \frac{f(1)}{f(0)} > 0$ and an empirically supported characteristic of non-stationary discounting of future noted in eq. (2).

The first order condition in eq. (5) resembles and generalizes an "equilibrium" concept introduced in Phelps and Pollak (1968, p. 193-4) where the "second-best optimum" is augmented with the anticipation of subsequent generation $i+t$'s recalculation of own optimal plan: The "second-best optimum" results from two key assumptions that subsequent generation $i+t$'s consumption is constant and beyond generation i's control. As the

		Consumption Rule	Consumption Path
Naive Dynasty	Plan	λ_i^* for $\tau = i$ λ_τ^* for $\tau \geq i+1$	$\mathbf{c^*_{i,t}} = \left\{ \begin{array}{l} \lambda_i^* A_i, \lambda_\tau^* A_i (1+r)(1-p\lambda_i^*), \\ \lambda_\tau^* A_i (1+r)^2 (1-p\lambda_i^*)(1-p\lambda_\tau^*), \\ \lambda_\tau^* A_i (1+r)^3 (1-p\lambda_i^*)(1-p\lambda_\tau^*)^2, \ldots \end{array} \right\}$
	Actual	λ_i^* for $\tau \geq i$	$\mathbf{c^*_{i,i}} = \left\{ \begin{array}{l} \lambda_i^* A_i, \lambda_i^* A_i (1+r)(1-p\lambda_i^*), \\ \lambda_i^* A_i (1+r)^2 (1-p\lambda_i^*)^2, \\ \lambda_i^* A_i (1+r)^3 (1-p\lambda_i^*)^3, \ldots \end{array} \right\}$
Sophisticated Dynasty		λ^{**} for $\tau \geq i$	$\mathbf{c^{**}_{i,t}} = \left\{ \begin{array}{l} \lambda^{**} A_i, \lambda^{**} A_i (1+r)(1-p\lambda^{**}), \\ \lambda^{**} A_i (1+r)^2 (1-p\lambda^{**})^2, \\ \lambda^{**} A_i (1+r)^3 (1-p\lambda^{**})^3, \ldots \end{array} \right\}$

Table 1: Generation i's Consumption Rule and Consumption Path

"equilibrium" reflects the generation i's best response given its "correct" anticipation of the future, the plan that meets eq. (5) will be actually followed by subsequent generation $i+t$. Thus, it mirrors the strategy of consistent planning in Strotz (1956, p. 173-6) that the intergeneration tussle is anticipated and gives no room for subsequent generation $i+t$ to deviate from i's plan. I define $\mathbf{c^{**}_{i,t}}$ and $V^{**}_{i,t}$ as the sophisticated dynasty's sequence of consumption and the utility that result from eq. (5), respectively.

As already indicated in Phelps and Pollak (1968, p. 196-7), however, the equilibrium from eq. (5) is "not Pareto optimal" and may be further improved if generation i finds the way in which subsequent generation $i+t$'s consumption becomes controllable. Indeed, many self-control technologies discussed in recent economics and psychology literature can be directly applied to the problem of intergenerational tussle: For example, just like a person joining the Christmas Club, generation i may set aside the fund that is costly to assess for a certain number of subsequent generations but is freed from the penalty when its maturity comes. The problem is, unlike the intrapersonal counterpart's, the control technologies applied to the intergenerational tussle usually cannot find a proper third-party enforcer – the social planner of generation i also disappears when subsequent generation $i+t$ occupies the society and is in control. It makes the 'external' control technologies incredible when applied to the intergenerational setting.

2 The Remaining Puzzle

The puzzle, thus, remains: If a dynasty of generations cannot find the third-party enforcer for the control technologies to reduce its intergenerational tussle, would the dynasty only resort to the strategy of consistent planning and give up on the strategy of precommitment? The answer would depend on the benefit and cost of employing the technology to control the behavior of subsequent generations. Thus, though Phelps and Pollak (1968) has already shown such the non-Pareto optimality with the strategy of consistent planning, we use this section to pinpoint such a suboptimality. Furthermore, this section shows that the associated consumption path needs to be augmented to approximate the Pareto optimality and that the strategy of consistent planning may even lead to the dynasty of exacerbating the consumption spree.

To see whether the further Pareto improvement is feasible, let us revisit the standard model of non-stationary discounting but now employ the quasi-hyperbolic discount function

$$f(t) = \frac{\beta}{(1+\rho)^t} \tag{6}$$

and the CRRA utility function

$$u(c_\tau) = \frac{c_\tau^{1-\sigma}}{1-\sigma} \tag{7}$$

This allows the characterization of consumption rules, the visualization of consumption paths, and the comparison of the associated utility that each type of dynasty plans and actualizes: Thus, we are able to highlight the possible benefit of controlling subsequent generations rather than only resorting to the strategy of consistent planning.

Using the derivations of consumption rules in Appendix A, I may summarize and specify the consumption rules and consumption path of naive and sophisticated dynasty in Table 1 where each consumption rule can be defined a s follows:

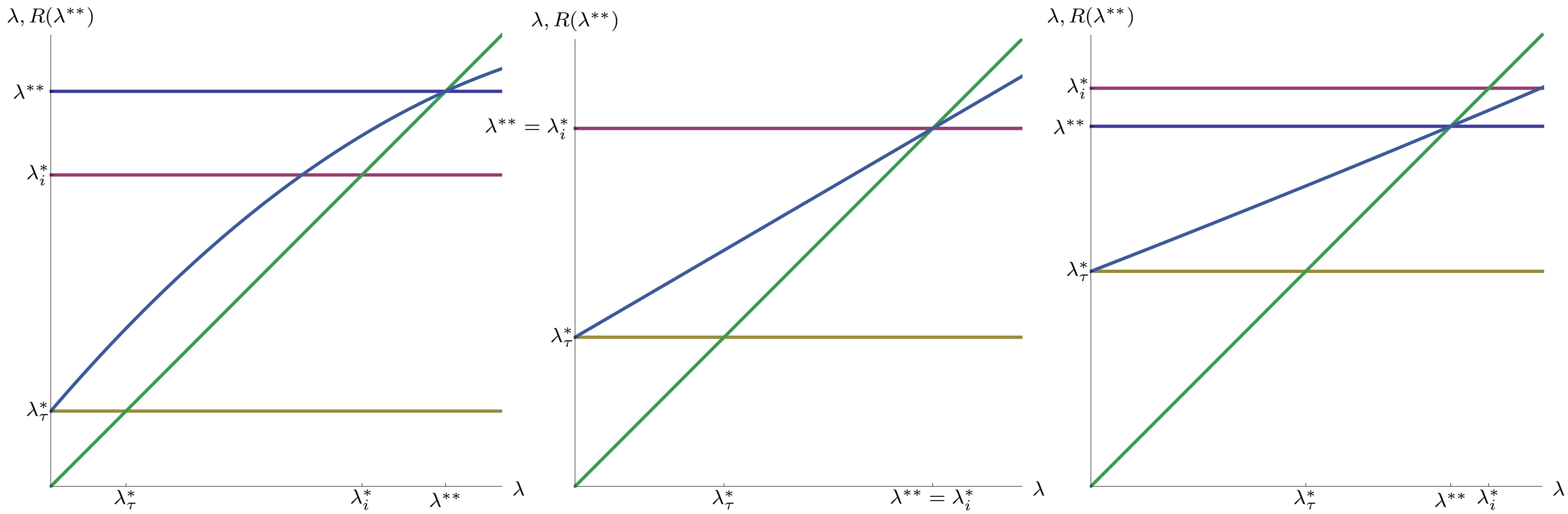
$$\lambda_\tau^* = \frac{1-((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}}}{p} \tag{8}$$

$$\lambda_i^* = \frac{1-((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}}}{p\left(1+((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}}(\beta^{\frac{1}{\sigma}}-1)\right)} > \lambda_\tau^* \text{ with } \beta < 1 \tag{9}$$

$$\lambda^{**} = \frac{1-((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1))^{\frac{1}{\sigma}}}{p} > \lambda_\tau^* \text{ with } \beta \leq 1,\ p\lambda^{**} \leq 1 \tag{10}$$

The naive dynasty uses λ_i^* for current generation i and λ_τ^* for subsequent generations $i+t$ in its planning, but as not being aware of the intergenerational tussle, continues to apply λ_i^* in actual decisions. On the other hand, the sophisticated dynasty who understands the problem of intergenerational conflict exercises the strategy of consistent planning and its plan via λ^{**} is actualized; the planned and actual decisions coincide for the sophisticated dynasty.

Let us first characterize each consumption rule, λ_i^*, λ_τ^*, and λ^{**}. As Figure 1 indicates and as intuitively anticipated, the intertemporal elasticity of substitution through σ dictates what kind of consumption rule is applied. For example, consider the case when $\sigma < 1$ or a higher substitutability of consumptions over generations. Understanding that subsequent generation $i+t$ will deviate from the plan at i, the sophisticated dynasty rather chooses a higher consumption rule than its naive counterpart such that $\lambda^{**} > \lambda_i^* > \lambda_\tau^*$ because a higher substitutability provides the flexibility to shift the consumptions intergenerationally: A high substitutability is analogous to the situation where anticipating the wasteful aid or charitable transfer via a (must) third-party, the society rather exacerbates own consumption



(a) when $\sigma < 1$ (b) when $\sigma = 1$ (c) when $\sigma > 1$

Figure 1: Consumption Rules; λ_τ^*, λ_i^*, and λ^{**}

Notes: The intersection of LHS and RHS in (10) is used to find the consumption rule of the sophisticated dynasty, λ^{**}.

$$\begin{aligned}
L(\lambda^{**}) &= \lambda^{**} & R(\lambda^{**}) &= \frac{1-\left((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1)\right)^{\frac{1}{\sigma}}}{p} \\
L'(\lambda^{**}) &= 1 & R'(\lambda^{**}) &= -\frac{(\beta-1)((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1))^{\frac{1}{\sigma}}}{((1-\beta)p\lambda^{**}+1)\sigma} \\
L''(\lambda^{**}) &= 0 & R''(\lambda^{**}) &= \frac{p(\beta-1)^2((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1))^{\frac{1}{\sigma}}(\sigma-1)}{((1-\beta)p\lambda^{**}+1)\sigma^2}
\end{aligned} \tag{11}$$

While LHS of eq. (10) is a linear 45 degree line, RHS is always increasing but concave when $\sigma < 1$, linear when $\sigma = 1$, and convex when $\sigma > 1$: λ^{**} is at the intersection of $L(\lambda^{**})$ (a 45 degree line) and $R(\lambda^{**})$ (a concave, linear, or convex curve), of which value depends on the shape of $R(\lambda^{**})$.

spree. When $\sigma > 1$, on the other hand, the sophisticated dynasty chooses the consumption rule such that $\lambda_i^* > \lambda^{**} > \lambda_\tau^*$ because though still understands and faces the same problem of intergenerational tussle, a lower substitutability implies the unexchangeability of intergenerational consumptions: A low substitutability is analogous to the situation where parents sacrifice to subsidize the children's consumption instead. When $\sigma = 0$, however, the sophisticated dynasty has no incentive to deviate from the naive counterpart's actual rule application.

Knowing that how each type of dynasty chooses and applies its consumption rule, Figure 2 now depicts consumption paths as well as their associated utilities where the dashed lines represent the first-best optimum where subsequent generations $i+t$ is assumed to consume as planned. Our main interest lies on the possibility for the further Pareto improvement towards the first-best-optimum: That is, we want to examine whether there is any additional utility available despite of the sophistication and the effort for the consistent planning. The direct observation of each type's utility in figure 2 suggests that the consumption path with the strategy of consistent planning is still suboptimal and further Pareto-improvement is possible: The area between the discounted subutility curve and τ-axis represents the dynasty's utility and it is easily observable that $V_{i,t}^* > V_{i,t}^{**}$.[5]

We formally summarize such an observation of the suboptimality as follows.

Proposition 1. *Given that any dynasty's utility may be restated as the function of current and subsequent λ's, $U^i = U(\lambda_i, \lambda_\tau)$,*

$$\frac{dU^i(\lambda_i, \lambda_\tau)}{d\lambda} = \frac{\partial U(\lambda_i, \lambda_\tau)}{\partial \lambda_i} + \frac{\partial U(\lambda_i, \lambda_\tau)}{\partial \lambda_\tau} < 0 \quad \textit{for some } \lambda < \lambda^{**} \textit{ and all } \lambda \geq \lambda^{**} \tag{12}$$

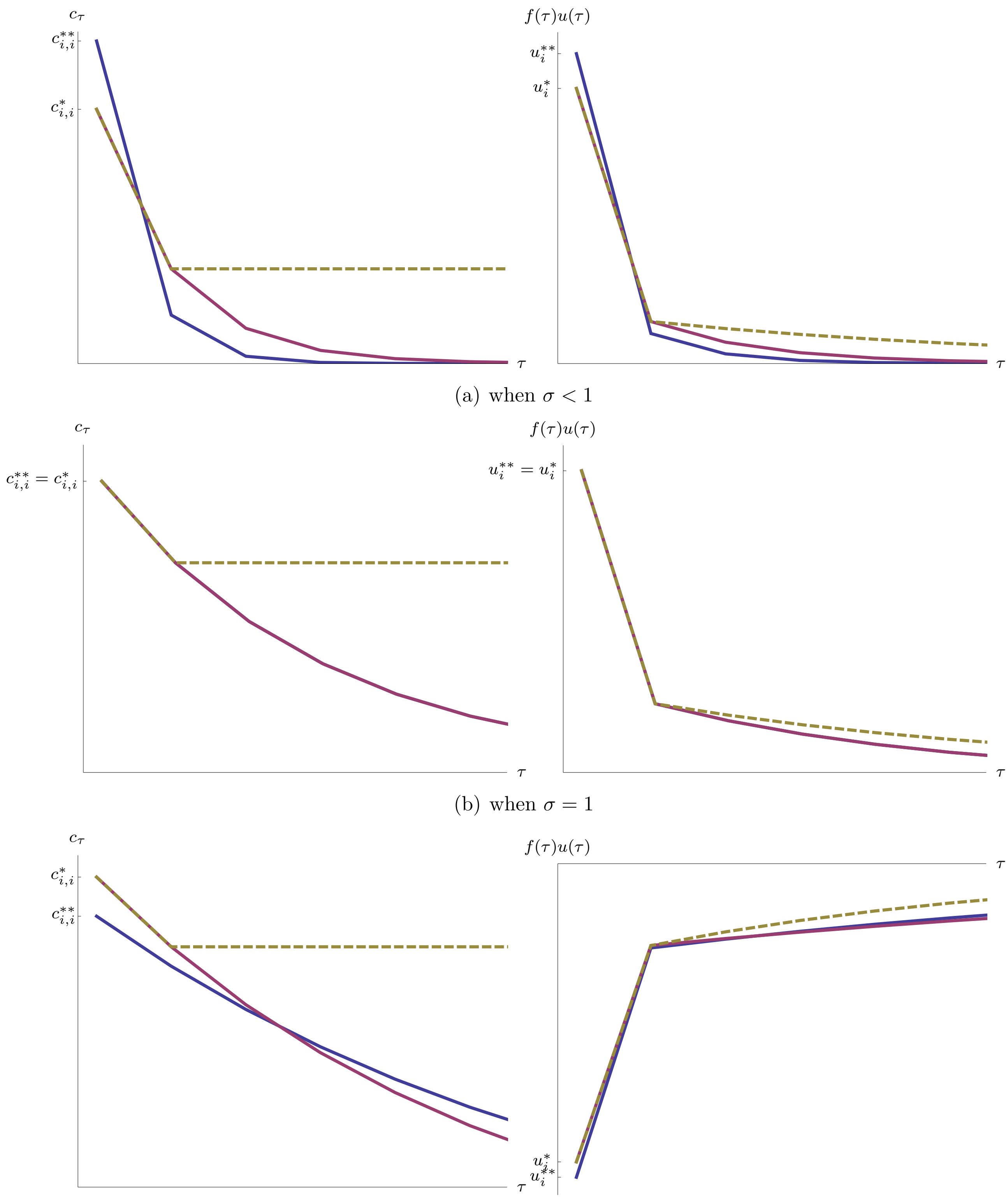
when evaluated at $\lambda = \lambda_i = \lambda_\tau$.

Proof. The proof is in Appendix B □

In other words, though any increase in the consumption rule beyond λ^{**} would make every generation in the dynasty worse-off, some decrease beyond λ^{**} would make every generation better-off, showing that λ^{**} is suboptimal and Pareto improvement is possible. Thus, every generation in the dynasty may be willing to lower the consumption rule if subsequent generations are held to do so. Another interpretation is that the wealth effect of leaving a higher stock of wealth to future generations may be large enough to compensate that from lowering the consumption rule, if every generation in the dynasty is forced to do so. In a sense that there are some other $\lambda < \lambda^{**}$ that dominates λ^{**}, the dynasty at the equilibrium of λ^{**} can be said to consume too much, and following corollary becomes immediately available.

Corollary 1. *The sophisticated dynasty, despite of its strategy of consistent planning, overconsumes by applying λ^{**} as its consumption rule.*

[5] When $\sigma > 1$, the utility is always negative where a larger area between the discounted subutility curve and τ-axis than the dashed curve's implies its suboptimality.



(a) when $\sigma < 1$

(b) when $\sigma = 1$

(c) when $\sigma > 1$

Figure 2: Consumption Path and Utility of a Dynasty

Proposition 1 and corollary 1 show that there is a room for the Pareto improvement. Moreover, as long as the cost-effective control technology is available, the dynasty with the intergenerational tussle is willing to implement it as the strategy of precommitment. In the next section, I extend the analysis by developing a simple model that the dynasty has an option of the strategy of precommitment to discipline its future generations. Such a control technology, however, is assumed to be 'endogenous' such that its effectiveness depends on the history of institutional investment.

3 The Model

We consider an augmented maximization problem of generation i of which the utility is now

$$\begin{aligned} U_i &= \sum_{t=0}^{T} f(t)u(c_{i+t}, H_{i+t}) \\ &= f(0)u(c_i, H_i) + f(1)u(c_{i+1}, H_{i+1}) + f(2)u(c_{i+2}, H_{i+2}) + \cdots \end{aligned} \tag{1'}$$

where H_{i+t} refers to the institutional setting that generation $i+t$ faces; or it can be interpreted as a reference point or norm that generation $i+t$ uses for its consumption decision. We assume that such institution evolves by following

$$H_{i+t+1} = (1-\delta)(h_{i+t} + H_{i+t}) \tag{13}$$

and is the outcome of the past generations' investment via h. However, the benefit from the investment in H_{i+t} is strictly from the future and has no value for the immediate gratification as h does not enter into the utility function directly. Furthermore, the equation of motion for the wealth is also augmented as

$$A_{i+t+1} = (1+r)(A_{i+t} - p_{i+t}c_{i+t} - h_{i+t}) \tag{3'}$$

Thus, though may not be perfect, generation i now has the ability to change the behavior of subsequent generations $i+t$ through H_{i+t}. Furthermore, the investment in the institution via h is 'voluntary' and 'endogenously' determined by evaluating its effectiveness and relative price to the consumption.

3.1 Augmented Naive Dynasty

Following Euler equation holds

3.2 Augmented Sophisticated Dynasty

$$f(1)\frac{\partial H_{i+1}}{\partial h} u_H(c_{i+1}, H_{i+1}) \tag{14}$$

Then using the linearized function that is normalized around zero

$$c = \lambda_A A + \lambda_H H \tag{15}$$

4 The Dynamics

[Write out the dynamics and results. Also talk about the debate between Jefferson and Madison about the 30-year revolution.]

[One of the motivations to develop and consider a type of model described above is to recognize a property of institution that Madison (1987 [1788], 1995 [1790]) emphasized. In arguing against a famous theme of Jefferson (1999 [1787], 1998 [1781], 1995 [1789]) on repetitive revolution and constitution convention of succeeding generations,[6] Madison (1987 [1788]) in Federalist Papers states that

> In the next place, it may be considered as an objection inherent in the principle, that as every appeal to the people would carry an implication of some defect in the government, frequent appeals would, in a great measure, deprive the government of that veneration which time bestows on every thing, and without which perhaps the wisest and freest governments would not possess the requisite stability. If it be true that all governments rest on opinion, it is no less true that the strength of opinion in each individual, and its practical influence on his conduct, depend much on the number which he supposes to have entertained the same opinion. The reason of man, like man himself, is timid and cautious when left alone, and acquires firmness and confidence in proportion to the number with which it is associated. When the examples which fortify opinion are ancient as well as numerous, they are known to have a double effect.

Furthermore, in his exchange of letters with Jefferson (1995 [1789]), on the earth belonging to the living and not to the dead, Madison (1995 [1790]) continues

> Would not a government, ceasing of necessity at the end of a given term, unless prolonged by some constitutional act previous to its expiration, be too subject to the casualty and consequences of an interregnum? Would not a government so often revised become too mutable and novel to retain that share of prejudice in its favor which is a salutary aid to the most rational government? Would not such a periodical revision engender pernicious factions that might not otherwise come into existence, and agitate the public mind more frequently and more violently than might be expedient?

It is clear that Madison (1987 [1788], 1995 [1790]) recognized the importance of cultivating the tradition intergenerationally around a stable governance and institution in a nation-building project: A role of H_{i+t} in eq. $(1')$ and (13) in the model is to allow the intergenerational transfer of institution in the growth theory; how does generation i decide to invest in institutions (i.e., bear the enforcement cost, h, of the rule of *de jure* law), from which generation i does not benefit but only succeeding generations at $i+1, \ldots, T-i$ do.

[6] The Jeffersonian theme is famously represented by "[g]od forbids we should ever be 20 years without such a rebellion " and "the earth belongs always to the living generation" (Jefferson, 1995 [1789], 1999 [1787]) – many contemporary authors interpret these remarks as "every generation needs a new revolution."

Of course, this type of property in institution is not new to economic literature. Becker (1992), in explaining how his interest in the theory of rational addiction developed, discusses the influence of past generations on current one through building the tradition around institutions in a society, and also refers to the assertion of Madison (1987 [1788]) made in Federalist Paper. Furthermore, a similar theme is promoted by Hayek (1973, p. 45) who states

> Some such rules all individuals of a society will obey because of the similar manner in which their environment represents itself to their minds, Others they will follow spontaneously because they will be part of their common cultural tradition. But there will be still others which they may have to be made to obey, since, although it would be in the interest of each to disregard them, the overall order on which the success of their actions depends will arise only if these rules are generally follows.

The purpose of this essay is to build upon this common theme on the intergenerational institution formation with its enforcement cost, and relates to economic development and growth. I am not aware of any work that employs and formalizes the institutional analysis in the growth theory. The first essay plans to do so around the model suggested previously.]

5 Conclusion

[Write the conclusion and contribution of the paper.]

References

Barro, R.J. 1999. "Ramsey Meets Laibson in The Neoclassical Growth Model*." *Quarterly Journal of Economics* 114 (4):1125–1152.

Becker, Gary S. 1992. "Habits, Addictions, and Traditions." *Kyklos* 45 (3):327–345.

Frederick, Shane, George Loewenstein, and Ted. O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* 40 (2):351–401.

Goldman, Steven M. 1980. "Consistent plans." *The Review of Economic Studies* :533–537.

Harris, Christopher and David Laibson. 2003. "Hyberbolic Discounting and Consumption." In *Advances in economics and econometrics: Theory and Applications, Eighth World Congress*, edited by E.L. Glaeser, J.A. Scheinkman, M. Dewatripont, LP Hansen, and S. Turnovsky. New York, NY: Cambridge University Press, 258–297.

Hayek, Friedrich A. 1973. *Law, Legislation, and Liberty, Vol. 1: Rules and Order*. Chicago: University of Chicago Press.

Jefferson, Thomas. 1995 [1789]. "Thomas Jefferson to James Madison, Paris, September 6, 1789." In *Jefferson and Madison: three conversations from the Founding*, edited by L. Banning. Madison House Pub.

———. 1998 [1781]. *Notes on the State of Virginia*. Penguin Books.

———. 1999 [1787]. "Jefferson Letter to William Stephens Smith, Paris, March 3, 1787." In *Jefferson: Political Writings*, edited by Joyce Appleby. Cambridge University Press.

Laibson, David. 1996. "Hyperbolic Discount Functions, Undersaving, and Savings Policy." Working Paper 5635, National Bureau of Economic Research.

———. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112 (2):443–477.

Loewenstein, George and Drazen Prelec. 1992. "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *Quarterly Journal of Economics* 107 (2):573–597.

Madison, James. 1987 [1788]. "Federalist 49." In *The federalist Papers*, edited by J. Madison, A. Hamilton, and J. Jay. Penguin Classics.

———. 1995 [1790]. "James Madison Letter to Thomas Jefferson, New York, February 4, 1790." In *Jefferson and Madison: three conversations from the Founding*, edited by L. Banning. Madison House Pub, 172–175.

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

———. 1999b. "Incentives For Procrastinators." *Quarterly Journal of Economics* 114 (3):769–816.

———. 2001. "Choice and Procrastination*." *Quarterly Journal of Economics* 116 (1):121–160.

Palacios-Huerta, Ignacio. 2003. "Time-Inconsistent Preferences in Adam Smith and David Hume." *History of Political Economy* 35 (2):241–268.

Peleg, Bezalel and Menahem E. Yaari. 1973. "On the existence of a consistent course of action when tastes are changing." *The Review of Economic Studies* 40 (3):391–401.

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

Pollak, Robert A. 1968. "Consistent Planning." *Review of Economic Studies* 35 (2):201–208.

Rabin, Matthew. 1998. "Psychology and Economics." *Journal of Economic Literature* 36 (1):11–46.

Schelling, Thomas C. 1984. “Self-Command in Practice, in Policy, and in a Theory of Rational Choice.” *American Economic Review: Papers and Proceedings* 74 (2):1–11.

Strotz, Robert H. 1956. “Myopia and Inconsistency in Dynamic Utility Maximization.” *Review of Economic Studies* 23 (3):165–180.

Thaler, Richard H. 1981. “Some Empirical Evidence on Dynamic Inconsistency.” *Economics Letters* :201–207.

A Derivation

This appendix provides the mathematical derivations that help understanding various discussions in the paper.

Standard Model of Non-Stationary Discounting, á la Phelps and Pollak (1968) and Laibson (1996): I first begin with a standard model of non-stationary discounting. The economic model of non-stationary discountning provided in the paper is more general than what have been presented in the literature, even more so than Phelps and Pollak (1968) and Laibson (1996). But it does not constitute the major contribution of this paper. However, the discussion in section 1 around the value of commitment technology and the voluntary motive to implement it is the direct consequence of derivations provided here.

Derivation of Naive Euler Equation, Eq. (4)*:* Assuming $u(c_\tau)$ is differentiable, standard method leads me to following Euler equation for generation i.

$$\begin{aligned} 0 &= f(0)u'(c_i) - f(1)\frac{\partial c_{i+1}}{\partial A_{i+1}}(1+r)p_i u'(c_{i+1}) \\ &\quad - f(2)\frac{\partial c_{i+2}}{\partial A_{i+2}}(1+r)^2 p_i \left(1 - p_{i+1}\frac{\partial c_{i+1}}{\partial A_{i+1}}\right) u'(c_{i+2}) \\ &\quad - f(3)\frac{\partial c_{i+3}}{\partial A_{i+3}}(1+r)^3 p_i \left(1 - p_{i+1}\frac{\partial c_{i+1}}{\partial A_{i+1}}\right)\left(1 - p_{i+2}\frac{\partial c_{i+2}}{\partial A_{i+2}}\right) u'(c_{i+3}) + \cdots \\ &= f(0)u'(c_i) - \sum_{k=1}^{T-t} f(k) \prod_{l=1}^{k-1}\left(1 - p_{i+l}\frac{\partial c_{i+l}}{\partial A_{i+l}}\right) p_i (1+r)^k \frac{\partial c_{i+k}}{\partial A_{i+k}} u'(c_{i+k}) \end{aligned} \tag{A.1}$$

The interpretation of eq. (A.1) is standard: It shows how much consumption of subsequent generations c_{i+t} need to be undone when there is a small permutation in c_i. The definition of naiveté in section 1 suggests that generation i is unaware of future generation $i+t$' recalculation of own optimal plans and despite of the existence of the conflict of interests, regards generation $i+t$ will follow generation i's optimal plan. In other words, generation i assumes that the time preference of future generation $i+t$ are same as how generation i views the future; formally letting $g_i(t) = \frac{f'(t)}{f(t)}$ be generation i's instantaneous discount rate applied to the well-being of future generation $i+t$, following equality holds

$$g_i(t_1) = g_j(t_2) \quad \text{where} \quad i < j,\ i + t_1 = j + t_2 \tag{A.2}$$

Thus, in evaluating the Euler equation for generation j, generation i in a naive dynasty assumes that the Euler equation for generation j holds in a such a manner that it is equivalent to how generation i evaluated the consumption of future generation $i+t$; that is, the optimal plan of generation j is assumed not to change from how generation i planned. With such an implication of the naiveté in mind, I use a similar method to derive the following Euler equation.

$$0 = f(1)u'(c_{i+1}) - \sum_{t=2}^{T-i} f(t) \prod_{k=2}^{t-1}\left(1 - p_{i+k}\frac{\partial c_{i+k}}{\partial A_{i+k}}\right) p_{i+1}(1+r)^{t-1}\frac{\partial c_{i+t}}{\partial A_{i+t}} u'(c_{i+t}) \tag{A.3}$$

Assuming the price of c_τ does not change over time $p = p_\tau$, multiplying $\left(1 - p\frac{\partial c_{i+1}}{\partial A_{i+1}}\right)(1+r)$ to (A.3) and subtracting it from eq. (A.1), I arrive at the Euler equation for the naive dynastic in

$$f(0)u'(c_i) = f(1)(1+r)u'(c_{i+1}) \tag{4}$$

Derivation of Sophisticated Euler Equation, Eq. (5)*:* If the dynasty is sophisticated, however, generation i correctly anticipates the intergenrational conflict of interests from generation $i+t$'s recalculation of own optimal plans; that is, the equality condition in instantaneous discount rates for the naiveté dynasty, eq. (A.2), does not hold anymore. Instead, a game-theoretic consideration of a strategic response becomes necessary where the correct Euler equation of generation $i+1$ is anticipated.

$$0 = f(0)u'(c_{i+1}) - \sum_{k=2}^{T-t} f(k-1) \prod_{l=2}^{k-1} \left(1 - p_{i+l}\frac{\partial c_{i+l}}{\partial A_{i+l}}\right) p_{i+1}(1+r)^{k-1}\frac{\partial c_{i+k}}{\partial A_{i+k}} u'(c_{i+k}) \tag{A.4}$$

Similar to the derivation of the naive Euler equation in eq. (4), assuming the price of c_τ does not change over time $p = p_\tau$, multiplying $\frac{f(1)}{f(0)}\left(1 + p\frac{\partial c_{i-1}}{\partial A_{i+1}}\right)(1+r)$ to eq. (A.4), and then subtracting it from eq. (A.1) leads to the Euler equation for the sophisticated dynasty in

$$\begin{aligned} f(0)u'(c_i) &= f(1)(1+r)u'(c_{i+1}) \\ &\quad + \sum_{t=2}^{T-i}\left(f(t) - \frac{f(1)}{f(0)}f(t-1)\right)\prod_{k=1}^{t-1}\left(1 - p\frac{\partial c_{i+k}}{\partial A_{i+k}}\right)p(1+r)^t\frac{\partial c_{i+t}}{\partial A_{i+t}}u'(c_{i+t}) \end{aligned} \tag{5}$$

Numerical Example of Standard Non-Stationary Model: To obtain the closed form solutions and graphical representation of the standard model of non-stationary discounting, I adopt the 'quasi-hyperbolic' discount function, which was first used in Phelps and Pollak (1968) and recently popularized by Laibson (1996)

$$U_i = u(c_i) + \beta\sum_{t=1}^{T-i}\frac{1}{(1+\rho)^t}u(c_{i+t}) \tag{1'}$$

where the dynasty of generations is also assumed to hold a following constant relative risk averse (CRRA) subutility function

$$u(c_\tau) = \frac{c_\tau^{1-\sigma}}{1-\sigma} \tag{7}$$

Backward induction suggests that c_τ is linear in A_τ; that is, the chosen consumption at τ is proportional to, but less than A_τ or $c_\tau = \lambda_\tau A_\tau$ where $\lambda_\tau < 1$ is a constant. As combining the linear function for c_{i+1} with eq. (3) suggests

$$c_{i+1} = \lambda_{i+1}A_{i+1} = \lambda_{i+1}(1+r)(1-pc_i) \tag{A.5}$$

I plug eq. (6), (7), and (A.5) into the Euler equation for either naive or sophisticated dynasty, which is then solved for c_i.

Numerical Example for Naive Dynasty: I first derive the closed form solution for the naive dynasty. As plugging in eq. (6), (7), and (A.5) into eq. (4) leads to

$$c_i^{-\sigma} = \beta\left(\frac{1+r}{1+\rho}\right)(\lambda_{i+1}(1+r)(1-pc_i))^{-\sigma} \tag{A.6}$$

I solve it for c_i

$$c_i = \lambda_i A_i \quad \text{where} \quad \lambda_i = h_i(\lambda_{i+1}) = \frac{\lambda_{i+1}}{((1+r)^{1-\sigma}\beta(1+\rho)^{-1})^{\frac{1}{\sigma}} + p\lambda_{i+1}} \tag{A.7}$$

λ_i in eq. (A.7) defines the consumption rule of generation i in the naive dynasty at i. What remains is to see how the recursive process of the naive dynasty takes place for its consumption planning.

Since the consumption rule of λ_i in eq. (A.7) only displays λ_i as the function of λ_{i+1} but not λ_{i+t} as the function of $\lambda_{i+(t+1)}$, a direct examination of eq. (A.7) is not proper to examine how the recursive convergence takes place in the naive dynasty's planning. It is because in generation i's calculation of consumption plan for future generation $i+t$, a respective dynasty of future generation $i+t$ is falsely assumed to be without the preference reversal where β has disappeared: Accordingly, the consumption rule of future generation $i+t$ as the function of its succeeding generation $i+t+1$, $\lambda_{i+t} = h_{(i+t)+1}(\lambda_{(i+t)+1})$, needs to be recalibrated as the problem of a stationarily discounting dynasty. Thus, the falsely-assumed consumption rule of generation $i+1$ from the standpoint of generation i with the naiveté is from the standard Euler equation of a succeeding dynasty without the preference reversal.

$$f(1)u'(c_{i+1}) = f(2)(1+r)u'(c_{i+2}) \tag{A.8}$$

where the same process of plugging in the terms and solving for c_{i+1} leads to

$$c_{i+1} = \lambda_{i+1}A_{i+1} \quad \text{where} \quad \lambda_{i+1} = h_{i+1}(\lambda_{i+2}) = \frac{\lambda_{i+2}}{((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}} + p\lambda_{i+2}} \tag{A.9}$$

and the same linear function for $c_\tau = h_\tau(\lambda_{\tau+1})$ also holds for $\tau \geq i+1$. We are interested in the convergence $\lambda_\tau \to \lambda_\tau^*$ from the recursive process. As dividing $\lambda_\tau = h_\tau(\lambda_{\tau+1})$ in eq. (A.9) by λ_τ leads to

$$1 = \frac{h(\lambda_{\tau+1})}{\lambda_\tau} = \frac{\lambda_{\tau+1}}{\lambda_\tau}\frac{1}{((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}} + p\lambda_{\tau+1}} \tag{A.10}$$

and the first term in RHS is one when converges, the denominator of the second term in RHS has to be one. Thus,

$$\lambda_\tau \to \lambda_\tau^* = \frac{1 - ((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}}}{p} \tag{8}$$

which is also the consumption rule for the stationarily discounting dynasty. As λ_τ^* in eq. (8) defines the falsely-assumed (i.e., without the preference reversal) consumption rule of subsequent generations $i+t$ from the standpoint of generation i, I may rewrite eq. (A.7)

$$c_{i,i}^* = \lambda_i^* A_i \quad \text{where} \quad \lambda_i^* = \frac{1 - ((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}}}{p\left(1 + ((1+r)^{1-\sigma}(1+\rho)^{-1})^{\frac{1}{\sigma}}(\beta^{\frac{1}{\sigma}} - 1)\right)} > \lambda_\tau^* \text{ with } \beta < 1 \tag{A.7$'$}$$

Using eq. (3), (A.7′) and (8), I can specify the planned sequence of consumption for the naive dynasty.

$$\begin{aligned}\mathbf{c_{i,t}^*} &= \{c_{i,i}^*, c_{i,i+1}^*, c_{i,i+2}^*, c_{i,i+3}^*, \ldots\} = \{\lambda_i^* A_i, \lambda_\tau^* A_{i+1}, \lambda_\tau^* A_{i+2}, \lambda_\tau^* A_{i+3}, \ldots\} \\ &= \left\{\lambda_i^* A_i, \lambda_\tau^* A_i(1+r)(1-p\lambda_i^*), \lambda_\tau^* A_i(1+r)^2(1-p\lambda_i^*)(1-p\lambda_\tau^*), \lambda_\tau^* A_i(1+r)^3(1-p\lambda_i^*)(1-p\lambda_\tau^*)^2, \ldots\right\}\end{aligned} \tag{A.11}$$

which, together with eq. (1), is used to identify the maximum obtainable utility, $V_{i,i+t}^* = \sum_{t=0}^{T-i} u(c_{i,i+t}^*)$. Furthermore, I now can write a naive dynasty's actual sequence of consumption.

$$\begin{aligned}\mathbf{c_{i,i}^*} &= \{c_{i,i}^*, c_{i+1,i+1}^*, c_{i+2,i+2}^*, c_{i+3,i+3}^*, \ldots\} = \{\lambda_i^* A_i, \lambda_i^* A_{i+1}, \lambda_i^* A_{i+2}, \lambda_i^* A_{i+3}, \ldots\} \\ &= \left\{\lambda_i^* A_i, \lambda_i^* A_i(1+r)(1-p\lambda_i^*), \lambda_i^* A_i(1+r)^2(1-p\lambda_i^*)^2, \lambda_i^* A_i(1+r)^3(1-p\lambda_i^*)^3, \ldots\right\}\end{aligned} \tag{A.12}$$

which I use to obtain the actual utility of a naive dynasty, $V_{i,i}^* = \sum_{t=0}^{T-i} u(c_{i+t,i+t}^*)$

Numerical Example for Sophisticated Dynasty: I now pick up the discussion for the sophisticated dynasty by plugging in eq. (6), (7), and (A.5) into a bit augmented sophisticated Euler equation from eq. (5) (to accomodate the 'quasi-hyperbolic' discount function).

$$c_i^{-\sigma} = \left(\frac{1+r}{1+\rho}\right)((\beta-1)p\lambda_{i+1} + 1)(\lambda_{i+1}(1+r)(1-pc_i))^{-\sigma} \tag{A.13}$$

which I solve eq. (A.13) for c_i

$$c_i = \lambda_i A_i \quad \text{where} \quad \lambda_i = h(\lambda_{i+1}) = \frac{\lambda_{i+1}}{((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda_{i+1}+1))^{\frac{1}{\sigma}} + p\lambda_{i+1}} \tag{A.14}$$

λ_i in eq. (A.14) now defines the consumption rule for the sophisticated dynasty. Similar to what was done for the naive dynasty, I now examine how the recursive process of the sophisticated dynasty takes place for its strategy of consistent planning. Note that unlike the naive dynasty, the strategy of consistent planning implies that generation i in the sophisticated dynasty correctly anticipates how its succeeding generation will behave and that actual and planned sequence of consumption coincide.

As dividing $\lambda_i = h(\lambda_{i+1})$ in eq. (A.14) by λ_i leads to

$$1 = \frac{h(\lambda_{i+1})}{\lambda_i} = \frac{\lambda_{i+1}}{\lambda_i} \frac{1}{((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda_{i+1}+1))^{\frac{1}{\sigma}} + p\lambda_{i+1}} \tag{A.15}$$

and the first term in RHS is one when converges, the denominator in the second term in RHS has to be one. Thus,

$$\lambda_i \to \lambda^{**} = \frac{1 - \left((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1)\right)^{\frac{1}{\sigma}}}{p} > \lambda^* \text{ with } \beta \leq 1, \ p\lambda^{**} \leq 1 \tag{10}$$

which can be further solved depending on the value of σ. For the simplicity, I learn about the characteristic of λ^{**} by examining the first and second derivatives of LHS and RHS of eq. (10) rather than directly solving it for λ^{**}. As

$$\begin{aligned}
L(\lambda^{**}) &= \lambda^{**} & R(\lambda^{**}) &= \frac{1 - \left((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1)\right)^{\frac{1}{\sigma}}}{p} \\
L'(\lambda^{**}) &= 1 & R'(\lambda^{**}) &= -\frac{(\beta-1)((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1))^{\frac{1}{\sigma}}}{((1-\beta)p\lambda^{**}+1)\sigma} \\
L''(\lambda^{**}) &= 0 & R''(\lambda^{**}) &= \frac{p(\beta-1)^2((1+r)^{1-\sigma}(1+\rho)^{-1}((\beta-1)p\lambda^{**}+1))^{\frac{1}{\sigma}}(\sigma-1)}{((1-\beta)p\lambda^{**}+1)\sigma^2}
\end{aligned} \tag{A.16}$$

I know that while LHS of eq. (10) is a linear 45 degree line, RHS is always increasing but concave when $\sigma < 1$, linear when $\sigma = 1$, and convex when $\sigma > 1$: λ^{**} is at the intersection of $L(\lambda)$ (a 45 degree line) and $R(\lambda)$ (a concave, linear, or convex curve), of which value depends on the shape of $R(\lambda)$. In any case, I may write a suphisticated dynasty's planned and actual sequence of consumption

$$\begin{aligned}
\mathbf{c}^{**}_{\mathbf{i,t}} &= \{\lambda^{**}A_i, \lambda^{**}A_{i+1}, \lambda^{**}A_{i+2}, \lambda^{**}A_{i+3}, \ldots\} \\
&= \left\{\lambda^{**}A_i, \lambda^{**}A_i(1+r)(1-p\lambda^{**}), \lambda^{**}A_i(1+r)^2(1-p\lambda^{**})^2, \lambda^{**}A_i(1+r)^3(1-p\lambda^{**})^3, \ldots\right\}
\end{aligned} \tag{A.17}$$

Comparison of Naive and Sophisticated Dynasties: With such a characteristic of λ^{**} in mind, I now compare the consumption rules of naive and sophisticated dynasties; λ^*, λ°, and λ^{**}. As shown in figure 1, $\lambda^{**} > \lambda^\circ > \lambda^*$ when $\rho < 1$, $\lambda^{**} = \lambda^\circ > \lambda^*$ when $\rho = 1$, and $\lambda^\circ > \lambda^{**} > \lambda^*$ when $\rho > 1$.

Augmented Model with Strategy of Endogenous Precommitment: We now derive the Euler equations for the augmented model. Though the augmented model may look as a simple extension of the standard model of non-stationary discounting, it is quite complex than it seems.

Derivation of Naive Euler Equation, Eq : Using the standard method, I note that the following Euler equation holds

$$
\begin{aligned}
0 = f(0)u_c(c_i, H_i) &- \sum_{t=1}^{T-i} f(t) \prod_{k=1}^{t-1} \left(1 - p_{i+k}\frac{\partial c_{i+k}}{\partial A_{i+k}} - \frac{\partial h_{i+k}}{\partial A_{i+k}}\right) p_i(1+r)^t \frac{\partial c_{i+t}}{\partial A_{i+t}} u_c(c_{i+t}, H_{i+t}) \\
&- \sum_{t=2}^{T-i} f(t) \sum_{k=2}^{t} \left((1-\delta)^{t-k+1} p_i (1+r)^{k-1} \frac{\partial h_{i+k-1}}{\partial A_{i+k-1}} \prod_{l=2}^{k-1} \left(1 - p_{i+l-1}\frac{\partial c_{i+l-1}}{\partial A_{i+l-1}} - \frac{\partial h_{i+l-1}}{\partial A_{i+l-1}}\right)\right) u_H(c_{i+t}, H_{i+t})
\end{aligned} \tag{A.18}
$$

$$
\begin{aligned}
&= f(0)u_c(c_i, H_i) \\
&\quad + f(1)\left(u_c(c_{i+1}, H_{i+1})\frac{\partial c_{i+1}}{\partial A_{i+1}}(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) + u_H(c_{i+1}, H_{i+1})(1-\delta)\frac{\partial h_i}{\partial c_i}\right) \\
&\quad + f(2)u_c(c_{i+2}, H_{i+2})\frac{\partial c_{i+2}}{\partial A_{i+2}}(1+r)\left(\left(1 - p_{i+1}\frac{\partial c_{i+1}}{\partial A_{i+1}} - \frac{\partial h_{i+1}}{\partial A_{i+1}}\right)(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) - \frac{\partial h_{i+1}}{\partial H_{i+1}}(1-\delta)\frac{\partial h_i}{\partial c_i}\right) \\
&\quad + f(2)u_H(c_{i+2}, H_{i+2})(1-\delta)\left(\frac{\partial h_{i+1}}{\partial A_{i+1}}(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) + \left(\frac{\partial h_{i+1}}{\partial H_{i+1}} + 1\right)(1-\delta)\frac{\partial h_i}{\partial c_i}\right) \\
&\quad + f(3)u_c(c_{i+3}, H_{i+3})\frac{\partial c_{i+3}}{\partial A_{i+3}}(1+r)\left(1 - p_{i+2}\frac{\partial c_{i+2}}{\partial A_{i+2}} - \frac{\partial h_{i+2}}{\partial A_{i+2}}\right)(1+r)\left(\left(1 - p_{i+1}\frac{\partial c_{i+1}}{\partial A_{i+1}} - \frac{\partial h_{i+1}}{\partial A_{i+1}}\right)(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) - \frac{\partial h_{i+1}}{\partial H_{i+1}}(1-\delta)\frac{\partial h_i}{\partial c_i}\right) \\
&\quad + f(3)u_c(c_{i+3}, H_{i+3})\frac{\partial c_{i+3}}{\partial A_{i+3}}(1+r)\frac{\partial h_{i+2}}{\partial H_{i+2}}(1-\delta)\left(\frac{\partial h_{i+1}}{\partial A_{i+1}}(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) + \left(\frac{\partial h_{i+1}}{\partial H_{i+1}} + 1\right)(1-\delta)\frac{\partial h_i}{\partial c_i}\right) \\
&\quad + f(3)u_H(c_{i+3}, H_{i+3})(1-\delta)\frac{\partial h_{i+2}}{\partial A_{i+2}}(1+r)\left(\left(1 - p_{i+1}\frac{\partial c_{i+1}}{\partial A_{i+1}} - \frac{\partial h_{i+1}}{\partial A_{i+1}}\right)(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) - \frac{\partial h_{i+1}}{\partial H_{i+1}}(1-\delta)\frac{\partial h_i}{\partial c_i}\right) \\
&\quad + f(3)u_H(c_{i+3}, H_{i+3})(1-\delta)\left(\frac{\partial h_{i+2}}{\partial H_{i+2}} + 1\right)(1-\delta)\left(\frac{\partial h_{i+1}}{\partial A_{i+1}}(1+r)\left(-p_i - \frac{\partial h_i}{\partial c_i}\right) + \left(\frac{\partial h_{i+1}}{\partial H_{i+1}} + 1\right)(1-\delta)\frac{\partial h_i}{\partial c_i}\right) + \cdots
\end{aligned}
\tag{A.19}
$$

B Proof

Proposition 1. *Given that any dynasty's utility may be restated as the function of λ's, $U^i = U^i(\lambda_i, \lambda_\tau)$,*

$$\frac{dU^i(\lambda_i,\lambda_\tau)}{d\lambda} = \frac{\partial U^i(\lambda_i,\lambda_\tau)}{\partial \lambda_i} + \frac{\partial U^i(\lambda_i,\lambda_\tau)}{\partial \lambda_\tau} < 0 \quad \text{for some } \lambda < \lambda^{**} \tag{12}$$

when evaluated at $\lambda = \lambda_i = \lambda_\tau$.

Proof. As λ^{**} satisfies the Euler equation (A.1), $\partial U_i/\partial \lambda_i = 0$ at λ^{**}. Together with $\partial^2 U/\partial \lambda^2 < 0$, $\partial U_i/\partial \lambda < 0$ for any $\lambda > \lambda^{**}$. Furthermore, restating the dynasty's utility in eq. (1) with the quasi-hyperbolic discount function in eq. (6)

$$U^i(\lambda_i,\lambda_\tau) = u(c_i) + \left(\frac{\beta}{1+\rho}\right)\left(u(c_{i+1}) + \frac{1}{1+\rho}u(c_{i+1}) + \cdots\right) = u(c_i) + \left(\frac{\beta}{1+\rho}\right)\overline{U}(c_\tau) \tag{B.1}$$

$\partial U/\lambda_\tau = \partial \overline{U}/\partial \lambda_\tau = 0$ at $\lambda_\tau^* < \lambda^{**}$ since λ_τ^* satisfies the Euler equation (A.3) or (A.8). Again with $\partial^2 \overline{U}/\partial \lambda^2 < 0$, $\partial U^i/\partial \lambda_\tau < 0$ for any $\lambda > \lambda_\tau^*$. Thus, as $\lambda^{**} > \lambda_\tau^*$, it can be verified that $dU/d\lambda < 0$ for any $\lambda > \lambda^{**}$. Furthermore, $\partial^2 U/\partial \lambda^2$ suggests that for some λ such that $\lambda = \lambda^{**} - d\lambda < \lambda^{**}$, $|\partial U/\partial \lambda_\tau| > |\partial U/\partial \lambda_i|$ and shows that $dU/d\lambda < 0$ for some $\lambda < \lambda^{**}$. □