

Bruers, Stijn

Working Paper

An intergenerational social welfare function that represents neutral-range utilitarianism

Suggested Citation: Bruers, Stijn (2026) : An intergenerational social welfare function that represents neutral-range utilitarianism, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/344076>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

An intergenerational social welfare function that represents neutral-range utilitarianism

Stijn Bruers

KULeuven, Department of Economics

Naamsestraat 69 - box 3565

3000 Leuven, Belgium

ORCID 0000-0003-0906-4468

Stijn.bruers@gmail.com

October 2026

Abstract

An intergenerational social welfare function is presented for a population ethical theory that generalizes critical-level utilitarianism and necessitarianism. Utilities of future people are discounted when these utilities lie within a neutral range. This neutral-range utilitarianism mitigates the repugnant and sadistic conclusions in population ethics, but entails a time inconsistency: the ranking of options available to the social planner can change depending on when decisions are made. A modified social welfare function is proven to be top choice dynamically consistent. As applications, specific distributions of utilities of future people are used to derive closed form expressions of the Ramsey rule for the social discount rate and the maximum bandwidth of the neutral range.

Keywords: social welfare function; utilitarianism; time discounting; time inconsistency; population ethics; repugnant conclusion

JEL: D60, I30

Introduction

Ramsey (1928) started the study of intergenerational welfare economics: how to optimize decisions such as investments and savings when taking the interests of future generations into account (Dasgupta, 2020). A classical utilitarian calculus was assumed, with a social welfare function given as the sum of utilities of all present and future people. As this sum can grow arbitrarily large for

arbitrarily long futures, making the calculus intractable, a discount factor was introduced to weigh the utilities (Samuelson, 1937). Utilities in the more distant future are discounted more, by weighting them with a smaller discount factor. This keeps the social welfare function bounded. The use of an intertemporal social welfare function with discounted utilities became mainstream in the sixties and seventies, to study optimal growth (Cass, 1965; Koopmans, 1965), optimal taxation (Pestieau, 1974) and cost-benefit analysis (Feldstein, 1964). A few decades later, the intertemporal social welfare function became crucial in optimal climate policy (Nordhaus, 1992; Stern, 2008). But the use of a discount factor, and especially its value, remains controversial (Frederick, Loewenstein & O'donoghue, 2002; Stern, 2008; Greaves, 2017; Drupp e.a. 2018; Mogensen, 2022).

How future generations are included in an intergenerational social welfare function, is a normative question. Population ethics, an area in normative philosophy that started in the seventies and eighties, studies which situations are better when different situations have different existing people (Narveson, 1973; Parfit 1984). There are different population ethical theories, such as total utilitarianism (the best situation is the one that has the highest sum of utilities of everyone who exists and will exist) and average utilitarianism (the best situation is the one that has the highest average utility). Policies can influence who will exist in the future and how many people will exist in the future. Hence, an intergenerational social welfare function represents a population ethical theory.

The most famous problem in population ethics, is the repugnant conclusion (Parfit, 1984; Arrhenius, 2003), which is very difficult to avoid (Blackorby, Bossert & Donaldson, 2003; Arrhenius, 2011; Spears & Budolfson, 2019; Zuber e.a. 2021). Many population ethical theories, including total and average utilitarianism, entail that it can be good to drastically reduce the average welfare of happy people by adding a huge population of people who have lives barely worth living. The classical utilitarian intergenerational social welfare function also faces the repugnant conclusion: current generations may have to drastically reduce welfare or utility in order to bring into existence huge populations of future generations who have lives with positive but very small utilities. This repugnant conclusion can be avoided by choosing an extremely high time discount rate, which means that the time discount factor becomes zero extremely fast. With a sufficiently high time discount rate, the utilities of future generations become negligible so fast that their utilities cannot compensate for the loss of utility of the current generations. But now we get a sadistic conclusion (Arrhenius, 2000): it might not be bad to add future people who have negative utilities, when these utilities are totally discounted. To escape the repugnant conclusion, Blackorby, Bossert and Donaldson (1995; 1997) presented another social welfare function that represents critical-level utilitarianism. Instead of discounting utilities by multiplying them with a discount factor, utilities are reduced by subtracting them with a positive

term called critical level. This theory avoids the repugnant conclusion when the critical level is chosen sufficiently high, but it entails another sadistic conclusion: increasing someone's negative utility to a positive value by adding a huge number of extra people who have positive utilities slightly below the critical level, may be bad.

This article presents a new population ethical theory, called neutral-range utilitarianism, and derives its corresponding intergenerational social welfare function. With appropriately chosen parameters, this population ethical theory avoids (or at least mitigates) the repugnant and sadistic conclusions. In some limiting cases, when the parameters in this social welfare function take specific values, we get the traditional theories such as total utilitarianism, time-discounted utilitarianism and critical-level utilitarianism. As with many social welfare functions that use discount factors, the major problem of this neutral-range utilitarian social welfare function, is its time inconsistency (Strotz, 1956; Bergström, 1966; Phelps & Pollak, 1968; Simaan & Cruz, 1973; Kydland & Prescott, 1977).

After a general formulation of the neutral-range utilitarian social welfare function in the next section, interpretations and justifications of this welfare function are given. Next, the Ramsey rule, that connects the social discount rate with the utility discount rate, is derived and calculated assuming a logistic distribution of utilities. Finally, a possible approach or solution to the time inconsistency problem is presented and an argument is made for the optimal size of the neutral range.

Formulation: the intergenerational social welfare function of neutral-range utilitarianism

Assume a social planner having to make a decision at time $t = 0$. N_0 is the number of individuals who exist (or are born) at this initial time, and for $t > 0$, the number of individuals born at that time is N_t . The lifetime utility of individual $i \in \{1, \dots, N_t\}$ born at time t is $u_{it} \in \mathbb{R}$, with u_{i0} the lifetime utility of an individual who already exists (or is born) at time 0.¹ The lifetime utility can be expressed in terms of for example the WELLBY (well-being adjusted life years; Layard & Oparina, 2021) or the QALE (quality-adjusted life-expectancy at birth; Gold et al. 1996; Jia, Zack & Thompson, 2011; Collins, 2013). Note that this lifetime utility u_{it} of a person born at time t can be written as a time integral over a (possibly time-discounted) momentaneous utility (flow utility) v_{it} , at time t' :

¹ In case of uncertainty, expected values for lifetime utilities and social welfare functions can be used.

$$u_{it} = \int_t^{t+\delta_i} d_{it'} v_{it'} dt',$$

with δ_i the lifespan of individual i and $d_{it'} \in [0,1]$ an individual utility discount factor that captures the individual's personal pure time preference.² Using this time integral means that the lifetime utility u_{it} of a person born at time t can depend on decisions made at a later time $t' \in [t, t + \delta_i]$.

The traditional discounted utilitarian social welfare function (Samuelson, 1937), in discrete time, is given by

$$W_{DU}(\{u_{it}\}) = \sum_{t=0}^{\infty} \sum_{i=1}^{N_t} \beta^t u_{it},$$

with $\beta^t \in [0,1]$ the social utility discount factor and $\rho = \frac{1}{\beta} - 1$ the social utility discount rate (also called rate of pure time preference).³

Define the neutral-range utilitarian social welfare function, in discrete time, as:

$$W_{NRU}(\{u_{it}\}) = \sum_{t=0}^{\infty} \sum_{i=1}^{N_t} (u_{it} - (1 - \beta^t)C(u_{it})),$$

and in continuous time:

² An arguably better approach is to consider u_{it} as a discounted lifetime utility for a person i existing (instead of being born) at time t . The discount factor $d_{it'}$ can be interpreted as a measure of psychological connectedness between a momentaneous mind i at time t with a future momentaneous mind at time t' . It measures how much the individual i at time t identifies him/herself with a future self at time t' . This discount factor represents the concept of personal identity over time (e.g. Parfit, 1984): the higher the discount factor, the more the momentaneous selves at times t and t' are psychologically connected with each other. When the discount factor is at its maximum value of 1, the two momentaneous selves fully belong to the same person. A zero discount factor means there is no personal identity between the two momentaneous minds. This approach allows to avoid the replaceability problem of utilitarianism (Singer, 1979; Pluhar, 1982; Parfit, 1984): a person with a constant positive flow utility and a lifespan of 100 years contributes more to the social welfare function than two people with the same flow utility but each having a lifespan of 50 years. Hence, killing a person at age 50 and replacing that person with an equally happy person that lives another 50 years, is worse than letting the first person live 100 years and not bringing into existence the second person. For example, when $v_{1t'} = v_{2t'}$ and $d_{1tt'} = d_{2tt'}$ are constant, then

$$\int_0^{100} \int_0^{100} d_{1tt'} v_{1t'} dt' dt > \int_0^{50} \int_0^{50} d_{1tt'} v_{1t'} dt' dt + \int_{50}^{100} \int_{50}^{100} d_{2tt'} v_{2t'} dt' dt.$$

³ Define $0^0 \equiv 1$.

$$W_{NRU}(\{u_{it}\}) = \int_{t=0}^{\infty} \sum_{i=1}^{N_t} (u_{it} - (1 - e^{-\rho t})C(u_{it}))dt,$$

with $\rho \geq 0$ the social utility discount rate, and $C(u_{it}) \in [u_{min}, u_{max}]$ the neutral-range or cut function given by:

$$C(u_{it}) = \min(\max(u_{min}, u_{it}), u_{max}) = \begin{cases} u_{max} & \text{if } u_{it} > u_{max}, \\ u_{it} & \text{if } u_{it} \in [u_{min}, u_{max}], \\ u_{min} & \text{if } u_{it} < u_{min}. \end{cases}$$

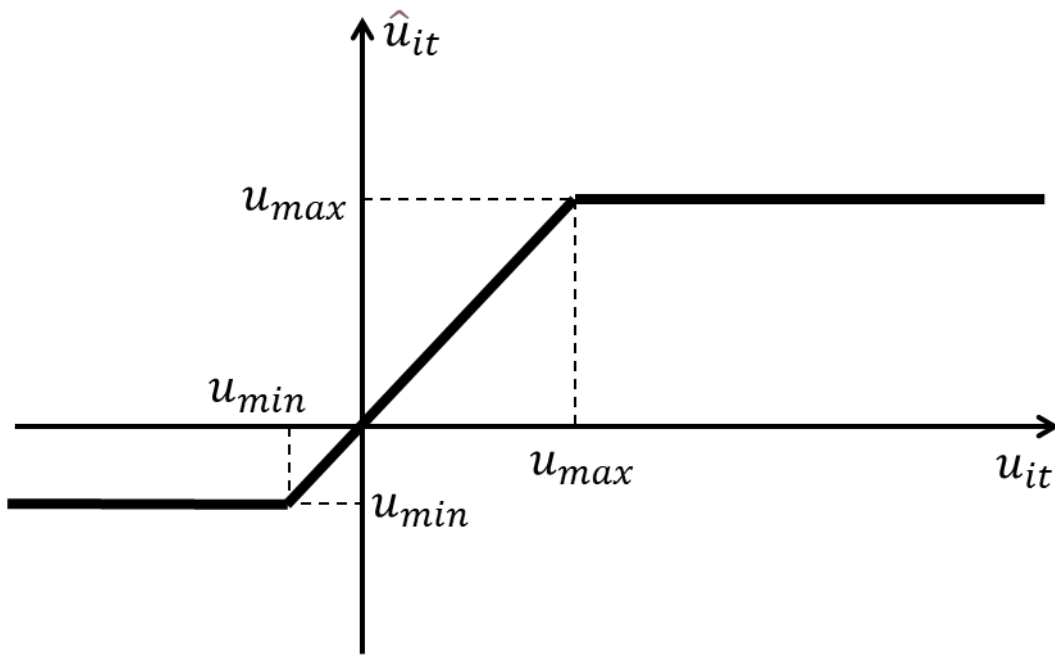


Figure 1: The cut function of the variable critical level used in the neutral range social welfare function.

The social planner, having to make a decision at time $t = 0$, can choose the option or state that maximizes the social welfare function. The discount factor β mitigates the repugnant conclusion. The smaller this discount factor, the more difficult it becomes to arrive at a repugnant conclusion. When the discount factor gets smaller, future people with lives barely worth living contribute less to the social welfare. This means their existence is less likely to compensate for the welfare loss of present generations. Present generations are less likely having to sacrifice themselves for the sake of those future people with lives barely worth living. The smaller the discount factor, the larger the population of future people have to be for the social welfare to increase. When the discount factor is zero, the repugnant conclusion can be fully avoided.

Interpretation: population ethical theories and the discount factor

In some limiting cases for the parameters β , u_{min} and u_{max} , the neutral-range social welfare function represents some influential population ethical theories.

- If $\beta = 1$, we get total or classical utilitarianism: the social planner should choose the option that maximizes the sum of everyone's lifetime utility.
- If $u_{min} = -\infty$ and $u_{max} = +\infty$, we get the social welfare function of time-discounted utilitarianism, which is most often used in economics: maximize the sum of discounted lifetime utilities, where the utilities are weighted by a discount factor.
- If $\beta = 0$, $u_{min} = -\infty$ and $u_{max} = +\infty$, we get presentism (Arrhenius, 2000): maximize the sum of lifetime utilities of everyone who exists in the present. This is a person-affecting theory (Narveson, 1973; Heyd 1988), because it will always be the case that if a state A is better (or worse) than a state B, then it is better (or worse) for at least someone who exists in both situations A and B. In other words, if the social welfare function W_{NRU}^A for option A is higher (or lower) than the social welfare function W_{NRU}^B for option B, there is an individual i who exists in the present and for whom $u_{i0}^A > u_{i0}^B$.
- When $\beta = 0$, $u_{min} = 0$ and $u_{max} = +\infty$, we get asymmetric presentism (Roberts, 2011; Algander, 2012). This theory entails the Procreation Asymmetry principle: adding an unhappy life in the future always makes things worse (all else equal), but adding a happy life not always makes things better.
- If $\beta = 0$ and $u_{min} = u_{max} > 0$, we get critical-level utilitarianism (Blackorby, Bossert and Donaldson, 1995; 1997): maximize the sum of relative lifetime utilities, where a relative utility is a person's utility relative to a positive critical level. Adding lives with utilities above the critical level is good, adding lives with utilities below the critical level is bad, all else equal.

In general, the inverse discounted cut function $(1 - e^{-\rho t})C(u_{it})$ can be interpreted as a variable critical level, as it varies with time and utility level. Different individuals can have different critical levels. The critical level becomes a neutral range, where future people with utilities in the range $[u_{min}, u_{max}]$ are discounted. If $\beta = 0$ or $\rho = \infty$, these people are fully discounted: their contribution to the social welfare function is neutral. The theory is neutral in the sense that it is indifferent between creating a life with a utility in the neutral range, and not creating such a life, all else equal. The addition to the global population of a life at a utility level in the neutral range 'has no positive or negative value in itself' (Broome, 2004, p145). This neutrality also means that the Pareto principle is violated: suppose one can increase the lifetime utility of a person born at time $t > 0$ from

u_{it} to u_{it}' . If those utilities lie in the neutral range, the social welfare function does not increase. Hence, making a future person better off without making anyone else worse-off, does not always improve the situation.

Necessitarianism (Arrhenius, 2000), a person-affecting utilitarian theory similar to presentism, offers a natural interpretation of the utility discount factor and discount rate. Necessitarianism makes a distinction between necessary and possible people. Necessary people exist necessarily in the sense that they exist in all eligible states (i.e. all states that are possible or feasible and can be chosen by the social planner). In contrast, possible people (synonyms: contingent or potential people) are individuals who do not exist in all eligible states. The necessary people not only include the present people, but also the people who will inevitably exist in the future once a state is chosen. A state should be considered as a world history that includes the whole future. Hence, the necessary people do not yet have to exist or be born right after the moment the state is chosen. It is possible that choosing a state will inevitably result in the existence of a determinate person in the far future, a long time after the moment the state is chosen. That person counts as a necessary person once the state is chosen.

The utility discount factor β^t can now be interpreted as the probability that individuals born at time $t > 0$ are necessary people. At time zero, the social planner has uncertainty about which future people are necessary people. The social planner may not know whether a future person would have existed in all the eligible states at time $t = 0$. The further in the future, the more likely it is to influence the existence of people by making a decision at time $t = 0$, and hence the more likely those people are possible people.

Similarly, the utility discount rate ρ is the rate at which future born people cease to be necessary people. This rate can be compared with the Lyapunov exponent in a dynamic system (Vulpiani, Cecconi & Cencini, 2009). The world is a dynamic system. If the Lyapunov exponent of a dynamic system is large, the system is highly chaotic, which means that a small change in initial conditions quickly causes a drastically different outcome. In our case, the outcomes are the people that are born in the future, and the different initial conditions correspond with the different decisions of the social planner at time $t = 0$. As the world is a chaotic system, its Lyapunov exponent is positive, which means the discount rate ρ is positive. It can be very large, if the world is a highly chaotic system.

Note that necessitarianism is a choice-set dependent theory: who counts as a necessary person depends on the set of eligible states that the social planner can choose. The more states are eligible, the less likely a future person exists in all of them, and hence the less likely a future person counts as

a necessary person. This means that when the social planner faces more options, the discount factor gets smaller.

This interpretation of the utility discount factor is a modification of the more common interpretation used by e.g. Stern (2008), that the discount factor measures the exogenous risk of extinction. The lower β^t , the more likely the whole population goes extinct before time t . If some potential future generations are not born because of an extinction event, their utilities should not be taken into consideration. If there is a 0,1% extinction risk per year, as assumed by e.g. Stern (2008), the discount rate ρ equals 0,1% per year. Necessitarianism extends this notion of extinction to the population of necessary people. As our world is a highly dynamically chaotic system, newborn people in the near future are likely possible instead of necessary people. That means the population of necessary people can go extinct rapidly, at a much higher probability rate than 0,1% per year.

To conclude: the person-affecting population ethical theory of necessitarianism offers an explanation for a positive utility discount rate. Hence, a positive discount rate cannot only be justified by agent-relative moral reasons such as kinship or partiality between generations, as argued by Beckerman & Hepburn (2007) and Mogensen (2022), but also by a necessitarian person-affecting view in population ethics.

Justification: avoiding counter-intuitive conclusions

As with positive critical-level theories, neutral-range utilitarianism can avoid the counter-intuitive repugnant conclusion when $u_{max} > 0$ is sufficiently high (higher than the utilities of lives barely worth living). Adding people with a utility below u_{max} will be neutral or negative and can never compensate for the loss of utility of necessary people.

Critical-level utilitarianism entails the very sadistic conclusion: “For any population of lives with negative welfare, there is a population of lives with positive welfare which is *worse*, other things being equal.” (Arrhenius, 2000, p73). Adding many people with positive utilities slightly lower than the critical level decreases the social welfare function more than adding one person with a very negative utility. Neutral-range utilitarianism can avoid this very sadistic conclusion by setting $u_{min} \leq 0$. But if this lower bound of the neutral range is strictly negative, another sadistic conclusion entails: adding people with negative utilities slightly above the lower bound, is neutral instead of bad, all else equal. To avoid this weak sadistic conclusion, we need $u_{min} \geq 0$. Hence, the lower bound of the neutral range should be set at zero.

Neutral-range utilitarianism resembles critical-range or incomplete utilitarianism (Blackorby, Bossert & Donaldson, 1996; Arrhenius, 2000, p74; Rabinowicz, 2009⁴; Gustafsson, 2020). This critical-range theory uses a range of critical levels and hence a set of critical-level utilitarian social welfare functions: one social welfare function for each critical level. According to this theory, an option A is better than an option B if and only if it is better with respect to all critical-level social welfare functions. The problem with critical-level utilitarianism is its incompleteness: it is not always possible to say whether option A is better than, equal to or worse than option B. Two states A and B are incommensurate (Qizilbash, 2007) when neither A is at least as good as B nor B is at least as good as A. Next to the evaluations 'good', 'bad' and 'neutral', that are measured in one dimension, there is a fourth evaluation 'undistinguished', measured in an orthogonal dimension (Gustafsson, 2020). The more a state is undistinguished, the less it can be said to be better than, worse than or equal to a neutral state. The advantage of neutral-range over critical-range utilitarianism, is that its ordering of options is complete. The social planner can always rank any two available options. This completeness property of neutral-range utilitarianism allows for a social welfare function representation, as shown in a previous section.

Neutral-range utilitarianism has also the properties of existence independence (Blackorby, Bossert & Donaldson, 2003) and replication invariance (Spears & Budolfson, 2019). Existence independence means that the ranking of two options is independent from the existence (and hence the number and utility levels) of unaffected people, i.e. of people who have (if they exist) the same utilities in both options. This property is important for the social planner, because the social planner does not have to know the numbers and utilities of all unaffected people to make a decision. People in the past and the far distance (think about aliens on another planet) cannot influence the decision of a social planner here and now. Therefore, this property is also called the independence of utilities of the dead (Blackorby, Bossert & Donaldson, 1995). This property is related to replication invariance: duplicating the world (making an exact copy of all people and their utilities) means the social welfare function is multiplied by two. Hence, if world A is better than world B, it can never be the case that the duplicated world 2xB becomes better than the duplicated world 2xA. This means the social welfare function is extensive (homogeneous of degree one, constant returns to scale) or additively separable (Blackorby, Bossert & Donaldson, 1995). The neutral-range social welfare function is simply the sum of utility functions over individuals. This distinguishes neutral-range utilitarianism from non-separable theories such as average utilitarianism where the discount factor is time constant and inversely proportional to the total number of individuals ($\beta^t = 1/\sum N_t$), time-average utilitarianism

⁴ Rabinowicz (2009) gave a new philosophical interpretation of critical-range utilitarianism described by Blackorby, Bossert & Donaldson (1996), but calls this theory neutral-range utilitarianism.

where the discount factor equals $\beta^t = 1/N_t$, variable-value or number-dampened utilitarianism where the discount factor is a higher order function of the population size (Hurka, 1983; Ng, 1989), restricted critical-level utilitarianism where the average utility is used if it falls in a neutral range (Blackorby, Bossert & Donaldson, 2003), number-sensitive critical-level utilitarianism where the critical level depends on the population size (Blackorby, Bossert & Donaldson, 2003) and rank-discounted utilitarianism where the discount factor depends not on time but on the rank according to utility, with the highest ranked utility getting the lowest discount factor (Asheim & Zuber, 2014).

Looking at the axiomatic approach to population ethics (Blackorby, Bossert & Donaldson, 1995; 2003), one axiom is violated by neutral-range utilitarianism: anonymity. This property says that a permutation of utilities over individuals does not change the social welfare. Due to the utility discount factor, switching the utilities of a present and a future person affects the social welfare function. Necessary people are treated different from possible people. Yet, a weaker version of anonymity remains: within-generation anonymity (or anonymity within the group of necessary people). Permuting individuals born at the same moment, does not affect the social welfare. Neither does permuting individuals who exist in the present.

Application: the Ramsey rule

In economic modeling, representative agents are often used in the social welfare function. These representative agents represent all current and future generations in the sense that they live infinitely long and at each moment in time have the average utility of the people alive (or born) at that time. When the population is constant, a single representative agent can be used. However, using the average of utilities is not valid when the social welfare function has a cut function. The appropriate utility of the representative agent, is not simply the average utility, but depends on the whole distribution of utilities.

As an example, suppose the utilities u_{it} are random draws from a logistic distribution, with cumulative distribution function

$$F_{\bar{u},\sigma}(u) = \frac{1}{1 + \exp \frac{\pi}{\sqrt{3}\sigma} (\bar{u} - u)},$$

whereby $\bar{u}$ and σ are respectively the mean and standard deviation. Then the average of the cut function $C(u_{it})$ becomes an elongated sigmoid function known as the Dombi-Gera squashing function (Dombi & Gera, 2005):

$$S(\bar{u}, \sigma) = \int_{-\infty}^{\infty} C(u) f_{\bar{u}, \sigma}(u) du = \frac{\sqrt{3}\sigma}{\pi} \ln \left(\frac{1 + e^{\frac{\pi}{\sqrt{3}\sigma}(\bar{u} - u_{\min})}}{1 + e^{\frac{\pi}{\sqrt{3}\sigma}(\bar{u} - u_{\max})}} \right) + u_{\min}.$$

This squashing function now becomes a smoothened (infinitely differentiable) cut function of the representative agent. It approaches the sharp cut function in the limit $\sigma \rightarrow 0$. In this limit, using the previously given neutral-range utilitarian social welfare function with the average utility of a representative agent, is valid. But once the utilities have a non-singular distribution, with non-zero variance, a smoothened cut function should be used. For each smooth distribution of u_{it} a smooth elongated sigmoid function is obtained. And as these smoothened functions are differentiable, they can be used for analytic purposes.

As an illustration, we can derive the Ramsey rule (as in Dasgupta, 2020)⁵:

$$r(\bar{u}(c(t)), t) = \bar{\rho}(\bar{u}(c(t)), t) + \eta(c(t))g(c(t))$$

with $\bar{u}(c(t))$ the utility stream of the representative agent (who represents all current and future generations), $c(t)$ the aggregate consumption level of the representative agent, $\eta(c(t)) = -c \frac{d^2 \bar{u}}{dc^2} / \frac{d\bar{u}}{dc}$ the consumption elasticity of marginal utility, $g(c(t))$ the consumption growth rate, $\bar{\rho}(\bar{u}(c(t)), t)$ the effective (or average) utility discount rate and $r(\bar{u}(c(t)), t)$ the social discount rate (Arrow & Kurz, 1970), also called the social rate of return or consumption rate of interest (Dasgupta, 2020). The effective utility discount rate equals

$$\bar{\rho}(\bar{u}(c(t)), t) = -\frac{1}{t} \ln \left(1 - (1 - e^{-\rho t}) \frac{\partial S(\bar{u}(c(t)), \sigma(t))}{\partial \bar{u}} \right).$$

When $u_{\min} = -\infty$ and $u_{\max} = +\infty$, we have $\frac{\partial S(\bar{u}, \sigma)}{\partial \bar{u}} = 1$ and the effective utility discount rate equals the constant ρ , and we get discounted utilitarianism. As $\frac{\partial S}{\partial \bar{u}} \leq 1$, the effective discount rate $\bar{\rho} \leq \rho$. As mentioned above, in a highly dynamically chaotic world, ρ can be large. But when the variance of the utilities σ is also large (and the neural range finite), the effective discount rate $\bar{\rho}$ can

⁵ Consider a small investment at time 0, which means a small decrease in consumption $dc(0)$ at time 0. This investment generates a return at time t such that the consumption at time t can increase with $dc(t) = e^{rt} dc(0)$. Write $v(t) = \bar{u}(c(t)) - (1 - e^{-\rho t}) S(\bar{u}(c(t)))$. Without this investment, consumption is chosen optimally, by maximizing the social welfare function given by the time integral of $v(t)$. Suppose at this optimal growth path, consumption grows at a rate g . If the investment is optimal (to first order), the (first order) decrease in consumption at time 0 is exactly offset by the increase in consumption at time t . Hence we have: $\frac{dv(\bar{u}(c(0)))}{d\bar{u}} \frac{d\bar{u}(c(0))}{dc} \approx \frac{dv(\bar{u}(c(t)))}{d\bar{u}} \frac{d\bar{u}(c(t))}{dc} \frac{dc(t)}{dc(0)} = \frac{dv(\bar{u}(c(t)))}{d\bar{u}} \frac{d\bar{u}(e^{gt} c(0))}{dc} e^{rt} = \frac{dv(\bar{u}(c(t)))}{d\bar{u}} \frac{d\bar{u}(c(0))}{dc} e^{-\eta g t} e^{rt}$. Taking the logarithm of both sides gives: $r = \eta g - \frac{1}{t} \ln \frac{v'(\bar{u}(c(t)))}{v'(\bar{u}(c(0)))} = \eta g - \frac{1}{t} \ln \frac{1 - (1 - e^{-\rho t}) S'(\bar{u}(c(t)))}{1 - (1 - e^{-\rho 0}) S'(\bar{u}(c(0)))}$.

be small. Or in other words: large differences in utilities between people justify a low effective discount rate.

Limitation: the time inconsistency problem

The many mathematically proven impossibility theorems in population ethics (e.g. Ng, 1989; Arrhenius, 2000; 2000b; 2003; 2011; Blackorby, Bossert & Donaldson, 2003) show that also neutral-range utilitarianism faces some problems. Arguably the most problematic feature of this theory, is its time inconsistency or dynamic inconsistency.⁶

To see this time inconsistency, suppose there are two options or states $s_1 = \{u_{it}\}$ and $s_2 = \{v_{it}\}$. Shift these two states a time period τ to the future (i.e. $u'_{it} = u_{it-\tau}$ for $t \geq \tau$ and $u'_{it} = 0$ for $t < \tau$). Then the social welfare function is:

$$\begin{aligned} W(\{u'_{it}\}) &= \int_{t=0}^{\infty} \sum_{i=1}^{N_t} (u'_{it} - (1 - e^{-\rho t})C(u'_{it}))dt = \int_{t=\tau}^{\infty} \sum_{i=1}^{N_t} (u_{it-\tau} - (1 - e^{-\rho t})C(u_{it-\tau}))dt \\ &= \int_{t=0}^{\infty} \sum_{i=1}^{N_t} (u_{it} - (1 - e^{-\rho(t+\tau)})C(u_{it}))dt, \end{aligned}$$

and hence

$$\frac{W(\{u_{it}\})}{W(\{v_{it}\})} \neq \frac{W(\{u'_{it}\})}{W(\{v'_{it}\})}$$

That means the social welfare ranking of states can change by simply shifting them to the future. At time $t = 0$, option A may be preferred over option B, but after choosing option A, option B could become better.

The time inconsistency is a consequence of another feature of neutral-range utilitarianism: its choice set dependence. Which option is preferred depends on the set of available or eligible options. Choosing an option may decrease the choice set of eligible options. For example suppose initially you can choose between three eligible states, two of them contain an extra individual. Once you choose one of these two states, it is no longer possible to choose the state where that extra individual does not exist. The choice set reduces from three to two eligible states. This reduction of the choice set can influence the betterness relationship between the eligible states. One state can be better than

⁶ For original work, see Strotz (1956), Phelps & Pollak (1968), Bergström (1966) and Simaan and Cruz (1973). For other famous examples of dynamic inconsistencies in e.g. macroeconomics, see Kydland & Prescott (1977).

another, if a third state is possible, but the order can switch once the third state is no longer an option. This is illustrated in the following example, that demonstrates the time inconsistency problem of neutral-range utilitarianism.

At time 0, a decision maker who has a very high lifetime utility level of 100 can decide to bring another individual into existence. Once that extra individual is born at time 1, there are two options: give the extra individual a low lifetime utility of 1, which consists of a life barely worth living, or a slightly higher utility of 4. The first options gives the decision maker the highest lifetime utility of 101, the second option a slightly lower utility level of 99. The decision-maker is a necessary person, who exists in all three eligible states. The extra individual is a possible person at time 0. The table below summarizes the lifetime utilities in the three possible states.

	State 1	State 2	State 3
Decision-maker (necessary person)	100	101	99
Extra individual (possible person)	Non-existence	1	4

Table 1: utilities of people in three available states, causing dynamic inconsistency

Suppose the upper bound of the neutral range is $u_{max} = 5$, and the lower bound is 0. As the extra individual, if she exists, gets a utility within the neutral range, her utility does not count in the social welfare function as long as she does not exist. That means state 2, with social welfare 101, is the best: it maximizes the neutral-range utilitarian social welfare function, and state 3 is the worst as long as person 2 is a possible person. However, choosing state 2 requires bringing into existence person 2. Once that person exists, that person becomes a necessary person, state 1 is no longer an option, and a choice between states 2 and 3 remains. As she became a necessary person, including the welfare of person 2 in the social welfare function now means that state 3 is the best, with total social welfare equal to $99+4=103$. That means state 2 becomes dominated by state 3 after the choice of bringing person 2 into existence.

If this reasoning means that we have to end up with state 3, we face the repugnant conclusion, because we can repeat the process. Suppose we can move from state 3 to a fourth state at time $t = 2$, by adding a third individual with utility 1. In state 4, the decision-maker may get utility 100. The social welfare function equals $100+4=104$, which is higher than 103 of state 3. But assume there is a fifth state where the decision maker gets utility 98 and the third individual gets utility 4. Once the third individual is brought into existence, that person becomes a necessary person, which means its utility counts and state 5 becomes the best. We see the utility of the decision-maker decreasing,

from 100 in state 1 to 99 in state 3 to 98 in state 5. After a hundred steps, we end up with an state in which the decision-maker gets zero lifetime utility and all the other people have low utilities in the neutral range (i.e. below 5). This example shows it is really hard to escape the repugnant conclusion (Spears & Budolfson, 2019).

The only way to escape this repugnant conclusion, is by not allowing state 2 to be a member of the initial choice set at time 0. If the initial choice is between the two permissible states 1 and 3, state 1 will be chosen and the second person will not become a necessary person. That second individual will not exist, and hence cannot complain against discounting her utility (if it lies in the neutral range).

In general, we can impose a constraint to avoid time inconsistency and the repugnant conclusion: if an eligible state, which initially seems the best (i.e. maximizes the neutral range utilitarian social welfare function at time 0), will later (when the state is chosen and the possible people become necessary people⁷) be dominated by another, still eligible state which initially (at time 0) has a lower social welfare than the initial second-best state, the initial best-seeming state should be excluded from the eligible options of the initial decision. Excluding from the initial choice set the states that will become dominated by other, less preferred states once chosen, is a forward-looking constraint or ‘later-no-regret’ condition to avoid time inconsistency.

We can formulate this constraint more formally. Define the social welfare function of state $s = \{u_{it}\}$ at future time $\tau \geq 0$ as

$$W_\tau(s) = \int_{t=\tau}^{\infty} \sum_{i=1}^{N_t} \left(u_{it} - (1 - e^{-\rho(t-\tau)})C(u_{it}) \right) dt,$$

and write the corresponding s -compatible choice set at time τ as C_τ^s , defined as the set of all states that can be chosen at time τ when state s is chosen at time 0. We have $C_\tau^s \subseteq C_0$, with C_0 the initial choice set when no state is chosen yet. The constraint now says that when $W_0(s) \geq W_0(s')$ for all $s' \in C_0$ (i.e. the state s is initially one of the best states) but there exists a time $\tau > 0$ and a state $s_\tau^* \in C_\tau^s$ for which $W_\tau(s_\tau^*) > W_\tau(s)$ (i.e. the state s becomes dominated by s_τ^*), then state s may not be chosen at time 0. This state is excluded from the set of permissible states. In the set of remaining states, we can again look at the best state and see if it should be excluded, by following the same process. All initially best states that become dominated, are excluded one by one. The best state of the remaining permissible states should be chosen.

⁷ Note that the necessary people do not yet have to exist or be born right after the moment the state is chosen. The choice of a state can result in the inevitable existence of a person in the far future.

The time consistency constraint can be incorporated in the social welfare function, by deducting a time inconsistency term. Let $C_t^{s(t)}$ be the set of all states (reachable trajectories) available at time t given that state or trajectory $s(t)$ was followed up to time t . The instantaneous dynamic inconsistency at time t can be measured⁸ by comparing the valuation of trajectory s against the best available alternative in $C_t^{s(t)}$:

$$i_t(s; C_t^{s(t)}) = \max_{y \in C_t^{s(t)}} \max\{0, W_{C_t^{s(t)}}(y) - W_{C_t^{s(t)}}(s)\}.$$

The total dynamic inconsistency penalty subtracted from the baseline neutral-range social welfare function $W_{C_0}(s)$ over continuous time is:

$$I(s; C_0) = \int_0^\infty \delta(t) \cdot i_t(s; C_t^{s(t)}) dt$$

where $\delta(t) > 0$ is a continuous weighting function (or discount factor) across time, chosen to make the integral $I(s; C_0)$ finite.

The overall ‘top choice dynamically consistent’ (TCDC) social welfare function in continuous time takes the form:

$$W_{C_0}^{TCDC}(s) = W_{C_0}(s) - I(s; C_0).$$

This social welfare function is top choice dynamically consistent, which means that any optimal trajectory chosen at an initial time $t = 0$ remains optimal at every subsequent time $\tau > 0$ along that trajectory, evaluated within the updated (reduced) choice set. The appendix contains a proof of top choice dynamic consistency of the above social welfare function.

Open question: the size of the neutral range

One crucial question remains for the social planner: what is the size of the neutral range? To avoid sadistic conclusions, the social planner can set $u_{min} = 0$. But what is an appropriate value for the upper-bound u_{max} ? This is a normative question. If the social planner chooses $u_{max} = 0$, then he chooses a total utilitarian social welfare function, which entails the repugnant conclusion. On the other hand, if the social planner wants to avoid the repugnant conclusion by choosing $u_{max} = \infty$,

⁸ In continuous time, $C_t^{s(t)}$ is infinite-dimensional. Proving that $i(s, t)$ is well-defined and continuous requires compactness conditions on the set of feasible trajectories under appropriate topologies.

she faces the extinction conclusion when the future will contain people with negative utilities. After all, those people with negative utilities always count and negatively contribute to the social welfare, whereas future possible people with positive utilities don't count. The future can only be bad. The total negative utility of all future generations can be much larger than the decrease in utility of the current generations, i.e. the necessary people, when they face extinction.

Perhaps the social planner wants to choose the highest possible upper-bound u_{max} that still avoids the extinction conclusion. That means the contribution of future, possible people to the social welfare has to be positive. Assume for simplicity that the utilities of future people have the same distribution as the utilities of current people at $t = 0$, with one free parameter that determines the mean utility. For example, in all eligible options or states, the utilities have a normal distribution, but the average utility depends on the state. Or the average utility can increase or decrease over time. Write N_0^x as the number of present people with utilities higher than x and N_0^- as the number of present people with negative utilities. Then the upper bound is given by the maximum over all eligible states s of the maximum over all levels x such that the social welfare of future people is not negative:

$$u_{max} = \max_s \max_x \left\{ x \mid \sum_{i=1}^{N_0^x} (u_{i0}^x(s) - x) + \sum_{i=1}^{N_0^-} u_{i0}^-(s) \geq 0, u_{i0}^x(s) > x, u_{i0}^-(s) < 0 \right\}.$$

Write the cumulative distribution function and probability density function as $F_s(u)$ and $f_s(u)$ respectively, then the condition in the above maximization reads:

$$\sum_{i=1}^{N_0^x} (u_{i0}^x(s) - x) + \sum_{i=1}^{N_0^-} u_{i0}^-(s) = N_0^+ (1 - F_s(x)) \left(\int_x^\infty u f_s(u) du - x \right) + \sum_{i=1}^{N_0^-} u_{i0}^-(s) \geq 0$$

If the distribution of positive utilities is exponential, i.e.

$$f_s(u_0^+) = \lambda_s e^{-\lambda_s u_0^+},$$

with the free parameter given by the inverse mean (the inverse of the expected utility of the positive utilities)

$$\lambda_s = \frac{1}{E(u_0^+(s))} \approx \frac{N_0^+}{\sum_{i=1}^{N_0^+} u_{i0}^+(s)},$$

then we get a simple closed form expression of the equation

$$\begin{aligned}
N_0^+ (1 - F_s(x_s)) \left(\int_{x_s}^{\infty} u f_s(u) du - x_s \right) + \sum_{i=1}^{N_0^-} u_{i0}^-(s) \\
= N_0^+ e^{-\lambda_s x_s} \left(\left(x_s + \frac{1}{\lambda_s} \right) e^{-\lambda_s x_s} - x_s \right) + \sum_{i=1}^{N_0^-} u_{i0}^-(s) = 0
\end{aligned}$$

given by:

$$x_s = \frac{E(u_0^+(s))}{2} \ln \left(-\frac{N_0^+ E(u_0^+(s))}{N_0^- E(u_0^-(s))} \right) = \frac{\sum_{i=1}^{N_0^+} u_{i0}^+(s)}{2N_0^+} \ln \left(\frac{\sum_{i=1}^{N_0^+} u_{i0}^+(s)}{\left| \sum_{i=1}^{N_0^-} u_{i0}^-(s) \right|} \right).$$

So the upper bound is the maximum over all eligible states s of one half times the average of the positive utilities times the logarithm of the ratio of total positive to total negative utility:

$$u_{max} = \max_s \frac{\sum_{i=1}^{N_0^+} u_{i0}^+(s)}{2N_0^+} \ln \left(\frac{\sum_{i=1}^{N_0^+} u_{i0}^+(s)}{\left| \sum_{i=1}^{N_0^-} u_{i0}^-(s) \right|} \right).$$

If there are no negative utilities, then the upper bound becomes infinite, which means the utilities of all future people can be discounted, no matter how high their utilities are. And if the total of positive utilities is lower than the (absolute value of) the total of negative utilities, the upper bound cannot be positive, there will be too much suffering in the future, and hence extinction would be preferred.

Conclusion

If a social planner can influence not only the utilities but even the existence of future people, and if that social planner uses a social welfare function to find the optimal choice, choosing an appropriate intergenerational social welfare function that deals with variable populations is crucial. Many population ethical theories that can be represented by a social welfare function, entail counter-intuitive implications such as the repugnant conclusion: sacrifice the current generations in order to create huge numbers of extra people in the far future who have lives barely worth living. Or the sadistic conclusion: adding people with negative utilities can be better than adding other people with positive utilities.

Neutral-range utilitarianism can avoid or at least mitigate such repugnant and sadistic conclusions. It discounts the utilities of future people when these utilities lie within a neutral range. If the neutral range becomes infinitely thin, the theory becomes a critical-level utilitarian theory. If the neutral range becomes infinitely wide, the theory becomes a time-discounted utilitarian theory often used in

economics. And if in the latter case the discount factor approaches zero, the theory becomes presentism, a person-affecting population ethical theory that only considers the utilities of people alive in the present.

To avoid sadistic conclusions, the lower bound of the neutral range has to be at zero utility. To avoid the repugnant conclusion, the upper bound of the neutral range has to be higher than the utility corresponding to a life barely worth living. And to avoid an extinction conclusion, which says that the social planner should prevent the existence of future generations, the upper bound cannot be higher than a positive value that depends on the absolute value of the negative utilities of future people who have lives not worth living.

The intergenerational social welfare function of neutral-range utilitarianism can be used to study optimal economic growth, optimal taxation and optimal climate policy, as well as perform cost-benefit analyses. A Ramsey rule can be derived that connects the social discount rate with the utility discount rate. However, the major difficulty with this neutral-range social welfare function is time inconsistency: the ranking of options available to the social planner can change depending on when decisions are made. This time inconsistency can be avoided by imposing a forward-looking constraint or ‘later-no-regret’ condition: if an eligible state, which initially seems the best, will later (when that state is chosen) be dominated by another, still eligible state which initially has a lower social welfare than the initial second-best state, the initial best-seeming state should be excluded from the eligible options of the initial decision.

Statements and declarations

The author declares no financial or non-financial conflicts of interest. The author did not receive support from any organization for the submitted work.

Appendix: proof of top choice dynamic consistency

Let $s \in C_0$ be a feasible continuous-time trajectory defined over $t \in [0, \infty[$, where C_0 is the initial set of feasible trajectories. If the decision-maker follows trajectory s up to time τ , the remaining set of feasible trajectories accessible from time τ onward is denoted $C_\tau^{s(\tau)} \subseteq C_0$. Note that $s \in C_\tau^{s(\tau)}$ always holds because s is one of the valid continuations from τ .

Let $W_{C_\tau^{s(\tau)}}(y)$ be the base choice-set-dependent social welfare of trajectory y evaluated under choice set $C_\tau^{s(\tau)}$.

For any trajectory $y \in C_\tau^{s(\tau)}$, the instantaneous deviation from optimality at time $t \geq \tau$ is:

$$i_t(y; C_t^{y(t)}) = \max_{z \in C_t^{y(t)}} \max\{0, W_{C_t^{y(t)}}(z) - W_{C_t^{y(t)}}(y)\}.$$

Evaluating a trajectory $y \in C_\tau^{s(\tau)}$ from the perspective of time τ yields:

$$W_{C_\tau^{s(\tau)}}^{TCD C}(y) = W_{C_\tau^{s(\tau)}}(y) - \int_\tau^\infty \delta(t) \cdot i_t(y; C_t^{y(t)}) dt,$$

where $\delta(t) > 0$ is a positive weighting function over time.

A social welfare function $W^{TCD C}$ is Top Choice Dynamically Consistent if, whenever x^* maximizes $W_{C_0}^{TCD C}(x)$ over C_0 at time $t = 0$:

$$x^* \in \arg \max_{y \in C_0} W_{C_0}^{TCD C}(y),$$

then for all subsequent times $\tau > 0$, x^* remains an optimal trajectory within the restricted choice set $C_\tau^{x^*(\tau)}$:

$$x^* \in \arg \max_{y \in C_\tau^{x^*(\tau)}} W_{C_\tau^{x^*(\tau)}}^{TCD C}(y).$$

Suppose x^* is chosen as the optimal trajectory at time $t = 0$:

$$W_{C_0}^{TCD C}(x^*) = \max_{y \in C_0} W_{C_0}^{TCD C}(y).$$

Because x^* is the top choice globally over C_0 at $t = 0$, the TCDC penalty term $I(x^*; C_0) = \int_0^\infty \delta(t) \cdot i_t(x^*; C_t^{x^*(t)}) dt$ subtracted from $W_{C_0}(x^*)$ at time $t = 0$ evaluates to zero. To see why, consider $t = 0$:

$$i_0(x^*; C_0) = \max_{z \in C_0} \max\{0, W_{C_0}(z) - W_{C_0}(x^*)\}.$$

If $i_t(x^*; C_t^{x^*(t)}) > 0$ over a set of non-zero measure, then $\int_0^\infty \delta(t) \cdot i_t(x^*; C_t^{x^*(t)}) dt > 0$, meaning $W_{C_0}^{TCD C}(x^*) < W_{C_0}(x^*)$.

By construction of the TCDC penalty, the trajectory x^* that maximizes $W_{C_0}^{TCD C}$ eliminates this inconsistency gap along its path:

$$i_t(x^*; C_t^{x^*(t)}) = 0 \quad \forall t \geq 0.$$

Therefore, for the optimal trajectory x^* :

$$W_{C_0}^{TCDC}(x^*) = W_{C_0}(x^*).$$

Now advance time to $\tau > 0$. The decision-maker evaluates x^* within the reachable choice set $C_\tau^{x^*(\tau)}$:

$$W_{C_\tau^{x^*(\tau)}}^{TCDC}(x^*) = W_{C_\tau^{x^*(\tau)}}(x^*) - \int_\tau^\infty \delta(t) \cdot i_t(x^*; C_t^{x^*(t)}) dt.$$

Since $i_t(x^*; C_t^{x^*(t)}) = 0$ for all $t \geq 0$ (and thus for all $t \geq \tau$):

$$W_{C_\tau^{x^*(\tau)}}^{TCDC}(x^*) = W_{C_\tau^{x^*(\tau)}}(x^*)$$

Take any alternative feasible trajectory $y \in C_\tau^{x^*(\tau)}$ available at time τ . Its evaluation under W^{TCDC} at time τ is:

$$W_{C_\tau^{x^*(\tau)}}^{TCDC}(y) = W_{C_\tau^{x^*(\tau)}}(y) - \int_\tau^\infty \delta(t) \cdot i_t(y; C_t^{y(t)}) dt.$$

By definition of $i_\tau(y; C_\tau^{x^*(\tau)})$, notice that:

$$i_\tau(y; C_\tau^{x^*(\tau)}) = \max_{z \in C_\tau^{x^*(\tau)}} \max\{0, W_{C_\tau^{x^*(\tau)}}(z) - W_{C_\tau^{x^*(\tau)}}(y)\}.$$

Since $x^* \in C_\tau^{x^*(\tau)}$, the maximum over $z \in C_\tau^{x^*(\tau)}$ is at least as large as $W_{C_\tau^{x^*(\tau)}}(x^*)$:

$$\max_{z \in C_\tau^{x^*(\tau)}} W_{C_\tau^{x^*(\tau)}}(z) \geq W_{C_\tau^{x^*(\tau)}}(x^*).$$

Thus:

$$i_\tau(y; C_\tau^{x^*(\tau)}) \geq \max\{0, W_{C_\tau^{x^*(\tau)}}(x^*) - W_{C_\tau^{x^*(\tau)}}(y)\} \geq W_{C_\tau^{x^*(\tau)}}(x^*) - W_{C_\tau^{x^*(\tau)}}(y)$$

Rearranging the inequality above:

$$W_{C_\tau^{x^*(\tau)}}(y) \geq W_{C_\tau^{x^*(\tau)}}(x^*) - i_\tau(y; C_\tau^{x^*(\tau)})$$

Now plug this into the TCDC formulation for alternative y :

$$W_{C_\tau^{x^*(\tau)}}^{TCDC}(y) = W_{C_\tau^{x^*(\tau)}}(y) - \int_\tau^\infty \delta(t) \cdot i_t(y; C_t^{y(t)}) dt.$$

Because $\delta(t) > 0$ and $i_t(y; C_t^{y(t)}) \geq 0$ for all $t \geq 0$:

$$\int_\tau^\infty \delta(t) \cdot i_t(y; C_t^{y(t)}) dt \geq 0$$

Combining the bounds gives:

$$W_{C_\tau^{x^*(\tau)}}^{TCD C}(y) \leq W_{C_\tau^{x^*(\tau)}}(x^*) - \int_\tau^\infty \delta(t) \cdot i_t(y; C_t^{y(t)}) dt \leq W_{C_\tau^{x^*(\tau)}}(x^*).$$

Since we know $W_{C_\tau^{x^*(\tau)}}^{TCD C}(x^*) = W_{C_\tau^{x^*(\tau)}}(x^*)$, we obtain:

$$W_{C_\tau^{x^*(\tau)}}^{TCD C}(y) \leq W_{C_\tau^{x^*(\tau)}}(x^*).$$

Conclusion: for all $\tau > 0$ and for all alternative feasible trajectories $y \in C_\tau^{x^*(\tau)}$, we have

$$W_{C_\tau^{x^*(\tau)}}^{TCD C}(y) \leq W_{C_\tau^{x^*(\tau)}}(x^*).$$

Thus:

$$x^* \in \arg \max_{y \in C_\tau^{x^*(\tau)}} W_{C_\tau^{x^*(\tau)}}^{TCD C}(y).$$

Subtracting the integrated instantaneous inconsistency gap penalizes any trajectory y by the exact amount it deviates from optimality relative to x^* . This guarantees that the top choice x^* selected at $t = 0$ remains globally optimal within $C_\tau^{x^*(\tau)}$ at every time $\tau > 0$, establishing Top Choice Dynamic Consistency in continuous time.

References

- Algander, P. (2012). A defense of the asymmetry in population ethics. *Res Publica*, 18(2): 145-157.
- Arrhenius, G. (2000). *Future Generations: A Challenge for Moral Theory*, PhD dissertation, Uppsala University.
- Arrhenius, G. (2000b). An impossibility theorem for welfarist axiologies. *Economics and Philosophy*, 16(2): 247-266.
- Arrhenius, G. (2003). The very repugnant conclusion. In: Segerberg, K. & Sliwinski R. (eds.), *Logic, law, morality: thirteen essays in practical philosophy in honour of Lennart Åqvist*, pp. 167–180. Uppsala philosophical studies 51. Uppsala: Department of Philosophy, Uppsala University.
- Arrhenius, G. (2011). The impossibility of a satisfactory population ethics. In E. N. Dzhaferov & L. Perry (Eds.), *Descriptive and Normative Approaches to Human Behavior*, 1–26. World Scientific Publishing Company.

- Arrow, K.J. & M. Kurz (1970), *Public Investment, the Rate of Return and Optimal Fiscal Policy*. Baltimore: Johns Hopkins University Press.
- Asheim, G. B., & Zuber, S. (2014). Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population. *Theoretical Economics*, 9(3), 629-650.
- Beckerman, W., & Hepburn, C. (2007). Ethics of the discount rate in the Stern Review on the economics of climate change. *World Economics-Henley on Thames*-, 8(1): 187-210.
- Bergström, L. (1966). *The alternatives and consequences of actions: An Essay on Certain Fundamental Notions in Teleological Ethics*. Stockholm: Almquist & Wiksell.
- Blackorby, C., Bossert, W., and Donaldson, D. (1995). Intertemporal population ethics: Critical level utilitarian principles. *Econometrica*, 65: 1303-1320.
- Blackorby, C., Bossert, W., & Donaldson, D. (1996). Quasi-orderings and population ethics. *Social Choice and Welfare*, 13(2): 129-150.
- Blackorby, C., Bossert, W., and Donaldson, D. (1997). Critical-level utilitarianism and the population-ethics dilemma. *Economics and Philosophy*, 13: 197-230.
- Blackorby, C., Bossert W., & Donaldson, D. (2003). The axiomatic approach to population ethics. *Politics Philosophy Economics*, 2(3): 342-381.
- Broome, J. (2004). *Weighing Lives*. Oxford: Oxford University Press.
- Cass, D. (1965). Optimum Growth in an Aggregative Model of Capital Accumulation. *Review of Economic Studies*. 32 (3): 233–240.
- Collins, B. (2013). Using a survey to estimate health expectancy and quality-adjusted life expectancy to assess inequalities in health and quality of life. *Value in health*, 16(4), 599-603.
- Dasgupta, P. (2020). Ramsey and Intergenerational Welfare Economics. *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.).
- Dombi, J., & Gera, Z. (2005). The approximation of piecewise linear membership functions and Łukasiewicz operators. *Fuzzy sets and Systems*, 154(2): 275-286.
- Drupp, M. A., Freeman, M. C., Groom, B., & Nesje, F. (2018). Discounting disentangled. *American Economic Journal: Economic Policy*, 10(4), 109-134.
- Feldstein, M. S. (1964). The social time preference discount rate in cost benefit analysis. *The economic journal*, 74(294): 360-379.

- Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2): 351-401.
- Greaves, H. (2017). Discounting for public policy: A survey. *Economics & Philosophy*, 33(3), 391-439.
- Gold, M. R., Siegel, J. E., Russell, R. B., & Weinstein, M. C. (1996). Cost-effectiveness in health and medicine. New York: Oxford University Press.
- Gustafsson, J. E. (2020). Population axiology and the possibility of a fourth category of absolute value. *Economics & Philosophy*, 36(1): 81-110.
- Heyd, D. (1988). Procreation and value: Can ethics deal with futurity problems? *Philosophia*, 18: 151-170.
- Hurka, T. (1983). Value and population size. *Ethics*, 93: 496–507.
- Jia, H., Zack, M. M., & Thompson, W. W. (2011). State quality-adjusted life expectancy for US adults from 1993 to 2008. *Quality of Life Research*, 20: 853-863.
- Koopmans, T. C. (1965). On the Concept of Optimal Economic Growth. *The Economic Approach to Development Planning*. Chicago: Rand McNally. pp. 225–287.
- Kydland, F. & Prescott, E. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–492.
- Layard, R., & Oparina, E. (2021). Living long and living well: The WELLBY approach. *World Happiness Report*, 191-208.
- Mogensen, A. L. (2022). The only ethical argument for positive δ ? Partiality and pure time preference. *Philosophical Studies*, 179(9): 2731-2750.
- Narveson, J. (1973). Moral problems of population. *The Monist*, 57: 62-86.
- Ng, Y. K. (1989). What should we do about future generations?: Impossibility of Parfit's Theory X. *Economics & Philosophy*, 5(2): 235-253.
- Nordhaus, W. D. (1992). An optimal transition path for controlling greenhouse gases. *Science*, 258(5086): 1315-1319.
- Parfit, D. (1984). *Reasons and Persons*, Oxford: Clarendon Press.
- Parfit, D. (2017). Future people, the non-identity problem, and person-affecting principles. *Philosophy & Public Affairs*, 45(2): 118-157.

- Pestieau, P. M. (1974). Optimal taxation and discount rate for public investment in a growth setting. *Journal of Public Economics*, 3(3), 217-235.
- Phelps, E. S., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, 35(2), 185-199.
- Pluhar, E. (1982). On replaceability. *Ethics and Animals* 3: 96-105.
- Qizilbash, M. (2007). The mere addition paradox, parity and vagueness. *Philosophy and Phenomenological Research*, 75: 129–151.
- Ramsey, F.P. (1928). A Mathematical Theory of Saving. *Economic Journal*, 38(4): 543–559.
- Rabinowicz, W. (2009). Broome and the intuition of neutrality. *Philosophical Issues*, 19: 389-411.
- Roberts, M. A. (2011). An asymmetry in the ethics of procreation. *Philosophy Compass*, 6(11): 765-776.
- Samuelson, P. A. (1937). A note on measurement of utility. *The review of economic studies*, 4(2): 155-161.
- Simaan, M. & Cruz, J. (1973). On the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11(5): 533–555.
- Singer, P. (1979). Killing humans and killing animals. *Inquiry* 22: 145-156.
- Spears, D., & Budolfson, M. (2019). Why variable-population social orderings cannot escape the repugnant conclusion: Proofs and implications. IZA Discussion Paper No. 12668.
- Stern, N. (2008). The economics of climate change. *American Economic Review*, 98(2): 1–37.
- Strotz, R. (1956) Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies*, XXIII: 165–80.
- Vulpiani, A., Cecconi, F., & Cencini, M. (2009). *Chaos: from simple models to complex systems* (Vol. 17). World Scientific.
- Zuber, S., Venkatesh, N., Tännsjö, T., Tarsney, C., Stefánsson, H. O., Steele, K., ... & Asheim, G. B. (2021). What should we agree on about the Repugnant Conclusion?. *Utilitas*, 1-5.