Journal of Environmental Economics and Management 71 (2015) 34–53

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

On sustainability and social welfare ☆

Marc Fleurbaey *

Princeton University, United States

ARTICLE INFO

Article history:
Received 7 August 2013
Available online 27 February 2015

JEL classification:
D6
D9
O2
Q5

Keywords:
Sustainability
Genuine savings
Ecological footprint
Social welfare
Overlapping generations

ABSTRACT

This paper proposes to define sustainability in terms of leaving it *possible* for future generations to sustain certain defined targets. It is shown that variants of genuine savings and the ecological footprint can then serve as indicators of sustainability. The link between sustainability and intergenerational welfare is examined, and it is shown how to incorporate indicators of sustainability into a social welfare measure, including risk in the analysis.

> "It is very hard to be against sustainability.
>
> In fact, the less you know about it, the better it sounds."
>
> Robert Solow (1991, p. 179)

Introduction

Ever since the Brundtland Commission characterized sustainable development as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission, 1987), sustainability has become a convenient slogan in the difficult exercise of pursuing the conflicting goals of bringing affluence to all human beings while preserving the capacity of the Earth to bear the human population.

It is exciting for economists to analyze such a notion and see if we can make sense of it in our theory. In this paper, I argue that, among the definitions adopted in economics, one conception of sustainability better captures the idea of sustainability than others, namely, the notion of giving future generations the *ability* (a word used in the Brundtland formulation) to sustain certain targets (Section "Definitions of sustainability in the literature"). But, although the literature has examined how to incorporate a priority for the future in intergenerational welfare objectives, it has not provided exact

☆ This paper, delivered as a keynote address at the EAERE Congress in Toulouse, 2013, has considerably benefited from advice and comments by G. Asheim and V. Martinet, though they should not be held responsible for any mistake. Comments by three referees and the editor are also gratefully acknowledged.

* Fax: +1 609 258 5974.
E-mail address: mfleurba@princeton.edu

http://dx.doi.org/10.1016/j.jeem.2015.02.005

indicators of sustainability (as defined here) and has not tried to make sustainability indicators a component of social welfare—this is what this paper studies. (The relationship with the literature is not explained further in this introduction and is discussed in detail at the beginning of each main section.)

The purpose of this paper is therefore threefold. (1) The ultimate goal is to see if sustainability indicators can be incorporated into a measure of social welfare. (2) For this purpose, a simple indicator of sustainability is needed, and therefore the paper first examines how sustainability can be defined and measured. (3) Finally, another goal of the paper is to make the analysis amenable to possible future applications, which requires showing that the analysis can incorporate uncertainty and can be applied in the more realistic context of overlapping generations.

To give a flavor of the main results, the sustainability analysis of social welfare that is made possible by this analysis takes the following form:

- When the future path is known for sure: intergenerational welfare can be approximated by the level of welfare of the present generation diminished by an indicator of unsustainability (equal to the fraction by which what future generations can sustain falls behind the present generation's welfare), or augmented by an indicator of sustainability (which is similar but not equal to the proportion in which what future generations can sustain is greater than the present generation's welfare), and further modified by a term measuring the difference between the sustainability capacities of future generations and their actual predicted achievements;
- When there is uncertainty about the future feasibility set and the future path: expected intergenerational welfare can be written in similar fashion as above, but the (un)sustainability terms are the product of the probability of (un) sustainability by the (un)sustainability indicators computed on the basis of the expected welfare level that the future generations can sustain.

This analysis shows that it is helpful to cast the design of sustainability indicators in the context of social welfare, because this gives useful guidance for the selection of the most relevant indicators. In Section "Indicators of sustainability under perfect information" we will see that at least three indicators could be considered (two variants of the ecological footprint, and a variant of genuine savings). Moreover, the study of the uncertainty context shows that it is not just the probability of unsustainability that one should worry about, but also the expected extent of the sustainability deficit. And a convenient combination of the two considerations is made possible by their incorporation into intergenerational welfare analysis.

The choice of framework is important (Section "Model"). Finding simple indicators of sustainability is made much easier by adopting a discrete time framework, because discrete time makes it possible to obtain a clear distinction between what the current period does and what the next periods can do given the capital stocks they inherit. The literature on sustainability indicators has often focused on continuous time models, in which optimization techniques are well developed. But sustainability is a concept that refers to the transmission of capital to the next periods, and a key message of this paper is that it is worthwhile analyzing sustainability in discrete time models (with successive generations and also in models with overlapping generations). As most economic data are provided in discrete time (annual), it is also practically useful to develop measures adapted to such data.

Therefore, starting with a discrete-time successive-generations model, I obtain simple indicators of sustainability, including a variant of genuine savings and two variants of the ecological footprint (Section "Definition of sustainability and first indicator" and Section "Sustainability and risk"). The Appendix provides a detailed discussion of the classical notion of genuine savings and compares it to the variant proposed here. It is explained there that the genuine savings indicator, as usually defined and measured, tells us very little about the *ability* of future generations to sustain the present generation's welfare level. In contrast, the notions and indicators proposed in this paper are flexible enough to make it possible to incorporate a variety of sustainability targets (growth, multiple objectives). A simple classification of sustainability configurations is provided (Section "Mapping out sustainability configurations"), as well as a numerical illustration of the concepts and indicators (Section "Example").

After these preliminaries, it is then possible to study how to incorporate sustainability indicators into a measure of social welfare, and clear decompositions are obtained which feature indexes of sustainability and unsustainability (Section "Sustainability as a component of intergenerational welfare"), as announced above.

Incorporating uncertainty in the analysis is possible, but raises a new set of issues having to do with whether the present generation just wants to check that its own welfare level is sustainable or whether a variety of welfare levels are likely to be sustainable (Section "Sustainability and risk"), with a special attention to catastrophic risks. As already mentioned earlier, a decomposition of intergenerational welfare then includes the double perspective of the risk (probability) of unsustainability and the magnitude of the setback. Another key issue is that the probability of sustaining a generational welfare level in the future must take account of the possibility for future generations to find sustained paths and to learn from the history before them. This issue is examined in detail in Section "Epistemic conditions of sustainability".

While most of the paper deals with successive generations, it is shown in Section "Overlapping generations" that the bulk of the analysis can be adapted to overlapping generations, which is of paramount importance for applications, because annual statistics are not obviously reconcilable with the successive-generation time frame. A short conclusion is given in Section "Conclusion".

Indicators of sustainability under perfect information

Definitions of sustainability in the literature

A vast literature has studied sustainability through the computation of optimal paths for social objectives incorporating a strong concern for future generations, such as the maximin (Solow, 1974) or the Chichilnisky (1996) criterion. This literature has in particular clarified the relationship between Hartwick's rule and a path of constant welfare. Excellent syntheses of this literature can be found in Heal (1998) and part II of Asheim (2007). The results of this literature are, however, of limited relevance to the task of checking if any, potentially suboptimal, management of resources by the present generation is sustainable in the Brundtland sense.

Two broad approaches to sustainability can be distinguished in the literature. In the first one, Arrow et al. (2010) define sustainable development as the set of "economic paths along which intergenerational well-being does not decline" (p. 2). More formally, their definition says that "economic development is sustained at t if $dV/dt \geq 0$," (p. 5) where V is the discounted utilitarian sum of generational utility. The definition is extended to an interval of time by integrating (i.e., looking at the evolution of V over that interval). Their analysis involves a prediction of how capital stocks determine V, alongside other exogenous factors (institutions, preferences, technology…): $V(t) = V(K(t), t)$. Therefore the evolution of V can be related to the total value of net investments at shadow prices, as crystallized in the notion of genuine savings).[1] In a similar vein, Heal and Kriström (2005) introduce what could be called Samuelsonian income, namely, the value of (discounted) consumption at supporting prices, and show that the variation of the value of consumption at supporting prices is the same as that of the discounted sum of utility. The paper by Dasgupta and Mäler (2000) was a pioneer of these approaches, making the key contribution of extending the framework to non-optimal paths.

An important weakness of this approach is that if fails to capture the key sustainability concern as expressed in the Bruntland quotation, which is about sustaining *generational* welfare *in the future*, rather than raising *intergenerational* welfare *in the present*. Indeed, one can have $dV/dt > 0$ at the present date even when generational utility and/or intergenerational welfare will necessarily decline in the future. In his review of the literature, Asheimch (2007, ch. 14), makes a clear distinction between the two branches of analyses, namely, intergenerational welfare versus sustainability of generational welfare.

The second approach is more in line with the Brundtland focus on the "ability" of future generations. Solow (1991) defines sustainability as referring to "an obligation to conduct ourselves so that we leave to the future the *option* or the *capacity* to be as *well off* as we are." (p. 181, emphasis added). A variant is proposed by Asheim (2007): "A generation's management of the resource base at some point in time is sustainable if it constitutes the first part of a feasible sustained development," (p. 2) i.e., development in which "the stream of well-being is nondecreasing" (p. 1). Two features of this definition are interesting. First, it bears on the current generation's management, and therefore highlights the important idea that sustainability of the current welfare level must take account of the constraints imposed on the future by the current generation's actions. Second, the condition that future generations should be able to follow a nondecreasing path is apparently more demanding than just sustaining the current level.[2] Pezzey (1997, p. 451), suggests that it may be too strong, as one could say that the present generation has done enough if all future generations can be at least as well off as itself.

Along this vein, other definitions are more directly connected to criteria for checking sustainability. Nordhaus (1995) (see also Cairns, 2000) distinguishes Hicksian income ("the maximum amount that can be consumed while leaving capital constant," p. 3) and Fisherian income ("sustainable income is the maximum amount that a national can consume while ensuring that all future generations *can* have living standards that are at least as high as those of the current generation, " p. 4, emphasis added), arguing in favor of the latter as a better notion. The idea of checking sustainability by comparing aggregate consumption to a unidimensional threshold may be too restrictive, because the impact of current actions is likely to depend on more detailed features (e.g., the composition of investments), as emphasized in Asheim (1994) and Aronsson et al. (1997).

Pezzey (1997, 2004) and Pezzey and Toman (2002) define sustainability as current welfare being below or equal to the maximum sustainable level (called the "maximin value"), and note that this inequality is a necessary, but not sufficient, condition for it being possible for welfare to follow a nondecreasing path forever.[3] Pezzey (2004) argues that the proposed definition makes sense only under efficiency assumptions.

In a related approach, Martinet (2011) defines sustainability in terms of the possibility of satisfying inequality conditions (thresholds) in the future. He proposes to maximize over the thresholds for some preference orderings. This generalizes the maximin, which consists in maximizing the threshold of generational welfare over which every future generation should stand, and the generalization occurs in two ways: first by going multidimensional, second by considering any suitable target instead of just welfare. This approach is similar to the optimizing approach mentioned in the beginning of this section.

---

[1] The idea of measuring the constant equivalent of a discounted sum (of consumption or utility) by the net national product goes back to Weitzman (1976).

[2] This introduces a nice recursive structure in the analysis. Generation 0 behaves sustainably if it makes it possible for generation 1 to behave sustainably (and so on).

[3] Pezzey (1997) actually requires utility at t to be below or equal to the sustainable level at t, for all t, for a "sustainable development", whereas the other references use this inequality at $t=0$ to define sustainability at time 0.

An interesting aspect of his analysis is the reference to viability theory (also mobilized in Martinet and Doyen, 2007; Baumgärtner and Quaas, 2009; Martinet, 2011), which defines the set of initial states of the economy that are compatible with respecting the thresholds over time, or, conversely, the set of thresholds that can be respected from any given initial state of the economy. On the basis of this theory, and focusing on a single welfare threshold, Doyen and Martinet (2012) propose a combination of two practical tests for checking whether the present generation's welfare can be sustained by future generations on the basis of this approach. The first test is that current utility should not exceed the maximin level (Pezzey and Toman's definition), and, when current utility is equal to the maximin level, one must also check that genuine savings evaluated at the shadow prices of the maximin program should be nonnegative, or equivalently that $dV/dt \geq 0$, where V now denotes the maximin value, i.e., the maximum sustainable level of generational welfare.[4]

Their approach covers non-optimal paths, which is quite valuable since in his chapter 14, Asheim (2007) noted that there were no known criteria of sustainability in the absence of optimality assumption. Regarding genuine savings (evaluated at suitable shadow prices), Asheim and Buchholz (2004) have a result of sufficient condition only when sustainability is taken as a constraint in the social maximization, and Pezzey (2004) has a necessary condition assuming optimality for discounted utilitarianism.

Cairns and Martinet (2012) propose to associate the two approaches distinguished here. They propose to retain the Pezzey–Toman definition of sustainability of current utility, but also to examine “sustainability improvement”, defined as $dV/dt \geq 0$. The interest of the latter notion, they argue, is that it captures development concerns which may be complementary to sustainability concerns. On the technical side, they provide a valuable map of all possible combinations of the sign of dV/dt with the sign of $W_0 - V$, where W_0 is current welfare and V the maximin value.

Let us take stock. It appears questionable to define sustainability in a way that is compatible with the future generations being actually unable to sustain the target level, due to the investment decisions of the present generation. We will therefore retain the idea that sustainability is about *making it possible for future generations to achieve some outcome*, either a level or a progression, *taking into account the constraints* generated by the management of resources by the current generation. Importantly, sustainability defined in this way can be ascertained *without making a precise prediction about the future generations' decisions*, since only their possibility set matters. So far, the literature has provided no exact indicator for sustainability defined in this way. The task of this section is to look for such indicators.

Model

The framework adopted in this paper is unusual in two ways. First, time is discrete $t = 0, 1, \ldots$, representing successive generations (overlapping generations are considered later), which is not uncommon in growth theory but more so in the sustainability literature.[5] As explained in the introduction, the motivation for discrete time is that it nicely allows to separate the possibilities for future generations from the achievements of the current generation, it simplifies the use of the maximin value as a criterion for sustainability, and makes the theoretical tools more amenable to applications to annual statistics.

A second original feature of this paper is that the horizon (denoted H) is finite.[6] This is not necessarily a forecast of the longevity of our species, but simply the horizon of planning. There are three reasons for adopting a finite horizon. First, it is more realistic and does not introduce a serious limitation in the analysis because the horizon of planning can go beyond the possible lifespan of stars if one wishes. Second, it eliminates from the analysis complications of dubious practical relevance that are solely due to what may happen in an infinite time. Such difficulties affect not only growth paths (possible non-existence of a regular maximin path, for instance), but also the definition of intergenerational welfare. With an infinite horizon, social welfare cannot be defined without introducing pure time preference (or dropping the Weak Pareto principle),[7] therefore it would seriously undermine the goal of this paper to incorporate sustainability indicators into interesting concepts of intergenerational welfare. Third, sustainability should arguably be defined *with respect to* a given horizon, and considering various horizons may be practically interesting. One can then say, for instance, that our current management makes the current welfare level sustainable for the next 50 years but not for a longer horizon. It is also likely that whatever we do, our current level of welfare cannot be sustained for more than a few tens of thousand years, due to natural catastrophes. Determining the horizon over which the current management is compatible with sustaining current welfare, for instance, might be a nice piece of analysis. The fixed planning horizon H adopted in this paper is therefore to be interpreted as an exogenous policy parameter. A drawback of the finite horizon framework is the non-stationarity of the analysis, but this is a cost that may be worth paying for the three reasons just listed. Another argument against a finite horizon is that it seems to single out a generation that would not care for its descendants. This exaggerates the meaning of the horizon (which can, if one wishes, be greater than any possible human history). It is only a working parameter of the

[4] When current utility is equal to the maximin value and the maximin value decreases, i.e., $dV/dt < 0$, then the current level of utility is sustainable but, given current actions, will not be sustained in any feasible path.

[5] As shown by Dasgupta and Mitra (1983), Hartwick's rule is not valid in discrete time. Other references that have studied sustainability in discrete time include Asheim and Brekke (2002), Llavador et al. (2011) and De Lara et al. (2015). Discrete time is common in intergenerational social choice (e.g., Asheim et al., 2012).

[6] This is not, however, the first paper to combine discrete time and finite horizon in the analysis of sustainability. See De Lara and Doyen (2008) and De Lara et al. (2015).

[7] See Basu and Mitra (2003) and Zame (2007).

present generation for its computation of how long any given state of affairs can be sustained. If the horizon is long enough, its precise value has little influence on the numerical analysis of the sustainability of present actions (see Section "Example" for an illustration of this point), and therefore does not single out more than a broad range of time perspectives.[8]

Otherwise, the model is similar to the literature. For $t=0,\dots,H-1$, a sequence $(x_t,\dots,x_H)$ is denoted x_t^+.

The stock of capital is $K_t \in \mathbb{R}_+^m$. The value of K_t is given in the beginning of period t. In addition to the usual components of capital, the vector K_t may include cosmic environmental factors (solar activity, gravitational influences, meteors, geomechanics, or world prices for a small country) that are not influenced by human activity. It may also include the population size.

Generation t's actions are described by the vector $A_t \in \mathbb{R}^\ell$. It includes all actions of generation t (consumption, production, investment, extraction, pollution, procreation).

Let the state of society in t be $S_t=(A_t,K_t)$. The capital at the beginning of the next period is determined by the function: $K_{t+1}=T(S_t)$. The technology may include natural laws that determine the evolution of cosmic factors.

The set of feasible actions A_t at t is delimited by the constraint $S_t \in \Phi$. The set Φ can either describe technical possibilities or also include certain political and behavioral constraints, such as difficulties to change lifestyles or reduce inequalities.[9] Note that if one is worried that a finite horizon means disregarding what happens after H, it is possible to incorporate in Φ the constraint that $K_{H+1}=T(S_H)$ must be above a minimum threshold.

Social welfare for generation t, denoted $W(S_t) \in \mathbb{R}$, may depend on all actions and also on capital stocks. Note that it only depends on the situation of the generation, not on the welfare of subsequent generations. While altruism is possible, it is better treated as an ethical issue in the definition of intergenerational welfare than in the definition of generational welfare.[10]

It is assumed throughout the paper that the functions T and W are continuous and that the set Φ is compact.

A feasible path from t on is a sequence S_t^+ such that K_t takes an exogenous value and for all τ such that $t \le \tau \le H$,

$$K_{\tau+1}=T(S_\tau),$$
$$S_\tau \in \Phi.$$

Let the set of feasible paths from t be denoted $\mathbf{\Phi}_t(K_t)$. The dynamic process is autonomous but as K_t is a comprehensive notion and T may include natural laws governing cosmic factors, it is compatible with exogenous factors making survival on Earth hard (e.g., when solar radiation increases, or meteors crash) independently of human activity.

The following two properties will be useful in the derivation of certain results in the next sections.

Pareto efficiency of a path S_0^+ means that it is not possible to improve the welfare of all generations without hurting some: it is not possible to find another path $S_0'^+$ such that $W(S_t') \ge W(S_t)$ for all $t=0,\dots,H$ and $W(S_t') > W(S_t)$ for some t.

Transferability means that it is always possible to increase (decrease) the welfare of the present generation at the expense (benefit) of another generation (if they are above the minimum possible): for every S_0^+, every t^*, such that $W(S_0) > \min W(\Phi)$ and $W(S_{t^*}) > \min W(\Phi)$, every $\varepsilon > 0$, there exists $S_0'^+$ such that $\varepsilon > W(S_{t^*}') - W(S_{t^*}) > 0 > W(S_0') - W(S_0) > -\varepsilon$ and for all $t \ne 0, t^*$, $W(S_t') = W(S_t)$; and there exists $S_0''^+$ such that $-\varepsilon < W(S_{t^*}'') - W(S_{t^*}) < 0 < W(S_0'') - W(S_0) < \varepsilon$ and for all $t \ne 0, t^*$, $W(S_t'') = W(S_t)$.

This condition is realistic but not satisfied in models in which capital cannot be consumed, so that when a generation does not save anything it is impossible to improve its fate at the expense of the future. With transferability, one obtains that a maximin path is always regular in the sense of having constant welfare.

Lemma 1. *For all t, all K_t, the set*

$$\Lambda_t(K_t) = \{\lambda \in \mathbb{R} \mid \exists S_t^+ \in \mathbf{\Phi}_t(K_t), \forall \tau \ge t, W(S_\tau) \ge \lambda\}$$

is closed and bounded from above. Under transferability, its maximum λ^* is obtained for S_t^+ such that $W(S_\tau) = \lambda^*$ for all $\tau \ge t$.

Proof. Continuity of T, W and compactness of Φ imply that $\mathbf{\Phi}_t(K_t)$ is compact for all K_t.[11] Therefore the set

$$B = \left\{ w \in \mathbb{R} \mid \exists S_t^+ \in \mathbf{\Phi}_t(K_t), \min_\tau W(S_\tau) = w \right\}$$

is compact. One has $\Lambda_t(K_t) =]-\infty, \max B]$.

Suppose that $\lambda^* = \max \Lambda_t(K_t)$ is such that there is S_t^+ for which $W(S_\tau) = \lambda^*$ for all $\tau \ge t$ and $W(S_{\tau^*}) > \lambda^*$ for some $\tau^* \ge t$. By transferability, it is possible to decrease $W(S_{\tau^*})$ and raise $W(S_\tau)$ for all $\tau \ge t, \tau \ne \tau^*$, thereby raising $\min_\tau W(S_\tau)$ above λ^*, which is impossible.

(Note that the transfer between τ^* and all other $\tau \ge t$ cannot be done in one stroke by transferability. One has to go through generation 0 and make transfers to one other generation at a time.) □

[8] Simulations of integrated assessment models often adopt a finite horizon and analysts discard the end of the scenarios, focusing on the results that do not depend much on the choice of the horizon.

[9] The extinction of the human species can be determined endogenously and is obtained when, given K_t and Φ, the only possible action A_t is inaction A^0. Note that as Φ is constant, all time dependent aspects of technical possibilities, political and behavioral constraints must be described by including appropriate capital stocks.

[10] On the measurement of generational welfare, see Fleurbaey and Blanchet (2013).

[11] When H is infinite, the set $\mathbf{\Phi}_t(K_t)$ may not be closed.

Definition of sustainability and first indicator

Assuming that intergenerational welfare is a function of the welfare of all the generations, it makes sense, in the spirit of the Brundtland definition, to examine the sustainability of generational welfare over time. We will primarily focus on sustaining the inequality $W(S_t) \geq W(S_0)$, which can be described in shorter parlance as "sustaining the current level of welfare."'

Definition 1. The welfare level $W(S_0)$ is sustainable given S_0 if there is a feasible path $S_1^+ \in \mathbf{\Phi}_1(T(S_0))$ such that for all $t \geq 1$, $W(S_t) \geq W(S_0)$.

The following proposition follows directly from this definition.[12] It says that $W(S_0)$ is sustainable given S_0 if and only if it is not greater than the "maximin value" of K_1, defined as the greatest value of $\min_{t \geq 1} W(S_t)$ over all paths made possible by K_1. This maximin value is well defined given Lemma 1.

Proposition 1. *$W(S_0)$ is sustainable given S_0 if and only if $W(S_0) \leq V_1(T(S_0))$, where*

$$V_1(K_1) = \max\left\{\lambda \in \mathbb{R} | \exists S_1^+ \in \mathbf{\Phi}_1(K_1), \forall t \geq 1, \ W(S_t) \geq \lambda\right\}.$$

This fact suggests that the ratio $W(S_0)/V_1(T(S_0))$ is an interesting indicator of (un)sustainability. In the next section, the indicator $1 - V_1(T(S_0))/W(S_0)$ and related notions will be introduced as components of social welfare.

Under transferability, sustaining the inequality $W(S_t) \geq W(S_0)$ is equivalent to sustaining the inequality $W(S_t) \geq W(S_{t-1})$, i.e., a zero growth rate.[13] It is also easy to extend the approach to sustaining other goals (e.g., a positive growth rate), and an adapted version of the function V_1 can be constructed for such cases.[14]

Other indicators and sustainability tests

As noted by Cairns and Martinet (2012), who allude to a similar ratio in the continuous-time model, the ratio $W(S_0)/V_1(T(S_0))$ bears similarity with the ecological footprint. However, something closer to the land-accounting form of the original footprint indicator (Wackernagel and Rees, 1995 for critical discussion, see Dietz and Neumayer, 2007; Mori and Christodoulou, 2012; Fleurbaey and Blanchet, 2013) can be imagined. Assume that K_t is divided in two components $\left(K^a, K^b\right)$, with K^a the relevant part for footprint evaluation (e.g., land). By the maximum theorem, $V_t(K_t)$ is continuous. If both $V_t(K)$ and $T^a(A,K)$ (the latter is the K^a part of $T(S)$) are increasing in K^a, then one can seek λ such that $W(S_0) = V_1\left(T\left(A_0, \left(\lambda K_0^a, K_0^b\right)\right)\right)$. Sustainability is then equivalent to $\lambda \leq 1$, because the latter inequality is then equivalent to $W(S_0) \leq V_1(T(S_0))$. The quantity λ is like the ecological footprint. It is the proportion of the current stock of capital K_0^a that would be needed to make $W(S_0)$ sustainable, given S_0. Note that, unlike the ecological footprint, it incorporates the possible future evolution of capital stocks (including technological know-how).

Another relevant magnitude for the measurement of sustainability is the maximin value that starts at $t=0$:

$$V_0(K_0) = \max\left\{\lambda \in \mathbb{R} | \exists S_0^+ \in \mathbf{\Phi}_0(K_0), \forall t \geq 0, \ W(S_t) \geq \lambda\right\}.$$

In a continuous time setting, there is no distinction between $V_0(K_0)$ and $V_1(K_1)$. Here the distinction is substantial. The inequality $W(S_0) \leq V_0(K_0)$ does not ensure sustainability in general. However, the following shows that it is a necessary condition.[15]

Proposition 2. *For all S_0 such that*

$$W(S_0) > V_0(K_0),$$

$W(S_0)$ is not sustainable given S_0.

Proof. By definition, $V_0(K_0) \geq \min\left\{W(S_0), V_1(T(S_0))\right\}$. If $W(S_0) > V_0(K_0)$, necessarily $V_0(K_0) \geq V_1(T(S_0))$, and therefore $W(S_0) > V_1(T(S_0))$, implying unsustainability given S_0. □

A final result provides a way to use both $V_0(K_0)$ and $V_1(T(S_0))$ to ascertain sustainability.

---

[12] The definition sets the sustainability problem as a viability problem: are current decisions compatible with keeping future welfare within a certain range? On viability, see the references provided in Section "Definitions of sustainability in the literature".

[13] Sustaining the latter obviously implies sustaining the former. Conversely, under transferability, by Lemma 1, sustaining the former can be done by a path of constant welfare level.

[14] If one wants to sustain the growth rate g for $W(S_t)$, one can simply compute

$$\hat{V}_1(K_1) = \max\left\{\lambda \in \mathbb{R} | \exists S_1^+ \in \mathbf{\Phi}_1(K_1), \forall t \geq 1, \ W(S_t) \geq (1+\lambda) W(S_{t-1})\right\}.$$

and compare it to g.

[15] It is a sufficient condition under very restrictive assumptions (see Fleurbaey, 2013 for details).

Proposition 3. *$W(S_0)$ is sustainable given S_0 if $V_0(K_0) < V_1(T(S_0))$. Under transferability, $W(S_0)$ is also sustainable given S_0 if $V_0(K_0) = V_1(T(S_0))$.*[16] *On the efficiency frontier for the pair $(W(S_0), V_1(T(S_0)))$, $W(S_0)$ is unsustainable given S_0 if $V_0(K_0) > V_1(T(S_0))$.*

Proof. As $V_0(K_0) \geq \min\{W(S_0), V_1(T(S_0))\}$, if $V_0(K_0) < V_1(T(S_0))$, then $W(S_0) \leq V_0(K_0)$, therefore $W(S_0) < V_1(T(S_0))$, implying sustainability.

Now assume that $V_0(K_0) = V_1(T(S_0))$. By definition, there is a feasible path S_1^+ such that $W(S_t) \geq V_1(T(S_0))$ for all $t \geq 1$, and thus $W(S_t) \geq V_0(K_0)$ for all $t \geq 1$. By transferability and Lemma 1, this requires $W(S_0) \leq V_0(K_0)$. As a consequence, since $V_0(K_0) = V_1(T(S_0))$ one has $W(S_0) \leq V_1(T(S_0))$, implying sustainability.

On the efficiency frontier for $(W(S_0), V_1(T(S_0)))$,

$$V_0(K_0) \leq \max\{W(S_0), V_1(T(S_0))\}.$$

This inequality is due to the fact that by definition, there exists a path $S_0'^+$ such that $W(S_t') \geq V_0(K_0)$ for all $t \geq 0$. On this path, one has $V_1(T(S_0')) \geq V_0(K_0)$. Therefore it would be below the efficiency frontier for $(W(S_0), V_1(T(S_0)))$ to have $W(S_0) < V_0(K_0)$ and $V_1(T(S_0)) < V_0(K_0)$.

If $V_0(K_0) > V_1(T(S_0))$, then (on the efficiency frontier) necessarily $W(S_0) \geq V_0(K_0)$, implying $W(S_0) > V_1(T(S_0))$, therefore unsustainability. □

The test $V_0(K_0) \leq V_1(T(S_0))$ can be used to construct a variant of the genuine savings indicator, as explained in the Appendix. There are two differences between this variant and the standard notion of genuine savings. First, the V_0, V_1 functions rely on the maximin rather than the discounted sum of utilities. Second, this variant does not assume that the growth path is optimal or even follows a particular plan or program (details on these issues are to be found in the Appendix).

Mapping out sustainability configurations

The results of the previous subsection allow us to identify the following configurations (equality cases are ignored, and simplified notations are adopted): (Table 1)

Intuitively, one can judge case 1 as a case of good (but not necessarily optimal) management, whereas case 4 achieves sustainability in a dubious way, because the present generation hurts itself more than it hurts the possibilities of future generations.[17]. Case 5 is a case of very bad management, as the future is harmed even more than the present generation. Case 6 suggests a selfish management in which the present generation deprives future generations from possibilities for its own benefit.

The impossibilities identified in cases 2 and 3 are due to the following observation.

Lemma 2. *The point $(V_0(K_0), V_0(K_0))$ is on the efficiency frontier*[18] *for the pair $(W(S_0), V_1(T(S_0)))$.*

Proof. The inequality

$$V_0(K_0) \leq \max\{W(S_0), V_1(T(S_0))\}$$

implies that $(V_0(K_0), V_0(K_0))$ is not above the frontier.

Moreover, it cannot be below the frontier since a path with $W(S_0) = V_1(T(S_0)) > V_0(K_0)$ would contradict the definition of $V_0(K_0)$. □

Fig. 1, which refines Fig. 1 from Howarth and Norgaard (1992, p. 474—only two generations are featured in their figure), illustrates the six cases listed in the previous subsection and summarizes the previous results. It also relies on Lemma 2.

The above list of cases can be compared to the table drawn by Cairns and Martinet (2012) for continuous time, which examines the possible combinations of $W_0 \gtrless V$ and $dV/dt \gtrless 0$. They show in particular that $W_0 > V$ and $dV/dt > 0$ are not compatible, but all other combinations are possible (though some, like $W_0 = V$ and $dV/dt > 0$, or $W_0 > V$ and $dV/dt = 0$, correspond to non-regular cases, which are due to non-transferability or to the infinite horizon). There is no equivalent to case 5 in continuous time.

As briefly alluded to in Section “Definitions of sustainability in the literature”, from Doyen and Martinet (2012), one can derive criteria of sustainability for the continuous case (and an infinite horizon). Let us say that $W(S_0)$ is sustainable given S_0 in continuous time if there exists a feasible path $(S_t)_{t \geq 0}$ such that $W(S_t) \geq W(S_0)$ for all $t \geq 0$.[19] Let $V(K_t)$ denote the

[16] This result is lost under an infinite horizon because with a non-regular maximin one may have $W(S_0) > V_0(K_0) = V_1(T(S_0))$.

[17] A war could illustrate this situation. It destroys capital, and hurts the present generation, but reconstruction is possible afterwards. A less dramatic illustration is provided in Section “Example”.

[18] This is the “Weak Pareto” efficiency frontier. Under non-transferability, or under an infinite horizon, the frontier may be horizontal at $(V_0(K_0), V_0(K_0))$, in which case there exists a feasible $(W(S_0), V_1(T(S_0)))$ such that $W(S_0) > V_0(K_0)$ and $V_1(T(S_0)) = V_0(K_0)$.

[19] Doyen and Martinet's definition is different, which is why some caution must be taken to make use of their results. In their definition, $W(S_0)$ is sustainable if there exists a feasible path $(S_t')_{t \geq 0}$ such that $W(S_t') \geq W(S_0)$ for all $t \geq 0$. But they also check whether a feasible path starting from S_0 can achieve the goal.

Table 1
Sustainability configurations.

Case 1: $W_0 < V_0 < V_1$	sustainable
Case 2: $V_0 < W_0 < V_1$	impossible
Case 3: $V_0 < V_1 < W_0$	impossible
Case 4: $W_0 < V_1 < V_0$	sustainable
Case 5: $V_1 < W_0 < V_0$	unsustainable
Case 6: $V_1 < V_0 < W_0$	unsustainable

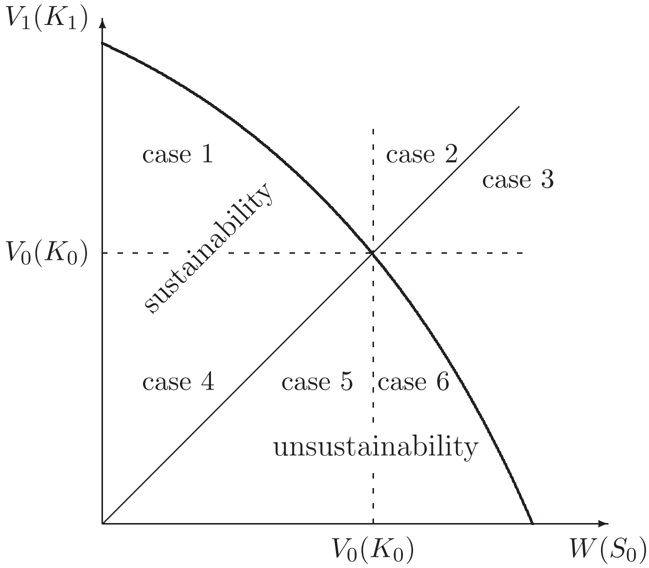

Fig. 1. Sustainability areas in the $(W(S_0), V_1(K_1))$ space.

maximum sustainable level of welfare from t on, given K_t. One then has that $W(S_0)$ is sustainable given S_0 if $W(S_0) \le V(K_0)$ and $dV/dt \ge 0$; $W(S_0)$ is not sustainable given S_0 if $W(S_0) > V(K_0)$ or if $W(S_0) = V(K_0)$ and $dV/dt < 0$. Note that by Cairns and Martinet's (2012) result, $dV/dt > 0$ is sufficient to guarantee sustainability of $W(S_0)$ given S_0 (because it implies $W(S_0) \le V(K_0)$ and $dV/dt \ge 0$), as in the discrete time setting. But $dV/dt \ge 0$ is, as in discrete time, neither sufficient nor necessary for sustainability.[20]

The main difference between discrete time and continuous time is therefore that the necessary and sufficient condition $W(S_0) \le V(K_1)$ has no equivalent in continuous time. This condition appears particularly useful for incorporating an indicator of sustainability into an intergenerational welfare function. Moreover, the ratio $W(S_0)/V(K_1)$ evaluates the magnitude of (un) sustainability.

Example

The notions of this section can be illustrated in a simple discrete-time version of the Dasgupta–Heal–Solow–Stiglitz model (similar to Dasgupta and Mitra, 1983).

There is one individual per generation, one produced good which can be used as consumption c or capital k, and an exhaustible resource s.

Let $\Delta k_{t+1} = k_{t+1} - k_t$ and $\Delta s_{t+1} = s_{t+1} - s_t$. The law of motion is $\Delta k_{t+1} = f(k_t, -\Delta s_{t+1}) - c_t$, with constraints $c_t \ge 0, k_t \ge 0, s_t \ge 0, \Delta s_{t+1} \le 0$.

In the notations of our model, $A_t = (c_t, \Delta k_{t+1}, \Delta s_{t+1})$, $K_t = (k_t, s_t)$,

$$\Phi = \left\{ ((c, \Delta k, \Delta s), (k, s)) \in \mathbb{R}^5 \;\middle|\; \begin{array}{c} c \ge 0, \\ -k \le \Delta k \le f(k, -\Delta s) - c, \\ -s \le \Delta s \le 0 \end{array} \right\},$$

$$T((c, \Delta k, \Delta s), (k, s)) = (k + \Delta k, s + \Delta s).$$

We define generational welfare as a function of consumption and stocks: $W((c, \Delta k, \Delta s), (k, s)) = U(c, s)$. We assume that $U(0, s) = \min U(\mathbb{R}^2_+)$. This makes the model satisfy transferability.

[20] Note also that $W(S_0) < V(K_0)$ is not a sufficient condition for sustainability.

By Lemma 1, a maximin path is regular, therefore one can compute $V_1(K_1)$ by solving the following program:

$$\max_{(k_t,s_t)_{t=2,\ldots,H}} U(f(k_H,s_H)+k_H,s_H)$$

under the constraint that for all $t=1,\ldots H-1$,

$$U(f(k_H,s_H)+k_H,s_H)=U\left(f\left(k_t,-\Delta s_{t+1}\right)-\Delta k_{t+1},s_t\right). \quad \text{(multiplier } \lambda_t\text{)}$$

With the notations $U_c^t=\partial U(c_t,s_t)/\partial c_t$, $U_s^t=\partial U(c_t,s_t)/\partial s_t$, $f_k^t=\partial f\left(k_t,-\Delta s_{t+1}\right)/\partial k_t$, $f_s^t=\partial f\left(k_t,-\Delta s_{t+1}\right)/\partial s_t$, the FOCs read

$$(1-\lambda_{H-1})U_c^H\left(f_k^H+1\right)=\lambda_{H-1}U_c^{H-1}$$
$$(1-\lambda_{H-1})\left(U_c^Hf_s^H+U_s^H\right)=\lambda_{H-1}U_c^{H-1}f_s^{H-1}$$
$$\vdots$$
$$\lambda_tU_c^t\left(f_k^t+1\right)=\lambda_{t-1}U_c^{t-1}$$
$$\lambda_t\left(U_c^tf_s^t+U_s^t\right)=\lambda_{t-1}U_c^{t-1}f_s^{t-1}$$

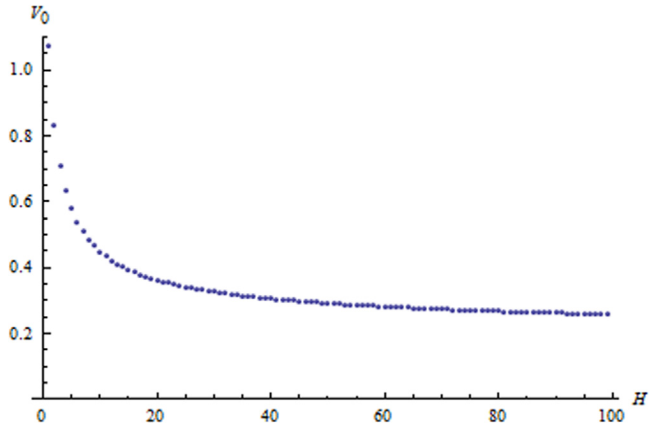

Fig. 2. $V_0(K_0)$ as a function of H, for $K_0=(1,1)$.

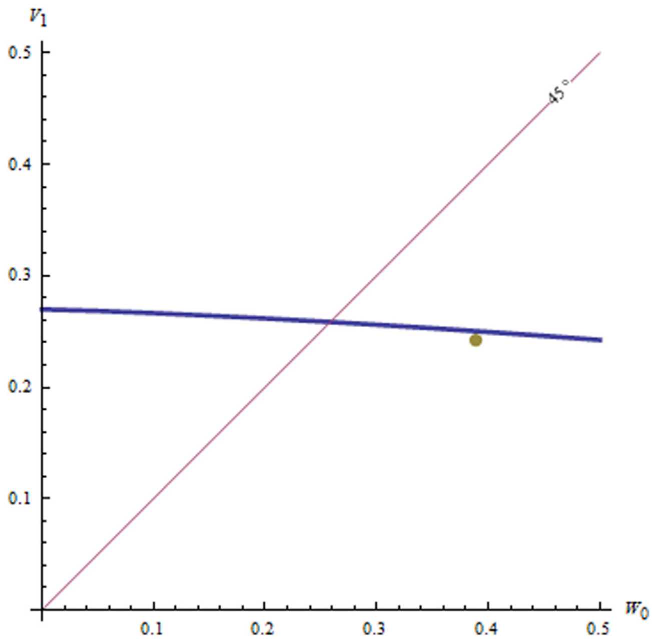

Fig. 3. The (W_0,V_1) possibility set, when $K_0=(1,1)$ and $H=100$.

implying the general condition for $t=2,\ldots,H$,

$$\frac{U_c^t\left(f_k^t+1\right)}{U_c^{t-1}}=\frac{U_c^t f_s^t+U_s^t}{U_c^{t-1}f_s^{t-1}},$$

which also reads as an augmented Hotelling rule:

$$f_k^t=\frac{f_s^t}{f_s^{t-1}}-1+\frac{U_s^t}{U_c^t f_s^{t-1}}.$$

Consider the following economy: $f(k,-\Delta s)=k^{0.5}(-\Delta s)^{0.4}$, $U(c,s)=c^{0.8}s^{0.2}$.

Fig. 2 illustrates how $V_0(1,1)$ depends on H, and shows that for a large H, the value of $V_0(K_0)$ does not depend much on H.

This curve can be used to answer the question: "for how long is the welfare level w^* sustainable?" For instance, for $w^*=.4$, at most 14 periods can sustain this level. For $w^*=.3$, the number of periods goes up to 43. This illustrates how the horizon can be used as a parameter in policy debates.

Let us now illustrate the computation of the sustainability indicators. Imagine that, endowed with $K_0=(k_0,s_0)=(1,1)$, generation 0 consumes $c_0=0.3$ and extracts $-\Delta s_1=0.1$, leaving $K_1=(k_1,s_1)=(1.10,0.9)$. One then has $W_0=U(c_0,s_0)=.382$, $V_0=.258$ and $V_1=.249$. This is a clear case of unsustainability (case 6 in the table of the previous section): $V_1<V_0<W_0$. The unsustainability indicator W_0/V_1 is equal to 1.53. The ecological footprint computed as the fraction of s_0 needed, given current actions $(c_0,\Delta k_1,\Delta s_1)$, to achieve sustainability is equal to 1.88.

Fig. 3 shows the application of Fig. 1 for this example. The (W_0,V_1) feasible set is quite asymmetric in favor of the present generation, due to the presence of exhaustible resources and to the large number of generations (H=100). The unsustainable point $(W_0,V_1)=(.382,.249)$, corresponding to the case highlighted in the previous paragraphs, is shown on the figure and is close to the frontier.

Now imagine that, compared to the case illustrated in the figure, generation 0 halves its consumption $c_0=0.15$ and double its extraction $-\Delta s_1=0.2$, leaving $K_1=(k_1,s_1)=(1.38,0.8)$. One then has $W_0=.219$, $V_0=.258$ and $V_1=.246$. This is a case of "dubious" sustainability (case 4 in the table): $W_0<V_1<V_0$.

Finally, let us briefly compare with the approach that assesses sustainability by the condition dV/dt where V is (discounted utilitarian) intergenerational welfare. In this model, it would examine the sign of

$$\sum_{t=1}^{\infty}\left(\frac{1}{1+\delta}\right)^{t-1}U(c_t,s_t)-\sum_{t=0}^{\infty}\left(\frac{1}{1+\delta}\right)^{t}U(c_t,s_t)=\sum_{t=0}^{\infty}\left(\frac{1}{1+\delta}\right)^{t}(U(c_{t+1},s_{t+1})-U(c_t,s_t)).$$

This requires information about the future path that is not needed for the indicators proposed in this paper. It is possible to estimate the sign of this expression from current data only under restrictive assumptions (i.e., assuming that the society follows a program determining the future path on the basis of the current stocks of capital).

Moreover, the sign of the expression can be positive even when W_0 is not sustainable in the sense adopted in this paper. (However, when it is negative one must have $U(c_{t+1},s_{t+1})<U(c_t,s_t)$ for some t.) The adaptation of this approach to a finite horizon and a more detailed discussion and comparison to the approach of this paper are provided in the Appendix on genuine savings.

Sustainability as a component of intergenerational welfare

We are now ready to examine how to make sustainability a component of intergenerational welfare. To simplify notations, let W_t denote $W(S_t)$ and V_t denote $V_t(K_t)$.

Sustainability and social welfare in the literature

There is a long tradition that seeks to avoid discounting future welfare and studies equitable criteria such as the maximin (Solow, 1974), or Ramsey's (1928) undiscounted sum, von Weizsäcker' (1965) overtaking criterion, Chichilnisky's (1996) additive criterion, and Asheim et al. 's (2012) sustainable recursive social welfare functions. In order to give an imperative force to the sustainability "obligation", one can in particular give absolute priority to the future when "it" (in some sense) is worse-off than the present (Asheim et al., 2012). The aim of this literature is primarily to come up with social objectives the maximization of which will produce nice development paths. But it is not easy to relate its results to the notion of sustainability studied here, in particular taking account of the fact that sustainability is about future possibilities, not actual outcomes, and has to be applicable in suboptimal contexts.

Moreover, these social welfare functions do not feature sustainability indicators as components of social welfare. Consider for instance the recursive intergenerational welfare function $F(W_0,F(W_1,\ldots))$ proposed by Asheim et al. (2012), where the F function satisfies the property $F(w,z)=z$ if $w\geq z$, reflecting the absolute priority of maintaining "the future" at the level of "the present". This property implies that an optimal policy will sustain welfare, whenever this is possible. But the formula $F(W_0,F(W_1,\ldots))$ does not reflect what the future generations *can* sustain, and the term $F(W_1,\ldots)$ is in general greater than $\min\{W_1,W_2,\ldots\}$, so that $F(W_0,F(W_1,\ldots))$ does not even clearly record the downturns that will occur in the future, unless they occur at the very end of history.

Another possibility is to use sustainability criteria as a constraint for policy decisions. The path may then be obtained from maximizing any social objective (including an objective that gives low priority to the future), or any other means.[21] One can interpret maximizing a social objective under a sustainability constraint as tantamount to maximizing a lexicographic objective in which achieving sustainability has absolute priority over the second-tier objective. This automatically introduces a sustainability indicator in the social objective, at the top place. But this is a rather extreme approach. In this section we explore more moderate options.

An exclusive sustainability focus?

Let us first examine an approach that only evaluates the future in terms of sustainability (i.e., ability of the future generations) and does not care about the future actual outcomes. It considers that it is enough to provide future generations with the means to have a good level of welfare. If instead they decide to make a special sacrifice for their descendants, this would not influence the current evaluation of the future. This would echo Solow's admonition: "we don't know what they will do, what they will like, what they will want. And, to be honest, it is none of our business" (1991, p. 182). This formulation is rather extreme, because if we do not know what future generations will like, we cannot even compute V_1. But it suggests an approach in which the current objective would focus on future opportunities rather than future outcomes.

One could think of letting intergenerational welfare be simply defined as a function $F(W_0, V_1)$, so that social welfare does not depend on what future generations will do but only on what they can do, and more specifically on the welfare level they are able to sustain.

This approach may generate an inefficient path if all generations seek to maximize the same objective $F(W_t, V_{t+1})$, because each generation bequeaths capital to the next generations for a maximin path (corresponding to V_{t+1}) that they may not follow. This inefficiency problem is avoided for sure only when F is the maximin criterion min $\{W_0, V_1\}$. One might perhaps also argue that potential inefficiency is the price to pay for a focus on opportunities rather than outcomes.

At any rate, let us note that if a function $F(W_0, V_1)$ is adopted, it is possible to decompose intergenerational welfare as follows:

$$\begin{aligned} F(W_0, V_1) &= F(W_0, W_0) + F(W_0, V_1) - F(W_0, W_0) \\ &= F(W_0, W_0)\left(1 + I^s - I^u\right), \end{aligned}$$

where indexes of sustainability and unsustainability are defined as

$$I^s = \max\left\{0, \frac{F(W_0, V_1)}{F(W_0, W_0)} - 1\right\},$$
$$I^u = \max\left\{0, 1 - \frac{F(W_0, V_1)}{F(W_0, W_0)}\right\}.$$

The indexes I^s and I^u have intuitive interpretations. When $I^u > 0$, it is equal to the fraction of intergenerational welfare that is lost due to the fact that future generations cannot sustain W_0 but only $V_1 < W_0$.

This approach is quite straightforward but the exclusive focus on sustainability in this definition of intergenerational welfare may be criticized in at least two ways. Both criticisms revolve around the focus on the opportunities of future generations.

First, it is not clear that it is fair to ignore the future generations' likely fate and only focus on their opportunities. This is an instance of collective responsibility that is even more dubious than individual responsibility. In the case of individual responsibility, one can save the idea of responsibility by relying on freedom and respecting preferences as an important foundation for letting individuals live what they want rather than catering to their welfare in a paternalistic way (Fleurbaey, 2008). But in the case of the collective of all future generations, it is quite farfetched to judge that we have done enough if we give them the means to achieve a great future, even if we can forecast that they will waste a lot of social welfare, just as our ancestors and ourselves have done so far by tolerating tremendous inequalities.[22]

Moreover, the measure of opportunities that is retained in the sustainability approach is especially questionable, because it is the maximum value that the responsible agent is able to obtain, and the presence of lower values in the opportunity set does not reduce the value of opportunities. One can also worry about opportunities being measured by the maximum value of the maximin criterion, when future generations may not share the same goal.[23] So this is not even the "elementary" evaluation of opportunity set discussed by Sen (1985), which consists in measuring the value of the opportunity set by the maximum utility the agent can obtain. Here the utility is a measure of opportunities by V_1 that pertains to an evaluator who just cares about the ability of future generations to sustain a certain level, rather than a more general notion of opportunities to allocating welfare among future generations in various ways, and this may not suit the collective agent (i.e., future generations) receiving these opportunities. For instance, if they are interested in growth, it may be more appropriate to focus on the sustainability of a growth rate, as defined earlier.

[21] Martinet (2012) argues against taking sustainability as a constraint, because if no efficient path achieves sustainability, the sustainability constraint induces inefficient choices.

[22] With an inequality averse generational welfare function W, the path corresponding to V_1 involves drastic reductions of inequalities.

[23] I thank G. Asheim for having raised this point.

Sustainability as a component of intergenerational welfare

Let us therefore, pace Solow (1991), come back to the classical approach to social welfare and examine if it is possible to make a sustainability component appear in a decomposition of intergenerational welfare.

Let $F(W_0, W_1, \ldots, W_H)$ be an intergenerational welfare function that can be quite general except that it evaluates the situation of each generation independently of the fate of the other generations. This approach makes it possible to take account of population change if this affects generational welfare, and it can accommodate the risk of extinction before H. No recursive form is imposed on F, and F may actually depend on the past history, though this can be kept implicit as the past history is fixed.

Let us assume that F satisfies the following principle, which says that social welfare is not harmed by progressive transfers (of welfare W) between generations:

Pigou–Dalton principle of transfer: $F(W'_0, W'_1, \ldots, W'_H) \geq F(W_0, W_1, \ldots, W_H)$ if $(W'_0, W'_1, \ldots, W'_H)$ is obtained from $(W_0, W_1, \ldots, W_H)$ by a transfer $\Delta > 0$ between generations t, t', i.e., $W'_t = W_t + \Delta \leq W_{t'} - \Delta = W'_{t'}$, while $W'_s = W_s$ for all $s \neq t, t'$.

Let the value of social welfare when future generations barely enjoy the sustainability opportunity as measured by V_1 be denoted:

$$F_0^* = F(W_0, V_1, \ldots, V_1).$$

This expression can be understood as the intergenerational welfare evaluation of the sustainability opportunities left to future generations, in view of the current level W_0. If future generations eventually decide to adopt a path in which some generations are below V_1, this does not affect F_0^*. This is akin to the approach explored in the previous subsection, except that here it will just be one component of intergenerational welfare.

We will assume that F is an equally distributed equivalent (EDE), so that $F(w, \ldots, w) = w$ (W_t may itself be an EDE over the generation's population). When the assumption that F is an EDE is not satisfied, the analysis of this subsection can be easily adapted but the formulae are less simple.

The following lemma provides an approximation formula for F_0^*, which implies that when $W_0 \geq V_1$, one has $F_0^* \simeq V_1$:

Lemma 3. *Under the EDE and Pigou–Dalton assumptions, if H is large one has*

$$F_0^* \simeq F(\min\{W_0, V_1\}, V_1, \ldots, V_1).$$

Proof. Under Pigou–Dalton, the EDE is always between the minimum and the average of the components of $(W_0, W_1, \ldots, W_H)$, so that one has

$$\min\{W_0, V_1\} \leq F(W_0, V_1, \ldots, V_1) \leq \frac{W_0 + HV_1}{H+1},$$

and when H is large, one has $(W_0 + HV_1)/(H+1) \simeq V_1$.

When $W_0 \geq V_1$, one therefore has $F(W_0, V_1, \ldots, V_1) \simeq V_1 = F(\min\{W_0, V_1\}, V_1, \ldots, V_1)$.

When $W_0 \leq V_1$, one directly has $F(W_0, V_1, \ldots, V_1) = F(\min\{W_0, V_1\}, V_1, \ldots, V_1)$. □

This lemma mimics the reasoning made in Asheim et al. (2012) in favor of Hammond Equity for the Future, an axiom that gives absolute priority to the future when all generations are less well-off than the current one, and H is infinite. When $W_0 > V_1$, equalizing between the current generation at W_0 and all future generations at V_1 produces a new distribution in which the level is close to V_1 when there are many future generations (they have to share the gift from W_0 among many recipients).

We finally obtain the desired decomposition of intergenerational welfare:

Proposition 4. *Let*

$$I^s = \max\left\{0, \frac{F_0^*}{W_0} - 1\right\} \quad \text{and} \quad I^u = \max\left\{0, 1 - \frac{V_1}{W_0}\right\}$$

measure the relative surplus of welfare above sustainability, and the relative deficit below sustainability. One then has, when H is large,

$$F(W_0, W_1, \ldots) \simeq W_0(1 + I^s - I^u) + (F - F_0^*). \tag{1}$$

Proof. One can write:

$$\begin{aligned} F(W_0, W_1, \ldots) &= W_0 + (F_0^* - W_0) + (F - F_0^*) \\ &= W_0\left[1 + \max\left\{0, \frac{F_0^*}{W_0} - 1\right\} - \max\left\{0, 1 - \frac{F_0^*}{W_0}\right\}\right] + (F - F_0^*). \end{aligned}$$

By Lemma 3, if $W_0 \geq F_0^*$, i.e., when $W_0 \geq V_1$, one has $F_0^* \simeq V_1$ when H is large. □

This decomposition can be compared to the mean-inequality decomposition that is classical since Atkinson (1970). This classical decomposition measures social welfare as $m(1-I)$, where m is average utility and I an inequality index.

For instance, an inequality index equal to .3 means that social welfare could be as great with just 70% of the current total utility if it were equally distributed.

Here, we similarly obtain lessons about intergenerational welfare in which inequality concerns take the special form of sustainability concerns. Consider an unsustainable situation as in the numerical example of Section "Mapping out sustainability configurations", in which $1-(V_1/W_0)=.348$, so that

$$F \simeq W_0(1-.348)+W_0\frac{F-F_0^*}{W_0}.$$

Suppose to fix ideas that $(F-F_0^*)/W_0=-.2$. Then one can say that, compared with a stationary path $(W_0,\ldots,W_0)$, the contemplated path is like a lower stationary path due to two factors. First, about 35% is lost due to unsustainability. Second, an additional 20% is lost due to the fact that the path will not even achieve the potential given by V_1.

Conversely, with a sustainable path, one can say that it is like a greater stationary path than $(W_0,\ldots,W_0)$ by $100I^s$%, and further modified (up or down) by $100(F-F_0^*)/W_0\%$. Note that this last term may be positive when future generations do not just follow V_1 but do something better by pursuing a growth path in which the sacrifice of earlier generations is more than redeemed, according to F, by the benefits enjoyed by later generations. While the I^s, I^u indexes capture some aspects of inequalities between generations, the $(F-F_0^*)$ term is also sensitive to inequalities between (and within, via W) generations and catches everything about inequalities that is not encapsulated in the sustainability indicators.

Sustainability and risk

Assuming perfect information for the analysis of sustainability is utterly unrealistic and, as a result, one may doubt about the practical relevance of the analysis presented so far in this paper. It is therefore essential to extend it to the context of risk.

Sustainability and risk in the literature

In the literature, three slightly different approaches to sustainability under risk can be distinguished. First, Baumgärtner and Quaas (2009) and Martinet (2011), inspired by the viability approach, define sustainability in a stochastic framework as the requirement that a sustainable trajectory (i.e., a trajectory keeping the system in a certain range) exists with a sufficiently high probability . In our model, assuming some uncertainty about the technology T, this would mean computing the probability that $V_1 \geq W_0$, and saying that sustainability is achieved if this probability is above a certain threshold (e.g., 95%).

The second approach, proposed in De Lara and Martinet (2009) and De Lara et al. (2015), is a different adaptation of the viability approach to a stochastic setting, and introduces rules that determine action A_t as a function $A_t(K_t)$. Assuming again uncertainty about T, for each such rule $(A_t(.))_{t\geq 1}$, one can compute the probability that T is such that the application of this rule enforces $W_t \geq W_0$ for all $t \geq 1$. One can then seek the rule that maximizes this probability, and say that sustainability is achieved if this probability is, once again, above a threshold.

The third approach, due to Asheim and Brekke (2002), similarly introduces a rule, now allowing the action in t to depend on the whole past history: $A_t = A_t(K_t, S_{t-1},\ldots,S_0)$. The path is sustained if at every t, the certainty-equivalent of welfare at $t+1$ is never lower than W_t. The situation is sustainable given S_0, then, if there exists a rule that guarantees this inequality at every t.

Each of these approaches is imperfect and can be improved. The first approach is too permissive because it does not pay attention to the epistemic possibility for the subsequent generations to find the suitable path.[24] For instance, suppose that only two T functions are considered possible, and that generation 0 makes sure that for each of the two functions, there is a path sustaining W_0. Then sustainability of W_0 is guaranteed for sure, according to the first approach. However, if generation 1 does not learn what the true T function is, it may be unable to guess what the appropriate action is at time 1, and its degree of confidence that its action is compatible with sustainability may be much less than 100%.

The second approach avoids this problem because if a rule is assessed at period 0, it is obviously possible for the subsequent generations to follow it. What is missing in the quoted literature is an explicit analysis of learning by future generations, which requires that the rule depends on the whole past history as in the third approach, not just the current stock of capital (unless risk takes the form of a sequence of mutually independent shocks).

An additional, common limitation of the two approaches is that they focus on given sustainability targets such as W_0, and ignore the risk of catastrophes and the probabilities associated with various catastrophes. In the context of risk, it appears insufficient to focus on the probability of sustaining W_0. It may happen, for instance, that a particular strategy implies a greater probability of sustaining W_0 than another strategy, but also a much greater probability of going below $W_0/10$, which may be quite worrisome.

The third approach allows for learning via Bayesian updating, but does not keep record of the probability of downturns and considers that sustainability is achieved even when it is highly possible that the path will pass below W_0 in the future. Indeed, sustainedness is defined as $W_t \leq CE(W_{t+1})$, where CE denotes the certainty-equivalent, and this can be achieved

[24] I thank V. Martinet for having drawn my attention to this issue.

even after a negative shock has induced $W_t < W_0$.[25] Note, however, that the magnitude $CE(W_{t+1})$ takes account of the whole pattern of possible outcomes in the next period, including catastrophes. But this approach does not measure the probability, at time 0, that W_0 (or any other level) will be sustained forever.

In conclusion, we need an approach that records the probabilities of sustaining various levels of welfare (not just W_0 or W_{t-1}), that allows for learning, and assesses whether the future generations are in a position to find the sustainable strategies. Most importantly, this approach must also extend (1) to the context of risk.

Accounting for the risk and extent of unsustainability

Assume that at time 0, there is uncertainty about T. Let $(\mathcal{T}, \Sigma)$ denote a measurable space on the set $\mathcal{T}$ of possible T, and μ a probability measure on the σ -algebra Σ, that represents the beliefs of the evaluator. This is a very general form of uncertainty, which encompasses structural uncertainty about the function but can also capture shocks on capital at every period. Indeed, as time can be a component of capital, the equation $K_{\tau+1} = T(S_\tau)$ can include a specific shock on T at this period.

We will proceed in two steps. In this subsection, we simply assume that, for every possible $T \in \mathcal{T}$, a suitable notion of sustainability of any given welfare level v has been defined. This notion should take account of risks and of the possibility of learning by future generations. It will be the topic of the next subsection.

Let $\mathcal{T}(v; S_0) \subset \mathcal{T}$ denote the subset of technologies T such that, given S_0, the appropriate notion of sustainability of $W_t \geq v$ is satisfied. The precise definition of $\mathcal{T}(v; S_0)$ depends on defining sustainability for the case of risk and is the topic of the next subsection. For the moment, we simply assume that $\mathcal{T}(v; S_0)$ defines the event in which sustainability of $W_t \geq v$ is guaranteed.

We assume that for every v, $\mathcal{T}(v; S_0) \in \Sigma$, and define the function

$$G(v; S_0) = 1 - \mu(\mathcal{T}(v; S_0)).$$

This function measures the probability that the sustainable level is less than v.

One can then define the probabilities of sustainability and unsustainability of W_0 as well as the expected values of any magnitude $h(v)$ conditional on sustainability or unsustainability of W_0:

$$p^s = 1 - G(W_0; S_0), \quad p^u = G(W_0; S_0),$$
$$E^s h = \frac{1}{p^s} \int_{W_0}^{+\infty} h(v) dG(v; S_0), \quad E^u h = \frac{1}{p^u} \int_0^{W_0} h(v) dG(v; S_0),$$

and apply these notions of conditional expected value to $F_0^*(v) = F(W_0, v, \ldots, v)$, in order to obtain a decomposition of welfare into sustainability outcomes and unsustainability outcomes, as follows.

We retain the assumption that F is an EDE.[26] The decomposition of expected intergenerational welfare now reads as follows.

Proposition 5. *Let*

$$I^s = \frac{E^s F_0^*}{W_0} - 1 \quad \text{and} \quad I^u = 1 - \frac{E^u v}{W_0}$$

measure the expected relative surplus of welfare above sustainability, and the expected relative deficit below sustainability,[27] *conditional on either sustainability or unsustainability being obtained. One then has, when H is large,*

$$EF(W_0, W_1, \ldots) \simeq W_0 \left(1 + p^s I^s - p^u I^u\right) + E\left(F - F_0^*\right). \tag{2}$$

Proof. One can write:

$$\begin{aligned} EF(W_0, W_1, \ldots) &= p^s E^s F_0^* + p^u E^u F_0^* + E\left(F - F_0^*\right) \\ &= W_0 + p^s E^s \left(F_0^* - W_0\right) - p^u E^u \left(W_0 - F_0^*\right) + E\left(F - F_0^*\right) \\ &= W_0 \left[1 + p^s E^s \left(\frac{F_0^*}{W_0} - 1\right) - p^u E^u \left(1 - \frac{F_0^*}{W_0}\right)\right] + E\left(F - F_0^*\right). \end{aligned}$$

By Lemma 3, $E^u F_0^* \simeq E^u v$ when H is large. □

This decomposition shows that one should care about p^u, i.e., the probability of unsustainability, but also about the expected gap between W_0 and the sustainable level v (this gap depends itself on the various probabilities of sustaining the levels $v \neq W_0$). In summary, one obtains a rather comprehensive and helpful decomposition: the expected value of

[25] A very simple example illustrates this issue. Suppose that actions have no impact on the path, which is such that, at $t=0$, there is 10% chance of having $W_t = 100W_0$ forever, and 90% chance of having $W_t = W_0/2$ forever. This may well satisfy $W_t \leq CE(W_{t+1})$ for all $t \geq 0$, but there is a 90% chance that $W_t < W_0$ for all $t \geq 1$.

[26] Fleurbaey (2010) advocates computing the expected value of an EDE in the context of risk.

[27] Notation: $E^u v = (1/p^u) \int_0^{W_0} v dG(v; S_0)$.

intergenerational welfare is equal to the welfare of the current generation, corrected by the prospects for sustainability and unsustainability (taking account of probability and magnitude of excess or deficit for each of them), and the expected gap between actual intergenerational welfare and welfare measured by the sustainability opportunities left to future generations. Even though the focus is still on sustaining W_0, it is clear that the probability of sustaining other levels becomes an important element in the evaluation. In particular, a significant probability of a catastrophe raises I^u and dampens the $W_0(1+p^sI^s-p^uI^u)$ term of the decomposition.

Epistemic conditions of sustainability

It remains to define a convincing mapping $\mathcal{T}(v;S_0)$. The first approach mentioned in Section "Sustainability and risk in the literature" would have

$$\mathcal{T}(v;S_0)=\left\{T\in\mathcal{T}\,|\,\exists S_1^+\in\Phi^H,\forall t\geq 1,K_t=T(S_{t-1}),W(S_t)\geq v\right\}.$$

The problem is that the mere existence of a sequence S_1^+ may be known without the future generations being able to find this particular sequence with reasonable confidence.

A more appealing approach, building on the second and third approaches listed in Section "Sustainability and risk in the literature", would focus on how the future generations can follow a sustainable strategy, learning from past history. We therefore must pay attention to the resolution of uncertainty along the path. We will make the specific assumption that the only way future generation t learns about T is by observing the realized values of $T(S_\tau)$, $\tau=0,\dots,t-1$. Generations have a common prior μ (identical of the beliefs of the evaluator) and do Bayesian updating along the path, as they observe the realized $T(S_\tau)$. This Bayesian updating simply consists in restricting μ to the sub-algebra of Σ defined over the subset of $\mathcal{T}$ that is compatible with the observations.[28] That is, if the sequence $h_0^t=(S_0,\dots,S_{t-1},K_t)$ is observed, one must restrict the σ -algebra to the subset

$$\mathcal{T}\left(h_0^t\right)=\{T\in\mathcal{T}\,|\,\forall\tau=0,\dots,t-1,\ K_{\tau+1}=T(S_\tau)\}.$$

Let $\mu\left(B|h_0^t\right)=\mu\left(B\cap\mathcal{T}\left(h_0^t\right)\right)/\mu\left(\mathcal{T}\left(h_0^t\right)\right)$ denote the updated measure conditional on observing the sequence h_0^t.

Here is the most sustainable strategy the future generations can follow, if sustaining $W(S_t)\geq v$ is the goal. It is defined by backward induction.

- (H) Generation H chooses A_H so as to maximize $W(S_H)$. This defines a function $A_H^*\left(h_0^H\right)$.(This generation has nothing to learn, as it observes K_H, which is the only relevant dimension of h_0^H for its own decision.)
- ($H-1$) Generation $H-1$ chooses A_{H-1} so as to maximize
$$\mu\left(\left\{T|K_H=T(S_{H-1}),A_H=A_H^*\left(h_0^H\right),W(S_H)\geq v\right\}|h_0^{H-1}\right)$$
under the constraint that $W(S_{H-1})\geq v$. This defines a function[29] $A_{H-1}^*\left(h_0^{H-1}\right)$.[30]
- ($H-2$) Generation $H-2$ chooses A_{H-2} so as to maximize
$$\mu\left(\left\{T|\forall t\geq H-1,K_t=T(S_{t-1}),A_t=A_t^*\left(h_0^t\right),W(S_t)\geq v\right\}|h_0^{H-2}\right)$$
under the constraint that $W(S_{H-2})\geq v$. This defines a function $A_{H-2}^*\left(h_0^{H-2}\right)$. ⋮
- ($t\geq 1$) Generation t chooses A_t so as to maximize
$$\mu\left(\{T|\forall\tau\geq t+1,K_\tau=T(S_{\tau-1}),A_\tau=A_\tau^*(h_0^\tau),W(S_\tau)\geq v\}|h_0^t\right)$$
under the constraint that $W(S_t)\geq v$. This defines a function $A_t^*\left(h_0^t\right)$.

It is natural to define the probability of sustainability of $W(S_t)\geq v$ as

$$G(v;S_0)=\mu\left(\left\{T|\forall t\geq 1,K_t=T(S_{t-1}),A_t=A_t^*\left(h_0^t\right),W(S_t)\geq v\right\}\right).$$

In words, $G(v;S_0)$ is the probability that $W(S_t)\geq v$ will be sustained when the future generations follow the decision rules $A_t^*\left(h_0^t\right)$.

Another approach, adopted in the quoted literature (e.g., De Lara et al., 2015), starts by considering an arbitrary sequence of decision rules $\left(A_t\left(h_0^t\right)\right)_{t\geq 1}$. One can then define the probability that $W(S_t)\geq v$ will be sustained under these rules:

$$\mu\left(\left\{T|\forall t\geq 1,K_t=T(S_{t-1}),A_t=A_t\left(h_0^t\right),W(S_t)\geq v\right\}\right).$$

[28] Other mechanisms for the resolution of uncertainty, involving exogenous learning, are conceivable, and would slightly complicate the analysis without altering it radically.

[29] The solution to the maximization program need not be unique, but it is of no consequence to pick an arbitrary selection from the optimal actions.

[30] For every generation t, we assume that if the constraint $W(S_t)\geq v$ cannot be met, it maximizes $W(S_t)$ without sustainability concern.

Then one seeks to find the decision rules that maximize this probability. This actually yields the same probability of sustainability, as stated below.[31]

Proposition 6. *One has*

$$G(v;S_0)=\max_{(A_t(.))_{t\geq 1}} \mu\left(\left\{T \mid \forall t\geq 1, K_t=T(S_{t-1}), A_t=A_t\left(h_0^t\right), W(S_t)\geq v\right\}\right).$$

Proof. To save space, let us introduce the notation

$$P\left((A_t(.))_{t\geq 1}\right)=\mu\left(\left\{T \mid \forall t\geq 1, K_t=T(S_{t-1}), A_t=A_t\left(h_0^t\right), W(S_t)\geq v\right\}\right).$$

By definition,

$$G(v;S_0)\leq \max_{(A_t(.))_{t\geq 1}} P\left((A_t(.))_{t\geq 1}\right).$$

Suppose that the inequality is strict and let $(A_t(.))_{t\geq 1}$ be a particular set of rules such that

$$G(v;S_0)<P\left((A_t(.))_{t\geq 1}\right).$$

Consider the revised rule $\left(A_t^H(.)\right)_{t\geq 1}=\left(A_1(.),\ldots,A_{H-1}(.),A_H^*(.)\right)$. Obviously one has

$$P\left(\left(A_t^H(.)\right)_{t\geq 1}\right)\geq P\left((A_t(.))_{t\geq 1}\right).$$

Now consider $\left(A_t^{H-1}(.)\right)_{t\geq 1}=\left(A_1(.),\ldots,A_{H-2}(.),A_{H-1}^*(.),A_H^*(.)\right)$. As the replacement of $A_{H-1}(.)$ by $A_{H-1}^*(.)$ does not affect the probability of having $W(S_t)\geq v$ for $t=1,\ldots,H-2$, and, conditional on every possible history h_0^{H-1}, maximizes the probability of having $W(S_t)\geq v$ for $t=H-1,H-2$, necessarily

$$P\left(\left(A_t^{H-1}(.)\right)_{t\geq 1}\right)\geq P\left(\left(A_t^H(.)\right)_{t\geq 1}\right).$$

By repetition of this argument, one obtains, for $t=2,\ldots,H$:

$$P\left(\left(A_t^{t-1}(.)\right)_{t\geq 1}\right)\geq P\left(\left(A_t^t(.)\right)_{t\geq 1}\right).$$

Therefore one obtains

$$P\left(\left(A_t^1(.)\right)_{t\geq 1}\right)\geq P\left((A_t(.))_{t\geq 1}\right).$$

Now, $\left(A_t^1(.)\right)_{t\geq 1}=\left(A_t^*(.)\right)_{t\geq 1}$, so that $P\left(\left(A_t^1(.)\right)_{t\geq 1}\right)=G(v;S_0)$, which yields a contradiction. □

Note that the algorithm introduced here to compute $\left(A_t^*(.)\right)_{t\geq 1}$ crucially depends on having a finite horizon.

One may criticize this approach for being excessively demanding on the future generations. But this is faithful to the approach to sustainability that has been adopted in this paper. In the absence of risk, sustainability was defined as the mere possibility for future generations to achieve a particular goal, and this is satisfied when the goal is barely achieved and the future generations have to do their best at achieving it (hence the central role of the V function). Similarly here, a possible technology T contributes to the probability of sustainability if the generations that do their best at learning from past history and at maximizing the probability of sustainability after them actually manage to achieve the goal. In the conclusion, perspectives for a less demanding notion of sustainability will be discussed.

Another potential worry is that T contributes to the probability of sustainability even if along the path that T and $\left(A_t^*(.)\right)_{t\geq 1}$ generate, the confidence of some generations about sustaining v after them is low. One might prefer paths in which the confidence about future sustainability remains as high as the initial confidence. Let us briefly explore this idea.

Formally, consider the goal of sustaining $W(S_t)\geq v$ and keeping confidence p about future sustainability. A technology T is then contributing to the probability of sustainability if, along the path that T and $\left(A_t^*(.)\right)_{t\geq 1}$ generate, one has: (i) for all $t\geq 1$, $W(S_t)\geq v$; (ii) for all $t\geq 1$, $\mu(\{T \mid \forall \tau\geq t+1, K_\tau=T(S_{\tau-1}), A_\tau=A_\tau^*(h_0^\tau), W(S_\tau)\geq v\} \mid h_0^t)\geq p$.

Condition (ii) means that along the path, the probability that sustainability is satisfied never goes below p. Let $\mathcal{T}(v,p;S_0)\subset\mathcal{T}$ denote the subset of technologies T such that, along the path that T and $\left(A_t^*(.)\right)_{t\geq 1}$ generate, conditions (i) and (ii) are satisfied. We assume that for all v,p, $\mathcal{T}(v,p;S_0)\in\Sigma$. One has, by definition,

$$G(v;S_0)=\mu(\mathcal{T}(v,0;S_0)).$$

It then seems natural to seek a level of p such that

$$\mu(\mathcal{T}(v,p;S_0))=p,$$

[31] There is a resemblance between this proposition and Prop. 7.10 in De Lara and Doyen (2008). Their model assumes i.i.d. shocks and involves no learning.

which means that the initial confidence in sustainability of v is also sustained along the path generated by a technology T that contributes to the initial confidence.

Proposition 7. *For every v there is a unique value p^* such that $\mu(\mathcal{T}(v,p;S_0)) \geq p^*$ whenever $p < p^*$, and $\mu(\mathcal{T}(v,p;S_0)) \leq p^*$ whenever $p > p^*$. This defines a function $\hat{G}(v;S_0)$ which is non-increasing in v.*

Proof. One has $0 \leq \mu(\mathcal{T}(v,p;S_0)) \leq 1$, so that $\mu(\mathcal{T}(v,p;S_0)) \geq p$ when $p=0$ and $\mu(\mathcal{T}(v,p;S_0)) \leq p$ when $p=1$.

Moreover, $\mu(\mathcal{T}(v,p;S_0))$ is non-increasing in p because condition (ii) becomes harder to satisfy when p rises, implying that the set $\mathcal{T}(v,p;S_0)$ weakly decreases in the sense of inclusion.

This proves that there is a unique value p^* such that $\mu(\mathcal{T}(v,p;S_0)) \geq p^*$ whenever $p < p^*$, and $\mu(\mathcal{T}(v,p;S_0)) \leq p^*$ whenever $p > p^*$.

The function $\mu(\mathcal{T}(v,p;S_0))$ is non-increasing in v because condition (i) becomes harder to satisfy when v rises, implying that the set $\mathcal{T}(v,p;S_0)$ weakly decreases in the sense of inclusion when v increases. This implies that $\hat{G}(v;S_0)$ is also non-increasing in v. □

It is arguable that the function $\hat{G}(v;S_0)$ defined in the proposition, being in general less than $G(v;S_0)$, is too conservative a measure of sustainability. After all, what matters is that future generations do sustain v, not that they never fear failure. $\hat{G}(v;S_0)$ is about sustaining not just a welfare level, but also a confidence level, which is a different, although perhaps legitimate, goal.

Overlapping generations

This section briefly checks that the analysis remains broadly valid when generations overlap. This is essential for any application of this analysis, because generations live about 75 years and overlap, whereas capital stocks and economic activities (consumption, investment) are recorded year by year. The successive-generations model can only be a toy model to make the analysis simpler in the beginning.

Suppose that generation t is born in period t and lives for L periods. So, in every period, L generations are coexisting. The welfare of generation t is now a function $W(S_t, \ldots, S_{t+L-1})$.

The technology constraints are still $S_t \in \Phi$ and $K_{t+1} = T(S_t)$.

In period 0 the generations $-L+1, \ldots, 0$ coexist, so that it is not clear whose welfare is to be the target of sustainability. The current generations' welfare is not even settled by A_0, except for generation $-L+1$. Let us simply generalize and consider any level w that may be considered suitable. It could be the average (predicted) welfare of the currently living people, or the welfare of generation $-L+1$, or any other similar value.

Definition 2. The level w is sustainable given S_0 if there is a feasible path $S_1^+ \in \Phi_1(T(S_0))$ such that for all $t \geq 1$, $W(S_{t-L+1}, \ldots, S_t) \geq w$.

We can still define a function determining the maximum sustainable level from t on:

$$V_t(K_t) = \max\{\lambda \in \mathbb{R} | \exists S_t^+ \in \mathbf{\Phi}_t(K_t), \forall \tau \geq t, W(S_{\tau-L+1}, \ldots, S_\tau) \geq \lambda\}.$$

We therefore obtain the criterion $w \leq V_1(T(S_0))$ as before. This is a criterion that works for any w. In contrast, $V_0(K_0) < V_1(T(S_0))$ is a sufficient condition only for sustainability of $w \leq W(S_{-L+1}, \ldots, S_0)$, the latter figure being the welfare level of the generation dying at the end of period 0.

A natural extension in this context is to evaluate the sustainability associated to a plan of action that extends over several periods.

Proposition 8. *Fix $t \geq 0$. Given a plan $(S_0, \ldots, S_t)$, the level of welfare w is sustainable (for all generations dying from $t+1$ on) if and only if $w \leq V_{t+1}(T(S_t))$, and the level of welfare $w \leq \min\{W(S_{-L+1}, \ldots, S_0), \ldots, W(S_{t-L+1}, \ldots, S_t)\}$ is sustainable if $V_0(K_0) < V_{t+1}(T(S_t))$.*

Proof. Fix $(S_0, \ldots, S_t)$. The inequality $w \leq V_{t+1}(T(S_t))$ means, by definition of $V_{t+1}(T(S_t))$ that there is $S_{t+1}^+ \in \mathbf{\Phi}_{t+1}(T(S_t)), \forall \tau \geq t+1, W(S_{\tau-L+1}, \ldots, S_\tau) \geq w$.

Assume $V_0(K_0) < V_{t+1}(T(S_t))$ and $w = \min\{W(S_{-L+1}, \ldots, S_0), \ldots, W(S_{t-L+1}, \ldots, S_t)\} > V_{t+1}(T(S_t))$. Then there is a feasible S_0^+ such that $\forall \tau \geq 0, W(S_{\tau-L+1}, \ldots, S_\tau) \geq V_{t+1}(T(S_t)) > V_0(K_0)$, contradicting the definition of $V_0(K_0)$. Therefore, if $V_0(K_0) < V_{t+1}(T(S_t))$ then $w \leq V_{t+1}(T(S_t))$. □

The analysis of intergenerational welfare extends directly without substantial modification, once W_t, the argument of the intergenerational welfare function $F(W_0, W_1, \ldots)$, is defined as the welfare of the generation dying at the end of period t.

Conclusion

This paper contains two key messages.

First, it is simple in principle, though much less in practice, to check sustainability defined as the mere *possibility* for future generations to achieve a certain outcome, and this can be done with revised genuine savings and footprint indicators.

The reason why this is actually difficult in practice is that the estimation of the maximin value (V_1) which is used both in genuine savings and in the footprint indicator is hard when the current management is far from the maximin policy.

Second, it is possible to incorporate sustainability indicators in a decomposition analysis of intergenerational social welfare. If one does not feel comfortable with the thought that what future generations will do "is none of our business", then sustainability is not enough and the gap between overall intergenerational welfare and its sustainability part is also relevant, as shown in Eqs. (1) and (2). This makes the sustainability approach all the more relevant on the negative side, because unsustainability then appears especially outrageous. But on the positive side, checking sustainability is then insufficient and forecasting what will actually happen to future generations becomes a relevant concern.

Observe, however, that the set of feasible states of society can incorporate technical constraints but also any kind of political or institutional constraints that one may want to put in the analysis. For instance, it may be politically very hard to control the size of the population, or to curb inequalities. The more constraints are introduced, the smaller the set (in the sense of inclusion), and the harder it is for the present generation to satisfy sustainability, while it becomes easier for future generations to actually sustain what is offered as sustainable to them. In the context of risk, one can similarly introduce bounded memory or bounded rationality in the learning process, which makes it possible to reckon with limited abilities of future generations to maximize the chance of sustaining the target. The concepts of sustainability developed in this paper are therefore quite flexible and can incorporate many concerns about the difficulties that the future generations are likely to face.

Appendix A. Genuine savings revisited

The literature has devoted a great deal of attention to the sign of genuine savings as an indicator of sustainability. For the sake of comprehensiveness and to clarify the links with the literature, this appendix examines how the standard genuine savings approach (based on discounted utilitarianism) is related to sustainability as defined in this paper. The literature adopts an infinite horizon and the adaptation to a finite H is worth examining (with an infinite horizon, the approximation issues discussed in this section vanish).[32]

First, observe that Proposition 3 implies that positive genuine savings at "maximin prices" (defined below) is a very general sufficient (but not necessary) condition for sustainability. Indeed, one has, by approximation to the first order,

$$\begin{aligned} V_1(K_1)-V_0(K_0) &= V_0(K_1)-V_0(K_0)+V_1(K_1)-V_0(K_1) \\ &\simeq p_0(K_1-K_0) \end{aligned}$$

if $p_0=\nabla V_0$ (i.e., the gradient vector—these are the maximin prices) and $V_1(K_1)\simeq V_0(K_1)$. The latter approximation is justified when H does not have a strong impact on V_0 (as illustrated in 2.6, Fig. 2, for a large H).

But the genuine savings literature has focused on discounted utilitarian prices instead of the maximin prices. Can this be linked to sustainability as defined here?

Following Dasgupta and Mäler (2000), consider a program $P: K_t \mapsto A_t$ determining the actions of generation t as a function of its initial capital. From this program and K_0, one is able to deduce A_0^+ such that for every $t=0,\dots,H$,

$$\begin{aligned} A_t &= P(K_t), \\ K_{t+1} &= T(S_t). \end{aligned}$$

Consider the discounted intergenerational objective:

$$\mathbf{W}_0(K_0)=\sum_{t=0}^{H}\frac{1}{(1+\delta)^t}W(S_t),$$

where A_0^+ is deduced from program P when K_0 depicts the initial conditions. Let $\kappa=\sum_{t=0}^{H}1/(1+\delta)^t$ equal the sum of the weights in this objective.

This criterion is linked to sustainability in three ways. First, there is a link with sustained development in Asheim's (2007) sense of nondecreasing welfare: $W(S_{t+1})\geq W(S_t)$. Indeed, one has:

$$\mathbf{W}_0(K_0)=\sum_{t=0}^{H}\frac{1}{(1+\delta)^t}W(S_t),\quad \mathbf{W}_0(K_1)=\sum_{t=0}^{H}\frac{1}{(1+\delta)^t}W(S_{t+1}),$$

so that

$$\mathbf{W}_0(K_1)-\mathbf{W}_0(K_0)=\sum_{t=0}^{H}\frac{1}{(1+\delta)^t}\left[W(S_{t+1})-W(S_t)\right].$$

When $\mathbf{W}_0(K_1)-\mathbf{W}_0(K_0)<0$, necessarily at some future date, $W(S_{t+1})-W(S_t)<0$. (This may occur at $t=H$, in which case the setback occurs after the planning horizon.) Note that a necessary condition for a sustained path is that $\mathbf{W}_0(K_1)-\mathbf{W}_0(K_0)$ be

[32] Earlier analysis of the limitations of genuine savings (and the related Hartwick rule) for the characterization of sustainability (in models with continuous time, with and without optimality) can be found in chapters 8 and 14 of Asheim (2007).

non-negative *for all* $\delta > 0$. In fact this result does not depend at all on the shape of the weights. One obtains a necessary and sufficient condition for sustainedness if all positive weights (normalized so that their sum is a constant) are considered.

Second, there is a link with sustaining the level $W(S_0)$. One has:

$$\mathbf{W}_0(K_0) = W(S_0) + \frac{1}{1+\delta}\mathbf{W}_0(K_1) - \frac{1}{(1+\delta)^{H+1}}W(S_{H+1}),$$

implying

$$\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) = (1+\delta)\left(\frac{\delta}{1+\delta}\mathbf{W}_0(K_0) - W(S_0)\right) + \frac{1}{(1+\delta)^H}W(S_{H+1})$$
$$\simeq (1+\delta)\left(\frac{1}{\kappa}\mathbf{W}_0(K_0) - W(S_0)\right),$$

where the approximation is obtained by omitting small terms involving $1/(1+\delta)^H$.[33]

When $\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) < 0$, necessarily, if the approximation is correct, $(1/\kappa)\mathbf{W}_0(K_0) < W(S_0)$, which implies that at some future date t, $W(S_t) < W(S_0)$ because $(1/\kappa)\mathbf{W}_0(K_0)$ is a weighted average of the $W(S_t)$, $t = 0, \ldots, H$. Again, one can check the condition for all $\delta > 0$. Contrary to the previous one, this result is valid only for discounted weights.

Third, under optimality conditions, a discrete time variant of Hamilton and Clemens' (1999) and Pezzey's (2004) result is obtained. This is the only result that really bears on sustainability as defined here.

Proposition 9. *Assume that the approximation* $\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) \simeq (1+\delta)\left((1/\kappa)\mathbf{W}_0(K_0) - W(S_0)\right)$ *is correct in the sense that the two sides have the same sign. If* S_0^+ *maximizes* $\sum_{t=0}^{H} 1/(1+\delta)^t W(S_t)$, *and* $\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) < 0$, *then* $W(S_0)$ *is not sustainable given* S_0.

Proof. Let $\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) < 0$. Assume that there is a path $S_1'^{+}$ such that $W(S_t') \geq W(S_0)$ for all $t = 1, \ldots, H$. One then has

$$W(S_0) + \sum_{t=1}^{H} \frac{1}{(1+\delta)^t} W(S_t') \geq \kappa W(S_0).$$

The approximation assumption implies $W(S_0) > (1/\kappa)\mathbf{W}_0(K_0)$, and therefore

$$W(S_0) + \sum_{t=1}^{H} \frac{1}{(1+\delta)^t} W(S_t') > \mathbf{W}_0(K_0),$$

which, in view of the fact that $\mathbf{W}_0(K_0) = \sum_{t=0}^{H} \frac{1}{(1+\delta)^t} W(S_t)$, implies that S_0^+ does not maximize this objective. □

There are three drawbacks with this discounted utilitarian genuine savings approach. First, even under optimality assumptions one obtains a necessary but not sufficient condition of sustainability. In particular, it may happen that $W(S_0)$ is unsustainable according to the previous definition in spite of $\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) > 0$ with the current program (Asheim, 1994).

Second, in the absence of the optimality assumption made in Proposition 9, it may also happen that the welfare level $W(S_0)$ can be sustained forever, even though $\mathbf{W}_0(K_1) - \mathbf{W}_0(K_0) < 0$. That is, the program may generate a non-sustained path whereas current welfare is sustainable (through a different program). Therefore one does not even have a necessary condition for sustainability given current decisions (as opposed to the current *and* future decisions contained in the program).

Third, defining $\mathbf{W}_0(K_0)$ as a function of capital stocks requires a program, which is a contingent strategy plan. No such complex policy object seems to exist in practice, because political platforms rarely describe what would be done in counterfactual worlds. There is a more parsimonious version of the approach that does not require a program, but only a predicted path S_0^+, in order to derive the three above results. By definition,

$$\sum_{t=0}^{H} \frac{1}{(1+\delta)^t} W(S_{t+1}) - \sum_{t=0}^{H} \frac{1}{(1+\delta)^t} W(S_t) = \sum_{t=0}^{H} \frac{1}{(1+\delta)^t} [W(S_{t+1}) - W(S_t)]$$
$$\simeq (1+\delta)\left(\frac{1}{\kappa}\sum_{t=0}^{H} \frac{1}{(1+\delta)^t} W(S_t) - W(S_0)\right),$$

so that

$$\sum_{t=0}^{H} \frac{1}{(1+\delta)^t} W(S_{t+1}) - \sum_{t=0}^{H} \frac{1}{(1+\delta)^t} W(S_t) < 0$$

is an indication that welfare will fall, and will fall below $W(S_0)$, in the future. But in the absence of a program, one cannot compute a function $\mathbf{W}_0(K)$, and therefore it is impossible to derive a genuine savings measure. In particular, accounting prices (the partial derivatives of $\mathbf{W}_0(K)$) cannot be computed.

[33] The omitted terms are $-\delta/((1+\delta)^{H+1} - 1)\mathbf{W}_0(K_0) + 1/(1+\delta)^H W(S_{H+1})$.

So, in the absence of an optimality assumption, and in the absence of a program, we are left with very little for checking *sustainability*. As far as checking whether a *sustained* path will be obtained, the last formulas presented here need a prediction of the path anyway, and with such prediction in hand it is immediate to check whether welfare is nondecreasing or remains above any given level, without having to compute weighted sums.

References

Aronsson, T., Johansson, P.O., Löfgren, K.G., 1997. Welfare Measurement, Sustainability and Green National Accounting. Edward Elgar, Cheltenham.
Arrow, K.J., Dasgupta, P., Goulder, L.H., Mumford, K.J., Oleson, K., 2010. Sustainability and the measurement of wealth. NBER Working Paper 16599.
Asheim, G.B., 1994. Net national product as an indicator of sustainability. Scand. J. Econ. 96, 257–265.
Asheim, G.B., 2007. Justifying, Characterizing and Indicating Sustainability. Springer, Dordrecht.
Asheim, G.B., Brekke, K.A., 2002. Sustainability when capital management has stochastic consequences. Soc. Choice Welf. 19, 921–940.
Asheim, G.B., Buchholz, W., 2004. A general approach to welfare measurement through national income accounting. Scand. J. Econ. 106, 361–384.
Asheim, G.B., Mitra, T., Tungodden, B., 2012. Sustainable recursive social welfare functions. Econ. Theory 49, 267–292.
Atkinson, A.B., 1970. On the measurement of inequality. J. Econ. Theory 2, 244–263.
Basu, K., Mitra, T., 2003. Aggregating infinite utility streams with intergenerational equity: the impossibility of being Paretian. Econometrica 71, 1557–1563.
Baumgärtner, S., Quaas, M.F., 2009. Ecological-economic viability as a criterion of strong sustainability under uncertainty. Ecol. Econ. 68, 2008–2020.
Cairns, R.D., 2000. Sustainability accounting and green accounting. Environ. Dev. Econ. 5, 49–54.
Cairns, R.D., Martinet, V., 2012. An environmental-economic measure of sustainable development, forth. In: European Economic Review.
Chichilnisky, G., 1996. An axiomatic approach to sustainable development. Soc. Choice Welf. 13, 231–257.
Dasgupta, P., Mäler, K.G., 2000. Net national product, wealth, and social well-being. Environ. Dev. Econ. 5, 69–93.
Dasgupta, S., Mitra, T., 1983. Intergenerational equity and efficient allocation of exhaustible resources. Int. Econ. Rev. 24, 133–153.
De Lara, M., Doyen, L., 2008. Sustainable Management of Natural Resources. Mathematical Models and Methods. Springer, Berlin.
De Lara, M., Martinet, V., 2009. Multi-criteria dynamic decision under uncertainty: a stochastic viability analysis and an application to sustainable fishery management. Math. Biosci. 217, 118–124.
De Lara, M., Martinet, V., Doyen, L., 2015. Satisficing versus optimality: criteria for sustainability. Bull. Math. Biol. 77 (2), 281–297.
Dietz, S., Neumayer, E., 2007. Weak and strong sustainability in the SEEA: concepts and measurement. Ecol. Econ. 61, 617–626.
Doyen, L., Martinet, V., 2012. Maximin, viability and sustainability. J. Econ. Dyn. Control 36, 1414–1430.
Fleurbaey, M., 2008. Fairness, Responsibility, and Welfare. Oxford University Press, Oxford.
Fleurbaey, M., 2010. Assessing risky social situations. J. Polit. Econ. 118, 649–680.
Fleurbaey, M., 2013. On sustainability and social welfare, mimeo. ⟨http://www.webmeets.com/files/papers/EAERE/2013/1285/Fleurbaey.pdf ⟩.
Fleurbaey, M., Blanchet, D., 2013. Beyond GDP. Measuring Welfare and Assessing Sustainability. Oxford University Press, New York.
Hamilton, K., Clemens, M., 1999. Genuine saving in developing countries. World Bank Econ. Rev. 13, 33–56.
Heal, G.M., 1998. Valuing the Future: Economic Theory and Sustainability. Columbia University Press, New York.
Heal, G.M., Kriström, B., 2005. National income and the environment. In: K.G., Mäler, J.R., Vincent (Eds.), Handbook of Environmental Economics, North-Holland, Amsterdam.
Howarth, R.B., Norgaard, R.B., 1992. Environmental valuation under sustainable development. Am. Econ. Rev. 82, 473–477.
Llavador, H., Roemer, J.E., Silvestre, J., 2011. A dynamic analysis of human welfare in a warming planet. J. Publ. Econ. 95, 1607–1620.
Martinet, V., 2011. A characterization of sustainability with indicators. J. Environ. Econ. Manag. 61, 183–197.
Martinet, V., 2012. Economic Theory and Sustainable Development: What can we Preserve for Future Generations?. Routledge, Oxon.
Martinet, V., Doyen, L., 2007. Sustainability of an economy with an exhaustible resource: a viable control approach. Resour. Energy Econ. 29, 17–39.
Mori, K., Christodoulou, A., 2012. Review of sustainability indices and indicators: towards a new City Sustainability Index (CSI). Environ. Impact Assess. Rev. 32, 94–106.
Nordhaus, W.D., 1995. How should we measure sustainable income?. Cowles Foundation Discussion Paper No 1101.
Pezzey, J.C.V., 1997. Sustainability constraints versus “optimality” versus intertemporal concern, and axioms versus data. Land Econ. 73, 448–466.
Pezzey, J.C.V., 2004. One-sided sustainability tests with amenities, and changes in technology, trade and population. J. Environ. Econ. Manage. 48, 613–631.
Pezzey, J.C.V., Toman, M.A., 2002. Progress and problems in the economics of sustainability. In: Tietenberg, T., Folmer, H. (Eds.), International Yearbook of Environmental and Resource Economics 2002/2003, Edward Elgar, Cheltenham.
Ramsey, F.P., 1928. A mathematical theory of savings. Econ. J. 38, 543–559.
Sen, A.K., 1985. Commodities and Capabilities. North Holland, Amsterdam.
Solow, R.M., 1974. Intergenerational equity and exhaustible resources. Rev. Econ. Stud. 41, 29–46.
Solow, R.M., 1991. Sustainability: an economist's perspective. The Eighteenth J. Seward Johnson Lecture. Woods Hole Oceanographic Institution, Woods Hole, MA.
Von Weizsäcker, C.C., 1965. Existence of optimal programs of accumulation for an infinite time horizon. Rev. Econ. Stud. 32, 85–104.
Wackernagel, M., Rees, W., 1995. Our Ecological Footprint: Reducing Human Impact on the Earth. New Society Publishers, Gabriola Island, BC.
Weitzman, M.L., 1976. On the welfare significance of net national product in a dynamic economy. Q. J. Econ. 90, 156–162.
World Commission on Environment and Development, 1987. Our Common Future. Oxford University Press, New York.
Zame, W.R., 2007. Can intergenerational equity be operationalized? Theor. Econ. 2, 187–202.