Econometrica Supplementary Material

SUPPLEMENT TO "SOCIAL DISCOUNTING AND INTERGENERATIONAL PARETO"

(*Econometrica*, Vol. 86, No. 5, September 2018, 1537–1567)

TANGREN FENG
Department of Economics, University of Michigan

SHAOWEI KE
Department of Economics, University of Michigan

THIS SUPPLEMENT consists of four parts: (i) the robustness of findings in the main paper with respect to several main assumptions, (ii) an alternative interpretation of intergenerational Pareto and its implication on quasi-hyperbolic discounting, (iii) a result with forward and backward individual exponential discounting, and (iv) a discussion of the choice domain of the main paper. The supplement uses definitions and notations from the main paper.

S1. DISCUSSION OF THE MAIN ASSUMPTIONS

Our main findings are built upon three sets of assumptions: (i) the assumptions about individual preferences, (ii) intergenerational Pareto and strong non-dictatorship, and (iii) the assumptions about the planner's preference. In the first, we have assumed that a parent's discount function and instantaneous utility function are inherited by his offspring. This assumption may or may not be realistic. It is helpful to understand how our results depend on it. In the second, intergenerational Pareto only has bite when all individuals from the current and future generations agree. It is useful to understand to what extent intergenerational Pareto can be strengthened. In the third, we have required that the planner have an EDU function. This assumption imposes restrictions on how the planner can aggregate individual preferences. We examine what results still hold if we drop this assumption.

We first state a more general version of Lemma 1, which also follows from Harsanyi (1955) and Fishburn (1984) directly.

LEMMA S1—Harsanyi (1955): *Suppose each generation-t individual i's utility function takes the following form*:

$$U_{i,t}(\mathbf{p}) = \sum_{\tau=t}^{T} \delta_{i,t}(\tau - t) u_i(p_\tau, \tau),$$

and the planner's utility function in period t takes the following form:

$$U_t(\mathbf{p}) = \sum_{\tau=t}^{T} \delta_t(\tau - t) u_t(p_\tau, \tau),$$

Tangren Feng: tangren@umich.edu
Shaowei Ke: shaoweik@umich.edu

 https://doi.org/10.3982/ECTA15011

in which $\delta_{i,t}(\cdot)$ *and* $\delta_t(\cdot)$ *are discount functions, and* $u_i(\cdot,\tau)$ *and* $u_t(\cdot,\tau)$ *are normalized instantaneous utility functions. The planner's preference* $(\succsim_t)_{t\in T}$ *is intergenerationally Pareto if and only if, in each period* $t \in T$, *there exists a finite sequence of nonnegative numbers* $(\omega_t(i,s))_{i\in N, s\geq t}$ *such that*

$$U_t = \sum_{i=1}^{N}\sum_{s=t}^{T}\omega_t(i,s)U_{i,s}.$$

We stick to the assumptions about individuals' and the planner's utility functions in the main paper, unless stated otherwise.

S1.1. *Inheriting Discount Functions and Instantaneous Utility Functions From Parents*

One maintained assumption about individual preferences is that each generation-t individual i's discount function δ_i and instantaneous utility function u_i are independent of t. We show in this subsection that this assumption can be removed without changing our main findings. We analyze two cases below. In the first case, for any $i \in N$ and finite t, suppose generation-t individual i's discount function is $\delta_{i,t}$ and instantaneous utility function is u_i; that is, we still assume that individual instantaneous utility functions do not depend on time. Fixing each generation-t individual i's discounting utility function for any $i \in N$ and *any natural number* t, our result may require us to vary the time horizon T. The result below shows that we can establish a positive result that is similar to Theorem 2.

THEOREM S1: *Suppose each generation-*t *individual* i*'s discounting utility function has an instantaneous utility function* u_i *and a discount function* $\delta_{i,t}$ *such that* (A2) *and* (A3) *hold and* $(u_i)_{i\in N}$ *is linearly independent. Let the planner's instantaneous utility function* u *be an arbitrary strict convex combination of* $(u_i)_{i\in N}$. *Then,*

1. *for each* $\delta > \max_{i,t}\delta^*_{i,t}$, *the planner is intergenerationally Pareto and strongly nondictatorial*;
2. *for each* δ *such that for some* i, t, $\delta < \delta^*_{i,t}$, *there exists some* $T^* > 0$ *such that if* $T \geq T^*$, *the planner is not intergenerationally Pareto*.

We will prove this theorem as a special case of Theorem S2 below. Theorem S1 shows that social discounting should still be more patient than the most patient individual's long-run discounting when individual discount functions may change across generations. Since generation-t individual i's discount function is now $\delta_{i,t}$ rather than δ_i, the cutoff for the social discount factor becomes $\max_{i,t}\delta^*_{i,t}$. The second part of the theorem can be understood as follows. Suppose the social discount factor δ is below some generation-t individual i's long-run discount factor. Then, as we increase T, this planner will eventually violate intergenerational Pareto.

One may wonder why we still assume that generation-t individual i's instantaneous utility function does not depend on t. Let us assume that generation-t individual i's instantaneous utility function is $u_{i,t}$. The example below shows that this assumption will lead to a trivial negative result that has nothing to do with discounting.

EXAMPLE S1: Suppose $N = 1$. Let generation-1 individual's instantaneous utility function be u_1, which is linearly independent of generation-2 individual's instantaneous utility function u_2. Since the planner has an EDU function, her instantaneous utility function should never change. In the first period, the planner's instantaneous utility function for period-1 consumption can only be u_1, because only the generation-1 individual cares about period-1 consumption. The planner's instantaneous utility function for period-2 consumption, however, must depend on both u_1 and u_2 due to strong non-dictatorship, which means that the planner's instantaneous utility function for period-2 consumption must differ from u_1. Therefore, it is impossible to require that the planner be intergenerationally Pareto and strongly non-dictatorial.

As can be seen in the example above, it seems inevitable that the planner's instantaneous utility function should depend on time; that is, the planner's instantaneous utility function for period-τ consumption should depend on τ. Indeed, one way to restore the positive result is to allow the planner's instantaneous utility function to be $u(\cdot, \tau)$.

However, there is another way to restore the positive result, which is the second case we want to analyze. For any $i \in N$ and finite t, suppose generation-t individual i's discount function is $\delta_{i,t}$ and instantaneous utility function for period-τ consumption is $u_i(\cdot, \tau)$; that is, if the planner's instantaneous utility function for period-τ consumption has to depend on τ, let us make the same assumption for individuals. Note that individual instantaneous utility functions now depend on time, but in a manner different from Example S1. The planner's discount function is again exponential.

These assumptions are particularly suitable in our setting. Recall that each individual only lives for one period, and he cares about future consumption based on altruism. Imagine that $u_i(\cdot, \tau)$ is generation-τ individual i's actual consumption utility—that is, the utility that generation-τ individual i derives by consuming rather than from altruism. Now, generation-t individual i's utility function is

$$U_{i,t}(\mathbf{p}) = \sum_{\tau=t}^{T} \delta_i(\tau - t) u_i(p_\tau, \tau),$$

which means that when the generation-t individual i altruistically cares about generation-τ individual i's consumption, he values the consumption in exactly the same way that generation-τ individual i will value it for himself.

THEOREM S2: *Suppose each generation-t individual i's discounting utility function has instantaneous utility functions $(u_i(\cdot, \tau))_{\tau \geq t}$ and a discount function $\delta_{i,t}$ such that* (A2) *and* (A3) *hold and $(u_i(\cdot, \tau))_{i \in N}$ is linearly independent for each $\tau \in T$. Suppose, for some positive $(\lambda_i)_{i \in N}$ such that $\sum_{i \in N} \lambda_i = 1$, the planner's $u(\cdot, \tau) = \sum_{i \in N} \lambda_i u_i(\cdot, \tau)$ for any $\tau \in T$. Then,*

1. *for each $\delta > \max_{i,t} \delta^*_{i,t}$, the planner is intergenerationally Pareto and strongly non-dictatorial*;
2. *for each δ such that for some i, t, $\delta < \delta^*_{i,t}$, there exists some $T^* > 0$ such that if $T \geq T^*$, the planner is not intergenerationally Pareto*.

PROOF: *Part I* This part is similar to Part I of Theorem S1. First, we prove a lemma for one-individual aggregation.

LEMMA S2: *Assume that $N = \{i\}$. Suppose each generation-t individual i's discounting utility function has instantaneous utility functions $u_i(\cdot, \tau)$ and a discount function $\delta_{i,t}$ such that* (A2) *and* (A3) *hold. Let the planner's instantaneous utility function be $u_i(\cdot, \tau)$ for any $\tau \in T$. For any $\delta > \max_t \delta^*_{i,t}$, the planner is intergenerationally Pareto and strongly non-dictatorial*.

PROOF: We want to show that for any $\delta > \max_{i,t} \delta^*_{i,t}$, there exists a finite sequence of positive numbers $(\omega_t(i,s))_{t\in T, s\geq t}$ such that

$$U_t(\mathbf{p}) = \sum_{\tau=t}^{T} \delta^{\tau-t} u(p_\tau, \tau) = \sum_{s=t}^{T} \omega_t(i,s) U_{i,s}(\mathbf{p})$$

for each $t \in T$. Given any $\delta > \max_t \delta^*_{i,t}$, we can construct $(\omega_t(i,s))_{t\in T, s\geq t}$ according to the following formula recursively:

$$\omega_t(i,s) = \begin{cases} 1 & \text{if } s = t, \\ \sum_{\tau=t}^{s-1} \big[\delta \cdot \delta_{i,\tau}(s-1-\tau) - \delta_{i,\tau}(s-\tau)\big]\omega_t(i,\tau) & \text{if } s > t. \end{cases} \tag{S1}$$

Note that by assuming $\delta > \max_t \delta^*_{i,t}$, $\omega_t(i,s) > 0$ for any $s \geq t$ and $t \in T$. Then,

$$\begin{aligned} U_t(\mathbf{p}) &= \sum_{s=t}^{T} \omega_t(i,s) U_{i,s}(\mathbf{p}) = \sum_{s=t}^{T} \omega_t(i,s) \sum_{\tau=s}^{T} \delta_{i,s}(\tau - s) u(p_\tau, \tau) \\ &= \sum_{\tau=t}^{T} \left(\sum_{s=t}^{\tau} \delta_{i,s}(\tau - s)\omega_t(i,s) \right) u(p_\tau, \tau). \end{aligned}$$

We want to prove that $U_t(p) = \sum_{\tau=t}^{T} \delta^{\tau-t} u(p_\tau, \tau)$. Clearly, for $\tau = t$, $\sum_{s=t}^{\tau} \delta_{i,s}(\tau - s)\omega_t(i,s) = \omega_t(i,t) = 1 = \delta^0$. Suppose, for some $\tau \geq t$, we have proven that $\sum_{s=t}^{\tau} \delta_{i,s}(\tau - s)\omega_t(i,s) = \delta^{\tau-t}$. We want to prove that for $\tau + 1$,

$$\sum_{s=t}^{\tau+1} \delta_{i,s}(\tau + 1 - s)\omega_t(i,s) = \delta^{\tau-t+1}. \tag{S2}$$

To prove (S2), we only need to notice that according to (S1),

$$\begin{aligned} &\sum_{s=t}^{\tau+1} \delta_{i,s}(\tau + 1 - s)\omega_t(i,s) \\ &\quad = \omega_t(i, \tau+1) + \sum_{s=t}^{\tau} \delta_{i,s}(\tau + 1 - s)\omega_t(i,s) \\ &\quad = \sum_{s=t}^{\tau} \big[\delta\delta_{i,s}(\tau - s) - \delta_{i,s}(\tau + 1 - s)\big]\omega_t(i,s) + \sum_{s=t}^{\tau} \delta_{i,s}(\tau + 1 - s)\omega_t(i,s) \\ &\quad = \delta \cdot \sum_{s=t}^{\tau} \delta_{i,s}(\tau - s)\omega_t(i,s) = \delta^{\tau-t+1}. \end{aligned}$$

By induction, we know that $\sum_{s=t}^{\tau} \delta_{i,s}(\tau - s)\omega_t(i,s) = \delta^{\tau-t}$ for all $\tau \geq t$, and hence $U_t(\mathbf{p}) = \sum_{\tau=t}^{T} \delta^{\tau-t} u_i(p_\tau, \tau)$. *Q.E.D.*

Next, for any social discount factor $\delta > \max_i \max_t \delta^*_{i,t}$, we can find $(\omega_{i,t}(s))_{t\in T, i\in N, s\geq t}$ such that

$$\sum_{s=t}^{T} \omega_{i,t}(s) U_{i,s}(\mathbf{p}) = \sum_{\tau=t}^{T} \delta^{\tau-t} u_i(p_\tau, \tau)$$

for each $i \in N$. Then, we know that

$$\begin{aligned} U_t(\mathbf{p}) &= \sum_{i=1}^{N} \sum_{s=t}^{T} \lambda_i \omega_t(i, s) U_{i,s}(\mathbf{p}) = \sum_{i=1}^{N} \sum_{\tau=t}^{T} \delta^{\tau-t} \lambda_i u_i(p_\tau) \\ &= \sum_{\tau=t}^{T} \delta^{\tau-t} \sum_{i=1}^{N} \lambda_i u_i(p_\tau, \tau) = \sum_{\tau=t}^{T} \delta^{\tau-t} u(p_\tau, \tau). \end{aligned}$$

Part II We prove it by contradiction. Suppose there exists an intergenerationally Pareto planner with social discount factor $\delta < \delta^*_{i,t}$ for some $i = i^*$ and $t = t^*$. By intergenerational Pareto, for each $t \in T$, there exists a finite sequence of nonnegative numbers $(\omega_t(i, s))_{i\in N, s\geq t}$ such that the following equality holds:

$$\delta^{\tau-t} u(p_\tau, \tau) = \sum_{i=1}^{N} \sum_{s=t}^{\tau} \omega_t(i, s) \delta_{i,s}(\tau - s) u_i(p_\tau, \tau) \tag{S3}$$

for any $\tau \geq t$. When $\tau = t$, the above equation reduces to

$$u(p_\tau, \tau) = \sum_{i=1}^{N} \omega_\tau(i, \tau) u_i(p_\tau, \tau) \tag{S4}$$

for any $\tau \in T$.

Since $u(\cdot, \tau) = \sum_{i\in N} \lambda_i u_i(\cdot, \tau)$ for any $\tau \in T$ and $(u_i(\cdot, \tau))_{i\in N}$ is linearly independent, $\omega_t(i, t) = \lambda_i > 0$, for any i and t. Multiply $\delta^{\tau-t}$ to both sides of equation (S4) and combine it with equation (S3). We obtain

$$\sum_{i=1}^{N} \omega_\tau(i, \tau) \delta^{\tau-t} u_i(p_\tau, \tau) = \sum_{i=1}^{N} \sum_{s=t}^{\tau} \omega_t(i, s) \delta_{i,s}(\tau - s) u_i(p_\tau, \tau).$$

Since $(u_i(\cdot, \tau))_{i=1}^{N}$ is linearly independent, the above equation is equivalent to

$$\omega_\tau(i, \tau) \delta^{\tau-t} u_i(p_\tau, \tau) = \sum_{s=t}^{\tau} \omega_t(i, s) \delta_{i,s}(\tau - s) u_i(p_\tau, \tau)$$

for any $i \in N$, $t \in T$, and $\tau \geq t$.

Let $i = i^*$ and $t = t^*$, and rearrange the above equations. We have

$$\delta^{\tau-t^*} = \frac{\sum_{s=t^*}^{\tau} \omega_{t^*}(i^*, s) \delta_{i^*,s}(\tau - s)}{\omega_\tau(i^*, \tau)}$$

$$= \frac{\omega_{t^*}(i^*, t^*)\delta_{i^*,t^*}(\tau - t^*) + \sum_{s=t^*+1}^{\tau} \omega_{t^*}(i^*, s)\delta_{i^*,s}(\tau - s)}{\omega_\tau(i^*, \tau)} \tag{S5}$$
$$\geq \frac{\lambda_i^* \cdot \delta_{i^*,t^*}(\tau - t^*)}{\lambda_i^*} = \delta_{i^*,t^*}(\tau - t^*)$$

for any $\tau > t^*$. However, we also know that $\delta < \lim_{\tau\to\infty} \sqrt[\tau]{\delta_{i^*,t^*}(\tau)}$, there exists T^* such that for any $\tau \geq T^*$, $\delta < \sqrt[\tau]{\delta_{i^*,t^*}(\tau)}$, which contradicts (S5). *Q.E.D.*

When we assume $u(\cdot, \tau) = \sum_{i\in N} \lambda_i u_i(\cdot, \tau)$, we have assumed that λ_i's do not depend on τ. In the social choice literature, some economists have argued that with normalized individual utility functions, equal utilitarian weights should be used (see Karni (1998), Dhillon and Mertens (1999), and Segal (2000)). To some extent, this is consistent with our assumption that λ_i's do not change over time, although in our case, λ_i's may not be $1/N$. In general, one may want λ_i's to depend on time. In that case, the fact that the planner's discount function is exponential will impose restrictions on how λ_i's may change over time.

S1.2. *Strengthening Intergenerational Pareto*

The premise of intergenerational Pareto requires that the current generation and future generations reach a consensus. A natural way to strengthen intergenerational Pareto may be to require that the planner prefer one consumption sequence over another if more than a certain fraction of current- and future-generation individuals agree.[1] However, in this case, how the planner aggregates individual preferences may differ somewhat from utilitarian aggregation.

Therefore, we strengthen intergenerational Pareto in the following simple way without deviating from standard utilitarianism. Let $I \subset N \times T$ be an arbitrary subset of individuals across generations. Let us weaken the premise of intergenerational Pareto by requiring that the planner prefer a consumption sequence $\mathbf{p}$ to $\mathbf{q}$ whenever individuals in I agree. Intergenerational Pareto and the strongly non-dictatorial property are adapted as follows.

DEFINITION S1: The planner's preference $(\succsim_t)_{t\in T}$ is I-intergenerationally Pareto if for any consumption sequences $\mathbf{p}, \mathbf{q} \in \Delta(X)^T$, in each period $t \in T$, $\mathbf{p} \succsim_{i,s} \mathbf{q}$ for all $(i, s) \in I$ with $s \geq t$ implies $\mathbf{p} \succsim_t \mathbf{q}$, and $\mathbf{p} \succ_{i,s} \mathbf{q}$ for all $(i, s) \in I$ with $s \geq t$ implies $\mathbf{p} \succ_t \mathbf{q}$.

DEFINITION S2: We say that the planner is I-strongly non-dictatorial if for each $t \in T$,

$$U_t(\mathbf{p}) = f_t\big(U_{1,t}(\mathbf{p}), \dots, U_{1,T}(\mathbf{p}), U_{2,t}(\mathbf{p}), \dots, U_{2,T}(\mathbf{p}), \dots, U_{N,T}(\mathbf{p})\big)$$

for some function f_t that is (strictly) increasing in $U_{i,s}$ for any $(i, s) \in I$.

It is straightforward to show that under I-intergenerational Pareto, the planner's utility function can be written as a weighted sum of the utility functions of individuals in I. Below, we show that under some assumption about I, positive results can still be established after strengthening intergenerational Pareto.

[1] This strengthening can certainly be applied to current-generation Pareto as well.

The following example shows why we need an additional assumption. Suppose $N = 2$ and individual instantaneous utility functions, u_1 and u_2, are linearly independent. Assume that $I = \{(2, 1), (1, 2)\}$; that is, the planner will give generation-1 individual 1 and generation-2 individual 2 zero weights. Then, the somewhat trivial negative result, as in Example S1, appears again. To see this, note that in period 1, the planner's instantaneous utility function for period-1 consumption must be equal to u_2, because only generation-1 individuals care about period-1 consumption and generation-1 individual 1 has been ignored. We have assumed that the planner has an EDU function, in which her instantaneous utility function never changes. Now, first, in period 1, the planner's instantaneous utility function for period-2 consumption is a strict convex combination of u_1 and u_2, which must differ from u_1; second, in period 2, following the same logic, the planner's instantaneous utility function for period-2 consumption must be equal to u_1, which is again different from u_1. Therefore, it is hopeless to derive any positive result.

The theorem below imposes a simple assumption to avoid the example above, which turns out to be strong enough for us to establish a positive result. For each $t \in T$, let $I_t := \{i \in N : (i, t) \in I\}$ be the set of generation-t individuals who may not be ignored by the planner, and let $\mathcal{I} := \bigcup_{t\in T} I_t$.

THEOREM S3: *Suppose $I \subset N \times T$, and each generation-t individual i's discounting utility function has an instantaneous utility function $u_i \in \{u^\theta\}_{\theta=1}^{\Theta}$ for some linearly independent Θ-tuple of instantaneous utility functions $(u^\theta)_{\theta=1}^{\Theta}$, and has a discount function δ_i such that* (A2) *and* (A3) *hold. Assume that* $\mathrm{co}(\{u_i\}_{i\in I_t}) = \mathrm{co}(\{u^\theta\}_{\theta=1}^{\Theta})$ *for any $t \in T$. Let the planner's instantaneous utility function u be a strict convex combination of $(u_i)_{i\in I_t}$. Then,*

1. *for each $\delta > \max_{\mathcal{I}} \delta_i^*$, the planner is I-intergenerationally Pareto and I-strongly non-dictatorial*;
2. *for each $\delta < \min_{\mathcal{I}} \delta_i^*$, there exists some $T^* > 0$ such that if $T \geq T^*$, the planner is not I-intergenerationally Pareto*.

PROOF: *Part I* With an abuse of notation, let $\Theta := \{1, \dots, \Theta\}$. For each $\theta \in \Theta$, let $I^\theta := \{i \in N : u_i = u^\theta\}$, which is the set of i's whose instantaneous utility function is u^θ. For each $\theta \in \Theta$ and $t \in T$, let $I_t^\theta := \{i \in I^\theta : (i, t) \in I\}$ be the set of generation-t individuals who may not be ignored by the planner and whose instantaneous utility function is u^θ. Let $\mathcal{I}^\theta := \bigcup_{t\in T} I_t^\theta$.

We prove this part in four steps. First, we aggregate individuals in each I_t^θ into a new "family" θ. Each generation-t family θ has instantaneous utility function $u^\theta(\cdot)$ and the following discount function:

$$\delta_t^\theta(\tau) = \frac{1}{|I_t^\theta|} \sum_{i\in I_t^\theta} \delta_i(\tau);$$

that is, if a generation-t individual i may not be ignored by the planner, his discount function $\delta_i(\cdot)$ enters family θ's generation-t discount function $\delta_t^\theta(\cdot)$ with a weight equal to that of other generation-t individual(s) in I_t^θ. Note that generation-t families' discount functions may change as t changes.

Next, we prove a lemma on one-family aggregation that is similar to Lemma S2.

LEMMA S3: *Assume $\Theta = \{\theta\}$. Suppose each generation-t family θ's discounting utility function has an instantaneous utility function $u^\theta(\cdot)$ and a discount function $\delta_t^\theta(\cdot)$. Let the planner's instantaneous utility function be $u^\theta(\cdot)$. For any $\delta > \max_{i\in\mathcal{I}^\theta} \delta_i^*$, the planner is intergenerationally Pareto and strongly non-dictatorial.*

PROOF: We want to show that for any $\delta > \max_{i\in\mathcal{I}^\theta} \delta_i^*$, there exists a finite sequence of positive numbers $(\omega_t^\theta(s))_{t\in T, s\geq t}$ such that

$$U_t(\mathbf{p}) = \sum_{\tau=t}^{T} \delta^{\tau-t} u^\theta(p_\tau) = \sum_{s=t}^{T} \omega_t^\theta(s) U_s^\theta(\mathbf{p})$$

for each $t \in T$. Given any $\delta > \max_{i\in\mathcal{I}^\theta} \delta_i^*$, we can construct $(\omega_t^\theta(s))_{t\in T, s\geq t}$ according to the following formula recursively:

$$\omega_t^\theta(s) = \begin{cases} 1 & \text{if } s = t, \\ \sum_{\tau=t}^{s-1} \big[\delta \cdot \delta_\tau^\theta(s-1-\tau) - \delta_\tau^\theta(s-\tau)\big]\omega_t^\theta(\tau) & \text{if } s > t. \end{cases} \tag{S6}$$

Note that if $\delta > \max_t \max_\tau \frac{\delta_t^\theta(\tau+1)}{\delta_t^\theta(\tau)}$, then $\omega_t^\theta(s) > 0$ for any $s \geq t$ and $t \in T$. We also know that

$$\begin{aligned} \frac{\delta_t^\theta(\tau+1)}{\delta_t^\theta(\tau)} &= \frac{\sum_{i\in I_t^\theta} \delta_i(\tau+1)}{\sum_{i\in I_t^\theta} \delta_i(\tau)} = \frac{\sum_{i\in I_t^\theta} \delta_i(\tau)\frac{\delta_i(\tau+1)}{\delta_i(\tau)}}{\sum_{i\in I_t^\theta} \delta_i(\tau)} \leq \frac{\sum_{i\in I_t^\theta} \delta_i(\tau) \max_{i\in I_t^\theta} \frac{\delta_i(\tau+1)}{\delta_i(\tau)}}{\sum_{i\in I_t^\theta} \delta_i(\tau)} \\ &\leq \max_{i\in I_t^\theta} \frac{\delta_i(\tau+1)}{\delta_i(\tau)} \leq \max_{i\in I_t^\theta} \delta_i^* \leq \max_{i\in\mathcal{I}^\theta} \delta_i^*. \end{aligned}$$

Therefore, $\max_t \max_\tau \frac{\delta_t^\theta(\tau+1)}{\delta_t^\theta(\tau)} \leq \max_{i\in\mathcal{I}^\theta} \delta_i^*$. Hence, by assuming $\delta > \max_{i\in\mathcal{I}^\theta} \delta_i^*$, $\omega_t^\theta(s) > 0$ for any $s \geq t$ and $t \in T$. The rest of the proof is the same as in Lemma S2. *Q.E.D.*

Thus, for any social discount factor $\delta > \max_{\theta\in\Theta} \max_{i\in\mathcal{I}^\theta} \delta_i^*$, we can find $(\omega_t^\theta(s))_{t\in T, \theta\in\Theta, s\geq t}$ such that

$$\sum_{s=t}^{T} \omega_t^\theta(s) U_s^\theta(\mathbf{p}) = \sum_{\tau=t}^{T} \delta^{\tau-t} u^\theta(p_\tau)$$

for each $\theta \in \Theta$. Consider any positive numbers $(\lambda^\theta)_{\theta=1}^{\Theta}$ such that $\sum_{\theta=1}^{\Theta} \lambda^\theta = 1$. Together with the weights $(\omega_t^\theta(s))_{t\in T, \theta\in\Theta, s\geq t}$ we have found above, the planner's utility function becomes

$$\begin{aligned} U_t(\mathbf{p}) &= \sum_{\theta\in\Theta} \sum_{s=t}^{T} \lambda^\theta \omega_t^\theta(s) U_s^\theta(\mathbf{p}) = \sum_{\theta\in\Theta} \sum_{\tau=t}^{T} \lambda^\theta \delta^{\tau-t} u^\theta(p_\tau) \\ &= \sum_{\tau=t}^{T} \delta^{\tau-t} \sum_{\theta\in\Theta} \lambda^\theta u^\theta(p_\tau) = \sum_{\tau=t}^{T} \delta^{\tau-t} u(p_\tau), \end{aligned} \tag{S7}$$

in which $u(p_\tau) = \sum_{\theta\in\Theta} \lambda^\theta u^\theta(p_\tau)$ can be any strict convex combination of $(u^\theta)_{\theta\in\Theta}$.

Last, we back out the weights $(\omega_t(i, s))_{t\in T, i\in N, s\geq t}$ and show that the planner has an EDU function, is I-intergenerationally Pareto, and is I-strongly non-dictatorial under these weights. We construct $(\omega_t(i,s))_{t\in T, i\in N, s\geq t}$ according to the following formula:

$$\omega_t(i,s) = \begin{cases} 0 & \text{if } (i,s) \notin I, \\ \lambda^\theta \dfrac{1}{|I_s^\theta|} \omega_t^\theta(s) > 0 & \text{if } (i,s) \in I. \end{cases}$$

Then,

$$\begin{aligned} \sum_{s=t}^{T}\sum_{i=1}^{N} \omega_t(i,s) U_{i,s}(\mathbf{p}) &= \sum_{s=t}^{T}\sum_{i=1}^{N} \omega_t(i,s) \sum_{\tau=s}^{T} \delta_i(\tau - s) u_i(p_\tau) \\ &= \sum_{s=t}^{T}\sum_{\theta\in\Theta}\sum_{i\in I_s^\theta} \lambda^\theta \frac{1}{|I_s^\theta|} \omega_t^\theta(s) \sum_{\tau=s}^{T} \delta_i(\tau - s) u_i(p_\tau) \\ &= \sum_{s=t}^{T}\sum_{\theta\in\Theta} \lambda^\theta \omega_t^\theta(s) \sum_{\tau=s}^{T}\sum_{i\in I_s^\theta} \frac{1}{|I_s^\theta|} \delta_i(\tau - s) u_i(p_\tau) \\ &= \sum_{s=t}^{T}\sum_{\theta\in\Theta} \lambda^\theta \omega_t^\theta(s) \sum_{\tau=s}^{T} \delta_s^\theta(\tau - s) u^\theta(p_\tau) \\ &= \sum_{s=t}^{T}\sum_{\theta\in\Theta} \lambda^\theta \omega_t^\theta(s) U_s^\theta(\mathbf{p}) = U_t(\mathbf{p}) = \sum_{\tau=t}^{T} \delta^{\tau - t} u(p_\tau). \end{aligned}$$

The first equality follows from the definition of $U_{i,s}$. The second equality follows the construction of $(\omega_t(i,s))_{t\in T, i\in N, s\geq t}$. The fourth equality follows the construction of $\delta_s^\theta(\cdot)$. The fifth equality follows from the definition of U_s^θ. The last two equalities follow equation (S7).

Part II We prove it by contradiction. Suppose there exists an I-intergenerationally Pareto planner with social discount factor $\delta < \min_{i\in\mathcal{I}} \delta_i^*$. By I-intergenerationally Pareto, there exists a finite sequence of nonnegative weights $(\omega_t(i,s))_{t\in T, i\in N, s\geq t}$ such that the following equality holds:

$$\delta^{\tau - t} u(p_\tau) = \sum_{s=t}^{\tau}\sum_{i\in I_s} \omega_t(i,s) \delta_i(\tau - s) u_i(p_\tau) \tag{S8}$$

for each $t \in T$ and $\tau \geq t$. Combining equation (S8) with the normalization assumption,

$$\delta^{\tau - t} = \sum_{s=t}^{\tau}\sum_{i\in I_s} \omega_t(i,s) \delta_i(\tau - s) \geq \sum_{i\in I_t} \omega_t(i,s) \delta_i(\tau - s) \tag{S9}$$

for each $t \in T$ and $\tau \geq t$.

We assume that $\arg\min_{i\in\mathcal{I}} \delta^* = \{i^*\}$. The proof can be easily extended to the case with multiple minima. The following two claims must hold:

1. $i^* \in \mathcal{I}$; that is, there exists $t^* \in T$ such that $i^* \in I_{t^*}$.
2. There exists T_1 such that for any $\tau \geq \max\{T_1, t^*\}$, $\delta_{i^*}(\tau - t^*) \leq \delta_i(\tau - t^*)$ for any $i \in \mathcal{I}$.

Consider the period-t^* planner. Let $t = t^*$ in equation (S9), and suppose $\tau \geq \max\{T_1, t^*\}$. We have

$$\delta^{\tau-t^*} = \sum_{s=t^*}^{\tau} \sum_{i \in I_s} \omega_{t^*}(i, s)\delta_i(\tau - s) \geq \sum_{i \in I_t^*} \omega_{t^*}(i, t^*)\delta_i(\tau - t^*)$$

$$\geq \sum_{i \in I_t^*} \omega_{t^*}(i, t^*)\delta_{i^*}(\tau - t^*) \geq \delta_{i^*}(\tau - t^*).$$

However, we know that $\delta < \delta^*_{i^*}$. Then, there exists T_2 such that for any $\tau \geq T_2$, $\delta < \sqrt[\tau]{\delta^*_{i^*}}$. Therefore, if $T \geq \max\{T_1, T_2, t^*\}$, there must be a contradiction. *Q.E.D.*

Note that we assume $\text{co}(\{u_i\}_{i \in I_t}) = \text{co}(\{u^\theta\}_{\theta=1}^{\Theta})$ for any $t \in T$. This is because we want $\text{co}(\{u_i\}_{i \in I_t})$ to remain constant across t to rule out the example we discuss before the theorem, and we want to assume that there is no redundant type.

The theorem seems different from our previous results that have only one cutoff for the social discount factor, but in fact it has a one-cutoff version that is similar to our previous positive results. However, the expression of the cutoff will become rather complicated.[2] The current version is easier to understand, and clearly shows that if the social discount factor is higher than the highest long-run discount factor among individuals who are not ignored in some generation, then we know that the planner is intergenerationally Pareto and strongly non-dictatorial. Again, this is not the only way to establish positive results. If the planner's instantaneous utility function is allowed to vary in a general way by taking the form of $u_t(\cdot, \tau)$, then the additional assumption we need can be weaker.

S1.3. *Utilitarianism and Long-Run Social Discounting*

The main question of this paper is, if a planner has an EDU function, under what conditions is she intergenerationally Pareto/utilitarian and strongly non-dictatorial? The fact that an intergenerationally Pareto/utilitarian planner has an EDU function certainly imposes restrictions on how the planner may aggregate individual preferences. On the one hand, economists often assume that a planner has an EDU function, and there are many reasons to believe that this is normatively appealing. Therefore, understanding the answer to our main question is important.

On the other hand, there are other ways to examine the planner's aggregation problem. For example, sometimes economists may believe that the planner's utility function should be equal to the simple average of individuals' discounting utility functions. However, because it is unlikely that the planner's discount function is exponential in this case, a choice about what to assume for the planner must be made.

A natural question arises: If we now want to allow the planner to aggregate individual preferences in a flexible way—in other words, we only require that the planner be intergenerationally Pareto/utilitarian and strongly non-dictatorial and do not require that her utility function be an EDU function—what insight from our main findings remains true? The following result shows that under this different requirement, the planner's "discount

[2]The cutoff for the social discount factor in the one-cutoff version should take the maximum across types and periods, and then for each type in each period, take the minimal individual long-run discount factor across all individuals who have the desired type and are not ignored in that period.

factor" should still be higher than the most patient individual's long-run discount factor. The result assumes that $T = +\infty$. Some notations and definitions for the case with $T = +\infty$ can be found in Section A.8.

THEOREM S4: *Suppose $T = +\infty$, each generation-t individual i's discounting utility function has an instantaneous utility function u_i and a discount function δ_i such that* (A2) *and* (A3) *hold, and the planner's utility function in any period $t \in T$ is $U_t = \sum_{\tau=t}^{\infty} \delta_t(\tau - t)u_t(p_\tau, \tau)$ for some discount function δ_t and* (*normalized*) *instantaneous utility function $(u_t(\cdot, \tau))_{\tau \geq t}$ such that $\delta_t^* = \lim_{\tau\to\infty} \frac{\delta_t(\tau+1)}{\delta_t(\tau)} = \lim_{\tau\to\infty} \sqrt[\tau]{\delta_t(\tau)}$ exists. If the planner is intergenerationally utilitarian and strongly non-dictatorial, $\delta_t^* \geq \max_i \delta_i^*$.*

PROOF: Since $U_t = \sum_{i=1}^{N} \sum_{s=t}^{\infty} \omega_t(i, s)U_{i,s}$, we know that

$$\delta_t(\tau - t)u_t(p_\tau, \tau) = \sum_{i=1}^{N} \sum_{s=t}^{\tau} \omega_t(i, s)\delta_i(\tau - s)u_i(p_\tau) \tag{S10}$$

for any $t \in T$ and $\tau \geq t$. Let $p_\tau = x^*$ in equation (S10). We have

$$\delta_t(\tau - t) = \sum_{i=1}^{N} \sum_{s=t}^{\tau} \omega_t(i, s)\delta_i(\tau - s) \geq \sum \omega_t(i, t)\delta_i(\tau - t) \geq \omega_t\big(i^*, t\big)\delta_{i^*}(\tau - t), \tag{S11}$$

in which $i^* := \arg\max_i \delta_i^*$. Let τ in (S11) go to infinity. We have $\delta_t^* \geq \max_i \delta_i^*$.

Thus, if the planner is intergenerationally Pareto and strongly non-dictatorial, her long-run discount factor should again be higher than the most patient individual's long-run discount factor. *Q.E.D.*

S2. AN ALTERNATIVE INTERPRETATION OF INTERGENERATIONAL PARETO AND A RESULT WITH QUASI-HYPERBOLIC DISCOUNTING

We say that the generation-t individual i has a *quasi-hyperbolic discounting utility* (QHDU) function if his discount function satisfies

$$\delta_i(\tau) = \begin{cases} 1 & \text{if } \tau = 0, \\ \beta_i \delta_i^\tau & \text{if } \tau \in \{1, \dots, T - 1\}, \end{cases}$$

for some $\beta_i \in (0, 1]$ and $\delta_i > 0$. In the literature of time inconsistency, economists sometimes ignore the β_i parameter and use an EDU function with a discount factor δ_i as the welfare criterion of individual i who has a QHDU function. The intuition is that because β_i is the cause of time inconsistency, β_i should not enter the welfare criterion. We show how our analysis provides some foundation for this practice.[3]

Consider Corollary 1. If we interpret the generation-$(t + 1)$ individual i in our model as the future self of the generation-t individual i, Corollary 1 provides some foundation for the use of this welfare criterion. Assume that individual i is the only individual ($N = 1$) and has a quasi-hyperbolic discount function. According to Corollary 1, we immediately know that any EDU function with a discount factor that is (strictly) greater than δ_i is a

[3]Recent papers by Drugeon and Wigniolle (2017) and Galperti and Strulovici (2017) introduce results similar to the one we present below.

welfare criterion that is consistent with intergenerational Pareto; that is, if individual i in every period t agrees that one consumption sequence is better than another, the welfare criterion says that the utility of the former is greater than the latter.

The following result is stronger than Corollary 1. It shows that δ_i is indeed the smallest discount factor such that the corresponding EDU function is consistent with intergenerational Pareto.

PROPOSITION S1: *Suppose each generation-t individual i has a QHDU function with an instantaneous utility function u, $\beta_i \in (0,1)$, and $\delta_i \in (0,1)$. Then,*

1. *for each $\delta \geq \min_i \delta_i$, the planner is intergenerationally Pareto and strongly non-dictatorial*;

2. *for each $\delta < \min_i \delta_i$, there exists some $T^* > 0$ such that if $T \geq T^*$, the planner is not intergenerationally Pareto*.

PROOF: The second part follows from Theorem 1. We only prove the first part.

LEMMA S4: *Assume that $N = \{i\}$. Suppose individual i has a QHDU function with parameters $\beta_i \in (0,1)$, $\delta_i \in (0,1)$, and u. Then, there exists a cutoff $\delta(T)$ for each T such that the planner is intergenerationally Pareto and strongly non-dictatorial if and only if $\delta > \delta(T)$. In addition, $\delta(T)$ is (strictly) increasing with a limit δ_i.*

PROOF: The planner is intergenerationally Pareto and strongly non-dictatorial if and only if there exists a finite sequence of positive weights $(\omega_t(i,s))_{t\in T, s\geq t}$ such that the following equation holds:

$$\omega_t(i,\tau)u(p_\tau) + \sum_{s=t}^{\tau-1} \omega_{i,t}(s)\beta_i \delta_i^{\tau-s} u(p_\tau) = \delta^{\tau-t} u(p_\tau) \tag{S12}$$

for any $t \in T$ and $\tau \geq t$. We can solve $(\omega_t(i,s))_{t\in T, s\geq t}$ from (S12) as follows:

$$\omega_t(i,t+m) = \begin{cases} 1 & \text{if } m = 0, \\ \delta^m - \dfrac{\beta_i}{1-\beta_i} \displaystyle\sum_{h=1}^{m} (1-\beta_i)^h \delta_i^h \delta^{m-h} & \text{if } 1 \leq m \leq T-t. \end{cases} \tag{S13}$$

Note that $\omega_t(i,t) = 1 > 0$, and the planner is intergenerationally Pareto and strongly non-dictatorial if and only if $(\omega_t(i,t+m))_{t\in T, 1\leq m\leq T-t}$ is positive.

We can rewrite the second equation of (S13) as $\omega_t(i,t+m) = F_m(\delta|\beta_i,\delta_i)$, in which F is a degree-m polynomial of a single indeterminate δ with parameters β_i, δ_i. Define

$$S(\beta_i,\delta_i,T) := \{\delta \in \mathbb{R}_+ : F_m(\delta|\beta_i,\delta_i) > 0 \text{ for any } 1 \leq m \leq T-1\}.$$

Therefore, the planner's preference is intergenerationally Pareto and strongly non-dictatorial if and only if $\delta \in S(\beta_i,\delta_i,T)$.

We want to show that $S(\beta_i,\delta_i,T)$ is an interval that (strictly) shrinks to $[\delta_i,+\infty)$ as T increases. First, we prove that there exists a unique root/cutoff $x_m \in (0,\delta_i]$ for $F_m(\delta|\beta_i,\delta_i)$ such that $F_m(x_m|\beta_i,\delta_i) = 0$, $F_m(\delta|\beta_i,\delta_i) < 0$ for $\delta < x_m$, and $F_m(\delta|\beta_i,\delta_i) > 0$ for $\delta > x_m$. We know that $F_m(0|\beta_i,\delta_i) = -(1-\beta_i)^{m-1}\delta_i^m < 0$, $F_m(\delta_i|\beta_i,\delta_i) = (1-\beta_i)^m \delta_i^m > 0$, and F_m is continuous. Therefore, the existence of x_m is guaranteed by Bolzano's theorem.

Also note that the function $G_m(\delta|\beta_i, \delta_i) := \delta^{-m} F_m(\delta|\beta_i, \delta_i)$ has the same root as $F_m(\delta|\beta_i, \delta_i)$, and $G_m(\delta|\beta_i, \delta_i)$ is (strictly) increasing in δ because

$$\frac{dG_m(\delta)}{d\delta} = \frac{\beta_i}{1-\beta_i} \sum_{k=1}^{m} k \frac{(1-\beta_i \delta_i)^k}{\delta^{k+1}} > 0.$$

By Rolle's theorem, there cannot be more than one root. Hence, the uniqueness is proved.

Second, we prove that the cutoff sequence $(x_m)_m$ is (strictly) increasing and converges to δ_i. Noting that

$$G_{m+1}(\delta|\beta_i, \delta_i) - G_m(\delta|\beta_i, \delta_i) = -\frac{\beta_i}{1-\beta_i} \left[\frac{(1-\beta_i)\delta_i}{\delta} \right]^{m+1} < 0,$$

we have $G_{m+1}(x_m|\beta_i, \delta_i) - G_m(x_m|\beta_i, \delta_i) < 0$. By the definition of $(x_m)_m$, $G_m(x_m|\beta_i, \delta_i) = G_{m+1}(x_{m+1}|\beta_i, \delta_i) = 0$. Therefore, $G_{m+1}(x_m|\beta_i, \delta_i) < G_m(x_m|\beta_i, \delta_i) = G_{m+1}(x_{m+1}|\beta_i, \delta_i)$. We also know that $G_m(\delta|\beta_i, \delta_i)$ is (strictly) increasing. Hence, $x_{m+1} > x_m$. Now that $(x_m)_m$ is bounded and (strictly) increasing, the convergence follows from the monotone convergence theorem.

The only remaining part is to prove that the limit of the cutoff sequence is δ_i. Suppose $\lim_{m\to\infty} x_m = x$. Then, $x_m < x$ for all $m > 1$. Since $G_m(\delta|\beta_i, \delta_i)$ is (strictly) increasing, we have

$$\begin{aligned} & G_m(x_m|\beta_i, \delta_i) < G_m(x|\beta_i, \delta_i) \\ \Leftrightarrow \quad & 0 < 1 - \frac{\beta_i}{1-\beta_i} \sum_{h=1}^{m} (1-\beta_i)^h \delta_i^h x^{-h} \\ \Leftrightarrow \quad & \frac{\beta_i}{1-\beta_i} \sum_{h=1}^{m} (1-\beta_i)^h \delta_i^h x^{-h} < 1 \\ \Leftrightarrow \quad & \sum_{h=1}^{m} \left[\frac{(1-\beta_i)\delta_i}{x} \right]^h < \frac{1-\beta_i}{\beta_i} \end{aligned} \tag{S14}$$

for any $m > 1$.

Given that $\frac{(1-\beta_i)\delta_i}{x} > 0$, we must have $\frac{(1-\beta_i)\delta_i}{x} < 1$; otherwise, $\sum_{h=1}^{m} [\frac{(1-\beta_i)\delta_i}{x}]^h$ diverges as m increases. Now, let m in (S14) go to infinity. We have

$$\begin{aligned} & \sum_{h=1}^{+\infty} \left[\frac{(1-\beta_i)\delta_i}{x} \right]^h \leq \frac{1-\beta_i}{\beta_i} \\ \Leftrightarrow \quad & \frac{(1-\beta_i)\delta_i}{x} \frac{1}{1 - \frac{(1-\beta_i)\delta_i}{x}} \leq \frac{1-\beta_i}{\beta_i} \\ \Leftrightarrow \quad & \delta_i \leq x. \end{aligned} \tag{S15}$$

In addition, since $x_m < \delta_i$ for all m, we have $x \leq \delta_i$. Therefore, $x = \delta_i$. *Q.E.D.*

Lemma S4 states that for any finite T, in each period t, the planner can aggregate each individual i from the tth generation to the Tth generation so that the aggregated utility function is an EDU function with a discount factor that is slightly below δ_i. Then, we can apply the if part of Proposition 2 for N exponential discounting individuals, and obtain a social discount factor $\delta \geq \min_i \delta_i$. *Q.E.D.*

When $T = +\infty$, we can assume in Proposition 4 that individuals have QHDU functions and obtain a similar result.

S3. THE CASE WITH BACKWARD DISCOUNTING

The result we introduce below shows that if individuals exponentially forward and backward discount consumption, our main results continue to hold. Before proceeding, it should be noted that backward discounting has no revealed-preference foundation. Whenever we observe an individual choosing, the past is sunk; there are no choices (yet) that allow the individual to alter the past. Therefore, we do not know how individuals think about the past from actual choice data.

However, economists have considered the possibility that individuals backward discount (see Strotz (1955), Caplin and Leahy (2004), and Ray, Vellodi, and Wang (2017)). Below, we analyze our aggregation problem with exponential discounting individuals who backward discount. Instead of assuming that $U_{i,t}(\mathbf{p})$ does not depend on past consumption, we assume that the generation-t individual i discounts both past and future by the same discounting factor δ_i.

DEFINITION S3: The generation-t individual i has an exponential forward and backward discounting utility function if his utility function has the following form:

$$U_{i,t}(\mathbf{p}) = \sum_{\tau=1}^{T} \delta_i^{|\tau-t|} u_i(p_\tau), \tag{S16}$$

in which the discount factor $\delta_i \in (0, 1)$, and u_i is the individual i's instantaneous utility function.

Note that the negative result, obviously, would continue to hold if we had assumed that each generation-t individual i's utility function was

$$U_{i,t}(\mathbf{p}) = \sum_{\tau=1}^{T} \delta_i^{\tau-t} u_i(p_\tau).$$

In that case, the individual i's offspring has exactly the same preference as the indivdual i. This is problematic, however, because the generation-2 individual i will value period-1 consumption even more than his own period-2 consumption.

The result below demonstrates that the assumption that the planner has an EDU function and intergenerational Pareto are compatible when individuals exponentially forward and backward discount consumption. The typical negative result in the literature only considers the planner's aggregation problem in period 1. The following result also focuses on the period-1 aggregation problem to highlight the difference.

PROPOSITION S2: *Suppose each generation-t individual i has an exponential forward and backward discounting utility function with discount factor δ_i and instantaneous utility function u_i such that $\bar{\delta} := \max_i \delta_i < 1$. Let the planner's instantaneous utility function u be an arbitrary strict convex combination of $(u_i)_{i\in N}$. Then, for each $\delta \in (\bar{\delta}, \bar{\delta}^{-1})$, the planner in period 1 is intergenerationally Pareto and strongly non-dictatorial.*

PROOF: To prove the proposition, we consider the one-individual case first.

LEMMA S5: *Assume that $N = \{i\}$. Suppose each generation-t individual i has an exponential forward and backward discounting utility function with discount factor $\delta_i \in (0, 1)$ and instantaneous utility function u. Then, for each $\delta \in (\delta_i, {\delta_i}^{-1})$, the planner in period 1 is intergenerationally Pareto and strongly non-dictatorial.*

PROOF: We want to show that for any $\delta \in (\delta_i, {\delta_i}^{-1})$, there exists a finite sequence of positive weights $\vec{\omega} = (\omega(i, 1), \omega(i, 2), \dots, \omega(i, T))$ such that the following equation holds:

$$U_1(\mathbf{p}) = \sum_{\tau=1}^{T} \delta^{\tau-1} u(p_\tau) = \sum_{s=1}^{T} \omega(i, s) U_{i,s}(\mathbf{p}). \tag{S17}$$

Plugging in $U_1(\mathbf{p})$ and $U_{i,s}(\mathbf{p})$, equation (S17) becomes

$$\sum_{\tau=1}^{T} \delta^{\tau-1} u(p_\tau) = \sum_{s=1}^{T} \omega(i, s) \sum_{\tau=1}^{T} \delta_i^{|\tau-s|} u(p_\tau) = \sum_{\tau=1}^{T} \sum_{s=1}^{T} \omega(i, s) \delta_i^{|s-\tau|} u(p_\tau); \tag{S18}$$

that is, for each $\tau \geq 1$,

$$\delta^{\tau-1} = \sum_{s=1}^{T} \omega(i, s) \delta_i^{|s-\tau|}. \tag{S19}$$

Next, we can rewrite equation (S19) as follows:

$$A \cdot \vec{\omega} = \vec{\delta}, \tag{S20}$$

in which $\vec{\delta} = (1, \delta, \delta^2, \dots, \delta^{T-1})$ and

$$A = \begin{pmatrix} 1 & \delta_i & \delta_i^2 & \dots & \delta_i^{T-1} \\ \delta_i & 1 & \delta_i & \dots & \delta_i^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \delta_i^{T-1} & \delta_i^{T-2} & \delta_i^{T-3} & \dots & 1 \end{pmatrix}.$$

Note that A is invertible. In particular,

$$A^{-1} = \frac{1}{1 - {\delta_i}^2} \begin{pmatrix} 1 & -\delta_i & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ -\delta_i & 1+\delta_i^2 & -\delta_i & 0 & & & & \vdots \\ 0 & -\delta_i & 1+\delta_i^2 & -\delta_i & \ddots & & & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & & \ddots & -\delta_i & 1+\delta_i^2 & -\delta_i & 0 \\ \vdots & & & & 0 & -\delta_i & 1+\delta_i^2 & -\delta_i \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & -\delta_i & 1 \end{pmatrix}.$$

We have $\vec{\omega} = A^{-1} \cdot \vec{\delta}$. If we can show that $\vec{\omega} \gg 0$, the lemma is proved. Showing that $\vec{\omega} \gg 0$ is equivalent to showing that $\omega(i,1) = 1 - \delta_i\delta > 0$, $\omega(i,s) = \delta^{s-2}[-\delta_i + (1+\delta_i^2)\delta - \delta_i\delta^2] > 0$ for $2 \leq s \leq T-1$, and $\omega(i,T) = -\delta_i\delta^{T-2} + \delta^{T-1} > 0$, which can be verified because $\delta \in (\delta_i, \delta_i^{-1})$. *Q.E.D.*

Lemma S5 shows that we can aggregate each individual i from the tth generation to the Tth generation into an EDU function with any discount factor δ within $(\delta_i, \delta_i^{-1})$. Now we can prove Proposition S2. For any social discount factor $\delta \in (\bar{\delta}, \bar{\delta}^{-1})$, we can find $(\omega(i,s))_{i\in N, s\geq 1}$ such that

$$\sum_{s=1}^{T} \omega(i,s) U_{i,s}(\mathbf{p}) = \sum_{\tau=1}^{T} \delta^{\tau-1} u_i(p_\tau)$$

for each $i \in N$. Consider any positive numbers $(\lambda_i)_{i\in N}$ such that $\sum_{i\in N} \lambda_i = 1$. Together with the weights $(\omega(i,s))_{i\in N, s\geq 1}$ we have found above, the planner's utility function becomes

$$\begin{aligned} U_1(\mathbf{p}) &= \sum_{i=1}^{N} \sum_{s=1}^{T} \lambda_i \omega_1(i,s) U_{i,s}(\mathbf{p}) = \sum_{i=1}^{N} \sum_{\tau=1}^{T} \delta^{\tau-1} \lambda_i u_i(p_\tau) \\ &= \sum_{\tau=t}^{T} \delta^{\tau-1} \sum_{i=1}^{N} \lambda_i u_i(p_\tau) = \sum_{\tau=t}^{T} \delta^{\tau-1} u(p_\tau), \end{aligned}$$

in which $u(p_\tau) = \sum_{i\in N} \lambda_i u_i(p_\tau)$ is an arbitrary strict convex combination of $(u_i)_{i\in N}$. *Q.E.D.*

S4. RISK RESOLUTION

The main model's choice domain is $\Delta(X)^T$; that is, in each period, there is a lottery/probability measure over X. In many dynamic economic models with uncertainty, uncertainty resolves over time. Below, we discuss what may change if we let uncertainty resolve over time, maintaining our assumptions about individuals' and the planner's utility functions.

For simplicity, assume that $T = 2$ and $N = 1$. In period 2, the choice object is again a *lottery* over X. Sometimes, it will be called a period-2 lottery. To distinguish between choice objects in the main paper and in this section, here we call X *outcomes* and period-1 choice objects *dynamic lotteries*. A dynamic lottery is a lottery over $X \times \Delta(X)$. For example, with probability 1/2, a dynamic lottery $\tilde{p}_1$ yields a period-1 outcome $x \in X$ and a period-2 lottery $q_2 \in \Delta(X)$; with probability 1/2, $\tilde{p}_1$ yields a period-1 outcome x' and a period-2 lottery $r_2 \in \Delta(X)$.

Now, the set of dynamic lotteries is $\Delta(X \times \Delta(X))$, rather than $\Delta(X)^2$.[4] However, $\Delta(X)^2$ can be viewed as a subset of $\Delta(X \times \Delta(X))$ that consists of all dynamic lotteries whose period-2 lotteries are independent of (the realization of) period-1 outcomes.

The following simple example shows in what sense, in period 1, the planner's aggregation problem under $\Delta(X \times \Delta(X))$ is the same as under $\Delta(X)^2$. Continue our example of $\tilde{p}_1$, q_2, r_2 above. Let q_2 be a lottery that yields $y, y' \in X$ with equal probability. Let r_2 be

[4]For any metric space Y, let $\Delta(Y)$ denote the set of Borel probability measures on Y. We endow $\Delta(X)$ with the Prohorov metric and $X \times \Delta(X)$ with product topology.

a degenerate lottery that yields $z \in X$. First, consider the generation-1 individual. A natural way to extend our period-1 individual utility function on $\Delta(X)^2$ to the new domain $\Delta(X \times \Delta(X))$ is as follows:

$$\begin{aligned} V_1(\tilde{p}_1) &= \frac{1}{2}\big(v(x,1) + \delta v(q_2,2)\big) + \frac{1}{2}\big(v(x',1) + \delta v(r_2,2)\big) \\ &= \frac{1}{2}\left(v(x,1) + \delta\left(\frac{1}{2}v(y,2) + \frac{1}{2}v(y',2)\right)\right) + \frac{1}{2}\big(v(x',1) + \delta v(z,2)\big), \end{aligned}$$

in which δ is the individual discount factor and $v(\cdot,\tau)$ is the period-τ individual instantaneous utility function. Note that the above equation can be rewritten as

$$V_1(\tilde{p}_1) = \left(\frac{1}{2}v(x,1) + \frac{1}{2}v(x',1)\right) + \delta\left(\frac{1}{4}v(y,2) + \frac{1}{4}v(y',2) + \frac{1}{2}v(z,2)\right);$$

that is, the utility of $\tilde{p}_1 \in \Delta(X \times \Delta(X))$ is equal to the following dynamic lottery: In period 1, the individual consumes a 50–50 lottery between x and x', and in period 2, he consumes a lottery that yields y with probability $1/4$, y' with probability $1/4$, and z with probability $1/2$.

It is not difficult to see the logic behind this observation. In general, given any $\tilde{p}_1 \in \Delta(X \times \Delta(X))$, we compute the marginal probability distribution of period-1 outcomes and call it $p_1 \in \Delta(X)$, and compute the marginal probability distribution of period-2 outcomes and call it $p_2 \in \Delta(X)$. Then, (p_1, p_2) is a dynamic lottery whose period-2 lotteries are independent of period-1 outcomes. It must be the case that $V_1(\tilde{p}_1) = V_1((p_1, p_2))$, because V_1 is a time-additively separable expected utility function.

Second, consider the generation-2 individual. Because we are examining the period-1 planner's problem, which means the dynamic lottery's risk has not resolved, how does the planner evaluate the second generation's utility of $\tilde{p}_1$? Arguably,

$$V_2(\tilde{p}_1) = \frac{1}{2}\left(\frac{1}{2}v(y,2) + \frac{1}{2}v(y',2)\right) + \frac{1}{2}v(z,2) \tag{S21}$$

seems to be a reasonable evaluation—with probability $1/2$, the second generation's utility will be $\frac{1}{2}v(y,2) + \frac{1}{2}v(y',2)$, and with probability $1/2$, the second generation's utility will be $v(z,2)$. Now, again,

$$V_2(\tilde{p}_1) = V_2\big((p_1,p_2)\big) = \frac{1}{4}v(y,2) + \frac{1}{4}v(y',2) + \frac{1}{2}v(z,2).$$

Therefore, $\tilde{p}_1$ and (p_1, p_2) are equivalent for the planner in period 1. The planner's period-1 aggregation problem under $\Delta(X \times \Delta(X))$ is the same as under $\Delta(X)^2$—there is a bijection between time-addtively separable expected utility functions defined on the domain with and without correlation. As long as the period-1 planner uses the same utilitarian weights to aggregate individual utility functions, the planner's preference will be the same in both cases.

Move on to period 2 and continue our previous example of $\tilde{p}_1$ and (p_1, p_2). With either $\Delta(X \times \Delta(X))$ or $\Delta(X)^2$, the second generation's utility function is defined on $\Delta(X)$, because individuals do not care about past consumption. Therefore, there is again a (trivial) bijection between generation-2 individual utility functions defined on the domain with and without correlation. The planner's period-2 preference will be identical in both cases as long as she uses the same utilitarian weights for individuals.

The analysis above can be extended to the case with more periods and more individuals. In this sense, focusing on consumption sequences $\Delta(X)^T$ without modeling how uncertainty resolves over time is without loss of generality.

However, it should be noted that with $\tilde{p}_1$, the period-2 lottery is either q_2 or r_2. With (p_1, p_2), no matter what the first generation consumes, the period-2 lottery is p_2. Therefore, there will be some ex post difference between $\tilde{p}_1$ and (p_1, p_2) about which generation consumes what. However, this difference should not affect the period-2 planner's aggregation problem.

Another issue to be noted is that in either the case with correlation or the case without, we only study what the planner's objective should be if she aggregates individuals' preferences. This exercise does not require us to consider, for example, feasibility constraints. If the planner's problem is to maximize some objective under certain constraints, correlation may be important in the feasibility constraints. For example, if there is a technological advancement in the first period, we can anticipate a larger feasible set of consumption in the future. This requires correlation in the constraints.

REFERENCES

CAPLIN, A., AND J. LEAHY (2004): "The Social Discount Rate," *Journal of Political Economy*, 112 (6), 1257–1268. [14]

DHILLON, A., AND J.-F. MERTENS (1999): "Relative Utilitarianism," *Econometrica*, 67 (3), 471–498. [6]

DRUGEON, J.-P., AND B. WIGNIOLLE (2017): "On Time-Consistent Collective Choice With Heterogeneous Quasi-Hyperbolic Discounting," Working Paper, Paris School of Economics. [11]

FISHBURN, P. (1984): "On Harsanyi's Utilitarian Cardinal Welfare Theorem," *Theory and Decision*, 17 (1), 21–28. [1]

GALPERTI, S., AND B. STRULOVICI (2017): "A Theory of Intergenerational Altruism," *Econometrica*, 85 (4), 1175–1218. [11]

HARSANYI, J. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63 (4), 309–321. [1]

KARNI, E. (1998): "Impartiality: Definition and Representation," *Econometrica*, 66 (6), 1405–1415. [6]

RAY, D., N. VELLODI, AND R. WANG (2017): "Backward Discounting," Working Paper, New York University. [14]

SEGAL, U. (2000): "Let's Agree That All Dictatorships Are Equally Bad," *Journal of Political Economy*, 108 (3), 569–589. [6]

STROTZ, R. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23 (3), 165–180. [14]

Co-editor Itzhak Gilboa handled this manuscript.

Manuscript received 13 January, 2017; final version accepted 17 May, 2018; available online 25 May, 2018.