Irreversibility, ignorance, and the intergenerational equity-efficiency trade-off

Nikolai Hoberg* and Stefan Baumgärtner

Department of Sustainability Sciences and Department of Economics,
Leuphana University of Lüneburg, Germany

December 9, 2011

Abstract: We demonstrate the existence of an intergenerational equity-efficiency trade-off in policy-making that aims at Pareto-efficiency across generations and sustainability, i.e. intergenerational equity. Our model includes two salient characteristics of sustainability problems and policy: (i) *temporal irreversibility*, i.e. the inability to revise one's past actions; (ii) *unawareness* ("*closed ignorance*"), i.e. future consequences of present actions in human-environment systems may be "unknown unknowns". If initially unforeseen sustainability problems become apparent and policy is enacted after irreversible actions were taken, policy-making faces a fundamental trade-off between intergenerational Pareto-efficiency and sustainability.

JEL-Classification: D3, H23, Q01, Q38, Q56

Keywords: climate change, closed ignorance, intergenerational equity-efficiency trade-off, irreversibility, Pareto-efficiency, sustainability, unawareness

---

*Corresponding author: Sustainability Economics Group, Leuphana University of Lüneburg, P.O. Box 2440, D-21314 Lüneburg, Germany, phone: +49.4131.677-2715, fax: +49.4131.677-1381, email: hoberg@uni.leuphana.de, http://www.leuphana.de.

1 Introduction

There is an increasingly alarming development in fundamental environmental indicators such as biodiversity, climate change and non-renewable resource scarcity (e.g. Millenium Ecosystem Assessment 2005, IPCC 2007, UNEP 2007). This has intensified the discussion on intergenerational justice, and has led to invoking sustainability as a societal goal in the design of policy (e.g. WCED 1987). In this vein, advocates of "climate justice" demand the equitable distribution of the benefits and damages from CO_2-emissions between developing and industrialized countries as well as between historic and future emitters (Neumayer 2000). Against the background of an apparently unsustainable situation, sustainability policy has to achieve an efficient and intergenerationally just allocation of resources and their benefits.[1]

A different, but similar, attempt to achieve justice *within* a generation – *intra*generational equity – in social policy is known to be subject to an equity-efficiency trade-off: the quest for equal utility levels (equity) incurs a (first-best) Pareto-inefficient allocation (Putterman et al. 1998, Le Grand 1990), because of different mechanisms such as incentive distortions or administrative costs. With this *intra*generational trade-off in mind, the question emerges whether policies aiming at sustainability are likewise subject to an *inter*generational equity-efficiency trade-off.

Following the intuition from the second welfare theorem, Howarth and Norgaard (1990, 1992) show that in an overlapping-generations-model both equity and efficiency can be achieved intergenerationally, given a set of public policies such as Pigouvian taxes, intergenerational transfer payments and the assignment of resource rights between generations. Krautkraemer and Batina (1999) find that a non-decreasing-utility constraint in a model with a renewable resource can lead to Pareto-inefficient overaccumulation of the resource. In that case, all generations could be made better off by allowing decreasing utility over time. Gerlagh and Keyzer (2001) compare different policy instruments for the sustainable intergenerational distribution of resources and find that a trust fund

[1] Sustainability policy goes beyond mere internalization of intertemporal externalities but aims at intergenerational equity (Pezzey 2004, Baumgärtner and Quaas 2010).

in which all natural resources and ecosystem services are administered that can be produced sustainably, leads to a Pareto improvement compared to a zero-extraction policy. Considering uncertain future outcomes and preferences, Krysiak (2009) finds a trade-off between protecting future individuals from potential harm (sustainability) and thereby abstaining from actions that would have made everyone better off (efficiency).

In this paper, we investigate how an intergenerational equity-efficiency trade-off in sustainability policy emerges from the genuine character and mechanisms of intergenerational policy-making. Compared to *intra*generational policy-making, there are two salient characteristics of sustainability problems and policy: (i) temporal irreversibility (Baumgärtner 2005), i.e. the inability to revise one's past actions; (ii) "closed ignorance" (Faber et al. 1992) or "unawareness" (Dekel et al. 1998), i.e. future consequences of present actions may be "unforeseen contingencies" (Dekel et al. 1998), also known as "unknown unknowns" (Rumsfeld 2002).

Our two-non-overlapping-generations model combines an intragenerational production decision on the use of circulating capital and a non-renewable resource, with a negative intergenerational externality: resource use in production by the first generation causes damages to the second generation. Initially, there is unawareness of this externality, i.e. it is an "unknown unknown" in the sense that it is unforeseen and the first generation is unaware even of its own ignorance. Damages only become apparent after production has irreversibly taken place. There is a social planner who uses two policy instruments – distribution of resource rights and capital transfer – to achieve the two goals of Pareto-efficiency across generations and sustainability, i.e. intergenerational equity. She shares the informational base of the first generation and enacts policy initially but may re-adjust her policy mix after damages have become apparent.

An important case in point is the current discussion of "climate justice" where the first generation in the model represents historic emitters (e.g. Europe and North America) who irreversibly used non-renewable fossil fuels for the production of consumption goods and, in the process, emitted greenhouse gases that lead to significant climate change. These actions were taken under unawareness of the effects of greenhouse gases on climate change. The second generation in the model represents future emitters (e.g.

China and India) who find diminished stocks of fossil fuels and also suffer the damages from climate change. The crucial challenge now is that climate policy is being shaped and implemented after historic production and emissions have already irreversibly taken place. While the amount of fossil fuels used for production in the past is irreversible, it is still possible to invest part of the historic emitters' output in capital for future emitters in order to address the concern for distributional equity.

We demonstrate that the social planner faces a fundamental trade-off between intergenerational Pareto-efficiency and sustainability: she can achieve either one of these two goals, but not both, if policy-making is done initially under unawareness and can be adjusted only after irreversible actions were made. That is, under these conditions she falls short of capturing the potential maximal utility. For climate policy this means that any attempt to achieve climate justice between historic and future emitters necessarily leads to Pareto-inefficiency, and Pareto-efficient policies will not be equitable.

2 Model

There are two successive, non-overlapping generations $t = 1, 2$. Both have identical, monotonic preferences over consumption C_t represented by a monotonic and concave utility function $U_t = U(C_t)$. Generation 1 is endowed with stocks of circulating capital and a non-renewable natural resource, with both stocks normalized to 1. Both generations use amounts K_t and R_t of capital and resource for the production of some intermediate good, $Y_t = F(K_t, R_t)$, where F is concave and exhibits positive and decreasing marginal products of both capital and resource input, F_{KR},$F_{RK} > 0$, and capital is essential for production, $F(0, R_t) = 0$. Of course, in the absence of any regulation generation 1 will use its capital stock completely $K_1 = 1$. The intermediate good thus produced in $t = 1$ can either be directly consumed by generation 1, or it may be transferred to generation 2 as circulating capital K_2:

$$C_1 = F(1, R_1) - K_2 \qquad (1)$$

Generation 2 will use all it inherits from generation 1 in production, K_2 and $R_2 = 1 - R_1$, and it will consume the entire amount of the intermediate good produced in $t = 2$.

The first generation's use of the resource in production causes damages $D(R_1)$ to the second generation, i.e. it diminishes the availability of their social product for consumption,

$$C_2 = (1 - D(R_1))F(K_2, 1 - R_1) \ , \tag{2}$$

with marginal damages being positive and increasing, $D'(R_1) > 0$ and $D''(R_1) \geq 0$, and total damages in the range $0 < D(R_1) < 1$ for all $R_1 > 0$ and $D(0) = 0$. To account for uncertainty on this actual fact, let $\kappa \in \{0, 1\}$ denote the state of information on damages. Expected second-generation consumption, contingent upon (un)awareness, then is:

$$C_2 = (1 - \kappa D(R_1))F(K_2, 1 - R_1) \ . \tag{3}$$

Initially, i.e. before any production, generation 1 is unaware of any potential future damages, $\kappa = 0$, and is not even aware of its ignorance, but firmly believes that its resource use does not entail any future damages, i.e. they are in a state of "ignorance" (sensu Faber et al. 1992). Thus, future damages are what has been called "unforeseen contingencies" (Dekel et al. 1998) or "unknown unknowns" (Rumsfeld 2002). Only after production by generation 1 has taken place, this unawareness is resolved and the full extent of damages becomes apparent, $\kappa = 1$.

There is a social planner who aims at (1) Pareto-efficiency across generations and (2) sustainability, i.e. non-decreasing utility over time. She acts during the first generation's lifetime and shares the same information as the first generation. In order to achieve her two goals, the social planner has two policy instruments at hand: (1) she can restrict resource use of generation 1, R_1, by an upper limit r; (2) she can oblige generation 1 to transfer at least k out of its intermediate product, Y_1, to generation 2 as capital, K_2.

The exact time structure is as follows. There are three time stages: $t = 1a$, $t = 1b$ and $t = 2$. Generation 1 lives in the first two of these, generation 2 lives in the last one. In the first stage $t = 1a$, generation 1 chooses its capital input K_1, resource input R_1, and capital transfer K_2 so as to maximize its own expected consumption C_1 subject to restrictions imposed by technology and policy. At this stage, the social planner may restrict resource use by r and make generation 1 plan with a minimal capital transfer of k. Production takes place in this stage, so that the inputs are irreversibly sunk, but as production takes time, the output is not turned out before the next stage. In the second stage $t = 1b$, output Y_1 of the intermediate good becomes available for use. At the same time, uncertainty is resolved and the future damages $D(R_1)$ from using the resource in production become fully apparent. In reaction to this information, the social planner can adjust her policy at this stage. As production by generation 1 has already taken place and resources R_1 are irreversibly sunk, she cannot revise the restriction on resource use any more. However, she can still adjust her second policy instrument and force generation 1 to transfer a higher amount k out of its intermediate good, thereby reducing generation 1's consumption C_1 and increasing generation 2's consumption C_2. Generation 1 cannot revise its production decision anymore at this stage, as the inputs are irreversibly sunk. In the third stage $t = 2$, generation 2 derives utility from its production of the intermediate good Y_2, which is entirely consumed in this same stage, with a reduction due to the damages caused by generation 1's resource use.

3 Definitions

First, we need to distinguish three definitions of feasibility as there are irreversibility, unawareness, and at a later stage awareness about future damages in the model. Ex-ante (ex-post) feasibility refers to those allocations that are deemed feasible at $t = 1a$ under unawareness (awareness) of future damages. Reduced feasibility refers to those allocations that are feasible at $t = 1b$, i.e. under awareness, after one has acted irreversibly at $t = 1a$. In the following we denote an allocation by $X = (K_1, R_1, Y_1, C_1, K_2, R_2, Y_2, C_2)$.

Definition 1 (Feasibility)

An allocation X is called *ex-ante (ex-post) feasible* if

$$\begin{aligned} 0 \leq K_1 \leq 1,\ 0 \leq K_2 \leq F(K_1, R_1),\ R_1 + R_2 \leq 1,\ R_1, R_2 \geq 0, \\ C_1 = Y_1 - K_2,\ Y_1 = F(K_1, R_1), \\ C_2 = Y_2 = (1 - \kappa D(R_1))F(K_2, 1 - R_1) \ \text{ with } \kappa = 0 \ (\kappa = 1). \end{aligned} \tag{4}$$

For any $0 \leq \overline{K}_1 \leq 1, 0 \leq \overline{R}_1 \leq 1$, and thus $\overline{Y_1} = F(\overline{K_1}, \overline{R_1})$, realized at $t = 1a$, an allocation is called *reduced feasible* if

$$\begin{aligned} 0 \leq K_2 \leq F(\overline{K_1}, \overline{R_1})\ , R_2 \leq 1 - \overline{R_1}\ , R_2 \geq 0\ , \\ C_1 = \overline{Y_1} - K_2, \\ C_2 = Y_2 = (1 - D(\overline{R_1}))F(K_2, 1 - \overline{R_1}). \end{aligned} \tag{5}$$

We understand the terms "sustainability" and "efficiency" as follows. Sustainability is defined as equal utility over time – the minimum requirement for the usual notion of sustainability as non-decreasing utility over time (Howarth 1995). With appropriate specification of the state of information, the criterion is as follows.

Definition 2 (Sustainability)
An ex-ante (ex-post) feasible intergenerational allocation X is called *ex-ante (ex-post) sustainable* if and only if it is yields

$$U_2 = U_1 \tag{6}$$

$$\text{where } U_1 = U(F(K_1, R_1) - K_2) \tag{7}$$

$$\text{and } U_2 = U((1 - \kappa D(R_1))F(K_2, R_2)) \text{ with } \kappa = 0 \ (\kappa = 1)\ . \tag{8}$$

Similarly, efficiency is defined in an information-and-irreversibility-differentiated manner in the sense of Pareto-efficiency. *Ex-ante efficiency* means that one cannot make a generation better off without making the other worse-off under unawareness of the damages from resource use before any irreversibility in resource use has taken effect. This is the relevant efficiency criterion to guide policy-making in $t = 1a$. *Ex-post efficiency* refers to the hypothetical case where there is awareness of the damages initially, i.e. before any irreversibility has taken effect, so that policy can be fully adjusted to future damages. Thus, it indicates the maximal potential utility in the system that is obtainable under awareness of the inevitable damages of resource use.

As a consequence of irreversible resource use in $t = 1a$, there is a reduced set of feasible actions in $t = 1b$. This irreversibility has to be taken into account by the policy-relevant efficiency criterion at this stage which together with awareness of damages leads to *reduced-feasibility efficiency*.

Definition 3 (Efficiency)

a) An ex-ante (ex-post) feasible allocation X is called *ex-ante (ex-post) efficient* if and only if there exists no other ex-ante (ex-post) feasible allocation X' for which $U'_t \geq U_t$ for $t = 1, 2$ and $U'_t > U_t$ for at least one t.

b) Contingent on $K_1 = \overline{K_1}, R_1 = \overline{R_1}, Y_1 = \overline{Y_1}$, a reduced feasible allocation X is called *reduced-feasibility efficient* (for short: *RF-efficient*) if and only if there exists no other reduced-feasible allocation X' with $K'_1 = \overline{K_1}, R'_1 = \overline{R_1}, Y'_1 = \overline{Y_1}$ for which $U'_t \geq U_t$ for $t = 1, 2$ and $U'_t > U_t$ for at least one t.

With this definition, efficient allocations are characterized as follows.

Lemma 1

An ex-ante (ex-post, reduced) feasible intergenerational allocation X is ex-ante (ex-post, reduced) efficient if and only if it meets the following conditions:

(i) Ex-ante efficiency:

$$F_R(1, R_1) F_K(K_2, 1 - R_1) = F_R(K_2, 1 - R_1), \tag{9}$$

$$F(K_2, 1 - R_1) = \overline{C} \text{ with } \overline{C} \in [0, \overline{C}^{EA,max}], \tag{10}$$

(ii) Ex-post efficiency:

$$F_R(1, R_1) F_K(K_2, 1 - R_1) - D'(R_1) F(K_2, 1 - R_1)/(1 - D(R_1)) = F_R(K_2, 1 - R_1), \tag{11}$$

$$(1 - D(R_1)) F(K_2, 1 - R_1) = \overline{C} \text{ with } \overline{C} \in [0, \overline{C}^{EP,max}], \tag{12}$$

(iii) Reduced-feasibility efficiency:

$$(1 - D(\overline{R_1})) F(K_2, 1 - \overline{R_1}) = \overline{C} \text{ with } \overline{C} \in [0, \overline{C}^{RF,max}], \tag{13}$$

where $\overline{C}$ is an intergenerational distribution parameter with $\overline{C} \in [0, \overline{C}^{i,max}]$, $i \in \{EA, EP, RF\}$.

Proof. See Appendix A.1. □

As consumption is the only good in the model, the intergenerational distribution parameter $\overline{C}$ determines not only consumption, but also utility levels. It can attain any value between 0 (all potential utility in this system is with generation 1, none with generation 2) and some $\overline{C}^{max}$ (all potential utility is with generation 2, none with generation 1), so that there exist infinitely many efficient allocations satisfying these characterizations. Conditions (9), and (11) state that the marginal gain in consumption for generation 2 from either of the following two alternative uses of the resource should be equal: (LHS) giving the additional resource to generation 1 as input into production, and then transferring the entire additional amount of the intermediate good thus produced as capital into generation 2's production; (RHS) giving the additional resource directly to generation 2 as input into their production. While Condition (9) expresses this without taking damages into account ($\kappa = 0$), Condition (11) states the same for the case where the damages are known ($\kappa = 1$) and taken into account from the beginning.[2] Condition (13) states that varying $\overline{C}$ determines the RF-efficient capital transfer K_2 which generates the set of RF-efficient allocations.

One can illustrate the efficient allocations through continuous and monotonically decreasing utility frontiers $U_2(U_1)$ in utility space (Figure 1). Unawareness (awareness) at $t = 1a$ about the damages yields the ex-ante (ex-post) utility frontier which runs from $U_2^{EA,max}$ ($U_2^{EP,max}$) to U_1^{max}.

Lemma 2

The ex-post utility frontier is the envelope of the reduced-feasibility utility frontiers that result for all $\overline{R_1} \in [0, 1]$.

Proof. See Appendix A.2. □

The ex-post utility frontier forms the envelope of all RF-utility frontiers. The RF-utility frontier depends on the actual realization of $R_1 = \overline{R_1}$ at $t = 1a$ and runs from $U_2^{RF,max}$ to $U_1^{RF,max}$. Due to the envelope property of the ex-post frontier and unaware-

[2] As (11) differs from (9), ex-ante efficient allocations are, in general, ex-post inefficient.

ness about factual damages, the ex-ante and ex-post frontiers both Pareto-dominate the RF-utility frontier. In the same figure, sustainable allocations are represented as a 45°-line from the axes.

4 Results

Under a laissez-faire policy the first generation will use up its entire circulating capital $K_1^0 = 1$, exploit the total amount of the resource $R_1^0 = 1$, and will not provide a capital transfer to the next generation, $K_2^0 = 0$. As a consequence, the first generation is better-off than the second one $U_1^0 > U_2^0$, as $C_1^0 = F(1,1) > 0$ and $C_2^0 = F(0,0) = 0$ (illustrated by X^0 in Figure 1). While this laissez-faire allocation is efficient by any notion of efficiency, it is not sustainable. This motivates sustainability policy by the social planner.

Sustainability policy has to follow the time structure laid out in Section 2. At $t = 1a$, the social planner devises a policy mix of restrictions on resource use and a capital transfer which should lead to an intergenerational allocation that is both ex-ante efficient and ex-ante sustainable.

Proposition 1 (Ex-ante sustainable and ex-ante efficient policy)
At time $t = 1a$, there exists a unique policy mix $(\tilde{r}, \tilde{k})$ that leads to an allocation $\tilde{X}$ which is both ex-ante sustainable and ex-ante efficient. It is characterized by the following necessary and sufficient conditions:

$$F_R(1, \tilde{r}) F_K(\tilde{k}, 1 - \tilde{r}) = F_R(\tilde{k}, 1 - \tilde{r}) \; , \tag{14}$$

$$F(1, \tilde{r}) - \tilde{k} = F\left(\tilde{k}, 1 - \tilde{r}\right) \; . \tag{15}$$

Proof. See Appendix A.3. □

In climate policy, the policy mix $(\tilde{r}, \tilde{k})$ refers to the situation under which production and consumption of the historic emitters took place. The use of fossil fuels was thought to be harmless as the effects of greenhouse gases on the atmosphere were unknown unknowns. Future emitters were thought to be at least equally well-off due to capital transfers and the ability to use the remaining fossil fuels. Yet, with the first report of the IPCC (1990) there was strong evidence that anthropogenic climate change was real and of relevant magnitude, and estimates of the damages due to historic greenhouse-gas emissions became available.

Thus, in the second stage $t = 1b$ the damages from resource use $D(R_1)$ are known, $\kappa = 1$. Obviously, not adapting the policy mix $(\tilde{r}, \tilde{k})$ to the new findings would result in an ex-post unsustainable allocation, with generation 2 worse-off than generation 1, as $U(F(1, \tilde{r}) - \tilde{k}) > U((1 - D(\tilde{r}))F(\tilde{k}, 1 - \tilde{r}))$. Therefore, the social planner must adapt her policy mix to ensure ex-post sustainability. However, production of the first generation of $\tilde{Y}_1 = F(\tilde{K}_1, \tilde{R}_1)$ has already taken place and inputs are irreversibly sunk. The only viable instrument is therefore to increase the minimum transfer of capital from generation 1 to generation 2, $k > \tilde{k}$. As this does not allow the achievement of ex-post efficiency, the policy maker faces the following fundamental sustainability-efficiency trade-off.

Proposition 2 (Sustainability-efficiency trade-off)
In general, and in particular for the ex-ante efficient and ex-ante sustainable policy $(\tilde{r}, \tilde{k})$, policy-making at time $t = 1b$ faces the following trade-off between ex-post efficiency and ex-post sustainability:

(i) there exists no policy mix $(r = \tilde{r}, k)$ that yields an allocation that is both ex-post efficient and ex-post sustainable, but

(ii) there exists a unique policy mix $(\hat{r} = \tilde{r}, \hat{k})$ with $\hat{k} > \tilde{k}$ that yields an allocation $\hat{X}$ that is ex-post sustainable but not ex-post efficient, and

(iii) there exists another unique policy mix $(r^* = \tilde{r}, k^*)$ with $k^* < \tilde{k}$ that yields an allocation X^* that is ex-post efficient but not ex-post sustainable.

Proof. See Appendix A.4. □

The intuition behind this result is as follows. Despite the damages the social planner can still achieve an ex-post sustainable allocation at time $t = 1b$ by adjusting her policy mix (Proposition 2 ii). This requires generation 1 to transfer more of its intermediate product as circulating capital to the second generation than originally planned, $\hat{k} > \tilde{k}$. As this transfer would exceed the one originally deemed necessary for sustainability and generation 1's production is irreversible, Condition (11) for ex-post efficiency is violated. Still, sustainability is achieved in spite of the damages and the irreversibility of resource use, as $U(\hat{C}_1) = U(F(1, \tilde{r}) - \hat{k}) = U((1 - D(\tilde{r}))F(\hat{k}, 1 - \tilde{r})) = U(\hat{C}_2)$. In climate policy this corresponds to a higher transfer of capital from historic to future emitters to compensate them for damages from climate change.

Alternatively, an ex-post efficient allocation X^* can be achieved, by decreasing the capital transfer in $t = 1b$, i.e. $k^* < \tilde{k}$ and $r^* = \tilde{r}$ (Proposition 2 iii). It is, however, not ex-post sustainable as $U(C_1^*) = U(F(1, r^*) - k^*) > U((1 - D(r^*))F(k^*, 1 - r^*)) = U(C_2^*)$. For climate policy this would mean a lower capital transfer and therefore no compensation to future emitters for damages from climate change. As the minimum capital transfer k is the only remaining policy variable at time $t = 1b$, and $k = \hat{k}$ would ensure ex-post sustainability while any $k \neq \tilde{k}$ leads to ex-post inefficiency, there exists no k that achieves both ex-post sustainability and ex-post efficiency (Proposition 2 i).

Despite this fundamental trade-off in policy-making with the two policy instruments studied here – a limit r on resource use by generation 1 and a minimum capital transfer k from generation 1 to generation 2 – there exists, in principle, a reduced feasible allocation that is both ex-post efficient and ex-post sustainable. Yet, this allocation X^{Bliss} has to be seen in light of the intragenerational equity-efficiency trade-off where a first-best efficient and equitable allocation is feasible, but not achievable given incentive distorting instruments.

Proposition 3 (Bliss)
There exists a unique policy mix (r^{Bliss}, k^{Bliss}) that yields an ex-post efficient and ex-post sustainable allocation X^{Bliss} which is reduced feasible, that is, feasible under unawareness and irreversibility.

Proof. See Appendix A.5. □

The various policies are illustrated with regard to sustainability and efficiency in

Figure 1. The laissez-faire allocation X^0 without policy interference is unsustainable,

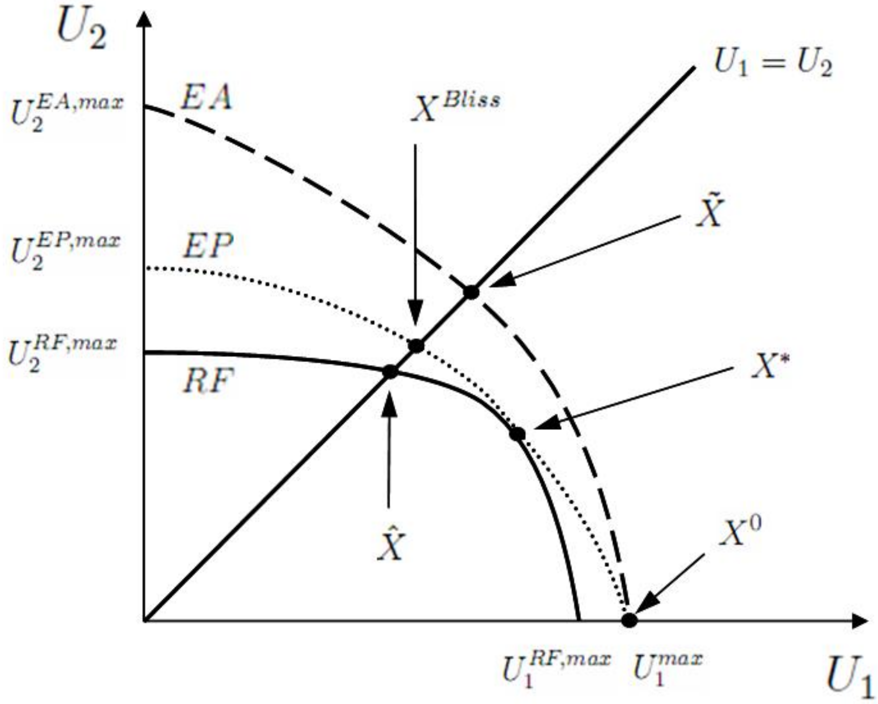

Figure 1: Illustration of sustainable allocations (45°-line), ex-ante efficient allocations (EA, dashed), ex-post efficient allocations (EP, dotted), and reduced-feasibility efficient allocations (RF, solid).

but ex-ante and ex-post efficient. The ex-ante efficient and ex-ante sustainable allocation $\tilde{X}$ that would result from policy $(\tilde{r}, \tilde{k})$ in $t = 1a$ lies at the intersection of the ex-ante utility frontier (dashed curve) with the sustainability line (45°-line). With damages to generation 2 becoming apparent in $t = 1b$ and capital transfer k^*, this allocation lies below the sustainability line and becomes the ex-post efficient, yet ex-post unsustainable allocation X^*.

All possible redistributions through a capital transfer k from generation 1 to generation 2 at $t = 1b$ after irreversibility has taken effect, i.e. starting from X^*, generate the RF-utility frontier (solid curve). Depending on whether the transfer is increased

$k > \tilde{k}$ or decreased $k < \tilde{k}$ one moves along the RF-utility frontier closer or further away from the sustainability line. The RF-utility frontier lies strictly below the ex-post utility frontier (dotted curve) except at X^* where both coincide (for $k = k^*$). This RF-utility frontier allows attaining the ex-post sustainable, yet ex-post inefficient allocation $\hat{X}$. Beyond the RF-utility frontier, there exists the ex-post sustainable and ex-post efficient allocation X^{Bliss}) at the intersection of the ex-post utility frontier and the sustainability line.

The trade-off between ex-post sustainability and ex-post efficiency at stage $t = 1b$ consists in the impossibility to reach both goals at once, i.e. the social planner must choose between the ex-post efficient allocation X^* and the ex-post sustainable allocation $\hat{X}$, or any combination of the two on the RF-utility frontier. Without irreversibility or unawareness there would be no such trade-off: if there were no irreversibility the ex-post utility frontier would be attainable in $t = 1b$ as the policy mix could easily be adjusted to the previously unknown damages. If there was no unawareness of future damages the social planner would choose a sustainable allocation on the ex-post utility frontier in $t = 1a$ and irreversibility would not be an issue as sustainability problems would be apparent from the very beginning.

5 Conclusion

We have studied the question of whether there exists a mechanism genuine to intergenerational policy-making that causes an intergenerational equity-efficiency trade-off. We found that sustainability policy that acts under a combination of temporal irreversibility and unawareness faces such a trade-off between efficiency across generations and intergenerational equity.

This result is relevant for current climate policy. Policies that want to achieve sustainability after damages were initially unknown (unawareness) must respect that past actions cannot be undone (temporal irreversibility), and that redistribution must therefore face a trade-off between efficiency and sustainability. For the case of climate justice – where climate policy is enacted after production and emissions have already irreversibly taken place – this means that there is a trade-off between equity and efficiency among historic and future emitters. Policymakers therefore need to be aware of the fact that pursuing sustainability as the overriding priority sacrifices efficiency, and that prudent policymaking requires a prior debate on how to balance these two conflicting goals.

Another conclusion concerns the timing of climate policy. As our analysis highlights, while uncertainty about damages is resolved over time, irreversibility reduces the set of feasible policy options and induces efficiency losses. This reduces the well-known quasi-option value of waiting under circumstances of uncertainty (Arrow and Fisher 1974, Henry 1974, Hanemann 1989) that has been cited to justify a procrastination of climate policy. Irreversibility, hence, provides an argument for enacting climate policy sooner rather than later, even while uncertainty about damages prevails.

Yet, there is still need for more research to analyze different instruments of sustainability policy with respect to their effect on efficiency (as e.g. in Gerlagh and Keyzer 2001). Describing and quantifying the trade-offs between sustainability and efficiency is necessary for the design of concrete policies. After all, we do not want to pay more for sustainability than necessary.

Acknowledgements

We are grateful to two anonymous referees from another journal and Geir Asheim for exceptionally insightful and helpful comments, to Maik Heinemann for critical and constructive discussion, and to the German Federal Ministry of Education and Research for financial support under grant 01UN1011A.

References

Arrow, K.J. and Fisher, A.C. (1974). Environmental preservation, uncertainty, and irreversibility. *Quarterly Journal of Economics*, 88(1):312–319.

Baumgärtner, S. (2005). Temporal and thermodynamic irreversibility in production theory. *Economic Theory*, 26(3):725–728.

Baumgärtner, S. and Quaas, M.F. (2010). Sustainability economics – general versus specific, and conceptual versus practical. *Ecological Economics*, 69:2056-2059.

Dekel, E., Lipman, L. and Rustichini, A. (1998). Recent development in modeling unforeseen contingencies *European Economic Review*, 42(3-5):523–542.

Faber, M., Manstetten, R. and Proops, J.L.R. (1992). Humankind and the environment: An anatomy of surprise and ignorance. *Environmental Values*, 1(3):217–242.

Gerlagh, R. and Keyzer, A.M. (2001). Sustainability and the intergenerational distribution of natural resource entitlements. *Journal of Public Economics*, 79:315–341.

Hanemann, W.M. (1989). Information and the concept of option value. *Journal of Environmental Economics and Management*, 16:23–37.

Henry, C. (1974). Investment decisions under uncertainty: the irreversibility effect. *American Economic Review*, 64:1006–1012.

Howarth, R.B. (1995). Sustainability under uncertainty: a deontological approach. *Land Economics*, 71(4):417–427.

Howarth, R.B. and Norgaard, R.B. (1990). Intergenerational resource rights, efficiency, and social optimality. *Land Economics*, 66(1):1–11.

Howarth, R.B. and Norgaard, R.B. (1992). Environmental valuation under sustainable development. *American Economic Review – Papers and Proceedings*, 82(2):473–477.

IPCC (1990). *IPCC First Assessment Report*. UNEP, New York.

IPCC (2007). *Climate Change 2007: The Physical Science Basis ; Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. UNEP, New York.

Krautkraemer, A.J. and Batina, G.R. (1999). On sustainability and intergenerational transfers with a renewable resource. *Land Economics*, 75(2):167–184.

Krysiak, F.C. (2009). Sustainability and its relation to efficiency under uncertainty. *Economic Theory*, 41(2):297–315.

Le Grand, J. (1990). Equity versus efficiency: the elusive trade-off. *Ethics*, 100(3):554-568.

Millenium Ecosystem Assessment (2005). *Ecosystems and Human Well-Being.* Island Press, Washington, DC.

Neumayer, E. (2000). In defence of historical accountability for greenhouse gas emissions. *Ecological Economics*, 33(2).

Pezzey, J.C.V. (2004). Sustainability policy and environmental policy. *Scandinavian Journal of Economics*, 106(2):339–359.

Putterman, L., Roemer, J.E. and Silvestre, J. (1998). Does egalitarianism have a future?. *Journal of Economic Literature*, 36(2):861–902.

Rumsfeld, D. (2002). US Department of Defense News Briefing on February 12, 2002. Available at http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636.

UNEP (2007). *Global Environment Outlook 4: Environment for Development.* UNEP, Kenya.

World Commission on Environment and Development (WCED) (1987). *Our Common Future.* Oxford University Press, Oxford.

Appendix

A.1 Proof of Lemma 1

As consumption is the only good in the model, Pareto-efficiency can be analyzed on the level of consumption. Thus, the intergenerational distribution parameter $\overline{C}$ equally determines the utility levels.

(i) An ex-ante feasible ex-ante efficient allocation is the solution to

$$\max_{K_1,R_1,K_2,R_2} C_1 \text{ s.t. } C_2 \geq \overline{C},\ C_1 = F(K_1,R_1) - K_2,\ C_2 = F(K_2,R_2),\ R_1 + R_2 = 1,\ K_1 = 1, \tag{A.16}$$

with the Lagrangian $\mathfrak{L} = F(1,R_1) - K_2 + \lambda_1(F(K_2, 1-R_1) - \overline{C})$. Obviously, in the optimal solution the constraint $C_2 \geq \overline{C}$ must hold with equality. The necessary first order conditions then are:

$$\frac{\partial \mathfrak{L}}{\partial R_1} = F_R(1,R_1) + \lambda_1 F_R(K_2, 1-R_1)(-1) = 0\ , \tag{A.17}$$

$$\frac{\partial \mathfrak{L}}{\partial K_2} = (-1) + \lambda_1 F_K(K_2, 1-R_1) = 0\ , \tag{A.18}$$

$$\frac{\partial \mathfrak{L}}{\partial \lambda_1} = F(K_2, 1-R_1) - \overline{C} = 0\ . \tag{A.19}$$

Rearranging (A.17) and (A.18) and eliminating λ_1 by dividing the two, one arrives at (9). (A.19) yields (10). These conditions are also sufficient, as the optimization problem (A.16) is strictly convex.

(ii) An ex-post feasible ex-post efficient allocation is the solution to

$$\max_{K_1,R_1,K_2,R_2} C_1 \text{ s.t. } C_2 \geq \overline{C},\ C_1 = F(K_1,R_1) - K_2,\ C_2 = (1-D(R_1))F(K_2,R_2),\ R_1 + R_2 = 1,\ K_1 = 1, \tag{A.20}$$

with the Lagrangian $\mathfrak{L} = F(1,R_1) - K_2 + \lambda_2((1-D(R_1))F(K_2,1-R_1) - \overline{C})$. Obviously, in the optimal solution the constraint $C_2 \geq \overline{C}$ must hold with equality. The necessary first order conditions then are:

$$\frac{\partial \mathfrak{L}}{\partial R_1} = F_R(1,R_1) + \lambda_2(-D'(R_1)F(K_2,1-R_1) + (1-D(R_1))F(K_2,1-R_1)(-1)) = 0, \tag{A.21}$$

$$\frac{\partial \mathfrak{L}}{\partial K_2} = (-1) + \lambda_2(1-D(R_1))F_K(K_2, 1-R_1) = 0\ , \tag{A.22}$$

$$\frac{\partial \mathfrak{L}}{\partial \lambda_1} = (1-D(R_1))F(K_2, 1-R_1) - \overline{C} = 0\ . \tag{A.23}$$

Rearranging (A.21) and (A.22) and eliminating λ_1 by dividing the two, one arrives at (11). (A.23) yields (12). These conditions are also sufficient, as the optimization problem (A.20) is strictly convex.

(iii) A reduced feasible RF-efficient allocation with given $K_1 = \overline{K_1}, R_1 = \overline{R_1}$ is the solution to

$$\max_{K_2} C_1 \text{ s.t. } C_2 \geq \overline{C},\ C_1 = F(\overline{K_1}, \overline{R_1}) - K_2,\ C_2 = (1 - D(\overline{R_1}))F(K_2, \overline{R_2}), \quad \text{(A.24)}$$
$$\overline{R_1} + \overline{R_2} = 1,\ \overline{K_1} = 1, 0 \leq K_2 \leq K_2^{RF,max} = F(\overline{K_1}, \overline{R_1})$$

with the Lagrangian $\mathfrak{L} = F(1, \overline{R_1}) - K_2 + \lambda((1 - D(\overline{R_1}))F(K_2, 1 - \overline{R_1}) - \overline{C})$. Obviously, in the optimal solution the constraint $C_2 \geq \overline{C}$ must hold with equality. The necessary first order conditions then are:

$$\frac{\partial \mathfrak{L}}{\partial K_2} = -1 + \lambda(1 - D(\overline{R_1}))F_K(K_2, 1 - \overline{R_1}) = 0\ , \quad \text{(A.25)}$$
$$\frac{\partial \mathfrak{L}}{\partial \lambda} = (1 - D(\overline{R_1}))F(K_2, 1 - \overline{R_1}) - \overline{C} = 0\ , \quad \text{(A.26)}$$

Choosing $\overline{C}$ determines K_2 in (A.26). K_2 determines λ in (A.25). Thus, (A.26) yields (13). Varying $\overline{C}$ between $\overline{C}^{min} = 0$ and $\overline{C}^{RF,max}$ by varying K_2 between 0 and $K_2^{max} = F(\overline{K_1}, \overline{R_1})$ generates the set of RF-efficient allocations.

A.2 Proof of Lemma 2

Decision problem (A.20) is the envelope to decision problem (A.24). Note, that this analysis holds equally for utility as consumption is the only good in the model. To prove this, define the value function $v(K_2)$ as the solution to the optimization problem (A.20):

$$v(K_2) = F(1, R_1(K_2)) - K_2. \quad \text{(A.27)}$$

The derivative of $v(K_2)$ is:

$$\frac{\partial v(K_2)}{\partial K_2} = F_R(1, R_1)\frac{\partial R_1}{\partial K_2} + (-1) \quad \text{(A.28)}$$

From (A.20) define constraint function $g(R_1(K_2), K_2) = (1 - D(R_1(K_2)))F(K_2, 1 - R_1(K_2)) - \overline{C}$ which is 0 for all K_2.

From (A.21) we know:

$$F_R(1, R_1) = \lambda_2 \frac{\partial g(R_1(K_2), K_2)}{\partial R_1} \quad \text{(A.29)}$$

Inserting this into (A.28) leads to

$$\frac{\partial v(K_2)}{\partial K_2} = \lambda_2 \frac{\partial g(R_1(K_2), K_2)}{\partial R_1} \frac{\partial R_1}{\partial K_2} + (-1) \tag{A.30}$$

As $dg(R_1(K_2), K_2)/dK_2 = 0$ this leads to:

$$\frac{\partial g(R_1(K_2), K_2)}{\partial R_1} \frac{\partial R_1}{\partial K_2} = -\frac{\partial g(R_1(K_2), K_2)}{\partial K_2} \tag{A.31}$$

Inserting this into (A.30) this leads to:

$$\frac{\partial v(K_2)}{\partial K_2} = -1 - \lambda_2 \frac{\partial g(R_1(K_2), K_2)}{\partial K_2} = -1 - \lambda_2((1 - D(R_1))F_K(K_2, 1 - R_1)) \tag{A.32}$$

For Pareto-efficiency the value function must be maximized which requires $\partial v(R_1^*, K_2)/\partial K_2 = 0$ and results the same condition as in (A.25). Therefore, the ex-post utility frontier forms the envelope of the RF-utility frontier.

A.3 Proof of Proposition 1

As consumption is the only good in the model, the analysis of consumption levels holds equally for utility. At time $t = 1a$, the social planner sets (r, k), expecting – given her unawareness, $\kappa = 0$ – that both generations, when maximizing their individual consumption subject to constraints from technology and policy, end up in an allocation $X = (K_1, R_1, Y_1, C_1, K_2, R_2, Y_2, C_2)$ with

$$C_1 = F(1, r) - k \tag{A.33}$$

$$C_2 = F(k, 1 - r) \ . \tag{A.34}$$

The social planner chooses (r, k) so that X is ex-ante efficient, i.e. it satisfies Conditions (9) and (10) (Lemma 1(i)), and ex-ante sustainable, i.e. it fulfills Condition (6) (Definition 2). With (A.33),(A.34) these conditions are

$$F(k, 1 - r) = F(1, r) - k \tag{A.35}$$

$$F_R(1, r)F_K(k, 1 - r) = F_R(k, 1 - r), \tag{A.36}$$

$$F(k, 1 - r) = \overline{C} \ . \tag{A.37}$$

There exists a unique value of $\overline{C} \in [0, \overline{C}^{EA,max}]$ so that this system can be solved for $(\tilde{r}, \tilde{k})$; with this value of $\overline{C}$, (A.35)–(A.37) reduce to Conditions (14),(15) and $(\tilde{r}, \tilde{k})$ is uniquely determined. To see this, note that $\overline{C}$ determines (r, k). Think of C_1 as a function of C_2 (defined by A.33, A.34, A.36, A.37 through variation of $\overline{C}$, where $\overline{C} = C_2$ as shown in Appendix A.1(i)) and consider first the minimal and maximal achievable consumption levels, indicated by $C_t^{EA,min}$ and $C_t^{EA,max}$, respectively. Setting $\overline{C} = 0$ implies $k^{EA,min} = 0$ and $r^{EA,min} = 1$, which yields $C_1^{EA,max} = F(1,1) > 0$ and $C_2^{EA,min} = F(0,0) = 0$. Setting $\overline{C} = \overline{C}^{EA,max}$ implies $k^{EA,max} = F(1, r^{EA,max})$. Inserting $k^{EA,max}$ into (A.36) uniquely yields $r^{EA,max}$, so that $C_1^{EA,min} = F(1, r^{EA,max}) - k^{EA,max} = 0$ and $C_2^{EA,max} = F(k^{EA,max}, 1 - r^{EA,max}) > 0$. By (A.37) we know that $dk/d\overline{C} = 1/F_K > 0$ and $dr/d\overline{C} = -1/F_R < 0$. As $F(\cdot,\cdot)$ is concave (by assumption) and C_1 is decreased linearly by increasing k as in (1), increasing k and reducing r by the ex-ante efficient mix (A.36) via increasing $\overline{C}$ from 0 to $\overline{C}^{EA,max}$ decreases C_1 monotonically from $C_1^{EA,max}$ to 0. As all functions involved are continuous $C_1(\overline{C})$ is continuous. Increasing $\overline{C}$ simultaneously increases C_2 continuously and monotonically from 0 to $C_2^{EA,max}$. Thus, by the intermediate value theorem and monotonicity, there exists a unique value of $\overline{C}$ so that the corresponding $(\tilde{r}, \tilde{k})$ yields $C_1 = C_2$, i.e. it fulfills (A.35).

A.4 Proof of Proposition 2

At time $t = 1b$, resource use $\tilde{r} = \tilde{R}_1$ is already irreversibly sunk in production and only the transfer of capital k can be adjusted. As consumption is the only good in the model, the analysis of consumption levels holds equally for utility.

(i) As noted in Appendix A.3 $(\tilde{r}, \tilde{k})$ meets Condition (9) so $F_R(1, \tilde{r}) F_K(\tilde{k}, 1 - \tilde{r}) = F_R(\tilde{r}, 1 - \tilde{r})$. Thus, $(\tilde{r}, \tilde{k})$ cannot meet Condition (11) for ex-post efficiency as $F_R(1, \tilde{r}) F_K(\tilde{k}, 1 - \tilde{r}) - D'(\tilde{r}) F(\tilde{k}, 1 - \tilde{r})/(1 - D(\tilde{r})) < F_R(\tilde{k}, 1 - \tilde{r})$. Furthermore, $(\tilde{r}, \tilde{k})$ does not meet Condition (10) for ex-ante efficiency due to the damages $D(R_1)$. Thus, no adaptation of the ex-ante efficient and ex-ante sustainable policy $(\tilde{r}, \tilde{k})$ yields an allocation that is ex-ante and ex-post inefficient and ex-post unsustainable as

$$F(1,\tilde{r}) - \tilde{k} > (1 - D(\tilde{r}))F(\tilde{k}, 1 - \tilde{r}) \tag{A.38}$$

Increasing k to some $k^b > \tilde{k}$ to move towards a sustainable allocation does not allow to meet Condition (11). This is due to positive decreasing marginal utility in both inputs, which changes the left and right hand side of (11) to: $F_R(1,\tilde{r})F_K(k^b, 1-\tilde{r}) - D'(\tilde{r})F(k^b, 1-\tilde{r})/(1-D(\tilde{r})) < F_R(1,\tilde{r})F_K(\tilde{k}, 1-\tilde{r}) - D'(\tilde{r})F(\tilde{k}, 1-\tilde{r})/(1-D(\tilde{r})) < F_R(\tilde{k}, 1-\tilde{r}) < F_R(k^b, 1-\tilde{r})$. Therefore, there exists no policy that is ex-post efficient and ex-post sustainable.

(ii) At time $t = 1b$ the social planner needs to resort to a higher capital transfer $k > \tilde{k}$ in her policy $(\tilde{r}, k)$ to achieve sustainability due to the damages $D(\tilde{r})$. As shown in (A.38) the second generation's consumption level is, $\hat{C}_2 = (1-D(\tilde{r}))F(k, 1-\tilde{r})$. With this condition for the behavior of the second generation and the irreversible production decision $\tilde{Y_1} = F(1,\tilde{r})$ the effect of the level of capital transfer on utility can be derived

$$\text{for generation 1:} \quad \hat{C}_1 = F(1,\tilde{r}) - k \ , \tag{A.39}$$

$$\text{for generation 2:} \quad \hat{C}_2 = (1 - D(\tilde{r}))F(k, 1-\tilde{r}) \ . \tag{A.40}$$

The social planner sets k so that X is ex-post sustainable, i.e. it fulfills Condition (6) (Definition 2) and is on the RF-utility frontier, i.e. it fulfills Conditions (A.39) and (A.40). As the production function is concave an increase of k monotonically decreases the generation 1's consumption/utility level (A.39) while monotonically increasing the generation 2's consumption/utility level (A.40). Think of $C_1 - C_2$ as a function of k and consider minimal and maximal achievable consumption levels $C_t^{RF,min}, C_t^{RF,max}$ along the RF-utility frontier in (A.39) and (A.40) with irreversible resource inputs $\tilde{r} = \tilde{R}_1$ and $1 - \tilde{r} = \tilde{R}_2$. Setting $k = 0$ leads to $C_1^{RF,max} = F(1,\tilde{r}) > 0$ and $C_2^{RF,min} = (1-D(\tilde{r}))F(0, 1-\tilde{r}) = 0$. Setting $k = k^{RF,max}$ leads to, $C_1^{RF,min} = F(1,\tilde{r}) - k^{RF,max} = 0$ and $C_2^{RF,max} = (1-D(\tilde{r}))F(k^{RF,max}, 1-\tilde{r}) > 0$. As $F(K_t, R_t)$ is concave and C_1 is reduced linearly by increasing k, increasing k decreases $C_1 - C_2$ monotonically from $C_1^{RF,max} - C_2^{RF,min} > 0$ to $C_1^{RF,min} - C_2^{RF,max} < 0$. As all functions involved are continuous $C_1 - C_2$ is continuous. Thus, by the intermediate value theorem and monotonicity there exists a unique policy mix $(\hat{r}, \hat{k})$ with $\hat{r} = \tilde{r}$ that yields the allocation $\left(\hat{K}_1, \hat{R}_1, \hat{Y}_1, \hat{C}_1, \hat{K}_2, \hat{R}_2, \hat{Y}_2, \hat{C}_2\right)$ that is both on the RF-utility frontier and ex-post sustainable, i.e. if fulfills $C_1 = C_2$. Therefore, there exists a $\hat{k}$ for which $\hat{C}_1 = F(1, \tilde{r}) - \hat{k} = (1 - D(\tilde{r}))F(\hat{k}, 1 - \tilde{r}) = \hat{C}_2$. As shown in Appendix A.4(i) for $\hat{k} > \tilde{k}$, Condition (11) for ex-post efficiency is not met.

(iii) As shown in Appendix A.4(i) $(\tilde{r}, \tilde{k})$ does not meet Condition (11) and $\tilde{R}_1 = \tilde{r}$ is irreversible. At time $t = 1b$ k can still be adapted in the range $k \in [0, k^{EP,max} = F(1, \tilde{r})]$.

The social planner sets k so that X is ex-post efficient, i.e. it fulfills Conditions (11) and (12) (Lemma 1(ii)) and is on the RF-utility frontier, i.e. it fulfills (A.39) and (A.40). This leads to the following system:

$$F_R(1, \tilde{r})F_K(k, 1 - \tilde{r}) - D'(\tilde{r})F(k, 1 - \tilde{r})/(1 - D(\tilde{r})) = F_R(k, 1 - \tilde{r}), \quad \text{(A.41)}$$

$$C_2 = (1 - D(\tilde{r}))F(k, 1 - \tilde{r}) = \overline{C} \quad \text{(A.42)}$$

There exists a unique value of $\overline{C} \in [0, \overline{C}^{EP,max}]$ so that this system can be solved for (r^*, k^*). To see this, note that $\overline{C}$ determines k and therefore C_2 in (A.40) and C_1 in (A.39). For ex-post efficiency consider the effect of minimal and maximal capital transfers on (A.41). From (A.41) define a function $\phi(k) = F_R(1, \tilde{r})F_K(k, 1 - \tilde{r}) - D'(\tilde{r})F(k, 1 - \tilde{r})/(1 - D(\tilde{r})) - F_R(k, 1 - \tilde{r})$.

Setting $\overline{C} = 0$ implies $k^{EP,min} = 0$. For $\phi(0)$ this yields $\phi(0) = F_R(1, \tilde{r})F_K(0, 1 - \tilde{r}) - D'(\tilde{r})F(0, 1 - \tilde{r})/(1 - D(\tilde{r})) - F_R(0, 1 - \tilde{r}) > 0$.

Setting $\overline{C} = \tilde{C}$ implies $\tilde{k}$ from Appendix A.3. As shown in Appendix A.4(i) this yields: $\phi(\tilde{k}) = F_R(1, \tilde{r})F_K(\tilde{k}, 1 - \tilde{r}) - D'(\tilde{r})F(\tilde{k}, 1 - \tilde{r})/(1 - D(\tilde{r})) - F_R(\tilde{k}, 1 - \tilde{r}) < 0$.

As a decreasing k monotonically increases $F_K(k, 1 - \tilde{r})$, monotonically decreases $F(k, \tilde{r})$ and monotonically decreases $F_R(k, 1 - \tilde{r})$, $\phi(k)$ is monotonically decreasing in k. Varying k from 0 to $\tilde{k}$ changes $\phi(k)$ from $\phi(0) > 0$ to $\phi(\tilde{k}) < 0$. As all functions involved are continuous $\phi(k)$ is continuous. Thus, by the intermediate value theorem and monotonicity, there exists a unique value of k and a corresponding value of $\overline{C}$ so that $\phi(k^*) = 0$ and (r^*, k^*) is ex-post efficient, i.e. it fulfills (A.41). Therefore, there exists a unique policy (r^*, k^*) with $r^* = \tilde{r}$ and $k^* < \tilde{k}$ that yields an allocation $X^* = (K_1^*, R_1^*, Y_1^*, C_1^*, K_2^*, R_2^*, Y_2^*, C_2^*)$ that is ex-post efficient.

A.5 Proof of Proposition 3

As consumption is the only good in the model, the analysis of consumption levels holds equally for utility. When the the social planner sets (r, k) at time $t = 1a$ under awareness of the damages $D(R_1)$ the generations' consumption levels at time $t = 1b$ are:

$$C_1 = F(1, r) - k \ , \tag{A.43}$$

$$C_2 = (1 - D(r))F(k, 1 - r) \ . \tag{A.44}$$

For an ex-post efficient allocation Conditions (11) and (12) must hold (Lemma 1(ii)), and Condition (6) for sustainability (Definition 2). With (A.43) and (A.44) these conditions are:

$$(1 - D(r))F(k, 1 - r) = F(1, r) - k \tag{A.45}$$

$$F_R(1, r)F_K(k, 1 - r) - D'(r)F(k, 1 - r)/(1 - D(r)) = F_R(k, 1 - r), \tag{A.46}$$

$$(1 - D(r))F(k, 1 - r) = \overline{C} \ . \tag{A.47}$$

There exists a unique value of $\overline{C} \in [0, \overline{C}^{Bliss,max}]$ so that this system can be solved for (r, k); with this value of $\overline{C}$ (r, k) is uniquely determined. To see this, note that $\overline{C}$ determines (r, k). Think of C_1 as a function of C_2 (defined by A.43, A.44, A.46, A.47 through variation of $\overline{C}$, where $\overline{C} = C_2$ as shown in Appendix A.1(i)) and consider first the minimal and maximal achievable consumption levels, indicated by $C_t^{Bliss,min}$ and $C_t^{Bliss,max}$, respectively. Setting $\overline{C} = 0$ implies $k^{Bliss,min} = 0$ and $r^{Bliss,min} = 1$, which yields $C_1^{Bliss,max} = F(1, r^{Bliss,min}) > 0$ and $C_2^{Bliss,min} = (1 - D(r^{Bliss,min}))F(0, 1 - r^{Bliss,min}) = 0$. Setting $\overline{C} = \overline{C}^{Bliss,max}$ implies $k^{Bliss,max} = F(1, r^{Bliss,max})$. Inserting $k^{Bliss,max}$ into Equation (A.46) uniquely yields $r^{Bliss,max}$, so that $C_1^{Bliss,min} = F(1, r^{Bliss,max}) - k^{Bliss,max} = 0$ and $C_2^{Bliss,max} = (1 - D(r^{Bliss,max}))F(k^{Bliss,max}, 1 - r^{Bliss,max}) > 0$. By (A.47) we know that $dk/d\overline{C} = 1/(1 - D(r))F_K > 0$ and $dr/d\overline{C} = -1/((1 - D(r))F_R - D'F(k, 1-r)) < 0$. As F is concave and C_1 is decreased linearly by increasing k as in (1), increasing k and reducing r by the ex-post efficient mix (A.46) via increasing $\overline{C}$ from 0 to $\overline{C}^{Bliss,max}$ decreases C_1 monotonically from $C_1^{Bliss,max}$ to 0. As all functions involved are continuous $C_1(\overline{C})$ is continuous. Increasing $\overline{C}$ simultaneously increases C_2 continuously and monotonically from 0 to $C_2^{Bliss,max}$. Thus, by the intermediate value theorem and monotonicity, there exists a unique value of $\overline{C}$ so that the corresponding (r^{Bliss}, k^{Bliss}) and allocation $X^{Bliss} = \left(K_1^{Bliss}, R_1^{Bliss}, Y_1^{Bliss}, C_1^{Bliss}, K_2^{Bliss}, R_2^{Bliss}, Y_2^{Bliss}, C_2^{Bliss},\right)$ ensure $C_1 = C_2$, i.e. fulfill (A.45).

This allocation is reduced feasible, that is feasible under irreversibility and unawareness as: at $t = 1a$ the social planner can choose any $r \in [0, 1]$. From the existence proof its clear that $r^{Bliss} \in [0, 1]$. So, in $t = 1a$ the social planner chooses r^{Bliss} and some matching k. At $t = 1b$ the damage becomes apparent and r^{Bliss} is fixed. Still, k can be adjusted $k \in [0, F(1, r^{Bliss})]$ generating a RF-utility frontier as in the existence proof. Therefore, the policy (r^{Bliss}, k^{Bliss}) which yields an ex-post efficient and ex-post sustainable allocation is reduced feasible, that is feasible under unawareness and irreversibility.