On Sustainability and Social Welfare*

Marc Fleurbaey†

June 2013

Abstract

This paper proposes to define sustainability in terms of making it possible for future generations to sustain whatever one wants to sustain. 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 emphasized that focusing only on the opportunities left to future generations is a rather extreme approach to social welfare. So, sustainability can at best be viewed as a minimal and arguably insufficient obligation for the present generation.

Keywords: sustainability, genuine savings, ecological footprint, social welfare, overlapping generations.

JEL Classification: D6, D9, O2, Q5.

"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)

1 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: p.70), 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 briefly review the various approaches that economists

---

*This paper has considerably benefited from advice and comments by G. Asheim, though he should not be held responsible for any mistake.
†Princeton University. Email: mfleurba@princeton.edu.
have adopted to incorporate the notion of sustainability, and argue that one of them, namely, the notion of giving future generations the *ability* (a word used in the Brundtland formulation) to sustain certain targets, better captures the idea of sustainability than others (section 2).

I propose a rather non-orthodox framework to analyze the approach (discrete time, finite planning horizon), and obtain simple criteria for the indication of sustainability, one of which involves a variant of genuine savings (section 3.2) and another a variant of the ecological footprint (section 3.4). A comparison with genuine savings is made in section 3.5.

The notion and criteria are flexible enough to make it possible to incorporate a variety of sustainability targets (growth, multiple objectives) and any relevant set of constraints that bear on the decisions of future generations (sections 4.1-3). A key aspect of the approach highlighted in this paper is the notion of opportunities given to future generations, and a variant of the definition of sustainability can also focus on the opportunities enjoyed by the current generation (section 4.4). 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 4.5).

Making the connection to social welfare reveals that sustainability can be a key part of social welfare analysis only under the extreme assumption that the current generation should only care about the opportunities it leaves to future generations, and not with what they are actually likely to achieve (section 5). While most of the paper deals with successive generations, it is shown in section 6 that the bulk of the analysis can be adapted to overlapping generations. A short conclusion is given in section 7.

2 Sustainability: a review of approaches

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 or the Chichilnisky 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 in a part of Asheim (2007). The results of this literature are, however, of limited relevance to the task of checking if the management of resources by the present generation is sustainable in the Brundtland sense.

An important feature of the Brundtland definition is indeed that sustainability is a matter of sufficiency rather than optimality. There are, however, three aspects of sufficiency in this definition, and they may be assessed differently. The first has to do with the idea that the current management can be sustainable without being optimal with respect
to any sensible intergenerational welfare objective. The second has to with a focus on “the ability of future generations”, rather than their likely fate (as predicted by an optimal or suboptimal plan). These two aspects will be retained in the notion of sustainability adopted in this paper. In contrast, focusing on needs as this definition does may seem too minimalist, as argued by Arrow et al. (2010), and we will adopt the orthodox view in economics that a more comprehensive notion of well-being should be used.

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). It combines the ability idea with well-being, which is appealing, but is vague about the status of sustainability. If it is an obligation, does it follow from a particular social objective or does it work as a constraint?

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. This preference is well in line with the current discussion, but his own implementation of the approach is not convincing, because he measures Fisherian income as the constant equivalent consumption of any given consumption path, for the discounted sum. The constant equivalent consumption can only be assimilated to opportunities if markets permit any intertemporal transfer at the corresponding prices. One can also question the whole idea of checking sustainability by comparing aggregate consumption to a unidimensional threshold, because the impact of current actions is likely to depend on more detailed features, as emphasized in Asheim (1994) and Aronsson et al. (1997).

Arrow et al. (2010) adopt a quite different vision of 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 \( \frac{dV}{dt} \geq 0 \),” (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 (genuine savings). In a similar vein, Heal

---

1 As he explains and is well known, the Fisherian notion also appeared as a potential definition of income in Hicks’ work.

2 The idea of measuring the constant equivalent of a discounted sum (of consumption or utility) by the net national product, studied in Weitzman (1976) and the subsequent literature, nicely connects the Hicksian income (as well as the Fisherian income as defined by Nordhaus) to genuine savings.
and Kilström (2008) 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 discounted utility. The paper by Dasgupta and Mäler (2000) was the pioneer of these approaches, making the key contribution of extending the framework to non-optimal paths.

An important weakness of this alternative approach is that one can have \( dV/dt > 0 \) even when generational utility and/or intergenerational welfare will necessarily decline in the future. The criterion is too weak. Moreover, it is about intergenerational welfare much more than about sustainability. In his review of the literature, Asheim (2007, ch.14) makes a clear distinction between the two branches of analyses.

A crystal clear definition of sustainability, more in line with the first notions introduced earlier, 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) and “development is sustained if 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 more demanding than just sustaining the current level. This introduces a nice recursive structure in the analysis. 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.\(^3\)

Pezzey (1997, 2004) and Pezzey and Toman (2002) define sustainability as current welfare being below or equal to the maximum sustainable level, namely, the maximin value, and note that this is implied by the possibility for welfare to follow a nondecreasing path forever.\(^4\) Pezzey (2004, p. 616) notes that in continuous time, this inequality (current welfare below or equal to the maximin value) does not guarantee sustainability of current welfare when the economy is inefficient, and this suggests that the possibility for future generations to sustain a level or a nondecreasing path better captures the idea of sustainability.

In a related approach, Martinet (2011) defines sustainability in terms of the possibility of satisfying inequality conditions (thresholds) in the future. A special feature of this

\(^3\)There may be a paradox in the logical possibility for the present generation to have a sustainable management (understood in terms of sustaining the current welfare level only) that does not leave future generations the option of having a sustainable management (because their own welfare level is greater). But it is debatable how much effort the current generation should make to avoid this situation.

\(^4\)Pezzey (1997) actually requires \( W(S_t) \leq V_t(K_t) \) for all \( t \) for a “sustainable development” whereas the other references use this inequality at \( t = 0 \) to define sustainability at time 0.
paper is the proposal to maximize preferences over the thresholds. 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. It is not clear, however, why one should maximize some preference over the thresholds.  

This approach is similar to the optimizing approach mentioned in the beginning of this section.

An interesting aspect of his analysis is the reference to viability theory (also mobilized in Baumgärtner and Quaas 2009), which defines the set of thresholds that can be respected from any given initial state of the economy, or, conversely, the set of initial states of the economy that are compatible with respecting the thresholds over time. On the basis of this theory, and focusing on a single welfare threshold, Doyen and Martinet (2012) propose a combination of two practical criteria for checking sustainability of the present generation’s welfare on the basis of this approach. The first criterion is that current utility should not exceed the maximin level (Pezzey and Toman’s criterion), and the second criterion is that genuine savings evaluated at the shadow prices of the maximin program should be nonnegative, or equivalently that $\frac{dV}{dt} \geq 0$, where $V$ now denotes the maximin value.

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 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. For the viability-maximin approach, Cairns and Martinet (2012) reintroduce a weak efficiency condition in order to simplify the Doyen-Martinet double criteria. They show that the $dV/dt \geq 0$ criterion is satisfied if and only if current welfare is below the maximin value $V$, under the assumption that the current policy maximizes $dV/dt$, given current utility.

Let us take stock. The concept of sustainability appears quite different from the current evolution of intergenerational welfare, and is about making it possible for future generations to achieve some outcome, either a level or a progression. Such possibility is determined by the management of resources by the current generation, and can be ascertained without making a precise prediction about the future generations’ decisions.

---

Martinet argues that “this approach is close to the Rawlsian theory of justice since it can be interpreted as a way to address the sustainability issue focusing on minimal rights that all generations should have” (p. 190). But in Rawls’ theory one should not maximize the minimal rights, so there seems to be a lack of justification for this particular feature of the approach. (Another issue that he identifies is the potential time inconsistency of the approach, as the maximization problem may yield different solutions as time goes on, when there are multiple thresholds.)
The viability approach seems appropriate for this purpose, and it is interesting to explore the construction of simple indicators of sustainability in the specific context of economic models.

Establishing a link with the evolution of the maximin value, as in Martinet and co-authors’ work, may seem to restore a direct connection between the evolution of intergenerational welfare and sustainability, but that would be misinterpreting the role of the maximin value here. It may be useful to check sustainability even if one does not want to adopt the maximin approach to intergenerational welfare (a utilitarian could still use the maximin value to check sustainability), and even if intergenerational welfare is not maximized (the predicted level of the lowest utility among future generations may be well below the maximin value).

With this background, this paper examines how to check sustainability in an economic model, and how to incorporate a concern for sustainability in the definition of social welfare. The main conclusion drawn at the end of this paper is that although sustainability indicators are easy to conceive in theory (though not in practice), focusing on sustainability in intergenerational policy issues may be quite questionable.

3 Indicators of sustainability under perfect information

3.1 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 sustainability literature. The motivation for discrete time is that it nicely allows to separate the possibilities for future generations from the achievements of the current generation, and this simplifies the use of the maximin value as a criterion for sustainability.

Second, the horizon (denoted $H$) is finite. This is not 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 given current cosmological theories, 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. Third, more importantly, 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

---

6 As shown by Dasgupta and Mitra (1983), Hartwick’s rule is not valid in discrete time.
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 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 cost 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.

Otherwise, the model is similar to the literature. For $t = 0, ..., H - 1$, a sequence $(x_t, ..., x_H)$ is denoted $x_t^+$. Vector inequalities are denoted $\leq, <, \ll$.

The stock of capital is $K_t \in \mathbb{R}^m$. The value of $K_t$ is given in the beginning of period $t$. 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.

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). It may include the population size.

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 $\Phi$ can either describe technical possibilities or also include certain political and behavioral constraints (this will be developed in section 4.3). The extinction of the human species can be determined endogenously when, given $K_t$ and $\Phi$, the only possible action $A_t$ is inaction.

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 left for the definition of intergenerational welfare (an ethical issue) than for the definition of generational welfare (a psychological issue).

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

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

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

Let the set of feasible paths from $t$ be denoted $\Phi_t(K_t)$. The dynamic process is autonomous but as $K_t$ is a comprehensive notion, 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 properties will be useful in the derivation of certain results in the next sections.

Free disposal means that it is always possible to decrease the welfare of any particular generation, provided it is above the minimum possible: for every feasible $S^+_0$, every $t^*$ such that $W(S_{t^*}) > \inf W(\Phi)$, every $w \in (\inf W(\Phi), W(S_{t^*}))$, there is a feasible $S^+_0$ such that $W(S^+_0) = w$ and for all $t \neq t^*$, $W(S^+_0) = W(S_t)$.

Transferability means that it is always possible to increase the welfare of a generation at the expense of another generation (if the latter is above the minimum possible): for every $S^+_0$ such that $W(S_0) > \inf W(\Phi)$, every $t^* > 0$ such that $W(S_{t^*}) > \inf W(\Phi)$, every $\varepsilon > 0$, there exists $S^{+_0}$ such that $\varepsilon > W(S^{+_0}) - W(S_{t^*}) > 0 > W(S^+_0) - W(S_0) > -\varepsilon$ and for all $t \neq 0, t^*$, $W(S^+_0) = W(S_t)$; and there exists $S^{+_0}$ such that $-\varepsilon < W(S^{+_0}) - W(S_{t^*}) < 0 < W(S^+_0) - W(S_0) < \varepsilon$ and for all $t \neq 0, t^*$, $W(S^+_0) = W(S_t)$. 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 this property, 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

$$\Lambda_t(K_t) = \{\lambda \in \mathbb{R} \mid \exists S^+_t \in \Phi_t(K_t), \forall \tau \geq t, W(S_\tau) \geq \lambda\}$$

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

Proof. Continuity of $T, W$ and compactness of $\Phi$ imply that $\Phi_t(K_t)$ is compact for all $K_t$. Therefore the set

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

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

Suppose that $\lambda^* = \max \Lambda_t(K_t)$ is such that there is $S^+_t$ for which $W(S_\tau) = \lambda^*$ for all $\tau \geq t$ and $W(S_{\tau'}) > \lambda^*$ for some $\tau' \geq t$. By transferability, it is possible to decrease $W(S_{\tau'})$ and raise $W(S_\tau)$ for all $\tau \geq t, \tau \neq \tau'$, thereby raising $\min_\tau W(S_\tau)$ above $\lambda^*$, which is impossible. ■

Fungibility means that capital stocks can be ordered in terms of welfare opportunities: for all $t$, all $K_t, K'_t$, one has either $\mathcal{W}_t(K_t) \subset \mathcal{W}_t(K'_t)$ or $\mathcal{W}_t(K_t) \supset \mathcal{W}_t(K'_t)$, where

$$\mathcal{W}_t(K_t) = \{w \in \mathbb{R}^{H-t+1} \mid \exists S^+_t \in \Phi_t(K_t), \forall \tau \geq t, W(S_\tau) = w_\tau\}.$$
Fungibility is a restrictive property which means that capital stocks essentially behave as one single good. This property will not be satisfied if different capital goods do not contribute equally to the production of consumption (or, more precisely, social welfare) and the production of capital (a stock of capital containing a greater fraction of capital goods that contribute more to the production of capital gives greater opportunities to make transfers to future generations), if capital goods have different depreciation rates (a stock of capital containing a greater fraction of capital goods with a lower depreciation rate gives greater opportunities to make transfers to future generations), or if capital goods have different investment-replenishment possibilities, for instance with some being produced capital and some being exhaustible resources (a stock of capital containing a greater fraction of capital goods with lower replenishment possibilities gives greater opportunities to make transfers to future generations).

3.2 Sustainability of current welfare level

Let us first seek to check if \( W(S_0) \) can be enjoyed by the following generations.

**Definition 1** The welfare level \( W(S_0) \) is sustainable given \( S_0 \) if there is a feasible path \( S^+_1 \in \Phi_1(T(S_0)) \) such that for all \( t \geq 1 \), \( W(S_t) \geq W(S_0) \).

This definition allows for the actual path to fall below \( W(S_0) \) if some generations make a sacrifice for their descendants. What is important is that it is possible for every generation to be at least at \( W(S_0) \). It is also important to define sustainability for a given \( S_0 \) and in particular for a given \( A_0 \). A level of welfare \( W(S_0) \) could be sustainable provided another action \( A'_0 \) was taken in period 0. This is irrelevant if \( A_0 \) is already determined.

By definition, \( W(S_0) \) is sustainable if and only if \( W(S_0) \in \Lambda_1(T(S_0)) \). Let

\[
V_t(K_t) = \max_{\Lambda_t(K_t)}.
\]

One can identify three (partial, sometimes) criteria of sustainability. The first directly follows from our definition. Note that, though obvious from the definition, this criterion cannot be easily formulated in continuous time.

**Proposition 1** \( W(S_0) \) is sustainable given \( S_0 \) if and only if

\[
W(S_0) \leq V_1(T(S_0)).
\]

But two other criteria can be used. The following is formally akin to Pezzey and Toman’s definition for a continuous time setting. In its negative form, it provides a criterion of strong unsustainability because it does not depend on current actions.
Proposition 2 \( W(S_0) \) is not sustainable given any \( S'_0 = (A'_0, K_0) \) such that \( W(S'_0) = W(S_0) \) if
\[
W(S_0) > V_0(K_0).
\]
Under transferability and fungibility, if there is \( S^+_1 \) such that \((S_0, S^+_1)\) is Pareto-efficient, then \( W(S_0) \) is sustainable given \( S_0 \) if
\[
W(S_0) \leq V_0(K_0).
\]

Proof. By definition, \( V_0(K_0) \geq \min \{ W(S'_0), V_1(T(S'_0)) \} \). If \( W(S'_0) = W(S_0) > V_0(K_0) \), necessarily \( V_0(K_0) \geq V_1(T(S'_0)) \), and therefore \( W(S'_0) > V_1(T(S'_0)) \), implying unsustainability given \( S'_0 \) by Prop. 1.

Assume transferability and fungibility. By Prop. 5, the fact that \((S_0, S^+_1)\) is Pareto-efficient implies that \( (W(S_0), V_1(T(S_0))) \) is on the frontier of the set
\[
\{(w, v) \mid \exists S_0 \in \Phi, (w, v) = (W(S_0), V_1(T(S_0)))\}.
\]
Moreover, by Lemma 2, on the frontier, if \( W(S_0) \leq V_0(K_0) \), then \( V_1(T(S_0)) \geq V_0(K_0) \), therefore \( V_1(T(S_0)) \geq W(S_0) \), guaranteeing sustainability by Prop. 1. ■

Note that \( W(S_0) \leq V_0(K_0) \) does not ensure sustainability in general, as already recalled for the Pezzey-Toman notion in the previous section.

The last criterion relies on the maximin value criterion invoked by Cairns and Martinet (2011) and Doyen and Martinet (2012), and we first focus on what can be said without any efficiency requirement. What is striking, compared to the literature on continuous time, is how easily one obtains a sufficient criterion for sustainability.

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 \( V_0(K_0) = V_1(T(S_0)) \).

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 by Prop. 1.

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 by Prop. 1. ■

Positive genuine savings at maximin prices is then a very general sufficient condition for sustainability. One has
\[
V_1(K_1) - V_0(K_0) = V_0(K_1) - V_0(K_0) + V_1(K_1) - V_0(K_1) \sim p_0(K_1 - K_0)
\]
if \( p_0 = \nabla V_0 \) (i.e., the gradient vector) and \( V_1 (K_1) \simeq V_0 (K_1) \). The latter approximation is justified when \( H \) is large.

These results allow us to identify the following configurations (equality cases are ignored):

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

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. 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.

3.3 Efficiency issues

We have seen that \( V_0 (K_0) < V_1 (T (S_0)) \) guarantees sustainability. But \( V_0 (K_0) > V_1 (T (S_0)) \) is also compatible with sustainability. In particular, sustaining \( W (S_0) \) is compatible with inefficiency: \( S_0 \) may destroy capital and resources, but \( W (S_0) \) may be low enough so that sustainability afterwards is possible (a war followed by reconstruction illustrates this configuration).

The following proposition invokes an efficiency condition similar to that in Cairns and Martinet (2012).

Proposition 4 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. 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 \( S_0^t \) such that \( W (S_0^t) \geq V_0 (K_0) \) for all \( t \geq 0 \). On this path, one has \( V_1 (T (S_0^t)) \geq V_0 (K_0) \). Therefore it would be inefficient 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 necessarily \( W (S_0) \geq V_0 (K_0) \), implying \( W (S_0) > V_1 (T (S_0)) \), therefore unsustainability by Prop. 1. ■
Figure 1, which is reminiscent of Fig. 1 in Howarth and Norgaard (1992, p. 474 — only two generations are featured in their figure), summarizes Propositions 1, 2, 3, and 4. It also relies on the following lemma.

**Lemma 2** The point \((V_0(K_0), V_0(K_0))\) is on the efficiency frontier 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 \(V_0(K_0)\).

Proposition 4 may suggest the thesis that if \(V_0(K_0) > V_1(T(S_0))\), then one has either unsustainability or inefficiency. But this holds true only for efficiency in terms of the pair \((W(S_0), V_1(T(S_0)))\). In general, such efficiency corresponds to efficiency of the path \(S_0^+\) only if the future generations will follow a maximin policy. A strong assumption like fungibility appears needed to extend this reasoning to inefficiency of the path \(S_0^+\).

**Proposition 5** Under transferability and fungibility, if \((W(S_0), V_1(T(S_0)))\) is not on the frontier, then \(S_0^+\) is inefficient. Therefore, if \(V_0(K_0) > V_1(T(S_0))\), either \(W(S_0)\) is unsustainable given \(S_0\) or \(S_0^+\) is inefficient.
Proof. Under transferability, the frontier for \((W(S_0), V_1(T(S_0)))\) has a finite downward slope. Therefore, if \((W(S_0), V_1(T(S_0)))\) is not on the frontier, there is \(S'_0\) such that \(W(S'_0) > W(S_0)\) and \(V_1(T(S'_0)) > V_1(T(S_0))\). By fungibility, the latter inequality implies that \(W_1(T(S_0)) \subset W_1(T(S'_0))\), so that there is \(S'_+\) such that \(W(S'_+) \geq W(S_t)\) for all \(t \geq 1\). Therefore \(S'_+\) Pareto dominates \(S'_0\), proving that the latter is inefficient. ■

It is also useful to note the converse result.

Proposition 6 Under transferability and fungibility, if \((W(S_t), V_{t+1}(T(S_t)))\) is on the frontier for all \(t = 0, \ldots, H - 1\), then \(S'_0\) is efficient.

Proof. Suppose that \((W(S_t), V_{t+1}(T(S_t)))\) is on the frontier for all \(t = 0, \ldots, H - 1\), but \(S'_0\) is inefficient. Let \(S'_0\) be a dominating path.

By transferability, one can assume that \(W(S'_0) > W(S_t)\) for all \(t\). Moreover, the frontier for \((W(S_0), V_1(T(S_0)))\) has a finite downward slope. Therefore \(W(S'_0) > W(S_0)\) and \(V_1(T(S'_0)) < V_1(T(S_0))\).

Under fungibility, the fact that \(W(S'_0) > W(S_t)\) for all \(t \geq 1\) implies that \(W_1(T(S_0)) \subset W_1(T(S'_0))\). This is incompatible with \(V_1(T(S'_0)) < V_1(T(S_0))\). ■

3.4 Footprint: an alternative

The idea of computing the proportion of available resources which would be needed to sustain the current level of well-being (Wackernagel and Rees 1995) is valuable, but it is poorly implemented in the current measure. The indicator measures the surface of land needed to sustain current living standards, at the global and at the national level. The measure relies on actual yields of various land uses and therefore does not take account of the depletion of soils or of predicted changes in productivity. To the land surfaces needed to produce various consumer goods, it adds the surface of forest needed to absorb carbon emissions. Given that actual yields are used in the computations, the surface needed for production does not vary much over time at the global level, and carbon emissions make up most of the increase of the indicator. This indicator has been criticized for assuming substitutability between the various forms of natural capital, for adopting too crude a link between greenhouse gases and forest surface requirements (Dietz and Neumayer 2007), as well as for ignoring technical progress and the degradation of soils (Mori and Christodoulou 2012).

Assume that \(K_t\) is divided in two components \((K^a, K^b)\), with \(K^a\) the relevant part for footprint evaluation. 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 \(\lambda\) such that \(W(S_0) = V_1(T(A_0, (\lambda K^a_0, K^b_0)))\). 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 $\lambda$ is like the ecological footprint. It is the proportion of the current stock of capital $K^a$ that would be needed to sustain $W(S_0)$, given $S_0$. Summarizing, one has the following proposition.

**Proposition 7** Assume that $V_1(K)$ and $T^a(A,K)$ are increasing in $K^a$ and that there is $\lambda$ such that $W(S_0) = V_1(T(A_0, (\lambda K_0^a, K_0^b)))$. $W(S_0)$ is sustainable given $S_0$ if and only if $\lambda \leq 1$.

**Proof.** Under the assumptions of increasing $V_1$ and $T^a$, $\lambda \leq 1$ iff $W(S_0) \leq V_1(T(S_0))$.

Cairns and Martinet (2012) propose to interpret the ratio $W(S_0)/V_1(T(S_0))$ in terms of ecological footprint. The notion introduced here appears closer to the original notion.

### 3.5 Genuine savings revisited

Let us briefly examine how the genuine savings approach is related to sustainability as defined in this paper. In this section, $H = \infty$ is adopted to reduce the length of formulas. A finite $H$ would require putting error terms in many equations.

Following Dasgupta and Mäler (2000), the literature assumes that there is 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 $t \geq 0$,

$$A_t = P(K_t),$$

$$K_{t+1} = T(S_t).$$

Consider the discounted intergenerational objective:

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

(the weights are such that $\sum_{t=0}^{\infty} \frac{\delta}{(1+\delta)^{t+1}} = 1$), where $A_0^+$ is deduced from program $P$ when $K_0$ depicts the initial conditions.

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. Indeed, one has:

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

so that

$$W(K_1) - W(K_0) = \sum_{t=0}^{\infty} \frac{\delta}{(1+\delta)^{t+1}} [W(S_{t+1}) - W(S_t)].$$
When $W(K_1) - W(K_0) < 0$, necessarily at some future date, $W(S_{t+1}) - W(S_t) < 0$. Note that a necessary condition for a sustained path is that this expression be 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 adding up to one are considered.

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

$$W(K_0) = \frac{\delta}{1+\delta} W(S_0) + \frac{1}{1+\delta} W(K_1),$$

implying

$$W(K_1) - W(K_0) = \delta (W(K_0) - W(S_0)).$$

When $W(K_1) - W(K_0) < 0$, necessarily $W(K_0) < W(S_0)$, which implies that at some future date $t$, $W(S_t) < W(S_0)$. 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 8** If $S_0^+$ maximizes $\sum_{t=0}^{\infty} \frac{\delta}{(1+\delta)^{t+1}} W(S_t)$, and $W(K_1) - W(K_0) < 0$, then $W(S_0)$ is not sustainable given $S_0$.

**Proof.** Let $W(K_1) - W(K_0) < 0$. Assume that there is a path $S_t^+$ such that $W(S_t) \geq W(S_0)$ for all $t \geq 1$. One has

$$\frac{\delta}{1+\delta} W(S_0) + \sum_{t=1}^{\infty} \frac{\delta}{(1+\delta)^{t+1}} W(S_t^+) \geq W(S_0).$$

Recall that $W(S_0) = W(K_0) - \frac{1}{\delta} (W(K_1) - W(K_0))$, implying that $W(S_0) > W(K_0)$, and therefore

$$\frac{\delta}{1+\delta} W(S_0) + \sum_{t=1}^{\infty} \frac{\delta}{(1+\delta)^{t+1}} W(S_t^+) > W(K_0),$$

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

There are three drawbacks with this 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 $W(K_1) - W(K_0) > 0$ with the current program (Asheim 1994).

Second, in absence of the optimality assumption made in the third result, it may also happen that the welfare level $W(S_0)$ can be sustained forever, even though $W(K_1) - W(K_0) < 0$. 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 \( W(K_0) \) 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. There is a more parcimonious 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,

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

so that

\[
\sum_{t=1}^{\infty} \frac{\delta}{(1 + \delta)^{t+1}} W(S_t) - \sum_{t=0}^{\infty} \frac{\delta}{(1 + \delta)^{t+1}} W(S_t) < 0
\]

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

So, in absence of an optimality assumption, and in 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 at hand it is immediate to check whether welfare is nondecreasing or remains above any given level, without having to compute weighted sums.

### 4 Extensions and variants

#### 4.1 Sustaining growth

We now examine generalizations of the notion of sustainability. First, one may be concerned with the future generations enjoying not just the current welfare level, but at least a certain growth rate. In this subsection we assume that \( W(.) \) takes positive values.

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

A special case is \( g = 0 \), which corresponds to Asheim’s definition of sustainability, i.e., making it possible that welfare never decreases (which is more restrictive than requiring that welfare remains above \( W(S_0) \)).
Let
\[ \Lambda^g_t (K_t) = \left\{ \lambda \in \mathbb{R} \mid \exists S_t^+ \in \Phi_t (K_t), W (S_t) \geq (1 + g) \lambda \text{ and } \forall t \geq 0, W (S_{t+1}) \geq (1 + g) W (S_t) \right\}. \]
This set is lower comprehensive and, under the assumptions of the model, has a maximum.
By definition, \( g \) is sustainable given \( S_0 \) if and only if \( W (S_0) \in \Lambda^g_1 (T (S_0)) \). Let
\[ V^g_t (K_t) = \max \Lambda^g_t (K_t). \]
The previous criteria hold for this notion of sustainability, with a small modification.

**Proposition 9** \( g \) is sustainable given \( S_0 \) if and only if
\[ W (S_0) \leq V^g_1 (T (S_0)). \]
It is not sustainable given any \( S'_0 \) such that \( W (S'_0) = W (S_0) \) if
\[ W (S_0) > (1 + g) V^g_0 (K_0). \]
Assuming free disposal, it is sustainable given \( S_0 \) if
\[ (1 + g) V^g_0 (K_0) < V^g_1 (T (S_0)). \]

**Proof.** \( W (S_0) \in \Lambda^g_1 (T (S_0)) \) iff \( W (S_0) \leq V^g_1 (T (S_0)). \)

If \( W (S_0) > (1 + g) V^g_0 (K_0) \), by definition of \( V^g_0 (K_0) \) there is no \( S_0^+ \in \Phi_0 (K_0) \) such that \( W (S_0^+) \geq W (S_0) \) and \( \forall t \geq 0, W (S_{t+1}) \geq (1 + g) W (S_t) \).

Under free disposal, \( (1 + g) V^g_0 (K_0) \geq \min \{ W (S_0), V^g_1 (T (S_0)) \} \), because if \( (1 + g) V^g_0 (K_0) < \min \{ W (S_0), V^g_1 (T (S_0)) \} \), there is \( S^+_0 \) such that \( (1 + g) V^g_0 (K_0) < W (S^+_0) = V^g_1 (T (S_0)) \), in contradiction with the definition of \( V^g_0 (K_0) \). Therefore, if \( (1 + g) V^g_0 (K_0) < V^g_1 (T (S_0)) \), then \( W (S_0) \leq (1 + g) V^g_0 (K_0) \), therefore \( W (S_0) < V^g_1 (T (S_0)) \).

The last criterion in the proposition is simple, as it requires a growth rate of \( V^g_t \) greater than \( g \) at \( t = 1 \). For genuine savings, this means, approximately,
\[ \frac{p_0 \Delta K}{V^g_0 (K_0)} > g, \]
which is unfortunately not the same as \( \frac{p_0 \Delta K}{p_0 K} > g \) unless \( V^g_0 (K_0) \) is linear in \( K \).

Regarding the special case \( g = 0 \), the following observation is useful. Under free disposal, one has \( V_1 (K_1) = V^0_1 (K_1) \) and there is no difference between sustaining \( W (S_0) \) and sustaining a non-negative growth.
4.2 Multiple sustainability

The approach is easily applicable to multiple targets of sustainability. To fix ideas, let us consider a simple example. Suppose one is interested in sustaining the level for $W^a(S_0)$ and the growth rate $g$ for $W^b(S_0)$. A typical example is seeking to simultaneously sustain a stock of environmental capital and a non-negative growth rate for welfare ("strong sustainability").

The previous uni-dimensional analysis remains valid, provided the values of $W^a(S_0)$ and $W^b(S_0)$ are positive and one reasons in terms of vectors along the ray determined by $(W^a(S_0), W^b(S_0))$.

Let

$$\Lambda^{ab}_1(K_1) = \left\{ \lambda \in \mathbb{R} \mid \exists S^+_1 \in \Phi_1(K_1), \forall t \geq 1, (W^a(S_t), W^b(S_t)) \geq \lambda (W^a(S_0), W^b(S_0)) \quad \text{and} \quad W^b(S_{t+1}) \geq W^b(S_t) \right\}.$$ 

This set is lower comprehensive. By definition, $W^a(S_0)$ and $g = 0$ for $W^b(S_0)$ are jointly sustainable if and only if $1 \in \Lambda^{ab}_1(T(S_0))$. Let

$$V^{ab}_1(K_1) = \max \Lambda^{ab}_1(K_1).$$

One sees that $(W^a(S_0), W^b(S_0))$ is sustainable if and only if $1 \leq V^{ab}_1(K_1)$, and so on.

This can be straightforwardly generalized to more than two dimensions, and to positive growth rates for some dimensions.

4.3 Constrained sustainability

The set $\Phi$ that delineates feasible $S_t$ 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. The more constraints are introduced, the smaller the set (in the sense of inclusion), and the harder it is to satisfy sustainability.

Figure 2 illustrates a situation in which technical sustainability is satisfied but not constrained sustainability. Taking account of additional constraints does not change $W(S_0)$ but reduces $V_1(K_1)$.

Defining the constraints that are taken into account when checking sustainability is at least in part an ethical choice. More on this issue will be said in the later section on social welfare.

4.4 Sustainability and opportunities

The above analysis defines sustainability in terms of opportunities for future generations. One may want to adopt an approach in terms of opportunities for the current generation.
as well, interpreting sustainability as a condition “on changes in opportunities over time” (see Pezzey and Toman 2002). After all, why declare that what the current generation does is sustainable if this just comes from a low welfare level obtained by squandering its own opportunities and those of future generations?

The relevant opportunities for generation 0 could be the maximum it can obtain for itself even at the detriment of all future generations (e.g., by consuming all exhaustible resources). That might be somewhat farfetched and not very attractive. Why should it be desirable to pass on to one’s descendants the same possibilities to squander the Earth’s resources at the expense of their own descendants?

A more attractive notion, perhaps, is the maximum welfare it can obtain while leaving it possible for all future generations to obtain the same welfare. This is \( V_0(K_0) \), by Prop. 2.

This definition is in fact more consistent with the notion of opportunities defined earlier for future generations, because they were defined in terms of possibilities for a minimum (under some assumptions, an equal) level of welfare given to all of them. But it suggests redefining the opportunities of future generations, by singling out the opportunities of each generation \( t \), which would then be measured by \( V_t(K_t) \).

Sustainability is then defined as generation 0 leaving the possibility for future generations to enjoy at least the same opportunities.

**Definition 3** The action \( A_0 \) is a sustainable management of opportunities if there is a
feasible path $S^+_t \in \Phi_1(T(S_0))$ such that for all $t \geq 1$, $V_t(K_t) \geq V_0(K_0)$.

One obtains a simple criterion that confirms the importance of genuine savings at maximin prices.

**Proposition 10** The action $A_0$ is a sustainable management of opportunities if and only if $V_0(K_0) \leq V_1(T(S_0))$.

**Proof.** If $V_0(K_0) > V_1(T(S_0))$, the definition of sustainable management is clearly violated since for all $S^+_t \in \Phi_1(T(S_0))$, $K_1 = T(S_0)$, so that nothing can be done by generation 1 to improve $V_1(T(S_0))$.

It $V_0(K_0) \leq V_1(T(S_0))$, there is a feasible path $S^+_t$ such that $W(S_t) = V_1(T(S_0))$ for all $t \geq 1$, and along this path $V_t(K_t) \geq V_1(T(S_0)) \geq V_0(K_0)$ for all $t \geq 1$. ■

Another, more comprehensive approach would depict the opportunities of generation $t$ as the feasible set of pairs $(W(S_t), V_{t+1}(T(S_t)))$. Formally, let

$$O_t(K_t) = \{(W(S_t), V_{t+1}(T(S_t))) \mid S_t = (A_t, K_t) \in \Phi\}.$$ 

This set includes both $(V_t(K_t), V_t(K_t))$ (by Lemma 2) and a point containing the maximum $W(S_t)$ that is feasible given $K_t$. The following proposition says that the two earlier approaches boil down to almost the same thing under some (restrictive) conditions, because the opportunity sets are nested.

**Proposition 11** Assume that $V_0(.) = V_1(.)$ and $O_0(.) = O_1(.)$. Under fungibility, $O_0(K_0) \subset O_1(K_1)$ if $V_0(K_0) < V_1(K_1)$.

**Proof.** Under fungibility, for all $K, K'$, all $t$, either $W_t(K) \subset W_t(K')$ or $W_t(K) \supset W_t(K')$. This directly implies that $O_t(K) \subset O_t(K')$ or $O_t(K) \supset O_t(K')$. By Lemma 2, one has $V_t(K) \leq V_t(K')$ if $O_t(K) \subset O_t(K')$.

Therefore $O_0(K_0) \subset O_0(K_1)$ if $V_0(K_0) < V_0(K_1)$. Replacing $O_0(K_1)$ and $V_0(K_1)$ by $O_1(K_1)$ and $V_1(K_1)$ completes the proof. ■

The equalities $V_0(.) = V_1(.)$ and $O_0(.) = O_1(.)$ are unlikely to hold under the finite horizon assumption, but they can be almost true when $H$ is large, in which case the proposition holds approximately.

### 4.5 Sustainability under uncertainty

We first present the extension of our approach to uncertainty and then compare it to the literature.

Assume that at time 0, there is uncertainty about $T$. (There could also be uncertainty about $\Phi$, but it is simpler to put all uncertainty in $T$. Note that if there is uncertainty
about \( \Phi \), one might not be sure that \( S_0 \) is feasible.) For any given technology \( T \), one can compute the corresponding \( V_1 (T (S_0)) \). For a given \( S_0 \), let \( V^* (T; S_0) \) denote this function.

Let \((T, \Sigma)\) denote a measurable space on the set \( T \) of possible \( T \), and \( \mu \) a probability measure on the \( \sigma \)-algebra \( \Sigma \), that represents the beliefs of the evaluator. Assume that \( V^* (\cdot; S_0) \) is measurable and let

\[
G(v; S_0) = \mu (\{T \mid V^* (T; S_0) \geq v\}).
\]

\( G(v; S_0) \) is the probability that \( v \) is sustainable given \( S_0 \).

If one wants to compare two actions \( S_0, S'_0 \) such that \( W (S_0) = W (S'_0) \) at the bar of sustainability, one may prefer \( S_0 \) if \( G(W (S_0); S_0) > G(W (S_0); S'_0) \).

However, one may be worried that this inequality could be compatible with the risk of a catastrophe being greater with \( S_0 \), i.e., one could have \( G(v; S_0) < G(v; S'_0) \) for a low \( v \). The following proposition suggests that heroic assumptions are needed to assuage this worry.

**Proposition 12** Assume that for all \( T \in T \), \( V^* (T; \cdot) \) viewed as a function of \( S_0 \) yields the same ordering on the set of \( S_0 \). Then for all \( S_0, S'_0 \), either \( G(v; S_0) \geq G(v; S'_0) \) for all \( v \) or \( G(v; S_0) \leq G(v; S'_0) \) for all \( v \).

**Proof.** Under the assumption, \( V^* (T; S_0) \geq V^* (T; S'_0) \) iff \( V^* (T'; S_0) \geq V^* (T'; S'_0) \) for all \( T, T' \) and \( S_0, S'_0 \).

If \( V^* (T; S_0) \geq V^* (T; S'_0) \) for all \( T, T' \), then for all \( v \),

\[
\{T \mid V^* (T; S_0) \geq v\} \supset \{T \mid V^* (T; S'_0) \geq v\}
\]

and therefore

\[
\mu (\{T \mid V^* (T; S_0) \geq v\}) \geq \mu (\{T \mid V^* (T; S'_0) \geq v\}),
\]

i.e., \( G(v; S_0) \geq G(v; S'_0) \). \( \blacksquare \)

Therefore one may suggest that a reasonable test to assess what is bequeathed to future generations is to look at the whole graph of \( G(\cdot; S_0) \), which is, if it is continuous, equal to one minus the CDF of \( V^* (T; S_0) \).

Let us now relate this to the literature. The test \( G(W (S_0); S_0) > G(W (S_0); S'_0) \) is similar to what is proposed by Baumgärtner and Quass (2009) and Martinet (2011), who define sustainability in terms of making sure that the probability of maintaining the system in a certain range is above a certain level. This would suggest the test

\[
G(W (S_0); S_0) \geq p^*,
\]

for a desired probability threshold \( p^* \) (i.e., 95%).
They also suggest studying the probability of the system remaining within a certain range, which is similar to studying the function $G(\cdot; S_0)$; and studying the range that can be kept at a certain probability, which is like studying the inverse of $G(\cdot; S_0)$.

Cairns and Van Long (2006) and Asheim and Brekke (2002) explore a quite different approach which consists in translating uncertainty into the certainty framework, so that the standard certainty tools can then be applied. Cairns and Van Long focus on the expected value of next period welfare, while Asheim and Brekke consider the more general notion of a certainty-equivalent of next period welfare. A key difficulty with this approach is that what happens in time $t + 1$ depends on the action at time $t$, therefore a prediction of this action is needed. Cairns and Van Long solve a maximization program that defines conditional actions, while Asheim and Brekke, in similar fashion as Dasgupta and Mäler (2000), invoke a program that defines the likely $A_t$ given initial conditions $K_t$. Note that, unlike Dasgupta and Mäler, Asheim and Brekke do not assume that a program is given and instead define sustainability as the possibility for future generations to find and follow a program such that the certainty-equivalent of welfare is never lower than the welfare realized in the previous period.

In contrast, the CDF approach introduced here does not require any optimization or programmation about future decisions. The only thing that is relevant in the approach is the possibility of certain actions and consequences, and this is completely described by the probabilistic belief $\mu$.

5 Sustainability as a component of intergenerational welfare

How is sustainability related to intergenerational social welfare?

There is a long tradition that seeks to avoid discounting future welfare and considers more or less 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 alii’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 that the present (Asheim et al. 2012). The aim of this literature is 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.

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. Martinet (2012)
argues against taking sustainability as a constraint, because if no efficient path achieves sustainability, the sustainability constraint induces inefficient choices. One can interpret the maximization of social objective under a sustainability constraint as maximizing a lexicographic objective in which achieving sustainability has absolute priority over the second-tier objective.

So, the problem always boils down to incorporating a sustainability component in intergenerational welfare. In this section, I study the possibility of defining intergenerational welfare in terms of opportunities left to future generations. There is indeed an important difference in perspective between classical social welfare analysis and sustainability analysis. Classical social welfare analysis seeks to evaluate streams of consumption or utility and therefore needs to predict what future generations will do and enjoy. In contrast, sustainability analysis stops at the possibilities offered to future generations, as measured by $V_1(K_1)$, and is not interested in forecasting future behavior.

Is it possible to conceive an intergenerational welfare approach that would embody the sustainability perspective? Insofar as intergenerational welfare evaluation is meant to guide decisions, it would not make the decisions of the present generation depend on a precise prediction of what future generations will do. It would consider that it is enough to provide them with the means to have a good level of welfare. If instead they decide to make a special sacrifice for their descendants, this should not influence our current evaluation of the future. This echoes 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 don’t know what future generations will like, we cannot even compute $V_1(K_1)$. 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 $F(W_t, V_{t+1})$. This problem is avoided for sure only when $F$ is the maximin criterion $\min \{W_0, V_1\}$. By Proposition 6, it is also avoided under transferability and fungibility (if $F$ is increasing). 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:

\[ F(W_0, V_1) = F(W_0, W_0) + F(W_0, V_1) - F(W_0, W_0) \]

\[ = F(W_0, W_0) (1 + I^s - I^u), \]

where indexes of sustainability and unsustainability are defined as

\[ I^s = \max \left\{ 0, \frac{F(W_0, V_1) - F(W_0, W_0)}{F(W_0, W_0)} \right\}, \]

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

This approach is quite straightforward and may be taken as the most natural way of defining intergenerational welfare around sustainability as defined in this paper.

There is another, perhaps more ambitious approach that defines opportunities in a way that is consistent with a prior notion of social welfare. It departs from the direct sustainability approach, and is presented here as an intriguing possibility.

Let \( F(W_0, W_1, ..., W_H) \) 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.

The function \( F \) is calibrated so that the evaluation in the presence of risk is done by taking the expected value of \( F \).

We now construct a function derived from \( F \), but focuses on a form of sustainability. Let \( \tilde{W}_1(W_0, W_1, ...) \) be the solution to

\[ F(W_0, W_1, ..., W_H) = F(W_0, \tilde{W}_1, ..., \tilde{W}_1) \]

where the right-hand side has the same horizon \( H \) as the left-hand side. Let \( W_1^*(S_0) \) be the maximum feasible \( W_1 \) given \( S_0 \).

In general, \( W_1^*(S_0) = V_1(T(S_0)) \) is not guaranteed. By Lemma 1, under transferability,

\[ F(W(S_0), V_1(T(S_0)), ..., V_1(T(S_0))) \]

is the maximum feasible \( F(W(S_0), v, ..., v) \). As this is necessarily less or equal to the maximum feasible \( F(W_0, W_1, ...) \), which is equal to the maximum feasible \( F(W_0, \tilde{W}_1, ..., \tilde{W}_1) \), one has \( W_1^*(S_0) \geq V_1(T(S_0)) \). In summary, we have the following proposition:

**Proposition 13** One has \( W_1^*(S_0) \geq V_1(T(S_0)) \), with equality when \( F(W_0, W_1, ...) = F(W_0, \min \{W_1, \ldots\}) \).
The difference between $W_1^*(S_0)$ and $V_1(T(S_0))$ is quite important conceptually and, potentially, quantitatively. If it is impossible to sustain $W(S_0)$ in the future, it may be sufficient to ensure the possibility of a future that is as good as $(W(S_0),..., W(S_0))$ even if $W(S_t)$ will necessarily fall below $W(S_0)$ in the future. At the other extreme, it may be possible to obtain an opulent future that is as good as a high $(v,...,v)$, although $v$ is greater than $V_1(T(S_0))$. More controversially, this means that the present generation can enjoy a high $W(S_0)$ close to $v$, but that the good future will obtain only if the next generations make a substantial sacrifice. Whether one likes these conclusions or not, it is important to recognize that defining sustainability by reference to $W_1^*(S_0)$ is more consistent with $F$ than using $V_1(T(S_0))$.

Let $F(w) = F(w,...,w)$ and $F^*(S_0) = F(W(S_0),W_1^*(S_0),...,W_1^*(S_0))$. We can define

$$F^*(W_0,W_1,...) = \min \{ F^*(S_0), F(W_1^*(S_0)) \}.$$ 

This function $F^*$ embodies an absolute priority to avoiding unsustainability as defined by the condition $W(S_0) \leq W_1^*(S_0)$. This is similar to the equity condition considered in Asheim et al. (2012). Note again that the value of the function depends only on $A_0$, given the exogenous conditions, and not on any prediction of future decisions. A very nonstandard feature of this function is that it depends on the feasibility set.

The function $F^*$ is not an arbitrary construction, it is the only function that embodies the two key elements of the sustainability approach, namely, the idea that we need only to provide opportunities to future generations and not worry about what they actually do, and the absolute priority given to avoiding that future generations be worse off than the present generation—in terms of their opportunities.

**Proposition 14** Fix the technology $(\Phi,T)$. Assume that $F$ is continuous, nondecreasing in every argument and increasing in $(w,...,w)$, and satisfies the following axioms:

**Opportunities for the future:** If $(W_0,W_1,...,W_H)$ is feasible for $(\Phi,T)$, then for all $(W'_1,...,W'_H) \leq (W_1,...,W_H)$, $F(W_0,W_1,...,W_H) = F(W_0,W'_1,...,W'_H)$.

**Priority for the deserving future:** $F(x - \varepsilon, y + \delta, ..., y + \delta) \geq F(x, y, ..., y)$ if $x - \varepsilon > y, \varepsilon, \delta > 0$, and $F(x,y,...,y) = \max_{(x,W_1,...)} \text{feasible } F(x,W_1,...)$.

Then one has $F = F^*$.

**Proof.** By Opportunities for the future, fixing $S_0$ one has

$$F(W(S_0),W_1,...) = \max_{(W(S_0),W_1,...)} \text{feasible } F(W(S_0),W_1,...)$$

$$= F(W(S_0),W^*_1(S_0),...,W^*_1(S_0))$$

$$= F^*(S_0)$$
by definition of $W^*_1(S_0)$ (which is well defined by continuity and nondecreasingness) and $F^*(S_0)$.

By a standard argument using Hammond equity for the deserving future, continuity and nondecreasingness, one has $F(x, y, ..., y) = F(y, y, ..., y)$ whenever $x \geq y$ and $F(x, y, ..., y) = \max_{x \leq y} \text{feasible } F(x, W_1, ...)$.

Therefore, when $W(S_0) \geq W^*_1(S_0)$, $F(W(S_0), W_1, ...) = F(W^*_1(S_0), ..., W^*_1(S_0)) = \tilde{F}(W^*_1(S_0))$; when $W(S_0) \leq W^*_1(S_0)$, $F(W(S_0), W_1, ...) = F^*(S_0)$.

One has $W(S_0) \leq W^*_1(S_0)$ iff

$$
F^*(S_0) = F(W(S_0), W^*_1(S_0), ..., W^*_1(S_0)) \leq F(W^*_1(S_0), ..., W^*_1(S_0)) = \tilde{F}(W^*_1(S_0)).
$$

Therefore $F(W(S_0), W_1, ...) = \min \{ F^*(S_0), \tilde{F}(W^*_1(S_0)) \}$. ■

In the presence of risk, described by the density function $f$ on $W^*_1$, one can define the probabilities of sustainability and unsustainability as well as the expected values conditional on sustainability or unsustainability:

$$p^s = \int_{W_0}^{+\infty} f(w)dw, \quad p^u = 1 - p^s$$

$$E^s h = \frac{1}{p^s} \int_{W_0}^{+\infty} h(w) f(w)dw, \quad E^u h = \frac{1}{p^u} \int_{W_0}^{+\infty} h(w) f(w)dw,$$

and apply these notions of conditional expected value to

$$F^*(S_0) = F(W(S_0), W^*_1, ..., W^*_1),$$

$$\tilde{F}(W^*_1) = F(W^*_1, ..., W^*_1)$$

viewed as random variables depending on $W^*_1$, in order to obtain a decomposition of welfare into sustainability outcomes and unsustainability outcomes, as follows.

**Proposition 15** Let

$$I^s = \left( \frac{E^s F^*(S_0) - \tilde{F}(W_0)}{F(W_0)} \right) \quad \text{and} \quad I^u = \left( \frac{\tilde{F}(W_0) - E^u \tilde{F}(W^*_1)}{F(W_0)} \right)$$

measure the expected relative surplus of welfare above sustainability, and the expected relative deficit below sustainability, conditional on either sustainability or unsustainability being obtained. One then has

$$EF^s(W_0, W_1, ...) = \tilde{F}(W_0) (1 + p^s I^s - p^u I^u).$$

**Proof.** One can write:

$$EF^s(W_0, W_1, ...) = p^s E^s F^*(S_0) + p^u E^u \tilde{F}(W^*_1)$$

$$= \tilde{F}(W_0) + p^s E^s (F^*(S_0) - \tilde{F}(W_0)) - p^u E^u (\tilde{F}(W_0) - \tilde{F}(W^*_1))$$

$$= \tilde{F}(W_0) \left[ 1 + p^s E^s \left( \frac{F^*(S_0) - \tilde{F}(W_0)}{F(W_0)} \right) - p^u E^u \left( \frac{\tilde{F}(W_0) - \tilde{F}(W^*_1)}{F(W_0)} \right) \right].$$
This decomposition shows that one should care about $p^u$, i.e., the probability of unsustainability, but also about the expected gap between $F(W_0)$ and $F(W_1^*)$.

The social welfare approach suggests two criticisms of the sustainability framework. Both 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 respect of preferences as an important foundation for letting individuals live what they want rather than catering to their welfare in a paternalistic way (Fleurbaey 2008, ch. 10). But in the case of the collectivity 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. Of course, one can put whatever political and rationality constraints one deems realistic in the definition of $W_1^*$ or $V_1$, but that still leaves the gap between possibility and likely outcome, unless the constraints taken into account are so comprehensive that there is no difference between opportunities and outcomes.

Moreover, the measure of opportunities that is retained in the sustainability approach is especially severe, 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 social welfare, whereas future generations may not share the same goal. 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 notion of intergenerational social welfare that pertains to the evaluator, not to the collective agent receiving these opportunities.

Second, if we evaluate the future in a different way than the maximin, it is inconsistent with our view of social welfare to look at sustainability defined through $V_1$, and we should instead rely on $W_1^*$, as argued above. But this makes it even more questionable to look at possibilities rather than likely realizations, because the future may have great possibilities only if some the next generations make a great sacrifice, a sacrifice we can avoid making because we can feel satisfied that the possibilities for the future are satisfactory.

So, only two options seem sensible (and minimally charitable to future generations): 1) Retain an interest for sustainability but adopt an egalitarian approach to intergenerational welfare, so that $W_1^* \equiv V_1$ — this still only looks at opportunities for future generations but at least it checks that it is possible to have $W_t \geq W_0$ for all future generations; 2) Make forecasts of generational social welfare taking account of the likely future outcomes, not
just the opportunities left to future generations. The latter route suggests to enrich the analysis of intergenerational welfare as in the following decomposition:

\[ EF(W_0, W_1, \ldots) = \tilde{F}(W_0) (1 + p^s I^s - p^u I^u) + E (F - F^s) (W_0, W_1, \ldots) \].

Unlike the previous decomposition, this one requires a probabilistic prediction of the future path. The last term captures the gap between opportunities and likely outcomes.

6 Overlapping generations

This section briefly checks that the analysis remains broadly valid when generations overlap. 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 4** 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} \mid \exists S^+_t \in \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 that extends over several periods.

**Proposition 16** 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 \(S^t_{t+1} \in \Phi_{t+1}(T(S_t)), \forall \tau \geq t + 1, W(S_{\tau-L+1}, \ldots, S_{\tau}) \geq w\).

Assume \(V_0(K_0) < V_{t+1}(T(S_t))\) and \(w = \min \{W(S_{-L+1}, \ldots, S_0), \ldots, W(S_{-L+1}, \ldots, S_t)\} > V_{t+1}(T(S_t))\). Then there is a feasible \(S^t_{t+1}\) 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 \(t\).

7 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 \(V\) function 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 rather difficult to relate sustainability to intergenerational social welfare, unless one is comfortable with assigning collective responsibility to future generations for making the best of the opportunities we leave them.

If one does not feel comfortable with the thought that what future generations will do “is none of our business”, ensuring sustainability is not enough. 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 the obvious concern. In such a context, as observed by G. Asheim,7 “our obligations to the future generations depend on what they actually will choose to do”.

Does this concern for what actually happens require maximizing an intergenerational social objective? In practice, just as computing the \(V\) function may prove difficult in practice, it is quite unlikely that a consensual and reliable computation of an intergenerational function \(F\) can be done, and maximizing such a function is unlikely to become part of a realistic global political agenda. What analysts can do, however, is to make a direct prediction of the likely evolution of generational welfare \(W\) over the future. As an evaluation exercise, aggregating flows of generational welfare into measures of intergener-

7Personal communication.
ational welfare is certainly also permissible but raises additional difficult issues. Making good predictions of future paths of generational welfare would already be a key tool for decision-making.

Truly, many economists are also skeptical about the computation of generational welfare, but they are overly pessimistic. Although one should not hope to find a single consensual measure of generational welfare (because there are different philosophical conceptions of the good society), there are methods that make it possible to compute notions of social welfare incorporating efficiency and equity considerations and covering multiple dimensions of quality of life of the populations (Fleurbaey and Blanchet 2013).

In conclusion, combining the computation of unsustainability warnings with a direct forecast of the likelihood of various future paths may therefore retain the most compelling components of our concern for future generations. In a nutshell, if one follows this line, while avoiding unsustainability is a top priority, ensuring sustainedness rather than sustainability is the goal.

References


