

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/381520539>

Intergenerational Discounting and Inequality

Preprint · June 2024

DOI: 10.13140/RG.2.2.18112.14080

CITATIONS

0

READS

498

2 authors:



Frikk Nesje

University of Copenhagen

29 PUBLICATIONS 581 CITATIONS

SEE PROFILE



Paolo G. Piacquadio

University of St.Gallen

28 PUBLICATIONS 106 CITATIONS

SEE PROFILE

Intergenerational Discounting and Inequality

Frikk Nesje¹ and Paolo G. Piacquadio²

June 19, 2024

Abstract

We study theories of justice that disentangle normative views on intergenerational discounting and intergenerational inequality. Any modular social welfare function is uniquely identified by a time-discounting function—capturing attitudes across generations—and an aggregator function—capturing attitudes towards inequality. The rich choice of such functions allows our theories to include the most common welfare criteria adopted in the literature as special cases and unveils yet unexplored families of alternative criteria. Our axiomatic characterization clarifies the properties and limits of disentangling discounting and inequality.

Keywords: Intergenerational justice, discounting, inequality.

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

²School 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 Rolf Aaberge, David Anthoff, Geir Asheim, Maximilian Auffhammer, Yaroslav Chechel, Moritz Drupp, Maya Eden, Johannes Emmerling, Eli Fenichel, Bård Harstad, Larry Karp, J. Paul Kelleher, Antony Millner, Morten Olsen, Riccardo Piacquadio, Peter Norman Sørensen, Philipp Strack, Christian Traeger, Kristen Vamsæter, Ulrich Wagner, and Peter Wakker for comments, in addition to seminar and conference audiences in Berkeley, Copenhagen, Hamburg, Munich, New Orleans, Oslo, Santa Barbara, Southampton, Stanford, St. Gallen, and online. This project has received funding from the Economic Policy Research Network (Grant agreement No. 124) and the European Research Council under the European Union's Horizon 2020 research and innovation program ERC Starting Grant VALURED (Grant agreement No. 804104). Frikk Nesje has received funding from the Independent Research Fund Denmark for a research stay at the University of California, Berkeley (Grant agreement No. 3126-00011B).

1 Introduction

Policies implemented today affect both present and future generations. Some generations may benefit from such policies, while others may lose. For example, earlier generations bear the costs of mitigation of climate change, research and development for technological change, as well as resource conservation, while later generations enjoy the returns of these sacrifices (Stern, 2007; Dasgupta, 2021). To aggregate the benefits and losses of generations and decide on policies, two aspects of intergenerational justice play a critical role: the attitude toward earlier versus later generations—i.e., intergenerational discounting—and the attitude toward better-off versus worse-off generations—i.e., intergenerational inequality. Yet, their overlap and interaction remain unclear.

This paper develops a new, large family of theories of intergenerational justice. An important feature of these theories is that the two attitudes—discounting and inequality—are independent of each other. This orthogonality property proves surprisingly malleable and provides tractable criteria that capture a large variety of ethical views. For example, one can consider non-constant social impatience (such as hyperbolic discounting) and positional attitudes to inequality between generations (that is, prioritizing worse-off generations for the fact of being worse off).

We formalize and axiomatically characterize the theories of intergenerational justice that disentangle discounting and inequality attitudes. These theories consist of two modules: one module deals exclusively with discounting (independently of inequality attitudes), while the other deals exclusively with inequality (independently of the discounting attitudes). By varying the discounting and inequality modules, many theories of intergenerational justice emerge. Importantly, our characterization is based on an axiom that is implicit in existing criteria, such as exponentially-discounted utilitarianism and its generalizations. For this reason, the most common theories of intergenerational justice adopted in the literature are “modular” and are special cases of our approach.

Our interest is in ranking intergenerational consumption streams.¹ The cen-

¹We follow convention and interpret consumption at an instance of calendar time either as the lifetime equivalent consumption of the generation born at that time or the equivalent

tral step of our representation result is to transform consumption streams in *calendar time* to consumption streams in *equivalent time*. The duration of each generation is expanded/compressed up to the point where all consumption has equal social value (and is thus expressed in equivalent time). Since each instant of consumption in equivalent time has the same value, we can apply any standard *anonymous* inequality averse aggregation of consumption in that space (see, among others, d’Aspremont and Gevers (2002)). Thus, each modular theory is fully described by the *time-discounting function*—that transforms the consumption streams to equivalent time—and the *aggregator function*—that aggregates these transformed consumption streams.

Our theories of intergenerational justice offer a number of valuable features for policy applications. First, by adopting novel definitions of social impatience and inequality aversion, we can compare the discounting and inequality attitudes across social welfare functions. Next, by relying on anonymous aggregator functions, we can measure and compare the levels of inequality of any consumption stream as well as decompose social welfare into their efficiency and equality components (independently of discounting). Such a decomposition exercise has the potential to simplify and clarify cross-country comparisons of social welfare and better inform policy choices.

Relatedly, the discounting and inequality attitudes provide a much more general (and yet easy to interpret) version of the “Ramsey rule”, 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.

Most intergenerational welfare criteria in economics rely on convenient functional forms, such as exponentially-discounted utilitarianism and other additively separable formulations. However, these social welfare functions impose very specific restrictions that might not well represent wide-held ethical views. Our results clarify that there exist tractable and transparent social welfare functions that relax these restrictions along the disentanglement of discounting and inequality

consumption enjoyed by society at that instant. This allows us to interpret consumption streams as “intergenerational consumption”. For the purposes of this paper, we need not take a stance on how consumption is aggregated to a real number at each instant of calendar time (see Li et al. (2024)), including how this could be derived from an underlying model.

attitudes. We combine the discounting and inequality modules to uncover classes of theories of intergenerational justice.

Specifically, we axiomatically characterize this disentanglement property by imposing the novel axiom of *nested time split*. It requires that any interval of calendar time can be divided into “equally valuable” sub-intervals. Intuitively, two sub-intervals are equally valuable when permuting their constant consumption is a matter of social indifference. The axiom also requires that the split is independent of the specific consumption stream considered. To illustrate, exponential discounting at a 1% rate would assess the next 38 years as equally valuable as the subsequent 62 years (splitting the next 100 years), and the next 69 years as equally valuable as all later years (splitting the entire time horizon).

The difference between the discounting and inequality attitudes remains unclear in the literature. Within the boundaries of the workhorse theory of intergenerational justice—exponentially-discounted utilitarianism—there has been a large debate on the discount rate and the concavity of the utility function (Stern, 2007; Nordhaus, 2007; Dasgupta, 2008). Yet, no parameters seem to entirely match individuals’ views on these important determinants of social welfare and policy (Weitzman, 2001; Drupp et al., 2018; Millner and Heal, 2023).² For these reasons, the literature tends to rely on *ad hoc* formulations of inequality aversion and employ sensitivity analysis (Del Campo et al., 2023; Nesje et al., 2023b).

Our approach enriches the economic toolbox of sensible theories of intergenerational justice and provides tractable criteria that can better match wide-held ethical views. Attitudes toward intragenerational inequality are also formulated through the lenses of largely debated theories of *intratemporal* distributive justice.

Another reason why the literature has limited the focus to discounting is the conundrum of combining Pareto efficiency with equality in infinite horizon

²Moreover, some confounding might occur as the concavity of the utility function captures inequality aversion, but might also be interpreted as the intertemporal elasticity of substitution as well as individual or social attitudes towards risk. Extensions to accommodate intergenerational risk and uncertainty exist. The seminal contributions are Epstein and Zin (1989, 1991), which disentangle attitudes toward risk and intergenerational inequalities with an isoelastic formulation. Further contributions include Fleurbaey and Zuber (2013, 2015), Traeger (2014), and Piacquadio (2020). Here, we restrict our analysis to settings without risk.

settings (Diamond, 1965). With minimal regularity conditions, efficiency implies some type of discounting (see Asheim (2010)). The very weak and natural equality requirement adopted—anonymity—means that changing the order of the consumption of generations is a matter of social indifference. Since anonymity, as an impartiality axiom, is at the basis of any concept of inequality aversion, this conflict with efficiency seems to rule out the possibility to formalize a (continuous and complete) inequality-averse social welfare function. As a partial solution, the literature on intergenerational equity has often expressed the problem in terms of ranking utility streams, rather than consumption streams. Then, the utility function, which generally remains unspecified, may capture some notion of inequality aversion. Other contributions instead propose weakenings of anonymity, continuity, or completeness to achieve compatibility with efficiency (Chichilnisky, 1996; Basu and Mitra, 2003; Asheim et al., 2012).

Our contribution clarifies that, while anonymity rules out discounting (for continuous and complete social welfare functions), discounting and efficiency are still compatible with a plethora of inequality attitudes. The key is to impose anonymity in “equivalent time” rather than in “calendar time”.³ Thus, our approach can accommodate the call of both economists and philosophers for more discussion of inequality attitudes for intergenerational justice (Nesje et al., 2023a).

A perspective on inequality attitudes is provided by the *leaky bucket experiment*, popularized by Okun (1975). He suggested to measure how much loss can be justified when transferring a dollar’s worth of income from a richer to a poorer individual. We call this “local inequality aversion” to clarify that this is inherently a local concept, related to the marginal rate of substitution. Moreover, rather than looking at richer and poorer individuals, we identify time intervals that—based on discounting attitudes—have equal social value.

Social welfare functions on income distributions define how inequality atti-

³A recent survey of the criteria that satisfy some form of anonymity in calendar time is Pivato and Fleurbaey (2024). An alternative is to reject efficiency and satisfy anonymity in “equivalent time”. The seminal contribution along these lines is by Zuber and Asheim (2012), which avoids discounting and focuses on a positional account of inequalities by weakening efficiency. We show how our representation result can be generalized to include no discounting, as implied for these cases, in Section 6.

tudes vary with the incomes of the donor and recipient and, in some cases, with the entire distribution. For example, the Atkinson criterion imposes that local inequality aversion only depends on the ratio between the incomes of the donor and recipient (see Atkinson and Brandolini (2015)). In contrast, the Gini criterion imposes that the local inequality aversion depends only on the income ranks of the donor and recipient. More general attitudes are also possible.

Our results clarify how to capture a wide range of inequality attitudes within theories of intergenerational justice.⁴ We also provide novel tools to compare criteria in terms of their social impatience and inequality aversion, and derive conditions similar to the Arrow-Pratt measure of risk aversion.

The remainder of the paper is organized as follows. In Section 2, we present the framework and introduce modular theories of intergenerational justice. Sections 3 and 4 discuss the discounting and aggregating modules, respectively. A central aspect is the formalization of attitudes towards social impatience and inequality aversion and the comparison of criteria based on these attitudes. Section 5 showcases the rich family of theories of intergenerational justice obtained by combining different choices of time-discounting and aggregator functions. The modular theories are axiomatically characterized in Section 6. Concluding remarks are collected in Section 7. Proofs are in the Appendix.

2 Welfare, calendar time, and equivalent time

2.1 Consumption streams

Calendar time is continuous and denoted by $\mathbf{t} \in \mathbf{T} \equiv [0, \infty)$. An intergenerational consumption stream is a bounded function $\mathbf{c} : \mathbf{T} \rightarrow \mathbb{R}_+$ that assigns an instantaneous consumption $\mathbf{c}(\mathbf{t})$ to each time $\mathbf{t} \in \mathbf{T}$.⁵ Let $\mathbf{C}$ denote the set of

⁴An increasing number of contributions aims at identifying inequality attitudes for income distributions (Clark and d’Ambrosio, 2015; Hvidberg et al., 2023; Fehr et al., 2023). Based on this evidence (albeit without a time dimension), we expect our more comprehensive toolkit for inequality attitudes to be needed to match individuals’ views.

⁵The 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. These issues are well known in the study of individuals’ risk preferences:

all consumption streams. A **social welfare function** $W : C \rightarrow \mathbb{R}$ represents a complete and transitive ranking of intergenerational consumption streams. We assume the social welfare function is continuous.

Consumption streams admit representations on the unit interval $I \equiv [0, 1]$. Let $\phi : T \rightarrow I$ be a continuous, bounded and strictly increasing transformation function mapping calendar time to the unit interval. Then, for each consumption stream $c \in C$, the unit-interval representation for ϕ is $y : I \rightarrow \mathbb{R}_+$ such that $y(\phi(t)) = c(t)$ for each $t \in T$ or, with a slight abuse of notation, $y(\phi) = c$. 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.⁶ 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 level of consumption smaller or equal to x . Then, for any pair of consumption streams such that $F_y(x) = F_{y'}(x)$ for each $x \in \mathbb{R}_+$, $W(y(\omega)) = W(y'(\omega))$.

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

A social welfare function W is **modular** if consumption streams can be represented in equivalent time and, thus, there exist a time-discounting function ω and an aggregator function V induced by W . Note that not all social welfare functions are modular. In Section 6, we axiomatically characterize this family of criteria in terms of existence and uniqueness. Importantly, the most common welfare criteria adopted in the literature are modular. Thus, our approach is sufficiently general to include these as special cases.

with unbounded consumption, continuity restricts risk attitudes.

⁶Section 6.2 relaxes several requirements introduced here, including boundedness of the time-discounting function ω .

It immediately follows that a social welfare function W is uniquely identified by the time-discounting function ω and the aggregator function V . Thus, one can construct a wide range of criteria by selecting and combining alternatives within these two modules.

The opposite is also true. Provided a social welfare function W admits an equivalent-time representation, W uniquely pins down the time-discounting function ω and (up to an order-preserving transformation) the aggregator function V .

2.2 Graphical illustration

We now illustrate the above 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, $i = \omega(t) = 1 - e^{-\rho t}$, is shown in Figure 1. The aggregator function in the time-discounted consumption space, V , 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 implied by the curvature of the increasing and concave function f .

Consider a consumption stream in calendar time, c , as illustrated in Figure 1:

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

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

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

with boundaries $\iota = 1 - e^{-\rho\tau}$ and $\iota' = 1 - e^{-\rho\tau'}$. Finally, the (ordinally equivalent) level of social welfare is given by

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

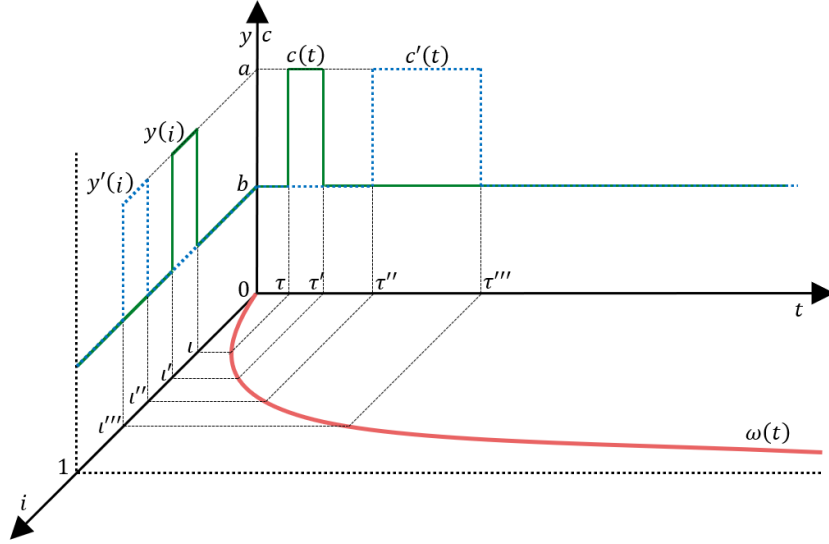


Figure 1: Consumption streams, c (in calendar time) and y (in equivalent time), and permuted consumption streams, c' and y' . Consumption is transformed between calendar time t and equivalent time i by the time-discounting function ω .

Consider next a permutation of the consumption stream in equivalent time, y' ; Figure 1:

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

It follows that $\tau'' = \ln(1/(1-\iota''))/\rho$, $\tau''' = \ln(1/(1-\iota'''))/\rho$, and $\tau' - \tau < \tau''' - \tau''$. This gives the boundaries for the corresponding consumption stream in calendar time, c' :

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

The resulting (ordinally equivalent) level of social welfare is

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

Because the total consumption at $\mathbf{y}$ and $\mathbf{y}'$ coincide (formally, $F_{\mathbf{y}}(\mathbf{x}) = F_{\mathbf{y}'}(\mathbf{x})$ for each $\mathbf{x} \geq 0$), ω and V ensure that W' obtains the same level as W . This means that exponentially-discounted utilitarianism is modular.

The insight that ω leads to anonymity with respect to $\mathbf{y}$ is powerful and generalizes to the most commonly used criteria. Next, we discuss these 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 needs to be continuous, strictly increasing, and satisfy $\omega(0) = 0$ and $\lim_{t \rightarrow \infty} \omega(t) = 1$. 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. The remaining requirements are without loss of generality.⁷

When the time-discounting function ω is linear over an interval $[\underline{t}, \bar{t})$, all periods of time in this interval have the same social value. However, such an interval cannot be the entire time horizon as the time-discounting function would either have to be constant (making informational equivalence impossible) or would not have a bounded image. 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 under minimal regularity conditions. As elaborated on in the Introduction, this tension is well-known in the literature on the aggregation of infinite streams (Diamond, 1965). We further discuss the informational requirement of the time-discounting function and how to relax it in Section 6.

Next, we discuss how the social welfare function can capture different ethical

⁷To clarify, assume that function f is piece-wise continuous, with a jump at $\tau \in \mathbb{T}$. Then, one can define a continuous function $\bar{f}$ such that $\bar{f}(t) = f(t)$ for each $t \in [0, \tau)$ and $\bar{f}(t) = f(t) - \lim_{k \rightarrow \tau^+} f(k)$ for each $t \geq \tau$. Clearly, the anonymity property for equivalent time is unaffected by this change, as it is the case for any rescaling of the time-discounting function.

views about the timing of consumption. This aspect of social preferences is central to the debate on intergenerational justice.⁸

3.2 Social impatience

We now formalize the concept of social impatience and discuss how this provides comparisons of time-discounting functions. We also relate our concept to existing definitions of impatience.

In plain words, we consider disjoint intervals in calendar time and switch their constant consumption. Assume the switch is a matter of social indifference. Then, the social welfare function exhibits **social impatience** *only* if the later interval lasts longer.

The idea is that social impatience pushes an evaluator to consider earlier time periods more important than later ones. As such, an impatient evaluator assigns the same social value to two intervals when the later one lasts longer.

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 $T_1 \equiv [\underline{t}_1, \bar{t}_1)$, the constant consumption $\tilde{c}$ to the interval $T_2 \equiv [\underline{t}_2, \bar{t}_2)$, and the consumption stream defined in c to the remaining times T^- .

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 $\underline{t}_2 \geq \bar{t}_1$, it holds that $\bar{t}_2 - \underline{t}_2 \geq \bar{t}_1 - \underline{t}_1$.

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

Previous definitions of impatience—such as Böhm-Bawerk (1889), Fisher (1930), and Koopmans (1960)—emphasize the preference for advancing rewards and postponing damages. For example, impatience is often taken to mean that interchang-

⁸It remains outside the scope of this paper to address how to compromise between different 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.

ing the consumption levels of two generations results in a strictly preferred plan if the later consumption was initially larger.

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.

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

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

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

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 T_2 , that is, $\bar{t}_2^* \geq \bar{t}_2$ and $\underline{t}_2^* = \underline{t}_2 \geq \underline{t}_1^* = \underline{t}_1 > \underline{t}_1 = \underline{t}_1$.

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*

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

This measure of social impatience resembles the Arrow-Pratt measures of risk aversion.

The comparison by levels of impatience is not new in economics. For example, Prelec (2004) 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 et al. (2023), and Nesje et al. (2023a) clarify that there is large disagreement among experts about social discount rates and impatience.

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, $\bar{t}_{2+\tau}^* \leq \bar{t}_2 + \tau$, and $\underline{t}_{2+\tau}^* = \underline{t}_2 + \tau \geq \bar{t}_{1+\tau}^* = \bar{t}_1 + \tau > \underline{t}_{1+\tau}^* = \underline{t}_1 + \tau$.

Interestingly, decreasing social impatience is only a matter of the time-discounting function. This emphasizes, once again, that the choice of how to discount inter-generational consumption streams is independent of how the discounted consumption streams are aggregated.⁹ For each $\tau \in T$, let $\omega_\tau : T \rightarrow I$ be the truncated time-discounting function from τ on 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 $\tau \in T$*

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

⁹As a technical remark, this follows from the definition of social impatience in terms of two consumption levels, meaning that inequality aversion plays no role in the discounting module. This is different from defining social impatience in terms of marginal changes in consumption from reference levels of consumption which are kept fixed, but might differ between time periods (see Asheim et al. (2001) for a similar definition).

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 and Pollak, 1968; Laibson, 1997; Barro, 1999; Montiel Olea and 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:

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

with $\kappa = 1/(\beta + (1 - \beta)e^{-\rho\tau})$.

Clearly, the special case of exponential discounting emerges when $\beta = 1$.

We represent these cases in Figure 2. The quasi-hyperbolic time-discounting function ω^{qh} and the exponential time-discounting function ω^e have the same discount factor δ . The difference is the additional weight placed on the interval $[0, \tau)$ due to $\beta > 1$ which leads to a kink.

The list of time-discounting functions discussed above is not exhaustive. Other candidates could be considered, such as variations of the hyperbolic time-discounting functions, capturing attitudes to absolute versus proportional discounting, as well as alternative formulations such as the gamma and Weibull time-discounting functions (see Laibson (1997), Weitzman (2001), Read (2001), Jamison and Jamison (2011), and Karp (2017)).

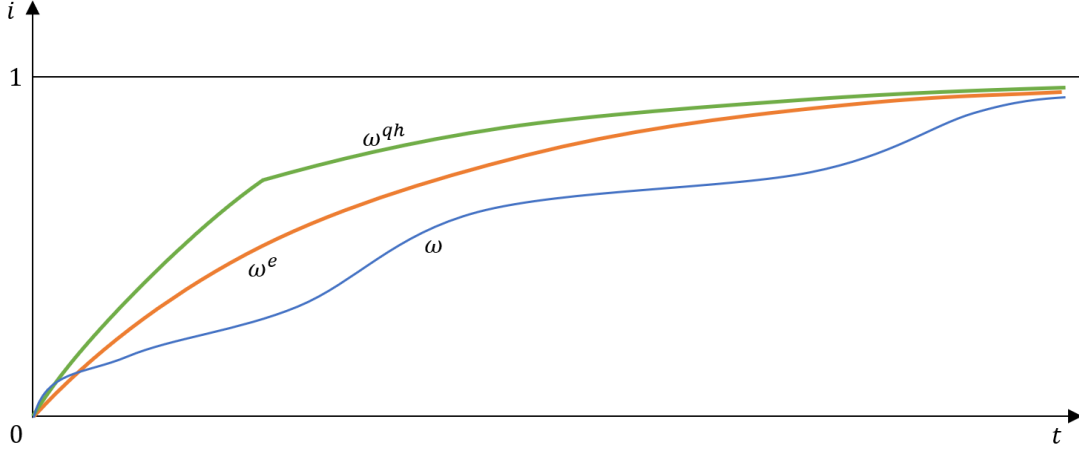


Figure 2: The time-discounting function ω and the special cases of exponential ω^e and quasi-hyperbolic ω^{qh} time-discounting function.

4 The aggregation module

4.1 The aggregator function

Up to now, the only restriction on the aggregator function V is anonymity with respect to the equivalent time representation of consumption streams.

We also impose, as standard in the context of ranking consumption streams, that the social welfare function is sensitive to the consumption of all generations.

A social welfare function W is **Paretian** if for each pair of consumption streams $c, c' \in C$ such that $c(t) \geq c'(t)$ 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 $\int_0^t y(i)di \geq \int_0^t y'(i)di$ for each $t \in I$ and $\int_0^1 y(i)di > \int_0^1 y'(i)di$ implies $V(y) > V(y')$.*

Next, we discuss how the social welfare function can capture aversion to inequalities.

4.2 Inequality aversion

Our novel axiom of inequality aversion imposes inequality aversion in a way that is consistent with time discounting. The Pigou-Dalton transfer principle requires that a progressive transfer—a non-leaky transfer from a richer donor to a poorer recipient—reduces inequality. 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 difference here is to identify equally-valued donor and recipient generations based on discounting attitudes.

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

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

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

implies

$$W(\{(1 - \varepsilon)\hat{c} + \varepsilon\tilde{c}\}_{T_1}, \{(1 - \varepsilon)\tilde{c} + \varepsilon\hat{c}\}_{T_2}, c_{T-}) \geq W(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-}).$$

The idea is the following. Assume that social preferences W are indifferent to changes in the order of a constant consumption $\hat{c}$ and $\tilde{c}$ between disjoint intervals T_1 and T_2 . Then, the evaluator weakly prefers that inequality in consumption is reduced. That is, a progressive transfer of consumption $\varepsilon \cdot |\hat{c} - \tilde{c}|$ between these intervals leads to weakly higher level of social welfare.¹⁰

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 $y, y' \in Y$ such that $\int_0^t y(i) di \geq$*

¹⁰The 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 weight of donors and recipients is by Berg and Piacquadio (2023), who suggest weighting individuals by their deservingness.

$\int_0^t y'(i) di$ for each $t \in I$ and $\int_0^t y(i) di > \int_0^t y'(i) di$ for some $t \in I$ implies $V(y) \geq V(y')$.

Consider an aggregator function V and a time-discounted consumption stream $y \in Y$ such that $y \equiv (\{\hat{y}\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I-})$ with $\hat{y} > \tilde{y}$ and $I_1, I_2 \subset I$ have the same cardinality. We define the **local inequality aversion** between I_1 and I_2 as:

$$LIA_{I_1 I_2}(y) = \frac{\lim_{\varepsilon \rightarrow 0^+} \frac{V(\{\hat{y}\}_{I_1}, \{\tilde{y} + \varepsilon\}_{I_2}, y_{I-})}{\varepsilon}}{\lim_{\varepsilon \rightarrow 0^-} \frac{V(\{\hat{y} + \varepsilon\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I-})}{\varepsilon}} - 1,$$

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 $y \equiv (\{\hat{y}\}_{I_1}, \{\tilde{y}\}_{I_2}, y_{I-}) \in Y$ the local inequality aversion between I_1 and I_2 is non-negative, that is, $LIA_{I_1 I_2}(y) \geq 0$.*

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 $T_2 \equiv [t, t + \Delta t)$ and $T_1 \equiv [0, \Delta t)$ when Δt goes to zero:

$$\frac{1}{t} \ln (ILIA_t(y)).$$

Intuitively, if the social welfare function is inequality averse and consumption increases over time, $\ln (ILIA_t(y)) \geq 0$.

Next, we introduce a ranking of social welfare functions in terms of the level of inequality aversion they exhibit. For each pair of consumption streams $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-}), (\{\hat{c}'\}_{T_1}, \{\tilde{c}'\}_{T_2}, c_{T-}) \in C$, we say that the latter is obtained 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 $c \equiv (\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-})$,

$\mathbf{c}' \equiv (\{\tilde{\mathbf{c}}\}_{T_1}, \{\hat{\mathbf{c}}\}_{T_2}, \mathbf{c}_{T-})$, and $\mathbf{c}''$ is obtained from $\mathbf{c}$ through a progressive leaky transfer, it holds that: *i)* $W^*(\mathbf{c}) = W^*(\mathbf{c}') = W(\mathbf{c}) = W(\mathbf{c}')$, and *ii)* $W(\mathbf{c}'') \geq W(\mathbf{c})$ implies $W^*(\mathbf{c}'') \geq W^*(\mathbf{c})$.

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 $LIA_{I_1, I_2}^*(\mathbf{y}) \geq LIA_{I_1, I_2}(\mathbf{y})$ for each $\mathbf{y} \in Y$.*

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 and generalized utilitarianism.¹¹ 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 consumption levels. In fact, an aggregator function is more inequality averse if and only if the transformation function is more concave.¹²

¹¹Atkinson (1970) has adopted this family for his seminal paper on inequality measurement. A recent generalization to multidimensional settings is Piacquadio (2017).

¹²A 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 .

A special case of the 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 of two distributions is invariant to their rescaling. Formally, the family of Atkinson aggregator functions is given by

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

where $\eta \geq 0$ 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 (see Drupp et al. (2018)).

Another aggregator function is the weighted sum of “ordered” consumptions: this family draws on the symmetry with choices under risk and reinterprets the risk preferences characterized by Yaari (1987, 1988).¹³ The idea is to proceed in three steps. First, reorder the levels of consumption in increasing order so that the consumption at (equivalent time) i is at least as large as that at j , that is, $y_i \geq y_j$. Next, define a weighting function $g : [0, 1) \rightarrow (0, \infty)$ that is non-increasing in i . Finally, take the sum of the weighted consumption stream. 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:

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

where $r(i)$ is the function that associates the rank to each $i \in I$.

An important special case is that of the **single-parameter Gini aggregator functions** (Donaldson and Weymark, 1980, 1983; Yitzhaki, 1983; Aaberge, 2001). In this case, the weighting function g follows from:

¹³This family is often referred to as “generalized Gini social welfare functions” and emphasize the relationship with the Gini inequality measure (see d’Aspremont and Gevers (2002)). We come back later to the relationship between social welfare and inequality.

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

where $\gamma \geq 0$ is a parameter measuring inequality aversion to differences in ranks. More precisely, γ determines the speed at which the weights decrease with the rank. The larger γ , the more inequality averse is the aggregation. 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 obtain. When f is the identity function, the rank-weighted criteria obtain.

The parametric forms above lead to:

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

where $\gamma \geq 0$ and $\eta \geq 0$ jointly determine social inequality aversion. Interestingly, γ and η influence inequality aversion in different ways. In fact, we can compare 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 3. On the two axis, 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, the rank-weighted aggregator functions have piece-wise 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

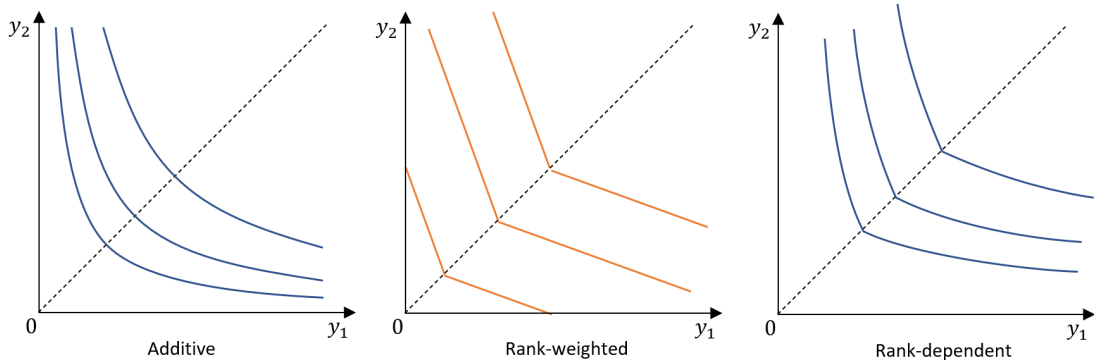


Figure 3: The special cases of the parameterized additive V^a , rank-weighted V^{rw} and rank-dependent V^{rd} aggregator functions.

can avoid that the indifference curves intersect the axis. The rank-dependent aggregator functions are more flexible and allow combining the different sensitivities to inequality of the other criteria.

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), and Chew and Nishimura (1992)). Further generalizations of rank-dependent aggregator functions have also been developed (e.g., Green and Jullien (1988)).

In the next section, we combine the insights from discounting and inequality aversion to illustrate the range of theories of intergenerational justice our modular approach can accommodate.

5 Social welfare functions

In this section, we combine the discounting and inequality aversion modules to uncover classes of theories of intergenerational justice.

Before that, we investigate the general properties of modular theories of intergenerational justice. First, we present the *efficiency-equity* representation of the criteria, disentangling discounting attitudes from the trade-off between efficiency and equity. Second, we identify the generalized “Ramsey rule,” clarifying how to

use our criteria to find sufficient statistics for cost-benefit analysis.

5.1 Efficiency-equity representation

The idea is that the level of social welfare can be expressed as a product of two components: the mean of the discounted consumption, measuring the efficiency of the consumption stream, and the equality of the discounted consumption, measuring equity.

Let $\mu : Y \rightarrow \mathbb{R}_+$ be the mean function. Then, the measure of **efficiency** can be expressed as $\mu(y)$ or, for a given time-discounting function ω , as $\mu(\omega(c))$. Importantly, the efficiency measure does not depend on the way consumption in equivalent time is aggregated.

The measure of **equity** is a function $E : Y \rightarrow [0, 1]$ implicitly defined as follows:

$$V(y) = V(k) \text{ where } k \equiv \mu(y) \cdot E(y).$$

By definition, the equity measure E only depends on the aggregation in equivalent time V . Moreover, when V is scale invariant—as for the special cases presented above— $1 - E$ is a relative measure of inequality (see Blackorby and Donaldson (1978)).

Finally, since $W(c) = V(\omega(c))$, we can now express the social welfare function in terms of the **efficiency-equity representation** (the proof is omitted and directly follows from Atkinson (1970)).

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

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

where $y(\omega) = c$.

Thus, social welfare can be decomposed into two main aspects: the time-discounting function and the efficiency-equity trade-off.

5.2 Generalized Ramsey rule

The question addressed next is that of social discounting with consumption 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 $\Delta x = e^{-\xi t} \Delta c$.¹⁴ Assume that the social welfare function is differentiable at the status quo consumption stream $c \in C$. Assume moreover 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 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

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

The intuition is simple. When two intervals have different duration 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 $\text{ILIA}_t(y)$ as the limit of the 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. Then,*

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

¹⁴To 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 $\Delta t \rightarrow 0$.

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 $\ln(\omega'(0)/\omega'(t)) > 0$. 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 $\ln(\text{ILIA}_t(y)) \geq 0$.

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 simplifies 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):

$$\text{SDR}_t^{ea} = \rho + \eta g_t^c.$$

We highlight the social discount rate for other criteria below.

5.3 Modular theories of intergenerational justice

By freely combining the 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 and quasi-hyperbolic ω^{qh} .

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

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

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

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

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

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

$$E^{rd}(y) \equiv \frac{\left(\int_0^1 (1 - r(i))^\gamma (y_i)^{1-\eta} di \right)^{\frac{1}{1-\eta}}}{\mu(y)}.$$

When $\gamma = 0$, $1 - E^{rd}(y)$ is a member of the Atkinson family of inequality measures. When $\eta = 0$, $1 - E^{rd}(y)$ 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:

$$SDR_t^{erd} = \rho + \eta g_t^c + \gamma g_t^r.$$

It consists of the pure time preferences ρ from the discounting module, the inequality aversion η towards consumption growth g_t^c , and the inequality aversion γ towards rank growth g_t^r , implicitly defined by setting $e^{tg_t^r} \equiv (1 - r_0)/(1 - r_t)$.

6 Characterization

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. 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.

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

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

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 $W(\underline{c}_{T_1}, \bar{c}_{T_2}) = W(\bar{c}_{T_2}, \underline{c}_{T_1})$ for some $\underline{c}, \bar{c} \in \mathbb{R}_+$. *Minimal time split* requires that indifference holds for all levels of consumption. While 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 stakes are larger. Habit formation is also ruled out.

The following axiom strengthens *minimal time split*. It imposes that one can split any interval of calendar time into “equally valuable” sub-intervals in the sense of indifference to permutations.¹⁵

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

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

This stronger axiom also excludes 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

¹⁵Our 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 sub-intervals is unaffected.

utility of each periods' consumption, weighted by a discount factor that depends on past consumption.¹⁶

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, thus, that the time-discounting function is somewhere flat. It thereby justifies our informational-equivalent requirement from Section 3.

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.¹⁷ We first introduce a more general class of social welfare functions.

¹⁶Formally, the criterion can be expressed as $W(c) = \int_0^\infty \psi(c, t)u(c(t))dt$, where $\psi(c, t) = e^{-\int_0^t \beta(c(s))ds}$, 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.

¹⁷We 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 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 continuous and complete social welfare functions.

Let the **generalized time-discounting function** $\bar{\omega} : T \rightarrow \mathbb{R}_+$ 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 leaves social welfare unchanged. The key gain in generality is to include anonymity in calendar time as a special case, which emerges when $\bar{\omega}$ is the identity function.

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 and Mitra, 2003; Asheim et al., 2022). Assume that the consumption stream $\mathbf{c}$ assigns to the interval $T_1 \subset T$ 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 $\mathbf{c} = (\{\hat{c}\}_{T_1}, \mathbf{c}_{T-}), \mathbf{c}' = (\{\tilde{c}\}_{T_1}, \mathbf{c}_{T-}) \in C$ 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 is that a 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 $\mathbf{c} = (\{\hat{c}\}_{T_1}, \mathbf{c}_{T-}), \mathbf{c}' = (\{\tilde{c}\}_{T_1}, \mathbf{c}_{T-}) \in C$ with $\hat{c} > \tilde{c}$ and $\liminf_t \mathbf{c}(t) > \tilde{c}$, it holds that $\bar{W}(\mathbf{c}) > \bar{W}(\mathbf{c}')$.

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 substituted for liminf-restricted dominance.*

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

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

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

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

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

In contrast to the previous literature, one can easily generalize the maximin and rank-discounted utilitarian criteria to accommodate any generalized time-discounting function.¹⁸

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 $T_1 \cup T_2 = T$ and each pair of consumption streams

¹⁸While one can also demand first- and second order stochastic dominance in generalized equivalent time, first-order stochastic dominance will no longer correspond to Pareto efficiency in calendar time.

$c, c' \in C$ with $c(t) = c'(t)$ for each $t \in T_1$, $W(c) \geq W(c')$ if and only if $W(c_{T_2}) \geq W(c'_{T_2})$.

Beyond exponentially-discounted utilitarianism, another criterion that satisfies *stationarity* is the Uzawa utility function discussed above. The latter, however, 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. 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.*

6.4 Further axioms: time consistency and time invariance

Next, we follow Halevy (2015) and introduce two further 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 than 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*.¹⁹ To formalize these axioms, we first extend our domain.

¹⁹The 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 concep-

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

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 $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$ with $\mathbf{c}(t) = \mathbf{c}'(t)$ for each $t \in [0, \tau)$, $W(\mathbf{c}) \geq W(\mathbf{c}')$ if and only if $W_{\tau}(\mathbf{c}_{\tau}) \geq W_{\tau}(\mathbf{c}'_{\tau})$, where $\mathbf{c}(t) = \mathbf{c}_{\tau}(t)$ and $\mathbf{c}'(t) = \mathbf{c}'_{\tau}(t)$ for each $t \geq \tau$.

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 $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$, $W(\mathbf{c}) \geq W(\mathbf{c}')$ if and only if $W_{\tau}(\mathbf{c}_{\tau}) \geq W_{\tau}(\mathbf{c}'_{\tau})$, where $\mathbf{c}(t) = \mathbf{c}_{\tau}(t + \tau)$ and $\mathbf{c}'(t) = \mathbf{c}'_{\tau}(t + \tau)$ 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 utilitarian social welfare function. However, stationarity 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 aggregator functions could change over time. Time consistency and 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.²⁰

tually 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.

²⁰Millner 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 since the past consumption can be taken into account when assessing alternatives. As an implication, the additive separability of the aggregator function does not follow.

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 $\omega_{\underline{t}}(\mathbf{t}) = 1 - (1 - \omega(\mathbf{t})) / (1 - \omega(\underline{\mathbf{t}}))$ for each $\underline{t} \in \mathbf{T}$ and each $\mathbf{t} \geq \underline{\mathbf{t}}$; and ii) the aggregator function is additive and satisfies $V_{\underline{i}} = V^a$, with $\underline{i} = \omega(\underline{\mathbf{t}})$, for each $\underline{t} \in \mathbf{T}$.

In contrast to stationarity and time consistency, time invariance imposes no restrictions on 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).²¹

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 $\underline{t}, \underline{t}' \in \mathbf{T}$.

7 Conclusion

Policies affect present and future generations differently. Any comparison of benefits and losses across generations—for instance, to identify optimal policy—depend on the theory of intergenerational justice adopted. Without choosing any specific theory, economics has the role of operationalizing these theories and showing their policy implications.

When comparing benefits and losses across generations, two aspects of intergenerational justice are central. These are the attitudes toward earlier versus later generations—i.e., intergenerational discounting—and the attitudes toward better-off versus worse-off generations—i.e., intergenerational inequality. In this paper, we show how to combine these attitudes, what theories of intergenerational justice these attitudes identify, and how these relate to existing commonly used criteria. Our approach includes as a special case the workhorse model of intergen-

²¹As 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.”

erational justice—exponentially-discounted utilitarianism—as well as many more novel criteria. These criteria emerge by combining discounting and inequality attitudes that emerged separately in the respective literatures. Our study therefore enriches the economic toolbox of theories of intergenerational justice.

Our main result is to propose and axiomatically justify theories of intergenerational justice that disentangle discounting and inequality. A distinctive feature is that this class of theories builds on and highlights a variety of attitudes towards discounting and inequality that are orthogonal to each other. Their combination is more likely to fit the wide range of views held by individuals. Another novel contribution of our approach is to explore intergenerational justice through the lenses of largely debated theories of *intratemporal* distributional justice. While the representation is very general and flexible, the parametric specifications of these theories remain suitable for investigation of optimal policies.

Estimation, application, and generalization of the modular theories also open many avenues for future research. A natural first step is to learn the intergenerational discounting and inequality attitudes behind individuals’ ethical views. Another further step is the application of these criteria to resource economics models, such as integrated models of the climate and economy, to detail optimal policy. The framework can be generalized along several dimensions. For example, population growth could be readily accommodated through a “population-weighted” time-discounting function: beyond discounting, the relative weight of any two intervals would be proportional to their population. Another important dimension, such as intratemporal inequality, could be accounted for by substituting the instantaneous consumption $c(t)$ with the equally-distributed equivalent consumption of the individuals living at t (or alternatively, those born at t). Finally, risk can be accommodated 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.
- Arrow, K. J. (1965), *Aspects of the Theory of Risk Bearing*, Helsinki: Yrjö Jahnssonin Saatio.
- Asheim, G. B. (2010), Intergenerational equity. *Annual Review of Economics*, **2**(1), 197–222.
- Asheim, G. B., Buchholz, W., and Tungodden, B. (2001), Justifying sustainability. *Journal of Environmental Economics and Management*, **41**, 252–268.
- Asheim, G. B., Kamaga, K., and Zuber, S. (2022), Maximal sensitivity under strong anonymity. *Journal of Mathematical Economics*, **103**, 102768.
- Asheim, G. B., Mitra, T., and Tungodden, B. (2012), Sustainable recursive social welfare functions. *Economic Theory*, **49**, 267–292.
- d’Aspremont, C., and Gevers, L. (2002), Social welfare functionals and interpersonal comparability. In *Handbook of Social Choice and Welfare 1* [K. J. Arrow, A. K. Sen, and K. Suzumura (eds.)], pp. 541–459, Elsevier.
- Atkinson, A. B. (1970), On the measurement of inequality. *Journal of Economic Theory*, **2**(3), 244–263.
- Atkinson, A. B., and Brandolini, A. (2015), Unveiling the ethics behind inequality measurement: Dalton’s contribution to economics. *Economic Journal*, **125**(583), 209–234.
- Barro, R. J. (1999), Ramsey meets Laibson in the neoclassical growth model. *Quarterly Journal of Economics*, **114**(4), 1125–1152.
- Basu, K., and Mitra, T. (2003). Aggregating infinite utility streams with intergenerational equity: the impossibility of being Paretian. *Econometrica*, **71**(5), 1557–1563.
- Berg, K. and Piacquadio, P. G. (2023), Fairness and Paretian social welfare functions. *Mimeo*.
- Blackorby, C., and 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*, New York: Stechert (Translation).

- Chew, S. (1983), A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais Paradox. *Econometrica*, **51**, 1065–1092.
- Chew, S., and 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., and 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 Welfare*, **13**(2), 231–257.
- Clark, A. E., and d’Ambrosio, C. (2015), Attitudes to income inequality: Experimental and survey evidence. In *Handbook of Income Distribution* (Vol. **2**, pp. 1147–1208). Elsevier.
- Dasgupta, P. (2008), Discounting climate change. *Journal of Risk and Uncertainty*, **37**, 141–169.
- Dasgupta, P. (2021), *The Economics of Biodiversity: The Dasgupta Review*, Cambridge and New York: Cambridge University Press.
- Del Campo, S., Anthoff, D., and Kornek, U. (2023), Inequality aversion for climate policy. *Review of Environmental Economics and Policy*, forthcoming.
- Diamond, P. A. (1965), The evaluation of infinite utility streams. *Econometrica*, **33**(1), 170–177.
- Donaldson, D., and Donaldson, J. A. (1980), A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, **22**(1), 67–86.
- Donaldson, D., and 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., and Nesje, F. (2018), Discounting disentangled. *American Economic Journal: Economic Policy* **10**(4), 109–134.
- Epstein, L. G. (1983), Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, **31**(1), 133–152.
- Epstein, L. G., and 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., and 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., and Senn, J. (2023), The fundamental properties, stability and predictive power of distributional preferences. *UZH working paper series*.
- Fishburn, P. C. (1983), Transitive measurable utility. *Journal of Economic Theory*, **31**(2), 293–317.
- Fisher, I. (1930), *The Theory of Interest*, New York: Macmillan.
- Fleurbaey, M. and Michel, P. (2003), Intertemporal equity and the extension of the Ramsey criterion. *Journal of Mathematical Economics*, **39**(7), 777–802.
- Fleurbaey, M. and Zuber, S. (2013), Inequality aversion and separability in social risk evaluation. *Economic Theory*, **54**, 675–692.
- Fleurbaey, M. and Zuber, S. (2015), Discounting, risk and inequality: A general approach. *Journal of Public Economics*, **128**, 34–49.
- Green, J. R. and Jullien, B. (1988), Ordinal independence in nonlinear utility theory. *Journal of Risk and Uncertainty*, **1**(4), 355–387.
- Gollier, C., van der Ploeg, F., and Zheng, J. (2023), The discounting premium puzzle: survey evidence from professional economists. *Journal of Environmental Economics and Management*, **122**, 102882.
- 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., and Stantcheva, S. (2023), Social positions and fairness views on inequality. *Review of Economic Studies*, **90**(6), 3083–3118.
- Jamison, D. T., and 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. M., and Wakker, P. P. (2023), The deceptive beauty of monotonicity, and the million-dollar question: row-first or column-first aggregation? *Mimeo*.
- Machina, M. J. (1982), “Expected utility” analysis without the independence axiom. *Econometrica*, **50**(2), 277–323.
- Montiel Olea, J. L., and Strzalecki, T. (2014), Axiomatization and measurement of quasi-hyperbolic discounting. *The Quarterly Journal of Economics*, **129**(3), 1449–1499.
- Millner, A., and Heal, G. (2018), Time consistency and time invariance in collective intertemporal choice. *Journal of Economic Theory*, **176**, 158–169.
- Millner, A. and Heal, G. (2023), Choosing the future: Markets, ethics, and rapprochement in social discounting. *Journal of Economic Literature*, **45**(3), 686–702.
- Nesje, F., Drupp, M. A., Freeman, M. C., and Groom, B. (2023a), 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., and Groom, B. (2023b), Sense and sensitivity: An argument against reporting multiple Net Present Values. *Mimeo*.
- Nordhaus, W. D. (2007), A review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, **61**(3), 1037–1087.
- Okun, A. M. (2015), *Equality and Efficiency: The Big Tradeoff*. Brookings Institution Press.
- Phelps, E. S., and 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. and Fleurbaey, M. (2024), Intergenerational equity and infinite-population ethics: a survey. *Mimeo*.
- 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. and 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.
- 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 and New York: 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), 628–652.
- Weitzman, M. L. (2001), Gamma discounting. *American Economic Review*, **91**(1), 260–271.
- 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. and 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 $\tau \in T$. Then, there exists $\varepsilon > 0$ and $\tau^* > \tau$ such that

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

Now consider a consumption stream such that $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-})$ with $\hat{c} \neq \tilde{c}$ and $T_1 \equiv [\tau - \varepsilon, \tau^*)$ and $T_2 \equiv [\tau^*, \tau + \varepsilon]$.

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

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

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 $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-})$, $(\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T-}) \in C$ 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 $\hat{c} \neq \tilde{c}$, $\underline{t}_2 \geq \bar{t}_1$, 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 $\tau \in T$ such that ζ is strictly convex at $\omega(\tau)$.

Consider the interval $T_1 \equiv [\tau - \varepsilon, \tau)$ for $\varepsilon > 0$. For a sufficiently small ε , there exist intervals $T_2 \equiv [\tau, \tau + k)$ and $T_2^* \equiv [\tau, \tau + k^*)$ 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 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 $(\{\hat{c}\}_{T_1}, \{\tilde{c}\}_{T_2}, c_{T-}), (\{\tilde{c}\}_{T_1}, \{\hat{c}\}_{T_2}, c_{T-}) \in C$ 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 (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 $t \geq \bar{t}_1^* = \bar{t}_1$

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

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

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

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 $\bar{c} \geq 0$ 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 immediately follows 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 $c(t) \geq c'(t)$ 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 $\varepsilon \in (0, 1)$,

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

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 $\mathbf{y} \in Y$ such that $\mathbf{y} \equiv (\{\hat{\mathbf{y}}\}_{I_1}, \{\tilde{\mathbf{y}}\}_{I_2}, \mathbf{y}_{I^-})$ with $\hat{\mathbf{y}} > \tilde{\mathbf{y}}$ and $I_1, I_2 \subset I$ have the same cardinality.

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 $\mathbf{y}_1, \mathbf{y}_2 \in Y$ such that $\mathbf{y}_1$ is majorized by $\mathbf{y}_2$, we have $V(\mathbf{y}_1) \geq V(\mathbf{y}_2)$. Because V is Schur concave, W exhibits inequality aversion if and only if for each $\mathbf{y} \equiv (\{\hat{\mathbf{y}}\}_{I_1}, \{\tilde{\mathbf{y}}\}_{I_2}, \mathbf{y}_{I^-}) \in Y$, $LIA_{I_1 I_2}(\mathbf{y}) \geq 0$.

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 $LIA_{I_1, I_2}^*(\mathbf{y}) < LIA_{I_1, I_2}(\mathbf{y})$. 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 $\text{LIA}_{I_1, I_2}^*(y) \leq \text{LIA}_{I_1, I_2}(y)$. ■

A.10 Proof of Theorem 1

($\Rightarrow$) Let $\bar{T} \subseteq T$. Construct a partition of $\bar{T}$ 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 $T_1 \cup T_2 = \bar{T}$ 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 $c \equiv (\underline{c}_{T_1}, \bar{c}_{T_2}, c_{T/\bar{T}})$, $c' \equiv (\bar{c}_{T_2}, \underline{c}_{T_1}, c_{T/\bar{T}}) \in C$, 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 $m \in M_k \equiv \{1, 2, \dots, 2^k\}$, 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) $W(\underline{c}_{T^m}, \bar{c}_{T^{m'}}, c_{T/\{T^m \cup T^{m'}\}}) = W(\bar{c}_{T_2}, \underline{c}_{T_1}, c_{T/\bar{T}})$ for each pair $m, m' \in M_k$ and for each consumption stream $c \in C$ with a pair of consumption levels $\underline{c}, \bar{c} \in \mathbb{R}_+$. 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 $c = y(\omega)$. 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, thus, 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 $t' > t$. 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 $t = t'$.

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

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 directly extends 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 exists a time-discounting function ω and an aggregator function V such that $W(\mathbf{c}) = V(\mathbf{y})$ for $\mathbf{y}(\omega) = \mathbf{c}$.

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 $s, t \in \mathbb{T}$ with $s \leq t$, 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 $T_1 \cup T_2$ must equal that of T'_2 with respect to $T'_1 \cup T'_2$. 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 general solution $1 - \omega(t) = (1 - \omega(1))^t$ for any rational $t \in \mathbb{T}$. Since ω is continuous and strictly increasing,

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

where $\rho \equiv -\ln(1 - \omega(1))$ and, 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 the consumption streams in calendar time. Thus, the aggregator function V satisfies **separable future**, that is, for each partition of equivalent time $I_1, I_2 \subset I$ with $I_1 \cup I_2 = I$ and each pair of consumption streams in equivalent time $\hat{y}, \tilde{y} \in Y$, if $V(\hat{y}_{I_1}, \hat{y}_{I_2}) \geq V(\hat{y}_{I_1}, \tilde{y}_{I_2})$ then $V(\tilde{y}_{I_1}, \hat{y}_{I_2}) \geq V(\tilde{y}_{I_1}, \tilde{y}_{I_2})$.

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 exists $\tau > 0$ and a pair of consumption streams $c, c' \in C$ with $c(t) = c'(t)$ for each $t \in [0, \tau)$ such that $W(c) \geq W(c')$ and $W_\tau(c_\tau) < W_\tau(c'_\tau)$, where $c(t) = c_\tau(t)$ and $c'(t) = c'_\tau(t)$ for each $t \geq \tau$. Let $y, y', y_\iota, y'_\iota \in Y$ be such that $y(\omega) = c$, $y'(\omega) = c'$, $y_\iota(\omega_\tau) = c_\tau$, $y'_\iota(\omega_\tau) = c'_\tau$, 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),

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

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 $t \in [0, \tau]$,

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

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

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

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

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

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 $[\tau, \infty)$ 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: $t_0 = \tau$; $t_K = \infty$; 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 $V(y) \geq V(y')$.

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, $V(y) \geq V(y')$ if and only if $V_{i_k}(y_{i_k}) \geq V_{i_k}(y'_{i_k})$ 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 $t \geq t_k$.

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. ■