Intergenerational Discounting and Inequality

Frikk Nesje[1] *and Paolo G. Piacquadio*[2]

April 7, 2025

Abstract

We study all theories of justice that disentangle ethical views on intergenerational discounting and intergenerational inequality. Each "modular" social welfare function is uniquely identified by a time-discounting function—capturing attitudes toward time—and an aggregator function—capturing attitudes toward inequality. Our axiomatic characterization identifies the properties and establishes the limits of disentangling discounting and inequality. Our theories include as special cases the most common welfare criteria adopted in the literature, while uncovering unexplored (and yet tractable) families of alternative criteria.

Keywords: Intergenerational justice, discounting, inequality.
JEL classification: D30, D60, H43, I31, Q50.

[1]Department of Economics, University of Copenhagen, and CESifo Research Network (frikk.nesje@econ.ku.dk).

[2]School of Economics and Political Sciences, University of St. Gallen, Department of Economics, University of Oslo, and CESifo Research Network (paolo.piacquadio@unisg.ch).

The authors are grateful to R. Aaberge, F. Annan, D. Anthoff, G. Asheim, M. Auffhammer, A. Bommier, Y. Chechel, F. Cherbonnier, J. Druedahl, M. Drupp, M. Eden, J. Emmerling, T. Fally, E. Fenichel, M. Fleurbaey, M. Fosgerau, C. Gollier, B. Groom, B. Harstad, A. Hill, L. Karp, J. P. Kelleher, L. Leijonklo, W. Marz, A. Millner, M. Olsen, R. Piacquadio, R. van der Ploeg, M. Rege, F. Salanié, E. Starkov, T. Sterner, P. Strack, P. B. Sørensen, P. N. Sørensen, V. Thivierge, R. Tol, C. Traeger, N. Treich, K. Vamsæter, U. Wagner, P. Wakker, D. Zilberman, and F. Zilibotti for comments, in addition to audiences in Bergen, Berkeley, Copenhagen, Gothenburg, Hamburg, Leipzig, Leuven, Munich, New Orleans, Oslo, Paris, Rome, Rotterdam, Santa Barbara, Southampton, Stanford, Stavanger, St. Gallen, Toulouse, and online. This project has received funding from the Economic Policy Research Network and the European Research Council under the European Union's Horizon 2020 research and innovation program ERC Starting Grant VALURED (Grant No. 804104). F. Nesje has received funding from the Independent Research Fund Denmark (Grant No. 3126-00011B) and Professor Wilhelm Keilhau's Memorial Fund for research stays at the University of California, Berkeley and the Toulouse School of Economics.

1 Introduction

Policies enacted today—such as climate change mitigation, public debt reduction, investment in technological innovation, and pension reform—typically involve trade-offs, where some generations gain while others lose. Any assessment of these policies must address two fundamental ethical choices. First, how to value different generations because they live at different times—intergenerational discounting. Second, how to value different generations when some are better off than others—intergenerational inequality.

In this paper, we formalize, axiomatically characterize, and study *all* theories of intergenerational justice that disentangle intergenerational discounting and inequality.[1] Such disentanglement provides a better understanding of existing criteria and their limitations, allows us to incorporate (tractable) discounting and inequality attitudes generally disregarded in the literature, and paves the way for transparent policy evaluations that reflect ethical views held by individuals.[2]

The key challenge is to reconcile discounting and inequality aversion. When a social welfare function satisfies Pareto efficiency, discounting of future consumption is difficult to avoid (Koopmans, 1960; Diamond, 1965; Asheim, 2010; Pivato & Fleurbaey, 2024). Moreover, as a byproduct of placing less weight on future generations, social preferences cannot prioritize equality of consumption across generations, making inequality attitudes hard to define.

Our approach addresses this challenge, regardless of whether social preferences satisfy Pareto efficiency and/or discount future generations. The central idea and

[1]The goal is neither to prescribe a specific theory of justice, nor to reject commonly used criteria—such as exponentially-discounted utilitarianism. However, a narrow focus on some welfare criteria can be problematic, especially when dictated by a lack of (tractable) alternatives. This concern becomes even more pressing if the discounting and inequality attitudes embodied in such criteria fail to capture compelling and/or widely held ethical views. Moreover, a misspecification of the welfare criterion could lead to large welfare losses (see Gollier (2024)).

[2]Other disentanglement results have proven extremely valuable in economics, such as the disentanglement between equity and efficiency in atemporal settings (Atkinson, 1970) or the disentanglement between time and risk attitudes (Epstein & Zin, 1989; Bommier, Kochov, & Le Grand, 2017; Dillenberger, Gottlieb, & Ortoleva, in press).

main novelty are to show that discounting attitudes can be conveniently expressed by extending or contracting calendar time, after which inequality attitudes follow naturally.[3]

More precisely, the *time-discounting function* converts consumption streams in calendar time to their "equivalent-time" representation. In equivalent time, more valuable intervals of calendar time last longer; more precisely, the duration of consumption is adjusted to the point where permuting constant consumption (in equivalent time) is a matter of indifference.[4] Then, inequality attitudes and their trade-offs with efficiency can be treated similarly to any standard inequality-averse criterion typically adopted for atemporal problems, here named *aggregator function* (see, among others, d'Aspremont and Gevers (2002)).

We refer to our theories of justice as "modular," since each theory is fully described by the module on time discounting and the module on aggregation. As we show, practically all existing theories of justice are modular—including the workhorse model of exponentially-discounted utilitarianism. More importantly, our results reveal a wealth of new alternative criteria (including many that are extremely tractable) by freely varying the discounting and inequality modules.

We characterize modular theories of intergenerational justice by imposing the novel axiom of *nested time split*. This axiom requires that any interval of calendar time can be split into "equally valuable" subintervals (possibly of unequal length). Intuitively, two subintervals are equally valuable if permuting their constant consumption is a matter of social indifference. The axiom also requires, as with standard separability axioms, that the partition be independent of the specific consumption stream under consideration. To illustrate, exponential discounting

[3]Attema, Bleichrodt, Gao, Huang, and Wakker (2016) build on a similar insight to experimentally investigate individuals' discounting attitudes without estimating their utility functions in a framework with additively separable preferences. Our results do not rely on time separability. In contrast, we show that time separability is particularly demanding with respect to inequality attitudes, since the priority attributed to a richer and a poorer generation cannot depend on how these fare compared to the other generations.

[4]To clarify, instead of comparing consumption over calendar time by their present value (thus adjusting the level of consumption), we compare them by their duration in equivalent time (thus expanding or compressing calendar time). The fundamental advantage is that the level of consumption remains unchanged and, thus, do not alter our understanding of inequality.

at a rate of 1% values the next 38 years as equally valuable as the following 62 years (splitting the next 100 years); furthermore, the next 69 years are seen as equally valuable as all later years (splitting the entire infinite time horizon).

Our theories of intergenerational justice offer a number of valuable insights for policy applications. First, by adopting novel definitions of social impatience and inequality aversion, we can compare discounting and inequality attitudes across social welfare functions under more general conditions than before (Böhm-Bawerk, 1889; Fisher, 1930; Koopmans, 1960; Atkinson, 1970). Second, by filtering out discounting attitudes, we can compare consumption streams by their level of inequality and decompose social welfare into their efficiency and equality components. Such decomposition extends the pioneering work by Atkinson (1970) and Blackorby and Donaldson (1978) to temporal settings.

Additionally, the flexible discounting and inequality attitudes provide a much more general (and yet easy to interpret) version of the "Ramsey rule" (Ramsey, 1928), identifying the value of a dollar's worth of consumption at different points in time. These formulations clarify how to use our criteria to find sufficient statistics for cost-benefit analysis.

Discounting and inequality attitudes remain understudied in the literature. For exponentially-discounted utilitarianism there has been much debate about the discount rate and the concavity of the utility function (Stern, 2007; Nordhaus, 2007; Dasgupta, 2008; Weitzman, 2007; Dietz & Stern, 2008; Gollier, 2013). However, no parameters seem to capture individuals' views on these attitudes (Drupp, Freeman, Groom, & Nesje, 2018; Millner & Heal, 2023). In addition, some confounding may occur because the concavity of the utility function captures inequality aversion, but could also be interpreted as the intertemporal elasticity of substitution (Eden, 2023), as well as individual or social attitudes toward risk.[5] For these reasons, parts of the literature rely on restrictive formulations of inequality aversion and employ sensitivity analysis for the discount factor (see

[5]Extensions to intergenerational risk and uncertainty exist. The seminal contributions are Epstein and Zin (1989, 1991), which disentangle attitudes to risk and intergenerational inequality with an isoelastic formulation. Other contributions include Fleurbaey and Zuber (2013, 2015), Traeger (2014), and Piacquadio (2020). Here, we restrict our analysis to risk-free settings.

Del Campo, Anthoff, and Kornek (2024); Nesje, Drupp, Freeman, and Groom (2024)).

However, the larger focus on discounting (rather than inequality) is mainly due to the difficulties of combining Pareto efficiency with an anonymous treatment of generations in infinite-time settings, referred to above (Diamond, 1965; Asheim, 2010). Here anonymity means that changing the order of consumption of generations is a matter of social indifference. As a partial solution, the literature on intergenerational equity has often framed the research question in terms of ranking utility streams rather than consumption streams. Then, the utility function, which is usually left unspecified, can capture some notion of inequality aversion.[6] Our paper clarifies that discounting and efficiency are compatible with a variety of inequality attitudes. The key is to introduce anonymity and inequality aversion in "equivalent time," rather than in "calendar time".

Our flexible perspective on inequality attitudes is borrowed from the *leaky bucket experiment* popularized by Okun (1975). He proposed to measure how much loss could be justified in transferring one dollar of income from a richer to a poorer individual. We formalize a similar measure here, called "local inequality aversion", and show how it feeds into the generalized Ramsey rule.[7] As a crucial difference, instead of applying this measure to individuals, we use intervals of equivalent time. Our results shed light on how to capture a wide range of inequality attitudes in theories of intergenerational justice.[8]

Our results respond to the call from both economists and philosophers for more discussion of inequality attitudes for intergenerational justice (Nesje, Drupp,

[6]To achieve compatibility with efficiency, other contributions suggest weakening anonymity, continuity, or completeness (Chichilnisky, 1996; Basu & Mitra, 2003; Asheim, Mitra, & Tungodden, 2012). For a recent survey of the criteria that satisfy efficiency and some form of anonymity in calendar time, see Pivato and Fleurbaey (2024). An alternative is to avoid discounting by weakening efficiency, as proposed by Zuber and Asheim (2012).

[7]Formally, we extend Chew and Mao (1995)'s equivalence between mean-preserving spread and Schur concavity to intergenerational inequality.

[8]A growing number of contributions aim to identify inequality attitudes for income distributions (Clark & d'Ambrosio, 2015; Hvidberg, Kreiner, & Stantcheva, 2023; Fehr, Epper, & Senn, 2023). Subgroup separability is also typically rejected in contexts with both inequality and risk (see Fleurbaey (2010); Fleurbaey and Zuber (2013)).

Freeman, & Groom, 2023). However, our results also hold for individual preferences. Inequality attitudes could be reinterpreted as individuals' willingness to smooth consumption over time. Then, our theorems provide novel insights for how to model individuals' behavior (see Frederick, Loewenstein, and O'Donoghue (2002); Prelec (2004); Quah and Strulovici (2013); Attema, Bleichrodt, Rohde, and Wakker (2010); Cohen, Ericson, Laibson, and White (2020)).

The structure of the paper is as follows. Section 2 presents the framework and introduces modular theories of intergenerational justice. Sections 3 and 4 discuss the discounting and aggregation modules, respectively. Section 5 presents the rich family of theories of intergenerational justice obtained by combining different choices of time-discounting and aggregator functions. These theories are characterized axiomatically and generalized in Section 6. Concluding remarks are collected in Section 7. Proofs are given in the Appendix.

2 Welfare, calendar time, and equivalent time

2.1 Consumption streams

Calendar time is continuous and denoted by $\mathsf{t} \in \mathsf{T} \equiv [0, \infty)$. An intergenerational consumption stream is a bounded function $\mathsf{c} : \mathsf{T} \to \mathbb{R}_+$ that assigns an instantaneous consumption $\mathsf{c}(\mathsf{t})$ to each time $\mathsf{t} \in \mathsf{T}$.[9] Let C denote the set of all consumption streams.[10] A **social welfare function** $W : \mathsf{C} \to \mathbb{R}$ represents a

[9]The restriction of bounded consumption streams avoids some technical difficulties. In particular, the results are less sensitive to the continuity requirement and do not exclude the most common criteria.

[10]We follow convention and interpret consumption at a point in calendar time as either the lifetime equivalent consumption of the generation born at that point in time or the equivalent consumption enjoyed by society at that point in time. While we consider calendar time to be continuous, discrete time can be accommodated simply by letting the consumption stream be a step function. For the purposes of this paper, we need not take a position on how consumption is aggregated to a real number at each instant of calendar time (see Li, Rohde, and Wakker (2024)), including how this might be derived from an underlying model. Because consumption is a scalar, the framework can be generalized along several dimensions, including population growth, intratemporal inequalities, multidimensional settings, and risk.

complete and transitive ranking of intergenerational consumption streams. We assume that the social welfare function is sup-norm continuous.

Consumption streams can be equivalently expressed on the unit interval $I \equiv [0, 1)$ as follows. Let $\phi : T \to I$ be a continuous, bounded, and strictly increasing transformation function that maps calendar time to the unit interval. Then, for each consumption stream $c \in C$, the unit-interval representation for ϕ is $y : I \to \mathbb{R}_+$ such that $y(\phi(t)) = c(t)$ for each $t \in T$; when this is the case, we will say that y is ϕ—equivalent to c, denoted $y \cong_\phi c$. The set of all unit-interval representations of consumption streams is denoted Y.

We say that consumption streams are **represented in equivalent time** if changes in the order of constant consumption on the unit interval leave social welfare W unchanged. The transformation function ω that achieves this property is called the **time-discounting function** and is a special case of the unit-interval representations (see Figure 1).[11] The idea is that ω contracts/extends consumption experienced in calendar time exactly to the point where each interval of these time-discounted consumption streams is given the same priority and the assessment is, thus, *anonymous*. Formally, let $F_y(x)$ denote the share of the unit interval for which the consumption stream defines a consumption level smaller than or equal to x. Then, for any pair of consumption streams $c, c' \in C$ with $y \cong_\omega c$ and $y' \cong_\omega c'$ such that $F_y(x) = F_{y'}(x)$ for each $x \in \mathbb{R}_+$, $W(c) = W(c')$.

The **aggregator function** measures social welfare in equivalent time. Given a social welfare function W and a time-discounting function ω, the aggregator function $V : Y \to \mathbb{R}$ is defined by setting $V(y) = W(c)$ for each $y \in Y$ with $y \cong_\omega c$, as shown in Figure 1.

A social welfare function W is **modular** if consumption streams can be represented in equivalent time and, thus, there exists a time-discounting function ω and an aggregator function V induced by W. Note that while almost all welfare criteria in the literature are modular, some are not. We axiomatically characterize modular social welfare functions in Section 6.

[11] Section 6.2 relaxes several requirements introduced here, including the boundedness of the time-discounting function ω.

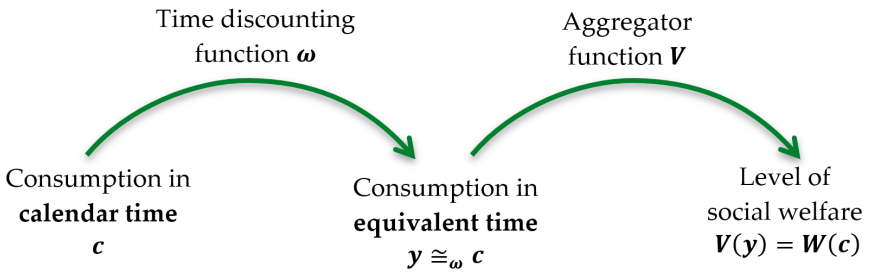


Figure 1: Modular theories of intergenerational justice derive the social welfare $W(c)$ through two modules. One module deals with discounting and identifies the equivalent-time representation y through the time-discounting function ω; the other deals with inequality and transforms y into social welfare through the aggregator function V.

2.2 Graphical illustration

We now illustrate the disentanglement property for the workhorse theory of intergenerational justice—exponentially-discounted utilitarianism. The time-discounting function that maps consumption streams in calendar time to consumption streams in equivalent time, $\omega^e(t) = 1 - e^{-\rho t}$, is shown in Figure 2. The aggregator function in the time-discounted consumption space, V^a, is additive. The resulting social welfare function takes the form

$$W = \int_0^\infty \omega'(t) f(c(t)) dt \propto \int_0^\infty e^{-\rho t} f(c(t)) dt,$$

with discount rate $\rho > 0$ and inequality aversion determined by the curvature of the increasing and concave function f.

In Figure 2, we illustrate the consumption stream in calendar time c, defined by

$$c(t) = \begin{cases} b, & \text{for } t \in [0, \tau) \\ a, & \text{for } t \in [\tau, \tau') \\ b, & \text{for } t \geqslant \tau'. \end{cases}$$

The corresponding representation in equivalent time y, also displayed, is

$$y(i) = \begin{cases} b, & \text{for } i \in [0, \iota) \\ a, & \text{for } i \in [\iota, \iota') \\ b, & \text{for } i \geqslant \iota', \end{cases}$$

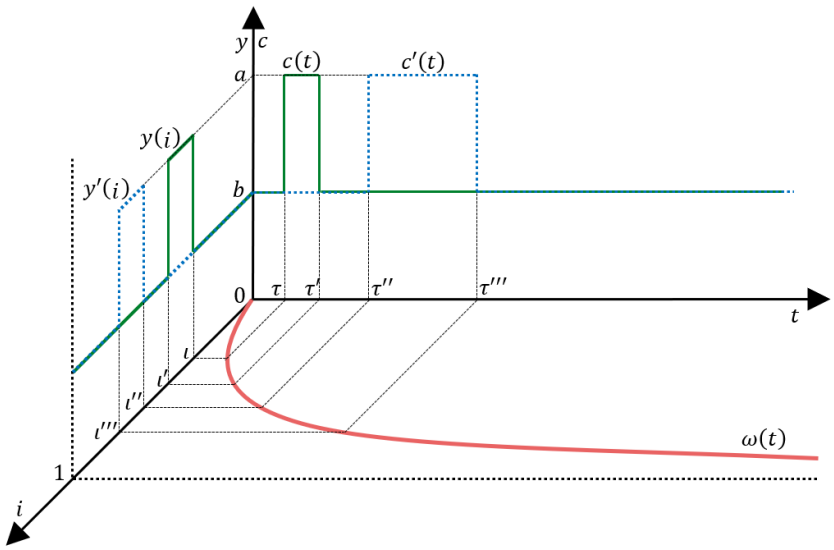


Figure 2: Consumption streams, $\mathfrak{c}$ (in calendar time) and $\mathfrak{y}$ (in equivalent time), and permuted consumption streams, $\mathfrak{c}'$ and $\mathfrak{y}'$. Consumption is transformed between calendar time $\mathfrak{t}$ and equivalent time $\mathfrak{i}$ by the time-discounting function ω.

with boundaries $\iota = 1 - e^{-\rho\tau}$ and $\iota' = 1 - e^{-\rho\tau'}$. Finally, the level of social welfare at $\mathfrak{c}$ is given by

$$W(\mathfrak{c}) = \int_0^{\tau} e^{-\rho t} f(b)\,dt + \int_{\tau}^{\tau'} e^{-\rho t} f(a)\,dt + \int_{\tau'}^{\infty} e^{-\rho t} f(b)\,dt.$$

The consumption stream in equivalent time $\mathfrak{y}'$ is obtained by a permutation of its consumption levels. As shown in Figure 2, the period with high consumption $\mathfrak{a}$ takes place later. Formally,

$$\mathfrak{y}'(\mathfrak{i}) = \begin{cases} \mathfrak{b}, & \text{for } \mathfrak{i} \in [0, \iota'') \\ \mathfrak{a}, & \text{for } \mathfrak{i} \in [\iota'', \iota''') \\ \mathfrak{b}, & \text{for } \mathfrak{i} \geqslant \iota''', \end{cases}$$

We next transform the permuted consumption $\mathfrak{y}'$ back to calendar time. It follows that $\tau'' = \ln(1/(1-\iota''))/\rho$, $\tau''' = \ln(1/(1-\iota'''))/\rho$, and $\tau' - \tau < \tau''' - \tau''$. Then, $\mathfrak{c}'$ is defined as

$$\mathfrak{c}'(\mathfrak{t}) = \begin{cases} \mathfrak{b}, & \text{for } \mathfrak{t} \in [0, \tau'') \\ \mathfrak{a}, & \text{for } \mathfrak{t} \in [\tau'', \tau''') \\ \mathfrak{b}, & \text{for } \mathfrak{t} \geqslant \tau'''. \end{cases}$$

The resulting level of social welfare is

$$W(c') = \int_0^{\tau''} e^{-\rho t} f(b)\,dt + \int_{\tau''}^{\tau'''} e^{-\rho t} f(a)\,dt + \int_{\tau'''}^{\infty} e^{-\rho t} f(b)\,dt.$$

Since the ordered consumptions of y and y' coincide (formally, $F_y(x) = F_{y'}(x)$ for each $x \geqslant 0$), ω and V ensure that c' achieves the same level of social welfare as c. Furthermore, since this is true for any consumption stream, exponentially-discounted utilitarianism is modular.

The insight that ω leads to anonymity with respect to y is powerful and generalizes to the most commonly used criteria. Next, we discuss these and other criteria through the lenses of our approach and show how the modular representation sheds new light on the assessment of intergenerational justice.

3 The discounting module

3.1 The time-discounting function

For the main results, the time-discounting function ω is continuous, strictly increasing, and satisfies $\omega(0) = 0$ and $\lim_{t\to\infty} \omega(t) = 1$. The function needs to be strictly monotone to ensure that the equivalent-time representation is informationally equivalent to the calendar-time representation of consumption streams.

When the time-discounting function ω is linear over an interval $[\underline{t}, \overline{t})$, all periods of time in this interval have the same social value. However, such an interval cannot cover the entire time horizon. If the time-discounting function is linear, it cannot have a bounded image; if it is constant, it is not informationally equivalent. As a result, if a social welfare function W is modular and, thus, admits an equivalent-time representation, it cannot be both anonymous and increasing in the consumption of all generations. As elaborated on in the Introduction, this tension is well-known in the literature on the aggregation of infinite streams (Diamond, 1965). We generalize time-discounting functions to include no discounting in Section 6.

Next, we discuss how the social welfare function can capture different ethical views about discounting.[12]

[12]It remains outside the scope of this paper to address how to compromise between different

3.2 Social impatience

We first formalize our concept of social impatience. For reference, previous definitions of impatience—such as Böhm-Bawerk (1889), Fisher (1930), and Koopmans (1960)—emphasize the preference for advancing rewards and postponing damages. In contrast, our definition involves switching the consumption levels of disjoint intervals of time. In plain words, assume that the switch is a matter of social indifference. Then, the social welfare function exhibits **social impatience** *only* if the later interval lasts longer.[13]

We introduce the following notation: a consumption stream $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-})$ assigns the constant consumption $\hat{c}$ to the interval $T_1 \equiv [\underline{t}_1, \overline{t}_1)$, the constant consumption $\tilde{c}$ to the interval $T_2 \equiv [\underline{t}_2, \overline{t}_2)$, and the consumption stream defined in c to the remaining times T^-, as shown in Figure 3.

A social welfare function W exhibits **social impatience** if for each pair of consumption streams $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-}), (\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T^-}) \in C$ such that

$$W(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-}) = W(\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T^-}),$$

with $\hat{c} \neq \tilde{c}$ and $\underline{t}_2 \geqslant \overline{t}_1$, it holds that $\overline{t}_2 - \underline{t}_2 \geqslant \overline{t}_1 - \underline{t}_1$. This definition is illustrated in Figure 3.

Proposition 1. *Assume that the social welfare function W is modular. Then, W exhibits social impatience if and only if ω is concave.*

The above also allows us to define the **pure time preference** between calendar time 0 and t of the social welfare function as

$$\frac{1}{t} \ln \left(\omega'(0) / \omega'(t) \right).$$

With social impatience, it follows that ω is concave and $\ln \left(\omega'(0) / \omega'(t) \right) > 0$.

views on discounting and, more generally, theories of intergenerational justice (see Millner and Heal (2023) and references therein). We leave such extensions, including whether the disentanglement might help aggregate different views by separately considering discounting from inequality attitudes, for future research.

[13]Our definition avoids taking a stand on aggregation, reflected by the implicit Paretian requirement that larger consumptions are more desirable, and, thus, is more general. In fact, rather than changing the consumptions of fixed generations, it changes the order of given consumptions for different durations by contracting/expanding time. The definitions are equivalent when assuming Pareto efficiency.

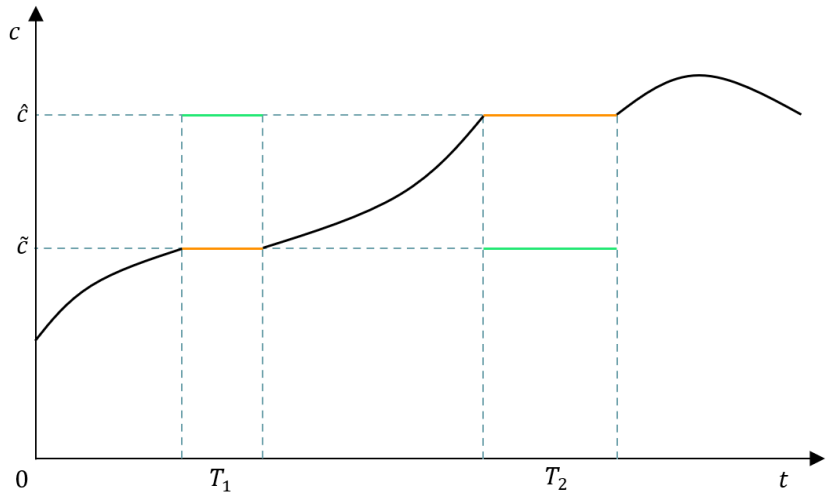


Figure 3: The green and orange consumption streams differ only during the intervals T_1 and T_2. Assume that the social welfare function W is indifferent between them. Then, W exhibits social impatience only if T_2 lasts longer than T_1.

A social welfare function W^* exhibits **more social impatience** than W if for each pair of consumption streams $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-}), (\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T-}) \in C$ such that

$$W(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-}) = W(\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T-}),$$

it holds that

$$W^*(\{\hat{c}\}_{T_1^*}, \{\tilde{c}\}_{T_2^*}, c_{T^*-}) = W^*(\{\tilde{c}\}_{T_1^*}, \{\hat{c}\}_{T_2^*}, c_{T^*-}),$$

where the later interval T_2^* is longer than than T_2, that is, $\overline{t}_2^* \geqslant \overline{t}_2$ and $\underline{t}_2^* = \underline{t}_2 \geqslant \overline{t}_1^* = \overline{t}_1 > \underline{t}_1^* = \underline{t}_1$.

Proposition 2. *Assume that the social welfare functions W and W^* are modular. Then, W^* exhibits more social impatience than W if and only if ω^* is everywhere more concave than ω.*

Corollary 1. *If ω and ω^* are differentiable, W^* exhibits more social impatience than W if and only if*

$$\frac{\omega''(t)}{\omega'(t)} \geqslant \frac{\omega^{*''}(t)}{\omega^{*'}(t)} \text{ for each } t \in T.$$

Comparisons by levels of impatience are not new in economics. For example, Prelec (2004), Attema et al. (2010), and Quah and Strulovici (2013) discuss

patience orderings of individual preferences that admit a time-separable representation. Weitzman (2001) and, more recently, Drupp et al. (2018), Gollier, van der Ploeg, and Zheng (2023), and Nesje et al. (2023) clarify that there is a large disagreement among experts about social discount rates and impatience. Our measure of social impatience highlights a close relationship to the Arrow-Pratt measures of risk aversion.

Next, we discuss how social impatience changes over time.

A social welfare function W exhibits **decreasing social impatience** if for each pair of consumption streams $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-}), (\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T^-}) \in C$ such that

$$W(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-}) = W(\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T^-})$$

and each $\tau \in T$ it holds that

$$W(\{\hat{c}\}_{T^*_{1+\tau}}, \{\tilde{c}\}_{T^*_{2+\tau}}, c_{T^{*-}}) = W(\{\tilde{c}\}_{T^*_{1+\tau}}, \{\hat{c}\}_{T^*_{2+\tau}}, c_{T^{*-}}),$$

where the later interval $T^*_{2+\tau}$ is shorter than T_2, that is, $\bar{t}^*_{2+\tau} \leqslant \bar{t}_2 + \tau$, and $\underline{t}^*_{2+\tau} = \underline{t}_2 + \tau \geqslant \bar{t}^*_{1+\tau} = \bar{t}_1 + \tau > \underline{t}^*_{1+\tau} = \underline{t}_1 + \tau$.

For each $\tau \in T$, let $\omega_\tau : T \to I$ be the truncated time-discounting function from τ onward to be such that $\omega_\tau(t) = \omega(t)/(1 - \omega(\tau))$.

Proposition 3*. Assume that the social welfare function W is modular. Then, W exhibits decreasing social impatience if and only if, for each $\tau \in T$, ω is everywhere more concave than ω_τ.*

Corollary 2*. If ω is differentiable, W exhibits decreasing social impatience if and only if for each $\tau \in T$*

$$\frac{\omega''(t)}{\omega'(t)} \leqslant \frac{\omega''_\tau(t)}{\omega'_\tau(t)} \text{ for each } t \in T.$$

3.3 Special cases

In the literature, the most common case is exponential discounting, which is typically adopted together with a time-additive aggregation of consumption streams (Samuelson, 1937; Koopmans, 1960).

The **exponential time-discounting function** is

$$\omega^e(t) = 1 - e^{-\rho t}, \text{ with } \rho > 0.$$

As expected, social impatience is higher when the discount rate ρ is larger.

Moreover, the exponential time-discounting function is the unique time-discounting function for which social impatience is constant. This provides a new characterization of exponential discounting that is independent of the time-additive aggregation of consumption streams.

Another case is that of quasi-hyperbolic discounting (Phelps & Pollak, 1968; Laibson, 1997; Barro, 1999; Montiel Olea & Strzalecki, 2014). This case generalizes exponential discounting by allowing a one-time change in the level of impatience at a specific point in time $\tau > 0$. The **quasi-hyperbolic time-discounting function** is then

$$\omega^{qh}(t) = \begin{cases} \beta\kappa(1 - e^{-\rho t}), & \text{for } t \in [0, \tau) \\ 1 - \kappa e^{-\rho t}, & \text{for } t \geqslant \tau, \end{cases}$$

with $\kappa = 1/(\beta + (1-\beta)e^{-\rho\tau})$. Clearly, the special case of exponential discounting emerges when $\beta = 1$.

Finally, we consider the hyperbolic time-discounting function proposed by Loewenstein and Prelec (1992), which allows for a smooth change of social impatience over time. The **hyperbolic time-discounting function** is

$$\omega^{h}(t) = 1 - (1 + \alpha t)^{1-\frac{\rho}{\alpha}},$$

with $\rho > \alpha > 0$. Exponential discounting emerges as the limit case when $\alpha \to 0$.

We represent these cases in Figure 4. The quasi-hyperbolic time-discounting function ω^{qh} and the exponential time-discounting function ω^{e} have the same discount factor. The difference is the additional weight placed on the interval $[0, \tau)$ due to $\beta > 1$ which leads to a kink. The difference between the hyperbolic time-discounting function ω^{h} and the exponential time-discounting function ω^{e}, illustrated for arbitrary discount factors, is the change in the weight over time due to $\alpha > 0$.

Note that we excluded the case of hyperbolic discounting where $\rho \leqslant \alpha$. For these parameters, the weight on future generations decreases too slowly for social welfare to admit an equivalent-time representation in infinite-time settings. To illustrate, consider simple hyperbolic discounting (Mazur, 1987), which emerges when $\rho = \alpha$. The weight attributed to consumption at time t follows a hyperbolic path and can be expressed as $(1 + \alpha t)^{-1}$. Then, the relative weight of any

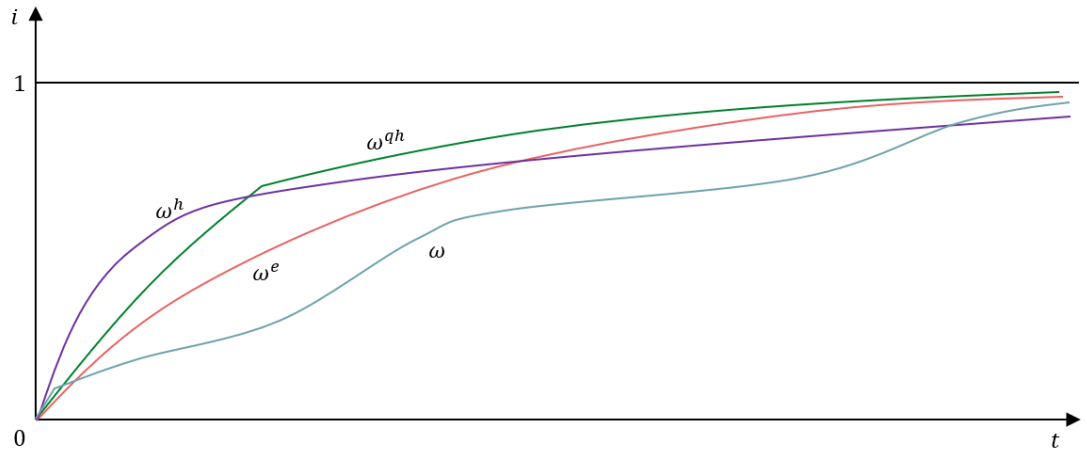


Figure 4: The time-discounting function ω and the special cases of exponential ω^{e}, quasi-hyperbolic ω^{qh} and hyperbolic ω^{h} time-discounting functions for arbitrary parameter values.

finite interval of calendar time $[\underline{t}, \bar{t})$ is infinitesimal compared to the total value assigned to the infinite time horizon (the integral of the weights is infinite for the entire time horizon). Since ultimately no finite interval matters in this setting, it is impossible to combine Pareto efficiency with a continuous and complete social welfare function that discounts the future by simple hyperbolic discounting. This difficulty extends to the case of hyperbolic discounting with $\rho < \alpha$ and is similar to no discounting. We deal with these cases in Section 6, where we extend our results to "generalized time-discounting" functions. Importantly, while the modular approach extends, these types of discounting attitudes require weakening either Pareto efficiency, continuity, or completeness.

The list of discounting attitudes discussed above is not exhaustive. With the extension of Section 6, any discounting attitude proposed in the literature is admissible, including unit-invariant, gamma, and Weibull discounting to name a few (Weitzman, 2001; Read, 2001; Ebert & Prelec, 2007; Jamison & Jamison, 2011; Attema et al., 2016; Karp, 2017).

4 The aggregation module

4.1 The aggregator function

As clarified, the defining property of an aggregator function V is its anonymity with respect to the equivalent-time representation of consumption streams. We next impose further restrictions that emerge when the social welfare function satisfies efficiency and inequality axioms.

First, we impose that the social welfare function is sensitive to the consumption of all generations. Formally, a social welfare function W is **Paretian** if for each pair of consumption streams $c, c' \in C$ such that $c(t) \geqslant c'(t)$ for each $t \in T$ and $\int_0^\tau c(t)dt > \int_0^\tau c'(t)dt$ for some $\tau \in T$, then $W(c) > W(c')$.

Proposition 4. *Assume that the social welfare function W is modular. Then, W is Paretian if and only if V satisfies first-order stochastic dominance, that is, if for each pair $y, y' \in Y$ such that $\int_0^\iota y(i)di \geqslant \int_0^\iota y'(i)di$ for each $\iota \in I$ and $\int_0^1 y(i)di > \int_0^1 y'(i)di$ implies $V(y) > V(y')$.*

Next, we discuss how social welfare functions capture aversion to inequalities.

4.2 Inequality aversion

Our definition of inequality aversion is inspired by the Pigou-Dalton transfer principle, which requires that a progressive transfer—a non-leaky transfer from a richer donor to a poorer recipient—reduces inequality (Atkinson, 1970). In the standard atemporal framework, all donors and all recipients are equally valued and justify the requirement that all progressive transfers increase social welfare. The key novelty here is to identify equally-valued donor and recipient generations based on discounting attitudes.[14]

A social welfare function W exhibits **inequality aversion**, in the discount-adjusted sense, if for each pair $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-}), (\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T-}) \in C$ and each

[14] Our definition first controls for social impatience by checking indifference to permutation. If indifference is satisfied, the progressive transfer is welfare improving. The definition thus avoids taking a stand on discounting. The standard Pigou-Dalton transfer principle emerges if there is no discounting.

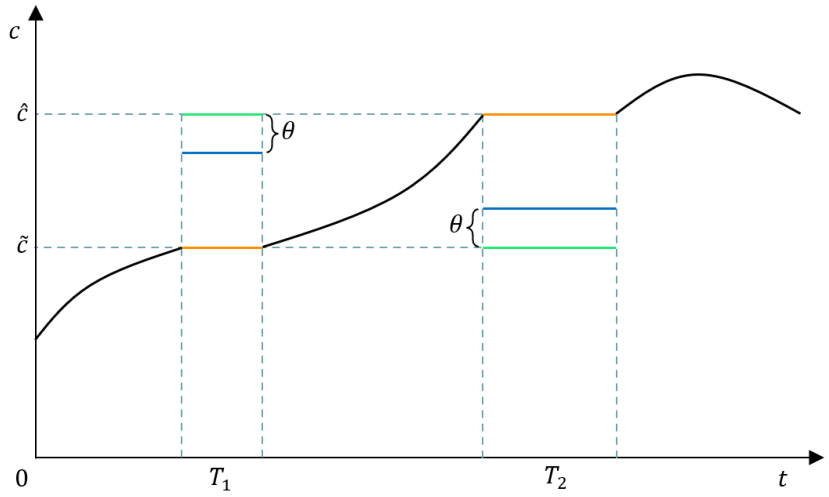


Figure 5: Consider three consumption streams that differ only during the intervals T_1 and T_2. Assume that the social welfare function W is indifferent between the orange and the green alternatives. Then, a progressive transfer reducing the consumption inequality between those intervals, as represented by the blue consumption levels, weakly increases social welfare. The size of the transfer is $\theta = \varepsilon\,(\hat{c} - \tilde{c})$ for some $\varepsilon \in (0,1)$.

$\varepsilon \in (0,1)$,

$$W(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) = W(\{\tilde{c}\}_{\mathsf{T}_1}, \{\hat{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-})$$

implies

$$W(\{(1-\varepsilon)\hat{c} + \varepsilon\tilde{c}\}_{\mathsf{T}_1}, \{(1-\varepsilon)\tilde{c} + \varepsilon\hat{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) \geqslant W(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}).$$

The idea is as follows and is illustrated in Figure 5. Assume that social preferences W are indifferent to changes in the order of constant consumptions $\hat{c}$ and $\tilde{c}$ between disjoint intervals T_1 and T_2. Then, the evaluator weakly prefers that the inequality in consumption is reduced. That is, a progressive transfer of consumption $\varepsilon \cdot |\hat{c} - \tilde{c}|$ between these intervals leads to a weakly higher level of social welfare.[15]

[15]The formulation in terms of a linear combination of consumption levels is similar to Piacquadio (2017) and avoids the need to specify the donor with larger consumption and the recipient with smaller consumption. Another approach that allows for different social weights of donors and recipients is by Berg and Piacquadio (2023), who suggest weighting individuals by their deservingness.

Proposition 5. *Assume that the social welfare function* W *is modular and Paretian. Then,* W *is inequality averse if and only if* V *satisfies second-order stochastic dominance, that is, if for each pair* $\mathsf{y}, \mathsf{y}' \in \mathsf{Y}$ *such that* $\int_0^{\mathsf{i}} \mathsf{y}(\mathsf{i})\mathrm{d}\mathsf{i} \geqslant \int_0^{\mathsf{i}} \mathsf{y}'(\mathsf{i})\mathrm{d}\mathsf{i}$ *for each* $\mathsf{i} \in \mathrm{I}$ *and* $\int_0^{\mathsf{i}} \mathsf{y}(\mathsf{i})\mathrm{d}\mathsf{i} > \int_0^{\mathsf{i}} \mathsf{y}'(\mathsf{i})\mathrm{d}\mathsf{i}$ *for some* $\mathsf{i} \in \mathrm{I}$ *implies* $\mathsf{V}(\mathsf{y}) \geqslant \mathsf{V}(\mathsf{y}')$.

Consider an aggregator function V and a time-discounted consumption stream $\mathsf{y} \in \mathsf{Y}$ such that $\mathsf{y} \equiv (\{\hat{\mathsf{y}}\}_{\mathrm{I}_1}, \{\tilde{\mathsf{y}}\}_{\mathrm{I}_2}, \mathsf{y}_{\mathrm{I}^-})$ with $\hat{\mathsf{y}} > \tilde{\mathsf{y}}$ and $\mathrm{I}_1, \mathrm{I}_2 \subset \mathrm{I}$ are equally-sized intervals. We define the **local inequality aversion** between I_1 and I_2 as

$$\mathsf{LIA}_{\mathrm{I}_1\mathrm{I}_2}(\mathsf{y}) = \frac{\lim_{\varepsilon \to 0^+} \frac{\mathsf{V}(\{\hat{\mathsf{y}}\}_{\mathrm{I}_1}, \{\tilde{\mathsf{y}}+\varepsilon\}_{\mathrm{I}_2}, \mathsf{y}_{\mathrm{I}^-})}{\varepsilon}}{\lim_{\varepsilon \to 0^-} \frac{\mathsf{V}(\{\hat{\mathsf{y}}+\varepsilon\}_{\mathrm{I}_1}, \{\tilde{\mathsf{y}}\}_{\mathrm{I}_2}, \mathsf{y}_{\mathrm{I}^-})}{\varepsilon}} - 1,$$

where the first term on the right-hand side is the the marginal rate of substitution between a high and low consumption in equivalent time.

Note that the above definition of local inequality aversion does not require differentiability. In fact, to accommodate a sufficiently rich variety of theories of justice, we cannot impose standard smoothness conditions—such as differentiability—on the social welfare function. For example, differentiability rules out theories of distributive justice where the rank of individuals matters (see d'Aspremont and Gevers (2002)).

Corollary 3. W *exhibits inequality aversion if and only if for each* $\mathsf{y} \equiv (\{\hat{\mathsf{y}}\}_{\mathrm{I}_1}, \{\tilde{\mathsf{y}}\}_{\mathrm{I}_2}, \mathsf{y}_{\mathrm{I}^-}) \in \mathsf{Y}$ *the local inequality aversion between* I_1 *and* I_2 *is non-negative, that is,* $\mathsf{LIA}_{\mathrm{I}_1\mathrm{I}_2}(\mathsf{y}) \geqslant 0$.

For the later discussion of the social discount rate, we also define the **instantaneous local inequality aversion** between calendar time 0 and t of the social welfare function W as the limit of the local inequality aversion between $\mathsf{T}_2 \equiv [\mathsf{t}, \mathsf{t} + \Delta\mathsf{t})$ and $\mathsf{T}_1 \equiv [0, \Delta\mathsf{t})$ when $\Delta\mathsf{t}$ goes to zero. Formally,

$$\frac{1}{\mathsf{t}} \ln\left(\mathsf{ILIA}_{\mathsf{t}}(\mathsf{y})\right).$$

Intuitively, if the social welfare function is inequality averse and consumption increases over time, $\ln\left(\mathsf{ILIA}_{\mathsf{t}}(\mathsf{y})\right) \geqslant 0$.

Next, we show how to compare social welfare functions in terms of the attitude to inequality they exhibit. For each pair of consumption streams $\mathsf{c} \equiv (\{\hat{\mathsf{c}}\}_{\mathsf{T}_1}, \{\tilde{\mathsf{c}}\}_{\mathsf{T}_2}, \mathsf{c}_{\mathsf{T}^-}), \mathsf{c}' \equiv (\{\hat{\mathsf{c}}'\}_{\mathsf{T}_1}, \{\tilde{\mathsf{c}}'\}_{\mathsf{T}_2}, \mathsf{c}_{\mathsf{T}^-}) \in \mathsf{C}$, we say that c' is obtained from

c through a **progressive leaky transfer** if there exist $\varepsilon, \varepsilon' \in (0,1)$ such that $\hat{c}' = \varepsilon\hat{c} + (1-\varepsilon)\tilde{c}$ and $\tilde{c}' = \varepsilon'\hat{c} + (1-\varepsilon')\tilde{c}$.

A social welfare function W^* exhibits **more inequality aversion** than W if for each triplet of consumption streams $c, c', c'' \in C$ such that $c \equiv (\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-})$, $c' \equiv (\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T^-})$ with $W(c) = W(c')$, and c'' is obtained from c through a progressive leaky transfer, it holds that: *i)* $W^*(c) = W^*(c')$, and *ii)* $W(c'') \geqslant W(c)$ implies $W^*(c'') \geqslant W^*(c)$.

The definition of "more inequality averse" is similar to that of "more risk averse" pioneered by Yaari (1969). Condition *i)* makes sure that the time-discounting function is equal across social welfare functions. Making a parallel with the individual risk setting, this ensures that the individuals being compared share the same beliefs about events: if violated, the willingness to take a gamble might be driven by different beliefs. Condition *ii)* sees to it that progressive leaky transfers are more acceptable, the more inequality averse social preferences are. Thus, if a progressive leaky transfer is welfare improving for W it must also be welfare improving for the more inequality concerned social welfare function W^*.

Proposition 6. *Assume that the social welfare functions W and W^* are modular and Paretian. Then, W^* exhibits more inequality aversion than W if and only if $\mathsf{LIA}^*_{I_1,I_2}(y) \geqslant \mathsf{LIA}_{I_1,I_2}(y)$ for each $y \in Y$ and each pair of disjoint intervals I_1, I_2.*

4.3 Special cases

In the literature, the most common aggregator functions are given by the sum of a concave transformation of each consumption. These include, as special cases, utilitarianism, generalized utilitarianism, and prioritarianism.[16] We refer to this family as **additive aggregator functions** and express it as

$$V^a = \int_0^1 f(y_i)di,$$

with f being an increasing concave transformation of each consumption. Here, the concavity of f dictates the degree of inequality aversion to differences in

[16] Atkinson (1970) has adopted this family for his seminal paper on inequality measurement. A recent generalization to multidimensional settings is Piacquadio (2017).

consumption levels. In fact, an aggregator function is more inequality averse if and only if the transformation function is more concave.[17]

A special case of additive criteria is the family of the **Atkinson aggregator functions**. These aggregator functions emerge when imposing *scale invariance*, that is, the requirement that the ranking is invariant to proportional rescaling of consumption streams. Formally, the family of Atkinson aggregator functions is given by

$$V^{a} = \int_0^1 \frac{(y_i)^{1-\eta}}{1-\eta} di,$$

where $\eta \geqslant 0$ is a parameter measuring inequality aversion to differences in consumption levels. This interpretation differs from the discounting literature where η often features as the elasticity of intertemporal substitution of individuals' utility functions (Drupp et al., 2018).

Another aggregator function is the weighted sum of "ordered" consumptions: this family draws on the similarity with choices under risk and reinterprets the risk preferences characterized by Yaari (1987, 1988).[18] The idea is to proceed in three steps. First, identify the rank $r(i)$ of consumption at i. By definition, $r(i) > r(j)$ means that $y_i \geqslant y_j$. Next, define a non-increasing weighting function $g : [0, 1) \to (0, \infty)$, attaching a weight to each rank. Finally, take the sum of the rank-weighted consumption stream. Clearly, when the weights are decreasing with the rank, a higher consumption is given a smaller weight, and social welfare is inequality averse. Thus, the slope of g is a measure of inequality aversion to differences in ranks. We refer to this family as **rank-weighted aggregator functions** and express it as

$$V^{rw} = \int_0^1 g(r(i))y_i di.$$

An important special case is that of the **single-parameter Gini aggregator**

[17] A function f' is more concave than a function f if the first can be expressed as a concave transformation of the second, that is, if $f' = \psi \circ f$ for some concave function ψ.

[18] This family is often referred to as "generalized Gini social welfare functions" and emphasizes the relationship with the Gini inequality measure (d'Aspremont & Gevers, 2002). We will return to the relationship between social welfare and inequality.

functions (Donaldson & Weymark, 1980, 1983; Yitzhaki, 1983; Aaberge, 2001). In this case, the weighting function g follows from

$$V^{rw} = (1+\gamma)\int_0^1 (1-r(i))^{\gamma} y_i \, di,$$

where $\gamma \geqslant 0$ is a parameter that measures inequality aversion to differences in ranks. More precisely, γ determines the speed at which the weights decrease with rank. The larger γ, the more inequality averse the aggregation is. The **Gini aggregator function** emerges when $\gamma = 1$.

A more general family includes the two above as special cases. The **rank-dependent aggregator function** allows for both a concave transformation of consumption and for a non-constant weighting of ranks. This aggregator function can be expressed as

$$V^{rd} = \int_0^1 g(r(i))f(y_i)di,$$

with g and f as defined above. When $g(r(i)) = 1$ for each $i \in I$, the additive criteria emerge. When f is the identity function, the rank-weighted criteria emerge.

The parametric forms above lead to

$$V^{rd} = (1+\gamma)\int_0^1 (1-(r(i))^{\gamma}\frac{(y_i)^{1-\eta}}{1-\eta}di,$$

where $\gamma \geqslant 0$ and $\eta \geqslant 0$ jointly determine social inequality aversion. Interestingly, γ and η influence inequality aversion in different ways. In fact, we can compare the aggregator functions within this class by their degree of inequality aversion *only* by dominance: one is more inequality averse if and only if both parameters are at least as large.

To illustrate the difference between the two, we represent the indifference curves of the above families of aggregator functions in Figure 6. On the two axes, we write the consumption levels y_1 and y_2 given, respectively, to $i \in [0, 1/2)$ and $i \in [1/2, 1)$. There are two main differences between the additive and rank-weighted aggregator functions. First, the additive criteria have smooth indifference curves at the 45 degree line of equality, while the rank-weighted ones exhibit a kink. Thus, the latter are more sensitive to "small" inequalities. Second,

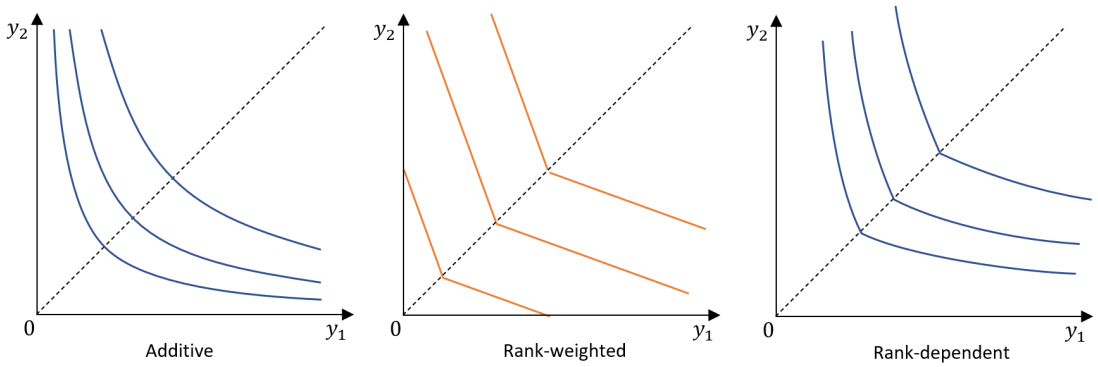


Figure 6: The special cases of the parameterized additive V^{a}, rank-weighted V^{rw} and rank-dependent V^{rd} aggregator functions for arbitrary parameter values.

the rank-weighted aggregator functions have piecewise-linear indifference curves, while those of additive criteria are strictly convex. Thus, the latter allow the local inequality aversion to increase with the income differences and thereby can avoid that the indifference curves intersect the axes. The rank-dependent aggregator functions are more flexible and allow the different sensitivities to inequality to be combined.

The list of aggregator functions discussed above is not exhaustive. Other aggregator functions could be considered, such as the Gâteaux smooth, implicit weighted and quadratic aggregator functions (see Machina (1982); Chew (1983); Fishburn (1983); Green and Jullien (1988); Chew and Nishimura (1992)).

5 Social welfare functions

In this section, we explore the theories of intergenerational justice that emerge by combining the discounting and inequality aversion modules.

We first present the *efficiency-equity* representation of modular criteria. We then identify the appropriate generalized "Ramsey rule." Finally, we illustrate these results for a tractable family of criteria.

5.1 Efficiency-equity representation

We next present a representation of social welfare that disentangles discounting attitudes from the efficiency-equity trade-off.

Let $\mu(y)$ be the mean of the consumption stream $y \in Y$. The mean is a measure of the **efficiency** of a consumption stream. Let the **equity** of y, denoted $E(y)$, be implicitly defined by setting $V(y) = V(\mu(y) \cdot E(y))$. The equity level $E(y)$ measures the share of consumption that would be needed if this were distributed equally. Note that when V is scale invariant—as for the special cases presented above—$1 - E$ is a relative measure of inequality (Blackorby & Donaldson, 1978). Since $W(c) = V(y)$ whenever $y \cong_\omega c$, a direct application of Atkinson (1970) implies that the social welfare function can be expressed in terms of the product of a measure of the efficiency and equity of y (the proof is omitted).

Proposition 7*. Assume that the social welfare function W is modular and Paretian. Then,*

$$W(c) = \mu(y) \cdot E(y),$$

where $y \cong_\omega c$.

5.2 Generalized Ramsey rule

Next, we discuss the assessment of marginal changes in consumption, when taking consumption at time 0 as the unit of account. Consider a project that has a cost in terms of consumption Δx during the time interval $T_1 \equiv [0, \Delta t)$ and a benefit Δc during the interval $T_2 \equiv [t, t + \Delta t)$. What is the lowest return that makes the project welfare improving? This return is the **generalized Ramsey rule** for our criteria.

For small Δt, the welfare-preserving return is approximated by the rate ξ such that $\Delta x = e^{-\xi t} \Delta c$.[19] Assume that the social welfare function is differentiable at the status quo consumption stream $c \in C$. Moreover, assume that $c = (\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T^-})$ assigns the constant consumption $\hat{c}$ to the interval T_1 and the constant consumption $\tilde{c}$ to the interval T_2. Then, the ratio $\Delta c / \Delta x$ is the

[19] To see this, note that the implicit return needs to satisfy the equation $\int_0^{\Delta t} e^{-\xi s} \Delta x \, ds = \int_t^{t+\Delta t} e^{-\xi s} \Delta c \, ds$. Then, take the limit for $\Delta t \longrightarrow 0$.

marginal rate of substitution of changes in consumption between the intervals T_2 and T_1.

Let I_1 and I_2 denote the corresponding intervals in equivalent time and y the equivalent-time representation of the status quo consumption stream. If these intervals had the same duration in equivalent time, the ratio $\Delta c/\Delta x$ would be equivalent to the local inequality aversion between I_2 and I_1. More generally, we can show that

$$\frac{\Delta c}{\Delta x} = \frac{\omega(\Delta t)}{\omega(t+\Delta t)-\omega(t)} \mathrm{LIA}_{I_2 I_1}(y).$$

The intuition is simple. When two intervals have different durations in equivalent time, their relative value at equal consumption is proportional to their duration. This immediately implies that we can weight the marginal rate of substitution in equivalent time—that is, the local inequality aversion—by the duration of the intervals.

The first fraction can be approximated by $\omega'(0)/\omega'(t)$ for small Δt. In Section 4, we defined the instantaneous local inequality aversion $\mathrm{ILIA}_t(y)$ as the limit of local inequality aversion when Δt goes to zero.

The generalized Ramsey rule identifies the "social discount rate" (the proof is omitted).

Proposition 8. *Assume that the social welfare function W is modular, Paretian, and differentiable at c. Then,*

$$\mathrm{SDR}_t = \frac{1}{t}\ln\left(\omega'(0)/\omega'(t)\right) + \frac{1}{t}\ln\left(\mathrm{ILIA}_t(y)\right).$$

It consists of two parts. The first part captures social attitudes to discounting, that is, the pure time preference of the welfare criterion. With social impatience, we have established that $\ln\left(\omega'(0)/\omega'(t)\right) > 0$. The second part captures the social attitudes to inequality at the status-quo consumption stream. If the social welfare function is inequality averse and consumption increases over time, it follows from Corollary 3 that $\ln\left(\mathrm{ILIA}_t(y)\right) \geqslant 0$.

We illustrate the generalized Ramsey rule by deriving the standard Ramsey rule for exponentially-discounted utilitarianism. For the exponential time-discounting function $\omega^e(t) = 1 - e^{-\rho t}$, the first component simplifies to ρ. For the Atkinson aggregator function $V^a = \int_0^1 \frac{(y_i)^{1-\eta}}{1-\eta} di$, the second component sim-

plifies to ηg_t^c, where g_t^c is the average growth in consumption between calendar time 0 and t.

Jointly, these special cases justify the well-known social discount rate identified in Ramsey (1928), defined by

$$SDR_t^{ea} = \rho + \eta g_t^c.$$

We highlight the social discount rate for other criteria below.

5.3 Modular theories of intergenerational justice

By freely combining special cases from the discounting and aggregation modules, several classes of theories of intergenerational justice emerge. Rather than reviewing this long list, we next highlight the efficiency-equity representation and the generalized Ramsey rule for the more general class of discounting-adjusted **rank-dependent social welfare functions**. Each social welfare function in this class can be represented by

$$W^{rd} = \int_0^\infty \omega'(t) g(r(\omega(t))) f(c_t) dt.$$

Here, g represents aversion to rank-inequality: additive criteria emerge as a special case when g is constant. In contrast, f represents aversion to consumption-level inequality: rank-weighted criteria emerge as a special case when f is the identity function. In our framework, this choice is orthogonal and compatible with any choice of time-discounting function, such as the exponential ω^e, quasi-hyperbolic ω^{qh}, and hyperbolic ω^h as defined in Section 3.

A simple parametric specification of the **exponentially-discounted rank-dependent social welfare functions** is

$$W^{rd} \propto (1+\gamma) \int_0^\infty e^{-\rho t} (1 - (r(\omega(t)))^\gamma \frac{(c_t)^{1-\eta}}{1-\eta} dt,$$

where $\gamma \geqslant 0$ measures *rank* inequality aversion and $\eta \geqslant 0$ measures *level* inequality aversion.

In terms of the efficiency-equity representation, these social welfare functions can be expressed as

$$W^{rd} = \mu(y) \cdot E^{rd}(y),$$

where the equity measure is given by the ratio between the equally-distributed equivalent consumption in equivalent time and its mean. Formally,

$$\mathsf{E}^{\mathrm{rd}}(\mathsf{y}) \equiv \frac{\left(\int_0^1 (1-\mathsf{r}(\mathsf{i}))^{\gamma}(\mathsf{y}_{\mathsf{i}})^{1-\eta}\,\mathrm{d}\mathsf{i}\right)^{\frac{1}{1-\eta}}}{\mu(\mathsf{y})}.$$

When $\gamma = 0$, $1-\mathsf{E}^{\mathrm{rd}}(\mathsf{y})$ is a member of the Atkinson family of inequality measures. When $\eta = 0$, $1-\mathsf{E}^{\mathrm{rd}}(\mathsf{y})$ is a member of the generalized Gini family of inequality measures. The Gini coefficient emerges when $\eta = 0$ and $\gamma = 1$.

The generalized Ramsey rule for the corresponding parametric social welfare function identifies the social discount rate as

$$\mathsf{SDR}_t^{\mathrm{erd}} = \rho + \eta \mathsf{g}_t^{\mathrm{c}} + \gamma \mathsf{g}_t^{\mathrm{r}}.$$

It consists of pure time preferences ρ from the discounting module, inequality aversion η towards consumption growth $\mathsf{g}_t^{\mathrm{c}}$, and inequality aversion γ towards rank growth $\mathsf{g}_t^{\mathrm{r}}$, implicitly defined by setting $\mathsf{e}^{t\mathsf{g}_t^{\mathrm{r}}} \equiv (1-\mathsf{r}_0)/(1-\mathsf{r}_t)$.

6 Characterization and generalizations

6.1 Existence and uniqueness of the modular representation

In this section, we characterize the family of modular theories of intergenerational justice.

The idea is to impose that there exist partitions of calendar time of equal social value.[20] The weakest requirement in this direction, satisfied by modular theories, is that it is possible to split the time horizon in two intervals for which permutations of constant consumption are a matter of social indifference.

[20]The use of midpoints has a long tradition in mathematics and have also been applied in economics. In particular, Attema et al. (2016) rely on midpoints to investigate individual's short-term discounting attitudes without estimating utility functions in a time separable framework. Our characterization results do not depend on time separability, but show that it imposes particularly demanding requirements on inequality attitudes.

A social welfare function W satisfies **minimal time split** if there exists a partition of calendar time $T_1, T_2 \subset T$ with $T_1 \bigcup T_2 = T$ such that for each pair of consumption levels $\underline{c}, \overline{c} \in \mathbb{R}_+$,

$$W(\underline{c}_{T_1}, \overline{c}_{T_2}) = W(\overline{c}_{T_2}, \underline{c}_{T_1}).$$

While *minimal time split* is too weak to characterize modular theories of intergenerational justice, this axiom already clarifies the nature of theories of justice that are not modular. For example, assume $W(\underline{c}_{T_1}, \overline{c}_{T_2}) = W(\overline{c}_{T_2}, \underline{c}_{T_1})$ for some $\underline{c}, \overline{c} \in \mathbb{R}_+$. *Minimal time split* requires that indifference holds for all levels of consumption. Although this is a plausible requirement for social welfare, this axiom might not fit all individuals' preferences. In fact, as first suggested by Ainslie (1975), some people might be less impatient when the stakes are larger. Habit formation is also ruled out (Constantinides, 1990; Campbell & Cochrane, 1999).

The following axiom strengthens *minimal time split*. It imposes that one can split any interval of calendar time into "equally valuable" subintervals in the sense of indifference to permutations.[21]

A social welfare function W satisfies **nested time split** if for each $\overline{T} \subseteq T$ there exists a partition of $\overline{T}$ in two time intervals $T_1, T_2 \subset T$ with $T_1 \bigcup T_2 = \bar{T}$ such that, for each consumption stream $c \in C$ and each pair of consumption levels $\underline{c}, \overline{c} \in \mathbb{R}_+$,

$$W(\underline{c}_{T_1}, \overline{c}_{T_2}, c_{T/\overline{T}}) = W(\overline{c}_{T_2}, \underline{c}_{T_1}, c_{T/\overline{T}}).$$

This stronger axiom also excludes the possibility that the discounting attitude between two subsequent intervals may depend on the consumption stream outside these intervals. This is the case of the utility function introduced by Uzawa (1969). His suggestion is to measure intertemporal utility of individuals by the sum of the utility of each periods' consumption, weighted by a discount factor that depends on past consumption.[22]

[21] Our axiom is similar in spirit to "separable present" and "separable future" (see Asheim (2010)). Rather than demanding that the welfare assessment is independent of the consumption assigned in the present and/or future tails, our axiom requires that the split into equally valuable subintervals is unaffected.

[22] Formally, the criterion can be expressed as $W(c) = \int_0^\infty \psi(c,t)u(c(t))dt$, where $\psi(c,t) =$

We can now state our characterization result.

Theorem 1*. Assume that the social welfare function* W *is Paretian. A theory of justice is modular if and only if the social welfare function* W *satisfies nested time split.*

Corollary 4. *Assume that the social welfare function* W *is Paretian. If* W *satisfies nested time split, the time-discounting function* ω *is unique, and the aggregator function* V *is unique up to an order-preserving transformation.*

Note that the Pareto requirement excludes that some time intervals have no social value and guarantees that the time-discounting function is everywhere strictly increasing. It thereby justifies our informational-equivalent requirement from Section 3.[23]

6.2 Generalized modular representations

We now explore the family of "modular" theories of intergenerational justice that emerge by satisfying *nested time split*, while relaxing the axiom of Pareto efficiency of the social welfare function. As we show, the disentanglement property is more general than implied above.[24] We first introduce a more general class of social

$e^{-\int_0^t \beta(c(s))ds}$, u is a continuous real valued function, and β is a continuous positive valued function. See the axiomatic characterization by Epstein (1983) in discrete time and by Hara (2016) in continuous time.

[23]It is standard to identify "local" discount rates by the ratio of the Volterra derivatives of social welfare at two instances of time with equal consumption (Ryder & Heal, 1973; Epstein, 1987; Bommier, Lanz, & Zuber, 2015). The axiom of *nested time split* implies that these local discount rates exist and are independent of the consumption streams. However, it is stronger since it also disciplines discounting attitudes when the consumption at these instants differs.

[24]We thank Geir Asheim for suggesting that the disentanglement property holds more generally, by relaxing Pareto efficiency. Other requirements may also be relaxed, such as continuity and completeness. Intuitively, relaxing continuity, there exist criteria that satisfy efficiency and anonymity in calendar time, and thus treat inequality as orthogonal to (no) discounting. However, the existence of such a criterion relies on non-constructive arguments (Svensson, 1980). As Fleurbaey and Michel (2003) conclude, "there exists no explicit description (that is, avoiding the axiom of choice or similar contrivances) of an ordering which satisfies the anonymity and weak Pareto axioms". Relaxing completeness by restricting the domain of consumption streams to the sub-domain of non-decreasing consumption streams is an alternative to relaxing efficiency within our framework. Here, we disregard these cases and continue working with

welfare functions.

Let the **generalized time-discounting function** $\bar{\omega} : \mathsf{T} \to \mathbb{R}_+$ be a continuous and strictly increasing transformation function mapping calendar time into **generalized equivalent time**, now possibly unbounded, defined by the positive real space. As before, the transformation function is such that changes in the order of constant consumption leave social welfare unchanged. An immediate advantage is to include anonymity in calendar time as a special case, which emerges when $\bar{\omega}$ is the identity function. Thus, our results speak to the large literature on combining Pareto efficiency and anonymity (see Asheim (2010); Pivato and Fleurbaey (2024)). More importantly, this extension allows for accommodating any type of discounting attitude, including negative discounting or "positive, but slow" discounting, such as the case of hyperbolic discounting with $\rho \leqslant \alpha$ discussed in Section 3.[25]

A social welfare function $\bar{W}$ is **generalized modular** if there exist a generalized time-discounting function $\bar{\omega}$ and an aggregator function $\bar{V}$ induced by $\bar{W}$.

The following axiom captures a very weak positional dominance principle (Basu & Mitra, 2003; Asheim, Kamaga, & Zuber, 2022). Assume that the consumption stream $\mathbf{c}$ assigns to the interval $\mathsf{T}_1 \subset \mathsf{T}$ the smallest consumption level $\hat{c} = \inf_t \mathbf{c}(t)$. Then, increasing the consumption of this interval improves social welfare.

A generalized social welfare function $\bar{W}$ satisfies **inf-restricted dominance** if for each pair $\mathbf{c} = (\{\hat{c}\}_{\mathsf{T}_1}, \mathbf{c}_{\mathsf{T}^-}), \mathbf{c}' = (\{\tilde{c}\}_{\mathsf{T}_1}, \mathbf{c}_{\mathsf{T}^-}) \in \mathbf{C}$ with $\hat{c} = \inf_t \mathbf{c}(t) > \tilde{c}$, it holds that $\bar{W}(\mathbf{c}) > \bar{W}(\mathbf{c}')$.

A stronger positional dominance principle extends the improvement to all intervals with finite rank, as suggested by Zuber and Asheim (2012). The idea

---

continuous and complete social welfare functions.

[25] To clarify, when discounting is "sufficiently fast," as for exponential discounting, this attitude is consistent with a continuous and complete Paretian social welfare function. The intuition is that every finite interval of calendar time has a non-negligible weight as compared to the infinite time horizon, allowing this weight to be represented on the unit interval by time-discounting functions. In contrast, when discounting slows down sufficiently (or becomes null or negative), this conclusion changes. Any finite interval of time has negligible weight compared to the infinite time horizon. As a result, there is a tension between the importance of such intervals required by Pareto efficiency and due to discounting.

is that larger consumption improves social welfare for all generations (here time intervals) with finite rank.

A generalized social welfare function $\bar{W}$ satisfies **liminf-restricted dominance** if for each pair $c = (\{\hat{c}\}_{T_1}, c_{T^-}), c' = (\{\tilde{c}\}_{T_1}, c_{T^-}) \in C$ with $\hat{c} > \tilde{c}$ and $\liminf_t c(t) > \tilde{c}$, it holds that $\bar{W}(c) > \bar{W}(c')$.

These weak dominance requirements impose sufficient sensitivity to ensure that nested time split characterizes the class of generalized modular social welfare functions.

Theorem 2*. Assume that the generalized social welfare function $\bar{W}$ satisfies inf-restricted dominance. A theory of justice is generalized modular if and only if the social welfare function $\bar{W}$ satisfies nested time split.*

Corollary 5. *Theorem 2 holds also when inf-restricted dominance is replaced by liminf-restricted dominance.*

Two well-known criteria belong to this family. The **maximin social welfare function**

$$\bar{W}^{maxmin} \equiv \inf_t c(t)$$

satisfies inf-restricted dominance. Here, the generalized time-discounting function $\bar{\omega}$ is the identity function, meaning that the criterion satisfies anonymity in calendar time, and the aggregator function is of the maximin type.

The maximin social welfare function exhibits infinite inequality aversion and focuses only on the worst-off generation. An alternative is to adopt a rank-weighted aggregator. This criterion, named a **rank-discounted utilitarian social welfare function**, has been proposed and axiomatically characterized by Zuber and Asheim (2012). It satisfies liminf-restricted dominance and can be expressed as

$$\bar{W}^{rdu} \equiv \int_0^\infty g(r(t))c_t\,dt,$$

where g is a decreasing function of the rank with $\lim_{r\to\infty} = 0$ (specifically, a generalized exponential Gini aggregator function).

Enriching the previous literature, we can easily generalize the maximin and rank-discounted utilitarian criteria to accommodate any generalized time-discounting function.

6.3 Further axioms: stationarity

Next, we look at the relationship between modular theories in Theorem 1 and standard, commonly used axioms of intergenerational justice.

A prominent requirement is "stationarity", first proposed by Koopmans (1960) to characterize exponentially-discounted utilitarianism. The idea is that, if two consumption streams have a common path until time $\bar{t}$, removing this common part leaves the ranking of the consumption streams unchanged. We formalize the axiom in our continuous-time setting as follows.

A social welfare function W satisfies **stationarity** if for each partition of calendar time $T_1, T_2 \subset T$ with $T_1 \bigcup T_2 = T$ and each pair of consumption streams $c, c' \in C$ with $c(t) = c'(t)$ for each $t \in T_1$, $W(c) \geqslant W(c')$ if and only if $W(c_{T_2}) \geqslant W(c'_{T_2})$.

Beyond exponentially-discounted utilitarianism, another criterion that satisfies *stationarity* is the Uzawa utility function discussed above. However, the latter violates *nested time split* and does not belong to the family of modular theories of intergenerational justice. Thus, one might wonder if there are other modular theories that satisfy stationarity. Our next result clarifies the point. Within the subset of modular theories of intergenerational justice, *stationarity* implies exponentially-discounted utilitarianism.

Theorem 3. *Assume that the social welfare function W is modular and Paretian. Then, W satisfies stationarity if and only if it is the exponentially-discounted utilitarian social welfare function.*

An intuition for this is that stationarity implies that the time-discounting function is exponential and aggregator function is additively separable.[26] Combining Theorems 1 and 3 immediately leads to the following corollary (the proof is omitted).

Corollary 6. *Assume that the social welfare function W is Paretian. Then, W satisfies nested time split and stationarity if and only if it is the exponentially-discounted utilitarian social welfare function.*

[26]Since V is anonymous, separable future implies the standard axiom of separability, that is, the ranking of two consumption streams is invariant to changes in the level of consumption of an unaffected interval. See Asheim, Banerjee, and Mitra (2021).

6.4 Further axioms: time consistency and time invariance

Next, we follow Halevy (2015) and introduce two additional properties of intertemporal criteria: "time consistency" and "time invariance". Time consistency is the requirement that higher-ranked consumption paths remain such, also as time passes. Time invariance is the requirement that the assessment of consumption streams is independent of the initial time.

Importantly, these properties are conceptually different from those discussed until now. Rather than imposing restrictions on the social welfare function W, these axioms impose some consistency requirement between social welfare functions at *different times*.[27] To formalize these axioms, we first extend our domain.

A consumption stream is now a bounded function $c_{\underline{t}} : [\underline{t}, \infty) \to \mathbb{R}_+$ that assigns an instantaneous consumption $c(t)$ to each time $t \in [\underline{t}, \infty)$ with $\underline{t} \geqslant 0$. Let $C_{\underline{t}}$ denote the set of consumption streams $c_{\underline{t}}$. A social welfare function at $\underline{t}$ is $W_{\underline{t}} : C_{\underline{t}} \to \mathbb{R}$.

Social welfare functions $W_{\underline{t}}$, one for each $\underline{t}$, satisfy **time consistency** if for each $\tau > 0$ and each pair of consumption streams $c, c' \in C$ with $c(t) = c'(t)$ for each $t \in [0, \tau)$, $W(c) \geqslant W(c')$ if and only if $W_\tau(c_\tau) \geqslant W_\tau(c'_\tau)$, where $c(t) = c_\tau(t)$ and $c'(t) = c'_\tau(t)$ for each $t \geqslant \tau$.

Social welfare functions $W_{\underline{t}}$, one for each $\underline{t}$, satisfy **time invariance** if for each $\tau > 0$ and each pair of consumption streams $c, c' \in C$, $W(c) \geqslant W(c')$ if and only if $W_\tau(c_\tau) \geqslant W_\tau(c'_\tau)$, where $c(t) = c_\tau(t + \tau)$ and $c'(t) = c'_\tau(t + \tau)$ for each $t \in T$.

The triplet of axioms—stationarity, time consistency, and time invariance—has the following property: satisfying any two of them implies the remaining one (see Halevy (2015)). Here, we implicitly redefine stationarity to be jointly imposed for each social welfare function $W_{\underline{t}}$. Then, a modular theory of intergenerational justice that satisfies both time consistency and time invariance must also satisfy stationarity and, by Theorem 3, must be the exponentially-discounted

[27] The difference is logically similar to single-profile versus multi-profile approaches for the aggregation of preferences (see Roberts (1980)). Results in a single-profile approach are conceptually stronger and extend to all profiles in the domain. Results in the multi-profile approach generally substitute equity and efficiency requirements with consistency of criteria to changes in the profile.

utilitarian social welfare function. However, stationarity alone is not equivalent to the combination of time consistency and time invariance. Stationarity implies that each social welfare function $W_{\underline{t}}$ is exponentially-discounted utilitarian, but the time-discounting function and the aggregator functions could change over time. Time consistency or time invariance further impose that these functions remain the same over time.

As we show next, time consistency permits any discounting attitude (provided these are respected when time passes), but imposes additive separability of the aggregator function.[28]

Theorem 4. *Assume that the social welfare functions* $W_{\underline{t}}$*, one for each* $\underline{t}$*, are modular and Paretian. Then,* $W_{\underline{t}}$*, one for each* $\underline{t}$*, satisfy time consistency if and only if: i) the time-discounting functions satisfy* $\omega_{\underline{t}}(t) = 1-(1-\omega(t))/(1-\omega(\underline{t}))$ *for each* $\underline{t} \in T$ *and each* $t \geqslant \underline{t}$*; and ii) the aggregator function is additive and satisfies* $V_{\underline{i}} = V^a$*, with* $\underline{i} = \omega(\underline{t})$*, for each* $\underline{t} \in T$*.*

In contrast to stationarity and time consistency, time invariance imposes no restrictions on the discounting and inequality attitudes of modular theories of intergenerational justice. Said differently, any modular theory of intergenerational justice is admissible. The only requirement is that the social welfare function—and thus the time-discounting and aggregator functions—does not change over time (the proof is omitted).[29]

Theorem 5. *Assume that the social welfare functions* $W_{\underline{t}}$*, one for each* $\underline{t}$*, are modular and Paretian. Then,* $W_{\underline{t}}$*, one for each* $\underline{t}$*, satisfy time invariance if and only if* $W_{\underline{t}} = W_{\underline{t}'}$ *for each* $\underline{t}, \underline{t}' \in T$*.*

---

[28] Millner and Heal (2018) adopt a more nuanced setting, where the ranking of consumption streams can be history specific. This approach significantly weakens the axiom of time consistency, as the past consumption can be taken into account when assessing alternatives. As an implication, the additive separability of the aggregator function does not follow.

[29] As Halevy (2015, p. 341) writes: "by itself, [time invariance] does not impose restrictions on the structure of preferences at any given time, but only implies that preferences are not a function of calendar time. If preferences are time invariant, only time relative to the evaluation period matters."

7 Conclusion

Policies affect present and future generations differently. Any comparison of benefits and losses across generations—for instance, to identify optimal policies—depends on the theory of intergenerational justice adopted. Without choosing a particular theory, the role of economics is to operationalize these theories and to show their policy implications.

When comparing gains and losses across generations, two aspects of intergenerational justice are central. These are the attitudes toward earlier versus later generations—intergenerational discounting—and the attitudes toward better-off versus worse-off generations—intergenerational inequality. In this paper, we show how these attitudes can be combined, identify the correspondence with theories of intergenerational justice, and clarify the relationship to existing commonly used criteria. Our approach includes as a special case the workhorse theory of intergenerational justice—exponentially-discounted utilitarianism—as well as many more novel criteria. These criteria, resulting from the combination of discounting and inequality attitudes that have emerged separately in their respective literature, enrich the economic toolbox of theories of intergenerational justice.

Our main result is to axiomatically justify all theories of intergenerational justice that disentangle discounting and inequality. A distinctive feature is that this class of theories builds on and emphasizes a variety of attitudes toward discounting and inequality. Their combination is more likely to match the wide range of views held by individuals. Another novel contribution of our approach is to explore intergenerational justice through the lens of widely debated theories of *intratemporal* distributional justice. Furthermore, while the representation is very general and flexible, the parametric specifications of these theories remain suitable for studying optimal policies.

Modular theories open up many avenues for future research. A natural first step is to identify the intergenerational discounting and inequality attitudes that underlie individuals' ethical views and choices. A further step is to apply these criteria to economic models, such as integrated models of the climate and economy, to guide optimal policies. The framework can also be generalized along several dimensions. For example, population growth could be readily accommodated

by a "population-weighted" time-discounting function: beyond discounting, the relative weight of any two intervals would be proportional to their population. Intratemporal inequalities can be accounted for by substituting the instantaneous consumption $\mathtt{c}(\mathtt{t})$ with the equally-distributed equivalent consumption of the individuals living at $\mathtt{t}$ (or alternatively, those born at $\mathtt{t}$). Multidimensional settings can be incorporated by replacing consumption with some measure of well-being, capabilities, or opportunities. Finally, risk can be accounted for by simply adding the expectation operator to the criterion, or, to disentangle social attitudes to risk and inequality, by introducing a risk-averse aggregation of the equally-distributed equivalent representation of welfare in each state.

References

Aaberge, R. (2001). Axiomatic characterization of the Gini coefficient and Lorenz curve orderings. *Journal of Economic Theory*, *101*(1), 115–132.

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463–496.

Arrow, K. J. (1965). Aspects of the Theory of Risk-Bearing. *Yrjo Hahnsson Foundation*.

Asheim, G. B. (2010). Intergenerational equity. *Annual Review of Economics*, *2*(1), 197–222.

Asheim, G. B., Banerjee, K., Mitra, T. (2021). How stationarity contradicts intergenerational equity. *Economic Theory*, *72*, 423–444.

Asheim, G. B., Kamaga, K., Zuber, S. (2022). Maximal sensitivity under strong anonymity. *Journal of Mathematical Economics*, *103*, 102768.

Asheim, G. B., Mitra, T., Tungodden, B. (2012). Sustainable recursive social welfare functions. *Economic Theory*, *49*(2), 267–292.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263.

Attema, A. E., Bleichrodt, H., Gao, Y., Huang, Z., Wakker, P. P. (2016). Measuring discounting without measuring utility. *American Economic Review*, *106*(6), 1476–1494.

Attema, A. E., Bleichrodt, H., Rohde, K. I., Wakker, P. P. (2010). Time-tradeoff sequences for analyzing discounting and time inconsistency. *Management Science*,

56(11), 2015–2030.

Barro, R. J. (1999). Ramsey meets Laibson in the neoclassical growth model. *Quarterly Journal of Economics*, *114*(4), 1125–1152.

Basu, K., Mitra, T. (2003). Aggregating infinite utility streams with intergenerational equity: The impossibility of being Paretian. *Econometrica*, *71*(5), 1557–1563.

Berg, K., Piacquadio, P. G. (2023). Fairness and Paretian social welfare functions. *Mimeo, University of St. Gallen*.

Blackorby, C., Donaldson, D. (1978). Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory*, *18*(1), 59–80.

Böhm-Bawerk, E. V. (1889). *The Positive Theory of Capital*. Stechert (Translation).

Bommier, A., Kochov, A., Le Grand, F. (2017). On monotone recursive preferences. *Econometrica*, *85*(5), 1433–1466.

Bommier, A., Lanz, B., Zuber, S. (2015). Models-as-usual for unusual risks? On the value of catastrophic climate change. *Journal of Environmental Economics and Management*, *74*, 1–22.

Campbell, J. Y., Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, *107*(2), 205–251.

Chew, S. H. (1983). A generalization of the quasilinear mean with application to the measurement of income inequality and decision theory resolving the Allais Paradox. *Econometrica*, *51*, 1065–1092.

Chew, S. H., Mao, M. H. (1995). A Schur concave characterization of risk aversion for non-expected utility preferences. *Journal of Economic Theory*, *67*(2), 402–435.

Chew, S. H., Nishimura, N. (1992). Differentiability, comparative statics and non-expected utility preferences. *Journal of Economic Theory*, *56*(2), 294–312.

Chichilnisky, G. (1996). An axiomatic approach to sustainable development. *Social Choice and Welfare*, *13*(2), 231–257.

Clark, A. E., d'Ambrosio, C. (2015). Attitudes to income inequality: Experimental and survey evidence. In *Handbook of Income Distribution* (Vol. 2, pp. 1147–1208). Elsevier.

Cohen, J., Ericson, K. M., Laibson, D., White, J. M. (2020). Measuring time preferences. *Journal of Economic Literature*, *58*(2), 299–347.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy*, *98*(3), 519–543.

Dasgupta, P. (2008). Discounting climate change. *Journal of Risk and Uncertainty*,

37, 141–169.

d'Aspremont, C., Gevers, L. (2002). Social welfare functionals and interpersonal comparability. In *Handbook of Social Choice and Welfare* (Vol. 1, pp. 459–541). Elsevier.

Del Campo, S., Anthoff, D., Kornek, U. (2024). Inequality aversion for climate policy. *Review of Environmental Economics and Policy*, *18*(1), 96–115.

Diamond, P. A. (1965). The evaluation of infinite utility streams. *Econometrica*, 170–177.

Dietz, S., Stern, N. (2008). Why economic analysis supports strong action on climate change: A response to the Stern Review's critics. *Review of Environmental Economics and Policy*, *2*(1), 94-113.

Dillenberger, D., Gottlieb, D., Ortoleva, P. (in press). Stochastic impatience and the separation of time and risk preferences. *Theoretical Economics*.

Donaldson, D., Weymark, J. A. (1980). A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, *22*(1), 67–86.

Donaldson, D., Weymark, J. A. (1983). Ethically flexible Gini indices for income distributions in the continuum. *Journal of Economic Theory*, *29*(2), 353–358.

Drupp, M. A., Freeman, M. C., Groom, B., Nesje, F. (2018). Discounting disentangled. *American Economic Journal: Economic Policy*, *10*(4), 109–134.

Ebert, J. E., Prelec, D. (2007). The fragility of time: Time-insensitivity and valuation of the near and far future. *Management Science*, *53*(9), 1423–1438.

Eden, M. (2023). The cross-sectional implications of the social discount rate. *Econometrica*, *91*(6), 2065–2088.

Epstein, L. G. (1983). Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, *31*(1), 133–152.

Epstein, L. G. (1987). A simple dynamic general equilibrium model. *Journal of Economic Theory*, *41*(1), 68–95.

Epstein, L. G., Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, *57*(4), 937–969.

Epstein, L. G., Zin, S. E. (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, *99*(2), 263–286.

Fehr, E., Epper, T., Senn, J. (2023). The fundamental properties, stability and predictive power of distributional preferences. *CESifo Working Paper No. 10727*.

Fishburn, P. C. (1983). Transitive measurable utility. *Journal of Economic Theory*, *31*(2), 293–317.

Fisher, I. (1930). *The Theory of Interest.* Macmillan.

Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy*, *118*(4), 649–680.

Fleurbaey, M., Michel, P. (2003). Intertemporal equity and the extension of the Ramsey criterion. *Journal of Mathematical Economics*, *39*(7), 777–802.

Fleurbaey, M., Zuber, S. (2013). Inequality aversion and separability in social risk evaluation. *Economic Theory*, *54*, 675–692.

Fleurbaey, M., Zuber, S. (2015). Discounting, risk and inequality: A general approach. *Journal of Public Economics*, *128*, 34–49.

Frederick, S., Loewenstein, G., O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

Gollier, C. (2013). *Pricing the Planet's Future: The Economics of Discounting in an Uncertain World*. Princeton University Press.

Gollier, C. (2024). The welfare cost of ignoring the beta. *Toulouse School of Economics Working Paper No. 1556*.

Gollier, C., van der Ploeg, F., Zheng, J. (2023). The discounting premium puzzle: Survey evidence from professional economists. *Journal of Environmental Economics and Management*, *122*, 102882.

Green, J. R., Jullien, B. (1988). Ordinal independence in nonlinear utility theory. *Journal of Risk and Uncertainty*, *1*(4), 355–387.

Halevy, Y. (2015). Time consistency: Stationarity and time invariance. *Econometrica*, *83*(1), 335–352.

Hara, K. (2016). Characterization of stationary preferences in a continuous time framework. *Journal of Mathematical Economics*, *63*, 34–43.

Hvidberg, K. B., Kreiner, C. T., Stantcheva, S. (2023). Social positions and fairness views on inequality. *Review of Economic Studies*, *90*(6), 3083–3118.

Jamison, D. T., Jamison, J. (2011). Characterizing the amount and speed of discounting procedures. *Journal of Benefit-Cost Analysis*, *2*(2), 1–56.

Karp, L. (2017). Provision of a public good with multiple dynasties. *Economic Journal*, *127*(607), 2641–2664.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica*, *28*(2), 287–309.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of*

Economics, *112*(2), 443–478.

Li, C., Rohde, K. I., Wakker, P. P. (2024). The deceptive beauty of monotonicity, and the million-dollar question: Row-first or column-first aggregation? *Mimeo, Erasmus School of Economics*.

Loewenstein, G., Prelec, D. (1992). Anomalies of intertemporal choice: Evidence and interpretation. *Quarterly Journal of Economics*, *107*, 573–597.

Machina, M. J. (1982). "Expected utility" analysis without the independence axiom. *Econometrica*, *50*(2), 277–323.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In *The Effect of Delay and of Intervening Events on Reinforcement Value* (pp. 55–73). Lawrence Erlbaum Associates, Inc.

Millner, A., Heal, G. (2018). Time consistency and time invariance in collective intertemporal choice. *Journal of Economic Theory*, *176*, 158–169.

Millner, A., Heal, G. (2023). Choosing the future: Markets, ethics, and rapprochement in social discounting. *Journal of Economic Literature*, *61*(3), 1037–1087.

Montiel Olea, J. L., Strzalecki, T. (2014). Axiomatization and measurement of quasi-hyperbolic discounting. *Quarterly Journal of Economics*, *129*(3), 1449–1499.

Nesje, F., Drupp, M. A., Freeman, M. C., Groom, B. (2023). Philosophers and economists agree on climate policy paths but for different reasons. *Nature Climate Change*, *13*(6), 515–522.

Nesje, F., Drupp, M. A., Freeman, M. C., Groom, B. (2024). Sense and sensitivity: An argument against reporting multiple Net Present Values. *Available at SSRN 4663124*.

Nordhaus, W. D. (2007). A review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, *45*(3), 686–702.

Okun, A. M. (1975). *Equality and Efficiency: The Big Tradeoff*. Brookings Institution Press.

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

Piacquadio, P. G. (2017). A fairness justification of utilitarianism. *Econometrica*, *85*(4), 1261–1276.

Piacquadio, P. G. (2020). The ethics of intergenerational risk. *Journal of Economic Theory*, *186*, 104999.

Pivato, M., Fleurbaey, M. (2024). Intergenerational equity and infinite-population ethics: A survey. *Journal of Mathematical Economics*, *113*, 103021.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*(1), 122–136.

Prelec, D. (2004). Decreasing impatience: A criterion for non-stationary time preference and "hyperbolic" discounting. *Scandinavian Journal of Economics*, *106*(3), 511–532.

Quah, J. K.-H., Strulovici, B. (2013). Discounting, values, and decisions. *Journal of Political Economy*, *121*(5), 896–939.

Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal*, *38*(152), 543–559.

Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty*, *23*(1), 5–32.

Roberts, K. W. (1980). Social choice theory: The single-profile and multi-profile approaches. *Review of Economic Studies*, *47*(2), 441–450.

Rothschild, M., Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, *2*(3), 225–243.

Ryder, H. E., Heal, G. M. (1973). Optimal growth with intertemporally dependent preferences. *Review of Economic Studies*, *40*(1), 1–31.

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

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Svensson, L.-G. (1980). Equity among generations. *Econometrica*, *48*, 1251–1256.

Traeger, C. P. (2014). Why uncertainty matters: Discounting under intertemporal risk aversion and ambiguity. *Economic Theory*, *56*, 627–664.

Uzawa, H. (1969). Time preference and the Penrose effect in a two-class model of economic growth. *Journal of Political Economy*, *77*(4, Part 2), 628–652.

Weitzman, M. L. (2001). Gamma discounting. *American Economic Review*, *91*(1), 260–271.

Weitzman, M. L. (2007). A review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, *45*(3), 703–724.

Yaari, M. E. (1969). Some remarks on measures of risk aversion and on their uses. *Journal of Economic Theory*, *1*(3), 315–329.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, *55*(1), 95–115.

Yaari, M. E. (1988). A controversial proposal concerning inequality measurement.

Journal of Economic Theory, *44*(2), 381–397.

Yitzhaki, S. (1983). On an extension of the Gini inequality index. *International Economic Review*, *24*(3), 617–628.

Zuber, S., Asheim, G. B. (2012). Justifying social discounting: The rank-discounted utilitarian approach. *Journal of Economic Theory*, *147*(4), 1572–1601.

A Appendix

A.1 Proof of Proposition 1

($\Rightarrow$) Assume by contradiction that ω is not concave at $\tau \in \mathsf{T}$. Then, there exist $\varepsilon > 0$ and $\tau^* > \tau$ such that

$$\frac{\omega(\tau - \varepsilon) + \omega(\tau + \varepsilon)}{2} = \omega(\tau^*) > \omega(\tau).$$

Now consider a consumption stream such that $(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-})$ with $\hat{c} \neq \tilde{c}$ and $\mathsf{T}_1 \equiv [\tau - \varepsilon, \tau^*)$ and $\mathsf{T}_2 \equiv [\tau^*, \tau + \varepsilon)$.

Since $\omega(\tau^*) - \omega(\tau - \varepsilon) = \omega(\tau + \varepsilon) - \omega(\tau^*)$, it must hold that

$$W(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) = W(\{\tilde{c}\}_{\mathsf{T}_2}, \{\hat{c}\}_{\mathsf{T}_1}, c_{\mathsf{T}^-}).$$

Since this is true for each $\hat{c}$ and each $\tilde{c}$ and since $\tau^* - (\tau - \varepsilon) > (\tau + \varepsilon) - \tau^*$, the later interval is shorter, and a contradiction emerges.

($\Leftarrow$) Assume by contradiction that the social welfare function W does not exhibit social impatience. Then, there exists a pair of consumption streams $(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-})$, $(\{\tilde{c}\}_{\mathsf{T}_1}, \{\hat{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) \in \mathsf{C}$ such that

$$W(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) = W(\{\tilde{c}\}_{\mathsf{T}_1}, \{\hat{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}),$$

with $\hat{c} \neq \tilde{c}$, $\underline{t}_2 \geqslant \bar{t}_1$, and $\bar{t}_2 - \underline{t}_2 < \bar{t}_1 - \underline{t}_1$.

Since W is modular, there exists a time-discounting function ω. By anonymity in equivalent time, there exists $k \in [0, 1)$ such that $k = \omega(\bar{t}_1) - \omega(\underline{t}_1) = \omega(\bar{t}_2) - \omega(\underline{t}_2)$. Thus, the slopes of the time-discounting function over the intervals T_1 and T_2 are, respectively, $k/(\bar{t}_1 - \underline{t}_1)$ and $k/(\bar{t}_2 - \underline{t}_2)$. Since the first is smaller than the second, the time-discounting function ω cannot be concave. A contradiction. ■

A.2 Proof of Proposition 2

($\Rightarrow$) Assume that W^* exhibits more social impatience than W and, by contradiction, that ω^* is not more concave than ω. Then, the function ζ such that $\omega^* = \zeta(\omega)$ is not concave. Thus, there exists some $\tau \in \mathsf{T}$ such that ζ is strictly convex at $\omega(\tau)$.

Consider the interval $\mathsf{T}_1 \equiv [\tau - \varepsilon, \tau)$ for $\varepsilon > 0$. For a sufficiently small ε, there exist intervals $\mathsf{T}_2 \equiv [\tau, \tau + k)$ and $\mathsf{T}_2^* \equiv [\tau, \tau + k^*)$ with $\omega(\tau + k) - \omega(\tau) = \omega(\tau) - \omega(\tau - \varepsilon)$ and $\omega^*(\tau + k^*) - \omega^*(\tau) = \omega^*(\tau) - \omega^*(\tau - \varepsilon)$.

Since the intervals have equal duration in equivalent time, these are indifferent for W and W^*, respectively. Thus, we satisfy the premise of the definition of exhibiting more social impatience. However, by the convexity of ζ, it must be true that, for sufficiently small ε, $\tau + k > \tau + k^*$. Thus, T_2 lasts longer than T_2^*. This contradicts that W^* exhibits more social impatience.

($\Leftarrow$) Assume that ω^* is more concave than ω and, by contradiction, that W^* does not exhibit more social impatience than W.

Then, there exist a pair of consumption streams $(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}), (\{\tilde{c}\}_{\mathsf{T}_1}, \{\hat{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) \in \mathsf{C}$ such that

$$W(\{\hat{c}\}_{\mathsf{T}_1}, \{\tilde{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}) = W(\{\tilde{c}\}_{\mathsf{T}_1}, \{\hat{c}\}_{\mathsf{T}_2}, c_{\mathsf{T}^-}),$$

it holds that

$$W^*(\{\hat{c}\}_{\mathsf{T}_1^*}, \{\tilde{c}\}_{\mathsf{T}_2^*}, c_{\mathsf{T}^{*-}}) = W^*(\{\tilde{c}\}_{\mathsf{T}_1^*}, \{\hat{c}\}_{\mathsf{T}_2^*}, c_{\mathsf{T}^{*-}}),$$

where the later (contiguous) interval T_2^* is *shorter* than T_2, that is, $\bar{t}_2^* < \bar{t}_2$, $\underline{t}_2^* = \underline{t}_2 = \bar{t}_1^* = \bar{t}_1 > \underline{t}_1^* = \underline{t}_1$.

By construction of the time-discounting functions, $\omega(\bar{t}_1) - \omega(\underline{t}_1) = \omega(\bar{t}_2) - \omega(\underline{t}_2)$ and $\omega^*(\bar{t}_1^*) - \omega^*(\underline{t}_1^*) = \omega^*(\bar{t}_2^*) - \omega^*(\underline{t}_2^*)$.

Since ω^* is more concave than ω, it must be true that, for each $t \geqslant \bar{t}_1^* = \bar{t}_1$

$$\frac{\omega(t) - \omega(\bar{t}_1)}{\omega(\bar{t}_1) - \omega(\underline{t}_1)} \geqslant \frac{\omega^*(t) - \omega^*(\bar{t}_1^*)}{\omega^*(\bar{t}_1^*) - \omega^*(\underline{t}_1^*)}.$$

However, this is contradicted at $t = \bar{t}_2^*$, where the left-hand side is smaller than 1 and the right-hand side equals 1. ■

A.3 Proof of Corollary 1

The corollary follows quite directly from Proposition 2. The key is to realize that, when ω and ω^* are differentiable, ω^* is more concave than ω if and only if

$$\frac{\omega''(t)}{\omega'(t)} \geqslant \frac{\omega^{*''}(t)}{\omega^{*'}(t)} \text{ for each } t \in T.$$

This result is analogous to Arrow (1965) and Pratt (1964) for comparative measures of risk aversion. ■

A.4 Proof of Proposition 3

Let $\tau \in T$. Define the truncated social welfare function W_τ by setting $W_\tau(c) = W(\bar{c}_{T_1}, c)$ for some $\bar{c} \geqslant 0$ and $T_1 \equiv [0, \tau)$. Then, the time-discounting function of the truncated social welfare function is $\omega_\tau(t) = \omega(t)/(1 - \omega(\tau))$. By definition, W has "decreasing social impatience" if and only if W exhibits more social impatience than W_τ for each $\tau \in T$. Then, the result immediately follows from Proposition 2. ■

A.5 Proof of Corollary 2

Given the relationship between Proposition 2 and Proposition 3, the result follows immediately from Corollary 1. ■

A.6 Proof of Proposition 4

The result follows from the definition of the representation in equivalent time. More precisely, W is Paretian if and only if for each pair of consumption streams $c, c' \in C$ such that $c(t) \geqslant c'(t)$ for each $t \in T$ and $\int_0^\tau c(t)dt > \int_0^\tau c'(t)dt$ for some $\tau \in T$, then $W(c) > W(c')$. By substituting c for $y(\omega)$, the definition of first-order stochastic dominance of V emerges. ■

A.7 Proof of Proposition 5

We first define the Pigou-Dalton transfer principle in our domain. The aggregator function V satisfies the **Pigou-Dalton transfer principle** if and only if for each

stream in equivalent time $(\{\hat{y}\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I^-}) \in Y$ with $\hat{y} \neq \tilde{y}$, $|I_1| = |I_2|$ and each $\varepsilon \in (0,1)$,

$$V(\{(1-\varepsilon)\hat{y} + \varepsilon\tilde{y}\}_{I_1}, \{(1-\varepsilon)\tilde{y} + \varepsilon\hat{y}\}_{I_2}, y_{I^-}) \geqslant V(\{\hat{y}\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I^-}).$$

By construction of the representation in equivalent time, a modular social welfare function W is inequality averse if and only if V satisfies the Pigou-Dalton transfer principle.

Thus, the proposition follows from the equivalence between second-order stochastic dominance and mean-preserving spread identified by Rothschild and Stiglitz (1970) and first applied to the measurement of inequality by Atkinson (1970). ■

A.8 Proof of Corollary 3

Consider an aggregator function V and a consumption stream in equivalent time $y \in Y$ such that $y \equiv (\{\hat{y}\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I^-})$ with $\hat{y} > \tilde{y}$ and $I_1, I_2 \subset I$ are equally-sized intervals.

We make use of the definition of the Pigou-Dalton transfer principle in our domain from Proposition 5 and the corresponding representation in equivalent time. As the aggregator function V satisfies the Pigou-Dalton transfer principle, it is Schur concave on Y, that is for all $y_1, y_2 \in Y$ such that y_1 is majorized by y_2, we have $V(y_1) \geqslant V(y_2)$. Because V is Schur concave, W exhibits inequality aversion if and only if for each $y \equiv (\{\hat{y}\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I^-}) \in Y$, $\mathsf{LIA}_{I_1 I_2}(y) \geqslant 0$.

Thus, the corollary follows from the equivalence between mean-preserving spread and Schur concavity identified for risk by Chew and Mao (1995), and first derived for intergenerational inequality here. In their setting, Schur concavity is characterized by the marginal rate of substitution between a high and low income state being higher than 1. ■

A.9 Proof of Proposition 6

($\Rightarrow$) Assume by contradiction that W^* is more inequality averse than W while $\mathsf{LIA}^*_{I_1,I_2}(y) < \mathsf{LIA}_{I_1,I_2}(y)$. Then there exists a leaky progressive transfer that is accepted by W, but not by W^*. A contradiction.

($\Leftarrow$) Assume by contradiction that W is more inequality averse than W^*. From Corollary 3, we know that V is Schur concave. The statements W is more inequality averse than W^* and V is more Schur concave than V^* are equivalent when applying Chew and Mao (1995) to our setting. This implies the contradiction $LIA^*_{I_1,I_2}(y) \leqslant LIA_{I_1,I_2}(y)$. ■

A.10 Proof of Theorem 1

($\Rightarrow$) Let $\overline{T} \subseteq T$. Construct a partition of $\overline{T}$ in two time intervals $T_1 \equiv [\underline{t}_1, \bar{t}_1), T_2 \equiv [\underline{t}_2, \bar{t}_2) \subset T$ with $T_1 \bigcup T_2 = \bar{T}$ such that $\omega(\bar{t}_1) - \omega(\underline{t}_1) = \omega(\bar{t}_2) - \omega(\underline{t}_2)$. By construction, for each pair of consumption streams $c \equiv (\underline{c}_{T_1}, \overline{c}_{T_2}, c_{T/\overline{T}}), c' \equiv (\overline{c}_{T_2}, \underline{c}_{T_1}, c_{T/\overline{T}}) \in C$, the associated streams in equivalent time $y, y' \in Y$ are permutations of one another, that is, $F_y(x) = F_{y'}(x)$ for each $x \in \mathbb{R}_+$. Thus, $W(c) = W(c')$ and nested time split holds.

($\Leftarrow$) We proceed in two steps. In Step 1, we construct the time-discounting function. In Step 2, we construct the aggregator function.

Step 1. Let $k \in \mathbb{N}_+$. Let the time interval $T^m \equiv [\underline{t}^m, \bar{t}^m) \subseteq T$ for each $m \in M_k \equiv \{1, 2, ..., 2^k\}$, be such that: i) $\bar{t}^m = \underline{t}^{m+1}$ for each $m < 2^k$; ii) $\underline{t}^1 = 0$ and $\bar{t}^{2^k} = \infty$; and iii) $W(\underline{c}_{T^m}, \overline{c}_{T^{m'}}, c_{T/\{T^m \bigcup T^{m'}\}}) = W(\overline{c}_{T_2}, \underline{c}_{T_1}, c_{T/\overline{T}})$ for each pair $m, m' \in M_k$ and for each consumption stream $c \in C$ with a pair of consumption levels $\underline{c}, \overline{c} \in \mathbb{R}_+$. By nested time split, these intervals exist. Since W is Paretian, these are unique.

For each $m \in M_k$, set $\omega(\underline{t}^m) = (m-1)/2^k$. Since this is true for each k, the function ω is defined for a dense subset of the domain. By continuity, ω is uniquely defined.

Since W is Paretian and continuous, ω is strictly increasing and continuous. By construction, it takes values between 0 and 1. Thus, it is a time-discounting function.

Step 2. Set $V(y) = W(c)$ for each $c \in C$, $y \in Y$, such that $y \cong_\omega c$. Since ω is an equivalent representation on the unit interval, V is well-defined. Moreover, by construction of the time-discounting function ω, V is invariant to permutations of consumption in equivalent time and is therefore an aggregator function. ■

A.11 Proof of Corollary 4

Assume that there are two different time-discounting functions ω and ω' for the same modular and Paretian social welfare function W.

Assume that $\omega(t) = \omega'(t') = 0.5$ for some $t, t' \in T$ with $t' > t$. Then, the social welfare function W is indifferent to permutations of consumption between the time intervals $T_1 = [0, t)$ and $T_2 = [t, \infty)$ and between the time intervals $T_1' = [0, t')$ and $T_2' = [t', \infty)$. This leads to a violation of Pareto: the interval $[t, t')$ must be irrelevant for social welfare. Thus, $\omega(t) = \omega'(t') = 0.5$ implies $t = t'$.

Following the same reasoning, further splitting the time intervals at $\omega(t) = \omega'(t') = 0.25$ and $\omega(t'') = \omega'(t''') = 0.75$ shows that the time-discounting functions need to be equal at these points, that is, $t = t'$ and $t'' = t'''$. Proceeding recursively by further splitting the time intervals shows that $\omega(t) = \omega'(t') = i$ implies $t = t'$ for each rational number $i \in \mathbb{Q}$ such that $0 < i < 1$. By continuity, the result extends to all $i \in I$ and, thus, $\omega = \omega'$.

Finally, since W is ordinal (unique up to an order-preserving transformation), so is V. ■

A.12 Proof of Theorem 2

($\Rightarrow$) The corresponding proof for Theorem 1 continues to hold for generalized modular theories.

($\Leftarrow$) We follow the proof of Theorem 1, where we replace Step 1 by the following. Step 2 continues to hold.

Step 1'. We construct the generalized time-discounting function $\bar{\omega}$.

Let $k \in \mathbb{N}_+$. Let the time interval $T^m \equiv [\underline{t}^m, \overline{t}^m) \subseteq T$ for each $m \in \mathbb{N}_+$, be such that: i) $\overline{t}^m = m/k = \underline{t}^m + 1/k$; and ii) $\bar{W}(\underline{c}_{T^m}, \overline{c}_{T^{m'}}, c_{T/\{T^m \bigcup T^{m'}\}}) = \bar{W}(\overline{c}_{T_2}, \underline{c}_{T_1}, c_{T/\overline{T}})$ for each pair $m, m' \in \mathbb{N}_+$ and for each consumption stream $c \in C$ with a pair of consumption levels $\underline{c}, \overline{c} \in \mathbb{R}_+$.

By nested time split, these intervals exist. By inf-restricted dominance, these intervals are unique. For each $m \in \mathbb{N}_+$, set $\bar{\omega}(\underline{t}^m) = (m-1)/k$. Since this is true for each k, we can identify the function $\bar{\omega}$ for a dense subset of the domain.

By continuity, $\bar{\omega}$ is uniquely defined.

Since $\bar{W}$ satisfies inf-restricted dominance and is continuous, $\bar{\omega}$ is strictly increasing and is continuous. By construction, it takes values between 0 and ∞ (but could be bounded above). Thus, it is a generalized time-discounting function. ■

A.13 Proof of Corollary 5

The proof of Theorem 2 extends directly when inf-restricted dominance is strengthened to liminf-restricted dominance. ■

A.14 Proof of Theorem 3

It is immediate that exponentially-discounted utilitarianism satisfies stationarity (Koopmans, 1960). We show the converse.

By definition, a modular social welfare function $W(\mathbf{c})$ admits a representation in equivalent time, that is, there exist a time-discounting function ω and an aggregator function V such that $W(\mathbf{c}) = V(\mathbf{y})$ whenever $\mathbf{y} \cong_\omega \mathbf{c}$.

Step 1. Stationarity implies that the time-discounting function is exponential ω^e.

Define the importance of a time interval $[\underline{t}, \bar{t})$ by $\omega(\bar{t}) - \omega(\underline{t})$ and the relative importance of two time intervals by their ratio.

For each $s, t \in T$ with $s \leqslant t$, stationarity requires that the relative importance of two time intervals $T_1 \equiv [s, s+t)$ and $T_2 \equiv [s+t, \infty)$ is unchanged, when anticipating these time intervals by s to $T_1' \equiv [0, t)$ and $T_2' \equiv [t, \infty)$. Thus, the relative importance of T_2 with respect to $T_1 \bigcup T_2$ must equal that of T_2' with respect to $T_1' \bigcup T_2'$. Formally,

$$\frac{1-\omega(s+t)}{1-\omega(s)} = 1-\omega(t).$$

or, equivalently, that $1-\omega(s+t) = (1-\omega(t))(1-\omega(s))$.

This functional equation has the general solution $1-\omega(t) = (1-\omega(1))^t$ for any rational $t \in T$. Since ω is continuous and strictly increasing,

$$1-\omega(t) = e^{-\rho t},$$

where $\rho \equiv -\ln(1-\omega(1))$ and, by rearranging, $\omega(t) = \omega^e(t) = 1 - e^{-\rho t}$.

Step 2. Stationarity implies that the aggregator function is additively separable V^a.

By stationarity, adding or removing a common interval in calendar time does not affect the ranking of these alternatives. By modularity, the same holds for consumption streams in equivalent time. Thus, the aggregator function V satisfies **separable future**, that is, for each partition of equivalent time $I_1, I_2 \subset I$ with $I_1 \bigcup I_2 = I$ and each pair of consumption streams in equivalent time $\hat{y}, \tilde{y} \in Y$, if $V(\hat{y}_{I_1}, \hat{y}_{I_2}) \geqslant V(\hat{y}_{I_1}, \tilde{y}_{I_2})$ then $V(\tilde{y}_{I_1}, \hat{y}_{I_2}) \geqslant V(\tilde{y}_{I_1}, \tilde{y}_{I_2})$.

Since V is anonymous, separable future implies the standard axiom of **separability**, that is, the ranking of two consumption streams is invariant to changes in the level of consumption of an unaffected interval.

Together with Pareto, separability implies that the aggregator function V can be written as a sum of some transformation of the consumption streams. Thus, $V(y) = V^a(y)$. ■

A.15 Proof of Theorem 4

($\Rightarrow$) Consider modular social welfare functions satisfying i) and ii) and, by contradiction, assume that these are not time-consistent. Then, there exist $\tau > 0$ and a pair of consumption streams $c, c' \in C$ with $c(t) = c'(t)$ for each $t \in [0, \tau)$ such that $W(c) \geqslant W(c')$ and $W_\tau(c_\tau) < W_\tau(c'_\tau)$, where $c(t) = c_\tau(t)$ and $c'(t) = c'_\tau(t)$ for each $t \geqslant \tau$. Let $y, y', y_\iota, y'_\iota \in Y$ be such that $y \cong_\omega c$, $y' \cong_\omega c'$, $y_\iota \cong_{\omega_\tau} c_\tau$, $y'_\iota \cong_{\omega_\tau} c'_\tau$, and $\iota = \omega(\tau)$.

By condition i),

$$y(i) = y_\iota \left(1 - \frac{1-i}{1-\iota}\right)$$

and

$$y'(i) = y'_\iota \left(1 - \frac{1-i}{1-\iota}\right).$$

By condition ii),

$$W(c) = V(y) = \int_0^1 f(y(i))di \geqslant \int_0^1 f(y'(i))di = V(y') = W(c'),$$

and, similarly,

$$W_\tau(c_\tau) = V_\iota(y_\iota) = \int_0^1 f(y_\iota(i))di < \int_0^1 f(y'_\iota(i))di = V_\iota(y'_\iota) = W_\tau(c'_\tau).$$

Since $c(t) = c'(t)$ for each $t \in [0, \tau)$,

$$\int_\iota^1 f(y(i))di \geqslant \int_\iota^1 f(y'(i))di.$$

Let's adopt the variable change $j = 1 - (1-i)/(1-\iota)$. Then, $dj = di/(1-\iota)$ and, substituting,

$$(1-\iota)\int_0^1 f(y(j))dj \geqslant (1-\iota)\int_0^1 f(y'(j))dj.$$

Since $y(j) = y_\iota(i)$ and $y'(j) = y'_\iota(i)$ whenever $i = j$,

$$\int_0^1 f(y_\iota(i))di \geqslant \int_0^1 f(y'_\iota(i))di,$$

leading to a contradiction.

($\Leftarrow$)

Step 1. Time consistency implies i).

Consider the social welfare functions W and W_τ for some $\tau > 0$. The corresponding time-discounting functions are denoted ω and ω_τ, respectively.

Divide the time horizon $[\tau, \infty)$ into $K \in \mathbb{N}_+$ intervals with equal duration in equivalent time. Formally, for each $k \in [1, K]$ the corresponding time interval is defined as $T_k \equiv [t_{k-1}, t_k)$ such that: $t_0 = \tau$; $t_K = \infty$; and $\omega(t_k) - \omega(t_{k-1}) = (1 - \omega(\tau))/K$. Note that the latter is equivalent to $\omega(t_k) = k(1 - \omega(\tau))/K$.

Because the social welfare functions are modular, changes in the order of constant consumption across these time intervals in equivalent time are a matter of social indifference. By time consistency, the indifference extends to W_τ. Thus, by the modular structure, the time-discounting function ω_τ must satisfy $\omega_\tau(t_k) - \omega_\tau(t_{k-1}) = 1/K$ for each $k \in [1, K]$.

Combining the previous ones gives,

$$\omega_\tau(t_k) - \omega_\tau(t_{k-1}) = \frac{\omega(t_k) - \omega(t_{k-1})}{1 - \omega(\tau)} = \frac{1}{K}.$$

Substituting recursively, and using the fact that $\omega_\tau(t_K) = \omega(t_K) = 1$, gives

$$1 - \omega_\tau(t_k) = \frac{1 - \omega(t_k)}{1 - \omega(\tau)}.$$

Since this is true for each time t_k such that $\omega_\tau(t_k) = k/K$ is a rational number, property i) follows by continuity of the time-discounting function.

Step 2. Time consistency implies ii).

Consider two consumption streams in equivalent time $y, y' \in Y$ with $V(y) \geqslant V(y')$.

Divide the time horizon in $K \in \mathbb{N}_+$ intervals with equal duration in equivalent time. Formally, for each $k \in [1, K]$ the corresponding time interval is defined as $T_k \equiv [t_{k-1}, t_k)$ such that: $t_0 = 0$; $t_K = \infty$; and $\omega(t_k) - \omega(t_{k-1}) = 1/K$. Note that the latter is equivalent to $\omega(t_k) = k/K$.

Because the social welfare functions are modular, changes in the order of constant consumption across these time intervals in equivalent time are a matter of social indifference. By time consistency, the indifference extends to W_{t_k} for each $k \in [1, K]$. Formally, $V(y) \geqslant V(y')$ if and only if $V_{i_k}(y_{i_k}) \geqslant V_{i_k}(y'_{i_k})$ with $i_k = \omega(t_k)$, where $y_{i_k}(\omega_{t_k}(t)) = y(\omega(t))$ and $y'_{i_k}(\omega_{t_k}(t)) = y'(\omega(t))$ for each $t \geqslant t_k$.

By anonymity in equivalent time and time consistency, the ranking of any of these intervals is independent of the level of consumption of the remaining intervals. Thus, the social welfare function satisfies separability (as for Theorem 3) and property ii) follows. ■