EFFICIENCY AND FUTURE GENERATIONS

JOHN BROOME*

Abstract: Standard lessons from economics tell us that an externality creates inefficiency, and that this inefficiency can be removed by internalizing the externality. This papers considers how successfully these lessons can be extended to intergenerational externalities such as emissions of greenhouse gas. For intergenerational externalities, the standard lessons involve comparisons between states whose populations of people differ, either in their identities or their numbers. Common notions of efficiency break down in these comparisons. This paper supplies a new notion of efficiency that allows the lessons to survive, but at the cost of reducing their practical significance.

Keywords: Externalities, Climate change, Non-identity effect, Population ethics, Resourcism

1. EFFICIENCY AND FOUR LESSONS FROM ECONOMICS

Emissions of greenhouse gas are a negative externality. Those who cause emissions do not pay the full cost of them; instead, they impose external costs on all the people around the world who suffer from the climate change they cause.

Economics teaches us some important lessons about externalities and efficiency. These lessons apply to both positive and negative externalities, but I shall concentrate on negative ones such as greenhouse gas.

Efficiency is defined in terms of dominance:

Efficiency. A possible state B is efficient if and only if there is no possible state A that dominates it.

* Corpus Christi College, Oxford OX1 4JF, UK. Email: john.broome@philosophy.ox.ac.uk.
URL: http://users.ox.ac.uk/~sfop0060/

221

This paper considers various notions of dominance, each of which implies a different notion of efficiency. Until Section 3 I assume the population of people is fixed and I consider only Pareto dominance:

> **Pareto dominance.** A state *A* Pareto dominates another *B* if and only if at least one member of the population prefers *A* to *B* and no member prefers *B* to *A*.

Pareto efficiency is defined on the basis of Pareto dominance. Note that it is defined in terms of people's preferences rather than what is good for people. When *A* Pareto dominates *B*, I shall sometimes say it is a 'Pareto improvement' on *B*. Until Section 3, 'efficiency' refers specifically to Pareto efficiency.

By a 'state' I mean the whole past, present and future of the world. By 'possible' I mean practically possible at a particular time, given the Earth's resources and technology. All states that are possible at a time share the same past up to that time. In this paper I deal only with fully determined states; I abstract from uncertainty.

The first lesson from economics is that, except in some unusual circumstances,[1] an economy that contains an externality is not efficient. A Pareto improvement is possible.

The second lesson is that, when a state contains an externality, there is an efficient possible state that dominates it. This means that the inefficiency caused by an externality can be completely eliminated, and not merely reduced, in a way that requires no sacrifice from anyone: efficiency can be achieved without anyone's ending up in a state she disprefers.

The third lesson is that an externality can be corrected by 'internalizing' it. An externality is internalized if those who cause it pay a price for doing so that is equal to the external cost they impose on others. If the externality is internalized, it is no longer a source of inefficiency; if there is no other source of inefficiency, the resulting state is efficient. This is so even if the price paid is not passed on to those who are harmed by the externality.

Although the state that results from internalizing an externality is efficient, it does not normally dominate the original state that contains the externality. Normally, some of the people who pay the internalized cost of the externality end up in a state they disprefer to the original one. But the fourth lesson of economics is that internalizing the externality can be done in a way that brings about an efficient state that dominates the original one. Achieving this happy state will normally require that the internalizing of the cost is supplemented by transfer payments made by some people to others. These transfers can compensate those who pay

---

[1] The appendix contains examples of unusual circumstances.

the internalized cost, in such a way that no one ends up in a state she disprefers to the original.

These four lessons from economics apply to greenhouse-gas emissions. First, these emissions cause inefficiency. Second, the inefficiency can be fully eliminated without requiring a sacrifice from anyone. Third, the externality can be corrected by imposing a carbon price equal to the external cost of emissions. Fourth, an efficient state that dominates the original inefficient one can be achieved by means of a carbon price along with suitable compensating payments made by some people to others.

This comes as a surprise to many people. Many people assume that climate change calls for a sacrifice from the present generation for the sake of future generations. But no sacrifice is required to correct the inefficiency of greenhouse-gas emissions. Correcting the inefficiency would create a great benefit that could be shared between people in all countries and all generations in such a way that no one bears a cost.

If this surprises you, think of it this way. As things stand, we are leaving to our successors an atmosphere polluted with greenhouse gas. We are also leaving to them beneficial resources such as economic infrastructure and other capital goods, and those natural resources we leave in the ground rather than consume ourselves. The mix of good and bad things we leave them is inefficient. Our successors would prefer us to shift our present investment away from conventional investment in infrastructure and capital goods towards green investment such as windmills and insulation. This would reduce the burden of greenhouse gas we impose on them, and it would more than make up for the reduction in the other resources we leave them. They would prefer us to make this change, and it need not be a sacrifice for us, since we can achieve it merely by shifting the direction of our investment. We could maintain our own aggregate consumption at its present level, though we would no doubt have to adjust the types of goods we consume.

Just because a Pareto improvement is possible, it does not immediately follow that we should make one. Climate change involves an injustice in that the emitters of greenhouse gas unjustly harm those who suffer from the effects of greenhouse gas. A Pareto improvement will not correct this injustice. Moreover, the world's distribution of wealth is far from ideal, and a Pareto improvement will do little to improve it. Cost-benefit analyses of climate change suggest that the best response to climate change is one where the present generation does make a sacrifice.[2]

---

[2] This is the conclusion of both Stern *et al.* (2007) and Nordhaus (2008: 180). However, Rezai (2010) argues for the opposite view that the best response does not require a sacrifice from the present generation.
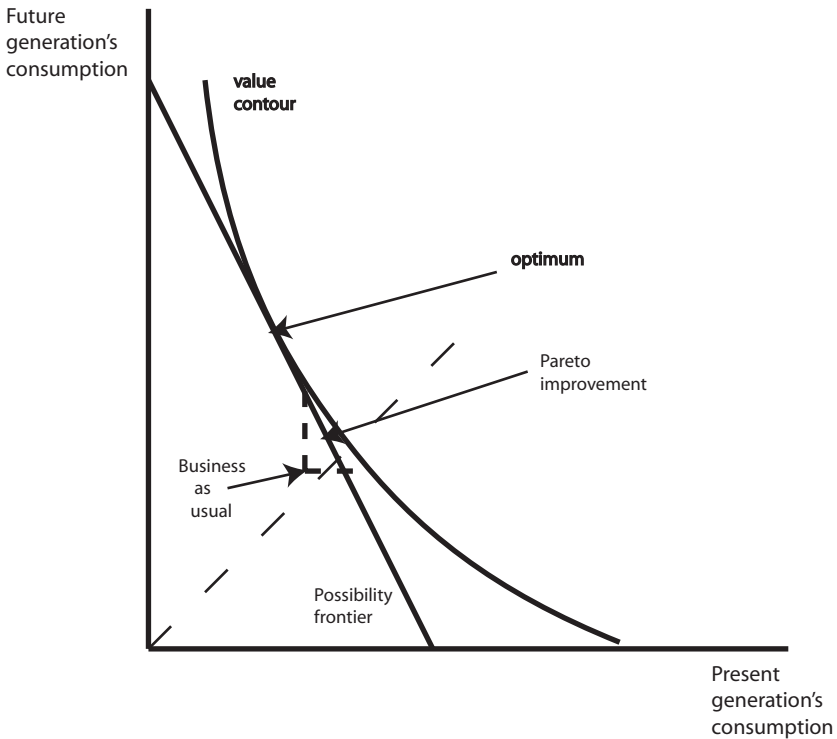
FIGURE 1. Pareto improvement and the best response.

Figure 1 illustrates this conclusion. It shows that economics can go beyond the four lessons I have mentioned in evaluating our response to climate change. Economics is not just about efficiency; its methods can also be used to make judgements about the overall value of alternative states. In particular, they can take account of distribution between generations. Nevertheless, this paper concentrates on efficiency for two reasons.

First, judgements of efficiency are less contentious than judgements of overall value, since they do not depend on interpersonal comparisons of value. Conclusions based on efficiency are therefore more robust than conclusions based on value.

Second, the best response to climate change, illustrated in Figure 1, is not attainable in practice. Attaining it would require the present generation to make a sacrifice for the sake of future people, and the world's failure over decades to respond adequately to climate change shows the present generation is unwilling to make this sacrifice. On the other hand, a Pareto improvement may be attainable because its attraction depends only on self-interest. To chase after the best response rather

than a Pareto improvement is to make the best the enemy of the good. The lessons of economics about externalities give us a new and hopeful approach to tackling climate change.[3]

But this paper is not about the importance of the lessons. It is about their truth. Most of the external harm done by our emissions will fall on generations of people who are not yet born. The very existence and identity of some of those people depend on what we do about our emissions. The standard theory of externalities fails when people's existence and identities are at stake. This paper considers how far the four lessons apply to this new context. I shall conclude that they can be made to apply, but only by altering the definition of efficiency. Making this alteration seriously weakens the lessons' significance.

Section 2 is a brief explanation of the four lessons within the standard theory of externalities. There is nothing new here. If you already know the economics, you could skip this section. The appendix sets out the lessons precisely, within an intertemporal general equilibrium model.

Intergenerational externalities can be included in the standard theory of externalities so they are covered to an extent in Section 2. They raise some problems that are well recognized within the theory. But the standard theory ignores problems of existence and identity. Those problems are explained in Section 3. Section 4 considers one approach to dealing with them, which fails. Section 5 describes the approach I favour, and Section 6 considers how far it succeeds. Section 7 draws my conclusions.

2.1. The first lesson

Suppose someone has a good that she is willing to sell at a particular price, and suppose someone else is willing to pay a higher price for it. If the first were to sell the good to the second at some intermediate price, both people would prefer the resulting state. No one else would be affected by this exchange, so it would be a Pareto improvement.

So long as one person is willing to pay for buying a good more than another person is willing to accept for selling it, a Pareto-improving exchange is possible. This means the situation is inefficient until an exchange takes place. If there is a well-functioning market for the good, exchanges like this will continue until all the opportunities for them are used up. At that point, people's willingnesses to pay and willingnesses to accept will all be equal. But if there is no such market, the inefficiency may persist.

---

[3] I was persuaded of this conclusion by Duncan Foley. He presents it in Foley (2009). See also Broome and Foley (2016) and Broome (2016).

An externality can be thought of as a good that is not marketed. Suppose you beautify your garden for your own enjoyment, but your neighbour also benefits from your work by enjoying her view of your garden. Let us suppose your neighbour would be willing to pay for an even better view, and that you would be willing to accept a payment to improve your garden still further. Then there is an opportunity for a Pareto-improving trade. But let us assume this sort of exchange is prohibited by the mores of friendly neighbourhood relations. The good of the view then counts as an externality just because your neighbour cannot pay you for it.

We are interested in the externality of greenhouse-gas emissions. This is a bad rather than a good. However, we can think of its opposite – a reduction in emissions – as a good, and it makes for a neater theory to deal in goods rather than bads. The quantity of this good can be measured by a negative number, which is minus the quantity of emissions.

A complication is that a reduction in emissions is a public good rather than a private one. A private good is one that benefits only a single person when it used. Examples are bicycles and sandwiches. A public good is one that many people benefit from, such as a city park. The good of reducing greenhouse-gas emissions is a public good that benefits almost everyone.

Like many other public goods, this one is produced by private companies and individuals. A company can produce more of this good by reducing its emissions. It would be willing to do so for a price. The consumers of the good are all the members of the public. They would be willing to pay for it. Suppose the total of what they would be willing to pay for a reduction comes to more than the company would be willing to accept for making a reduction. Then a Pareto improvement could be achieved by the company's reducing its emissions, and being paid for doing so by the public. In the case of a public good, a Pareto improvement is possible if the *total* of what the public is willing to pay for it is different from what a seller is willing to accept.

At present this is the situation in most countries. Emitters of greenhouse gas would be willing to reduce their emissions for a small price, but they are offered nothing for doing so. They emit freely at no cost. On the other hand, people around the would be willing to pay them to reduce their emissions. Unfortunately, they have no means of making the payment. So the situation is inefficient.

Much of the externality of greenhouse gas falls on future generations. To make a Pareto improvement in this case might require members of later generations to make payments to emitters of greenhouse gas in an earlier generation. These payments cannot be mere money or loans on paper; ultimately real goods would have to pass from later people to earlier ones. How is that possible? Because people's lives overlap. Present goods can be passed from a presently living young person Q to a presently living

old person *P*. *Q* can in turn receive goods later in her life from someone, *R*, who is still younger. *R* could fully compensate *Q* for what she earlier passed to *P*. In this way, real payment can pass from *R* to *P* even if *R*'s life does not overlap with *P*'s; *Q* serves as an intermediary. By iterating the process, real payment may be passed backwards across many generations.

2.2. The second lesson

The first lesson is that, if a state contains an externality, there is a possible state that Pareto dominates it. The second lesson is that there is in particular a possible state that Pareto dominates it and is itself Pareto efficient. Given the first lesson, the second seems obvious. Start with the original state and pick another possible state that Pareto dominates it. If that second state is not efficient, by definition there is a third possible state that Pareto dominates the second. This third state also Pareto dominates the first. And so on. Pick a sequence of Pareto dominating possible states until eventually you arrive at one that is not Pareto dominated by any other possible state. By definition, this is efficient.

This argument is not watertight, but the appendix contains a watertight proof.

2.3. The third lesson

In ideal conditions without externalities, the working of economic competition will ensure that each good has a single price, which is the same for everyone. No one can sell a good at a higher price than other sellers, because they will get no customers. Nor will anyone sell to a buyer who offers a lower price than other buyers. Furthermore, no individual has any control over the price. So long as there is a single price, there is no difference between what buyers are willing to pay for the good and what sellers are willing to accept for it. No Pareto-improving trades are possible. The outcome is therefore efficient. This is commonly known as 'the first theorem of welfare economics'.

This truth about competition depends on treating a good as a very specific thing. For example, rice is not one good but many. Rice at one place is not the same good as rice at another. Because it costs money to transport rice from one place to another, rice at different places may have different prices. Rice at one date is not the same good as rice at another. Rice at an earlier date can be converted into a greater quantity of rice at a later date by sowing it in the ground and reaping the crop that grows. So rice at different dates may have different prices. But if the economy is perfectly competitive, each specific good – each good at each place and each time – will have just one price.

One consequence of efficiency is that the production of goods is matched with demand for them. The market conveys information to

producers about consumers' preferences. Moreover, it conveys it in a way that provides incentives to the producers to satisfy those preferences. It also conveys information to consumers about the relative costs of production of different goods, and gives them an incentive to take these costs into account in their purchases.

How does this work between people who do not directly trade with each other? Tourists buying plastic nick-nacks in Europe do not trade directly with producers of plastic nick-nacks in China. So how do the producers know what particular nick-nacks to produce, and how are they incentivized to produce them? How do consumers take account of the costs of producing different sorts of nick-nacks? Through a chain of overlapping traders. Tourists in Europe buy nick-nacks from merchants, who in turn buy them from European importers, who buy them from Chinese exporters, who buy them from manufacturers. Incentivizing information is transmitted both ways along this chain in the form of the prices paid for the goods at each link.

In a way, the same mechanism of overlapping traders works intergenerationally. Suppose you are a forester today, wondering what species of trees to plant. Suppose your trees will not be felled for centuries. How do you decide between oaks and redwoods? If you plan to sell your forest before you die, you will choose a species on the basis of the price you expect to get for growing trees at the date when you sell. That will depend on the price the purchaser expects to receive when she in turn sells the forest on before she dies, and so on down the generations. Many generations in the future, someone will fell the trees and sell the timber to consumers. That person will receive a price that is determined by the preferences of the consumers at that time. When this final owner of the forest buys it, the price she is willing to pay is given by what she expects at the time to be the price she will eventually receive for timber from consumers. The previous owner will buy the forest on the basis of her expectation of the price the final owner will be willing to pay, and so on back through the generations. Incentivizing information about the preferences of the final consumers will be transmitted back to the present by this chain of owners in overlapping generations.

However, there is a difference. In the contemporaneous case, actual markets pass information back and forth along the chain of traders between Europe and China. In the intergenerational case, no information about future consumers' actual preferences can really come back through time along a chain of traders. Instead, we merely have a recursive chain of expectations. Present foresters have expectations about what the next generation will be willing to pay for forests containing particular species, and this depends on what they expect the next generation will expect of the subsequent generation, and so on. There is therefore no guarantee of genuine efficiency. It may turn out that foresters today plant species that

do not fit well with the preferences that future consumers will actually have. It might turn out that, had they planted different species, future consumers would have preferred the result, and no one would have dispreferred it. If so, a Pareto improvement would be possible, and the actual state is not truly efficient. For the intergenerational case, the market can achieve only a limited sort of efficiency that is relative to present people's expectations about future people's preferences.

Now back to externalities. Externalities prevent the market from working properly because, since the externality is not marketed, different people face different prices for it. If the emitters of a negative externality are made to pay a price for emissions that is equal to the value of the external harm they cause, the externality is internalized. The price they pay is in effect the price they receive for producing the public good that consists in reducing their emissions. A carbon price is an example. Internalizing the externality will achieve efficiency, so long as there is no other market failure. Efficiency requires that the price paid by emitters for emitting is equal to the total amount that members of the public would be willing to accept (at the margin) for giving up some of the good of greenhouse-gas reduction or – to put the same thing differently – what they would be willing to pay for receiving some of this good. The result will be efficient so long as emitters pay this price, even if the money is not passed on to the people who suffer from the emissions. This conclusion is a generalization of the first theorem of welfare economics. It is proved in the appendix. Internalization does not have to be achieved by a market. The price may be imposed on emitters by the government as a tax, through a cap and trade system or in some other way.

Internalizing the externality will remove the inefficiency. If there is no other source of inefficiency, it leads to a state where no Pareto improvement is possible.

2.4. The fourth lesson

The second lesson tells us there is a Pareto efficient state that is a Pareto improvement on the original state containing an externality. A version of the so-called 'second theorem of welfare economics' tells us that this efficient state can be the outcome of a market process in which the externality has a price that is equal to the value of the marginal harm the externality does. The appendix states this version of the theorem and cites a proof of it.

Achieving this state may well require some redistribution of income. In the case of climate change, if efficiency is achieved through a carbon price, many of those who pay the price – particularly the better-off among the present generation – will not directly benefit much from the slowing of climate change. So if there is to be a Pareto improvement, those

people will need to receive transfer payments from other people who do directly receive the benefit, to compensate them for paying the carbon price. The theorem tells us that transfer payments can indeed achieve full compensation, to the extent that no one need end up in a position she disprefers.

Some of this compensation can be achieved in practice by cutting other taxes. People who pay the carbon price, can be partially compensated by paying less of other sorts of tax. But some compensation will in practice probably have to be paid out of public debt, which is a practical means by which compensation can be transferred back from future generations. How full compensation can be made in practice is not the topic of this paper. (See Broome and Foley, 2016.)

The proof of the second theorem of welfare economics assumes that production processes and people's preferences are 'convex' in the mathematical sense. This is a real limitation on the theorem, since empirically not all production processes and preferences are convex. The theorem is true only in so far as nonconvexities do not stand in its way. If they do, the externality can in principle be corrected without a sacrifice on anyone's part, but this correction cannot be made by internalizing the externality in the way I have described. However, I know no particular reason for thinking nonconvexities will interfere with correcting the externality of greenhouse gas.

3. THE NON-IDENTITY PROBLEM FOR PARETO EFFICIENCY

The standard theory reviewed in Section 2 assumes that the same population of people exists in all possible states. But economic policy, including climate-change policy, affects the world's future population. We need to take this into account.

First, policy affects the identities of people who live in the future. A person's identity is determined by the particular sperm and egg that she grows from. Anyone who grew from a different sperm and egg would not be the same person. This means that a person's identity is sensitive to the very detailed circumstances of her conception; this is the 'non-identity effect'. Climate-change policy will make a difference to the way people live. It will change how much they travel, who meets whom, who has babies with whom, and when they have babies. Consequently, climate-change policy has a non-identity effect; it will cause different people to be born from those who would have been born had the policy not been implemented.[4]

Furthermore, economic policy may also affect the numbers of people who live in the future, as well as their identities.

---

[4] The moral significance of the non-identity effect was recognized by Parfit (1984: 351–80).

These effects on population pose problems for the notion of efficiency and for the economic theory of externalities. Some of the same problems arise even without a change in population. Even when a particular person exists in two different states, her preferences may differ between the states. The formal problems described below can arise equally from differing preferences as from the non-identity effect.

Given the possibility of a changing population or changing preferences, Pareto dominance splits into various versions. Here are some examples:

> **Pareto dominance from above.** A state *A* Pareto dominates a state *B* from above if and only if at least one person who exists in *A* prefers, in *A*, *A* to *B* and no one who exists in *A* prefers, in *A*, *B* to *A*.
> **Pareto dominance from below.** A state *A* Pareto dominates a state *B* from below if and only if at least one person who exists in *B* prefers, in *B*, *A* to *B* and no one who exists in *B* prefers, in *B*, *B* to *A*.
> **Narrow Pareto dominance.** A state *A* narrowly Pareto dominates a state *B* if and only if at least one person who exists in both *A* and *B* prefers, in both *A* and *B*, *A* to *B*, and no one who exists in both *A* and *B* prefers, either in *A* or in *B*, *B* to *A*.

To each version of dominance there is a corresponding version of efficiency.[5]

A useful notion of dominance needs to be asymmetric and acyclic.[6] It should not be possible both for *A* to dominate *B* and *B* to dominate *A*, or for *A* to dominate *B*, *B* to dominate *C* and *C* to dominate *A*. But none of these versions of Pareto dominance is guaranteed to be asymmetric and acyclic. Here are some examples.

Suppose that all those who exist in *A* prefer, in *A*, *A* to *B*, and all those who exist in *B* prefer, in *B*, *B* to *A*, then *A* Pareto dominates *B* from above, and also *B* Pareto dominates *A* from above. So Pareto dominance from above is not asymmetric.

The following example shows that narrow Pareto dominance can be cyclic. I represent states by a vector, where each place in the vector represents a particular person. If a person does not exist in a state, her place in the vector contains the symbol Ω. If she exists, her place contains a letter that denotes the sort of life she lives. Consider the three states (e, f, Ω), (Ω, e, f) and (f, Ω, e). Assume that each person, when she exists, prefers a state where she lives a life of sort *f* to one where she lives a life of sort *e*. Then

(e, f, Ω) narrowly Pareto dominates (Ω, e, f) because the only person who exists in both these states prefers the former. Similarly (Ω, e, f) narrowly Pareto dominates (f, Ω, e) and (f, Ω, e) narrowly Pareto dominates (e, f, Ω).

(Here is a case where narrow Pareto dominance is cyclic even though the population does not differ between states, but preferences do. There is one person and three states. In state A the person prefers C to A to B. In state B she prefers A to B to C. In state C she prefers B to C to A.)

The idea of Pareto dominance has little use if we can find no version of it that is asymmetric and acyclic. Possibly we can. We may be able to rescue Pareto dominance from below by making particular assumptions about the structure of people's preferences. These assumptions may be empirically plausible. It is an empirical fact that people tend to adapt to their condition. Their preferences tend to be biased towards the state they are in; people favour the state they are in more than people in other states favour it. Adaptation implies that Pareto dominance from below is asymmetric. If A Pareto dominates B from below, that means people in B favour A over their own state B. If B Pareto dominates A from below, that means people in A favour B over their own state A. It would be contrary to the empirical fact of adaptation for both of these things to be true.

The empirical fact of adaptation might perhaps also be made to support the claim that Pareto dominance from below is acyclic – though I do not know how to do this. If so, adaptation could plausibly make good sense of Pareto dominance from below. Still, adaptation is no more than a contingent feature of human psychology, and it would be a pity to have to rely on it in order to make sense of our concept of efficiency.

From this point on, I set aside problems that arise from the varying preferences of particular individuals, and concentrate on problems of identity only. To avoid complications, I shall assume that each person's preferences are the same in each state she exists in.

Even if Pareto dominance from below is asymmetric and acyclic, the non-identity effect falsifies the lessons of economics for Pareto efficiency from below. It falsifies even the first lesson that an externality leads to inefficiency. Take a possible state, B, with externalities. If we were to ignore the non-identity effect for a moment, the first lesson would tell us that B is Pareto inefficient. There would therefore be an alternative state A, having the same population as B, such that one person in B would prefer A to B and no one in B would prefer B to A. But to achieve this alternative state A would require some reallocation of goods among the people, and this might have a non-identity effect. The result would not be a state A having the same population as B, but some other, perhaps similar, state A' containing some different people. The one person in B who would prefer A might not even exist in A', and consequently may not prefer A' to B. We therefore have no reason to think that A' Pareto dominates B from below. B may indeed be Pareto efficient from below.

The conclusion of this section is that the non-identity effect throws into disarray the lessons economics teaches us about externalities and efficiency. In terms that will be recognized by moral philosophers, the problem is that Pareto efficiency is a 'person-affecting' criterion. It is concerned with how particular people are affected by a change. Person-affecting criteria do not stand up well to the non-identity effect.

4. BETTERNESS EFFICIENCY

Can the lessons of economics be reconstructed? They can, but only by radically revising the definition of efficiency.

Now we are dealing with cases where the population of people varies between states, we need to make comparisons between different people who exist in different states. Section 3 shows us that we cannot do this adequately in terms of preferences. You might at first think of switching to a definition of dominance and efficiency in terms of betterness rather than preferences, because interpersonal comparisons make good sense in terms of betterness: we can sensibly ask whether or not a person who exists in one state is better off than a different person who exists in another.

An attractive notion of dominance in terms of betterness is:

> **Suppes dominance.** A state *A* Suppes dominates another *B* if and only if there is a one-one correspondence between the people in *A* and the people in *B* such that at least one person in *A* is better off than the corresponding person in *B*, and every person in *A* is at least as well off as the corresponding person in *B*. (Suppes 1966.)

Suppes efficiency is defined correspondingly. If we want to understand efficiency in terms of betterness while recognizing the non-identity effect, Suppes efficiency seems the best way to do it.

However, the four lessons of economics do not apply to Suppes efficiency. The problem is that how well off a person is in a state depends only partly on her material conditions in that state. It also depends on her innate nature in the state. When we compare states in which different people exist, we have to recognize they may have different innate natures. This means the four lessons fail. For example, it might be that reducing greenhouse-gas emissions leads to the existence of some people who are innately miserable to an extent that cannot be sufficiently compensated for by giving them better material conditions. Then if *B* is a state with greenhouse-gas emissions, there may be no state *A* with reduced emissions that Suppes dominates it. *B* might be Suppes efficient. No doubt this is unlikely, but once we are comparing states containing different people, it cannot be ruled out.

Could we not say that, if reducing greenhouse-gas emissions leads to the existence of some innately miserable people, emissions have a positive externality that countervails against their negative externality? We could. But nevertheless they have a negative externality, which we need to recognize. Other things being equal, people are worse off in a state with a lot of emissions than they are in one with less. This is a fact we should be able to express using our notion of efficiency, and Suppes efficiency cannot express it.

The problem is compounded when we recognize that policies to reduce emissions may change the number of people as well as their identities. If the number of people changes, we cannot use a notion of efficiency that depends on one-to-one comparisons, as Suppes efficiency does. To make comparisons in terms of betterness in these cases, we would have to apply some theory of aggregate value. Indeed, we would need a population axiology, which takes account of the size of the population in assigning aggregate value to a state. But applying a theory of value takes us far beyond the notion of efficiency, which is the subject of this paper.

5. CONSTRAINED EFFICIENCY

If we are to salvage anything from the standard lessons, we need a notion of efficiency for future generations that is founded neither on people's preferences nor on what is better for people. I shall adopt *resourcism* for future generations. Resourcism is the idea that public ethics should be concerned with the resources that are available to people, and not directly with how successfully people use these resources to promote their good or satisfy their preferences. I suggest we build a definition of dominance and efficiency on the basis of the resources that are available to future generations, understood as all the goods that are available for their consumption.

For presently living people I suggest our definition should continue to embody the aim of satisfying their preferences. There is no non-identity effect among presently living people, so the difficulties I have described do not arise for them. I stretch the term 'present people' to cover everyone who is not subject to the non-identity effect: 'present people' includes people who exist in every possible state, even if they do not yet exist. 'Future people' covers everyone else.

Here is a crude version of the notion of dominance I have in mind:

Crudely constrained dominance. A state *A* crudely constrained dominates another *B* if and only if *A* Pareto dominates *B* for present people, and future people together consume at least as much of each good in *A* as future people together consume in *B*.

When I say '*A* Pareto dominates *B* for present people', I mean that at least one present person prefers *A* to *B* and no present person prefers *B* to *A*. The

idea is that present people are constrained not to diminish the resources that they leave to future people. Crudely constrained efficiency is defined correspondingly.

Notice that the resource constraint is on the total of goods that future people together consume. The constraint is that this total is at least as much in *A* as it is in *B*. It is distributed among whatever future people exist in the respective states, without any constraint on what each individual gets. Remember that the future people in the two states may differ in identity and also in number.

Remember too that some goods, such as reduction of greenhouse-gas emissions, are public goods. Public goods are consumed equally by everybody. When we work out the total of goods consumed, we add up everyone's consumption of private goods such as rice. So if 5 billion future people each consume a kilo of rice, 5 billion kilos of rice are included in the resource constraint. But we do not add up everyone's consumption of public goods or bads. If 20 billion tonnes of carbon dioxide are emitted, everyone suffers this particular quantity of this public bad. Twenty billion tonnes are included in the resource constraint.

Crudely constrained dominance is asymmetric and acyclic. If *A* crudely constrained dominates *B*, *A* Pareto dominates *B* for present people. Therefore *B* cannot Pareto dominate *A* for present people, and so *B* cannot crudely constrained dominate *A*. If *A* crudely constrained dominates *B* and *B* crudely constrained dominates *C*, then *A* Pareto dominates *B* for present people, and *B* Pareto dominates *C* for present people. It follows that *A* Pareto dominates *C* for present people. Therefore *C* cannot Pareto dominate *A* for present people, and so *C* cannot crudely constrained dominate *A*.

I deal only with externalities that impose some of their external costs on present people.[7] Greenhouse gas is one of these. When there is one of these externalities, present people could achieve a Pareto improvement for themselves by making exchanges among themselves. They would give up a certain amount of their consumption of private goods for the sake of increasing the amount of a public good such as greenhouse gas reduction. In doing this, they could leave the consumption of private goods by future people unaffected. So future people would end up with the same amount of private goods and more of the public good. The result would crudely constrained dominate the initial state. The initial state with the externality is therefore crudely constrained inefficient. This is the first lesson of economics applied to this sort of efficiency. The appendix demonstrates that the second lesson also holds.

Crudely constrained efficiency is therefore a way of salvaging some of the lessons of economics about externalities. However, it is a poor

---

[7] This is the effect of the 'substitution assumption' stated in the appendix.

way of doing so for long-lived externalities such as greenhouse gas. It recognizes only inefficiency among present people. But intuitively, the main inefficiency caused by greenhouse gas is not among present people but between present and future people. The mix of goods that we present people are leaving to future people is not the best. We are leaving them conventional capital and resources, which are good for them, and a dirty atmosphere, which is bad for them. We could leave them a mix that they would prefer, by switching some of our investment away from conventional goods to cleaning up the atmosphere instead. We could do this without reducing our own consumption, so the existing mix is intuitively inefficient. We need a more elaborate definition of constrained efficiency that recognizes inefficiency of this sort.

Here is a way to set one up. Define the *constraint set* of B as a set of bundles of goods. The intuitive idea is that the members of the constraint set are bundles of goods that are in some sense at least as beneficial to future people as those they get in B. At the moment I do not specify a criterion of beneficialness, so I leave the specification of the constraint set almost completely open. However, I do require that the constraint set includes every bundle of goods that contains at least as much of each good as future people consume in B itself.

> **Constrained dominance.** A state A constrained dominates another B if and only if A Pareto dominates B for present people, and future people in A consume some member of B's constraint set.

Present people are constrained not to worsen the resources that they leave to future people. Constrained dominance is asymmetric and acyclic for the same reasons as crudely constrained dominance is. Constrained efficiency is defined correspondingly.

The first lesson of economics applies to this sort of inefficiency. If there is an externality in a state B, it is constrained inefficient by this criterion. The argument is just as it was for crudely constrained inefficiency. Trading among present people can take them to a state that is a Pareto improvement for them, while leaving to future people the same private goods as they have in B and perhaps more of a public good. This bundle is within B's constraint set.

It may now be possible to go further. Present people could leave a different mix of goods to future people within the constraint set. They might be able to choose a different mix that is advantageous for present people: it might allow them to move to a state that Pareto dominates any state they can achieve just by trading among themselves. This captures the idea of inefficiency between present and future people.

That extends the first lesson to the new context. It applies whatever the constraint set is, so long as it includes every bundle of goods that is at least as big as the bundle future people get in the original state B. This

is an important aspect of this first lesson. We may use whatever criterion we think appropriate to judge the beneficialness of resources, and the first lesson still applies.

I cannot prove the truth of the second lesson for this sort of efficiency. I could run through the same argument as in Section 2.2, starting with a state that is constrained inefficient. There is another possible state that constrained dominates it. If this one is inefficient, there is another than in turn constrained dominates it, and so on. The difference from Section 2.2 is that I cannot guarantee this sequence of dominating states will come to an end. It will converge on some state – it will not go off to infinity – but the state it converges on may not be a possible one. The members of the sequence are each possible, and they get nearer and nearer to constrained efficiency – as near as we like – but they may never quite get there. This is mathematically unfortunate, but getting as near to efficiency as we like seems good enough in practice.

The third and fourth lessons of economics are true only if we specify the constraint set more closely, on the basis of future people's preferences. In a state B each future person consumes a bundle of goods including public goods. Putting together all their bundles we get a total bundle consumed by the future people in B. We specify the constraint set of B as all those bundles of goods that are sufficiently big that they could be distributed to the future people in B in such a way that none of those people ends up with a bundle she disprefers to her bundle in B. (This is known as the future people's 'Scitovszky set'; Scitovszky 1941.) If a state A constrained dominates B, its future people together consume a bundle of goods within this set. There is no constraint on what each individual gets.

When the constraint set is specified like this, I call it 'preference-based'. Constrained efficiency with a preference-based constraint set is plainly not entirely independent of preferences. Nevertheless, it is fair to call it a resourcist criterion for efficiency. The constraint is that future people get resources that are as beneficial as they are in the initial state B. It does not require that future people's preferences are satisfied or that future people are well off to any particular degree. It is just that the criterion for beneficialness is based on the preferences of those people who live in the initial state.

The appendix shows that the third and fourth lessons of economics apply to constrained efficiency with a preference-based constraint. The fourth lesson is once again subject to the assumption that non-convexities do not interfere. Since we now cannot rely on the second lesson, the fourth lesson is also subject to the assumption that there is an efficient state that dominates the initial state. The fourth lesson tells us that, if there is, it can be attained by internalizing the externality, together with suitable transfers.

6. INTERPRETATION

The upshot is that the lessons from economics can survive the difficulties created by changes in population, provided we adopt an appropriate version of resourcism for future people. We have to treat efficiency as constrained efficiency rather than Pareto efficiency. How damaging is this retreat? What significance should we attach to constrained efficiency?

First, constrained efficiency with a preference-based constraint set cannot be relative to future people's actual preferences. It is relative to what their preferences are expected to be. I explained this in Section 2.2. Suppose we are in a state that is efficient in this way, relative to our expectations. If our expectations are wrong, there could be another state that is preferred by future people. This is an inevitable weakness in applying the lessons of economics to future people. Since future people do not participate in the present market, any efficiency the market achieves cannot be relative to future people's actual preferences. This weakness has nothing to do with the special problems raised by the non-identity effect and changes in population. It applies equally to Pareto efficiency as to constrained efficiency.

Let us set this weakness aside. A different weakness applies to constrained efficiency in particular, as a consequence of the resourcism embedded in its definition. This sort of efficiency requires present people to leave to future people a collection of resources within the constraint set. Imagine we are in a state B with externalities. It is constrained inefficient, which means a constrained dominating state A is possible. The definition of constrained domination ensures that at least one present person prefers A to B, and no present person prefers B to A. But nothing guarantees that future people in B will not prefer B to A. After all, many of them would not exist in A, and they may prefer a state where they exist. Also, nothing guarantees that the future people who exist in B will not be better off than the future people who would exist in A. For example, the future people in A may be innately more miserable.

Remember especially that the two states may have different numbers of future people, so different numbers will be sharing in the resources that present people leave to them. This makes it possible, for instance, that every future person in A is much worse off than every future person in B. This could be raised as a severe objection to the idea of constrained efficiency.

There is another sort of objection. So long as we apply only the first lesson from economics, we are free to judge what resources are beneficial by any criterion we deem appropriate. But suppose we go further and apply the third lesson: we draw the conclusion that, when externalities are internalized, the state that results is constrained efficient. Then we are committing ourselves to a particular criterion of beneficialness: it

is determined by the preferences of the future people who exist in the efficient state, or more accurately by what we expect those future people's preferences to be. Unless we adopt that criterion, the third lesson is false. This limits the significance of this sort of efficiency. The beneficialness of a bundle of resources is judged, not by the expected preferences of the people who use these resources, but by the expected preferences of those who exist in the efficient state. Even though a state is constrained efficient, there might be another state that would be preferred by every future person who would live in that other state.

7. CONCLUSION

At the beginning of this paper I said that we can respond to climate change in a way that does not require a sacrifice from anyone. But that claim depended on the standard theory of externalities, which ignores the non-identity effect. How far is it true, now the non-identity effect has forced us to retreat to a resourcist notion of dominance and efficiency? The claim arose from only the first lesson of economics, so to answer the question, we have to attend to the interpretation of only the first lesson under the resourcist notion. Remember that this lesson is true however we judge the beneficialness of resources.

We learn from it that, when we are suffering from an externality such as climate change, a response can move us to an alternative state that constrained dominates the one we start in. In the alternative state:

(a) At least one present person is in a state she prefers to the initial one.
(b) No present person is in a state she disprefers to the initial one.
(c) Future generations inherit from us resources that are at least as beneficial, in a particular sense, as those they would have inherited in the initial state.

(b) confirms the point that no sacrifice is required of present people to respond to climate change; we can appeal to their self-interest instead.

However, in making (a), (b) and (c) true, it is not guaranteed that we can at the same time ensure the truth of any of these propositions:

No future person is in a state she disprefers to the initial one.
No future person is worse off than she is in the initial state.
Future people, taken as an aggregate in some way, will be no worse off than future people would have been had we not responded.

Indeed, it is possible that every future person is worse off than in the initial state.

At least, none of these things can be guaranteed by the bare-bones theory of efficiency I have been dealing with. It may be possible to arrive at stronger conclusions about efficiency by adding more concrete, empirical

assumptions. For example, we could add assumptions about the nature of people's preferences – that they are adaptive, say. We could also add empirical assumptions about the way population growth is affected by policies. So if you find the conclusions I have arrived at too weak to be useful, it might be possible to find stronger ones within the general approach that aims at efficiency, by making stronger assumptions.

Alternatively, you could abandon this aim, and instead aim at promoting overall goodness. In Section 1 I mentioned two problems with this approach. First, it is politically unattainable in the case of climate change. It asks present people and present governments to make sacrifices for the moral reason of promoting overall goodness. This they are unwilling to do. An appeal to self-interest is more likely to succeed.

Second, this approach is theoretically less robust than an approach aimed at efficiency. It demands a theory of value. To promote overall goodness we need an account of what goodness consists in. For one thing, we need an account of interpersonal comparisons of value. Moreover, in the intergenerational context in particular, we need a 'population axiology' – a theory about the value of increasing or decreasing the world's population. This is very hard to come by. After some decades of work in philosophy, we still lack even a basic consensus on population axiology.[8]

APPENDIX

The appendix appears as supplementary material for this article, at https://doi.org/10.1017/S0266267118000020

REFERENCES

Arrhenius, G. Forthcoming. *Population Ethics: the Challenge of Future Generations*. Oxford: Oxford University Press.
Broome, J. 2016. Do not ask for morality. In *Ethical Underpinnings of Climate Economics*, ed. A. Walsh *et al.*, 9–21. London: Routledge.
Broome, J. and D. Foley. 2016. A world climate bank. In *Institutions for Future Generations*, ed. I. González-Ricoy and A. Gosseries, 156–169. Oxford: Oxford University Press.

---

[8] For a survey of the difficulties, see Arrhenius (Forthcoming).

Foley, D. 2009. The economic fundamentals of global warming, in *Twenty-First Century Macroeconomics: Responding to the Climate Challenge*, ed. J. M. Harris and N. R. Goodwin, 115–126. London: Edward Elgar.

Golosov, M., L. E. Jones and M. Tertilt. 2007. Efficiency with endogenous population growth. *Econometrica* 75: 1039–1071.

Nordhaus, W. 2008. *A Question of Balance: Weighing Options on Global Warming Policies*. New Haven, CT: Yale University Press.

Parfit, D.1984. *Reasons and Persons*. Oxford: Oxford University Press.

Rezai, A. 2010. Recast the DICE and its policy recommendations. *Macroeconomic Dynamics*, 14 Supplement 2: 275–289.

Scitovszky, T. de. 1941. A note on welfare propositions in economics, *Review of Economic Studies* 9: 77–88.

Stern, N., *et al.* 2007. *The Economics of Climate Change: The Stern Review*. Cambridge: Cambridge University Press.

Suppes, P. 1966. Some formal models of grading principles. *Synthese* 16: 284–306.

BIOGRAPHICAL INFORMATION

John Broome is Emeritus White's Professor of Moral Philosophy and Emeritus Fellow of Corpus Christi College at the University of Oxford. He is also an Honorary Professor of the Australian National University. His most recent books are *Climate Matters: Ethics in a Warming World* (Norton, 2012) and *Rationality Through Reasoning* (Wiley-Blackwell, 2013). He was a Lead Author of the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Appendix to 'Efficiency and future generations' by John Broome

The arguments in this paper are based on a standard finite general equilibrium model containing public goods (e. g. Foley 1970).

The model

There are T public goods and N private goods. A typical vector of goods is $(e, g) = (e_1, \ldots e_T, g_1, \ldots g_N)$, where the public goods are e and the private goods g. Aggregate production is represented by a vector y of goods in which positive components represent outputs and negative components inputs. The production set Y is the set of aggregate productions that are technically possible. There are I people. A typical consumption vector of person i is $x^i = (e, g^i) = (e_1, \ldots e_T, g_1^i, \ldots g_N^i)$. Each person has a consumption set X^i of consumptions that are possible for her, and a preference relation $\succeq^i$ on X^i. She has an initial endowment of private goods $w^i = (w_1^i, \ldots w_N^i)$.

An *allocation* is a vector $x = (x^1, \ldots x^I) = (e, g^1, \ldots e, g^I)$, where each person has the same consumption of public goods e. An allocation is *feasible* if and only if $(x^i) \in X^i$ for all i and $(e, \sum_i (g^i - w^i)) \in Y$. (In the text of the article, I use 'possible' for feasible.) Let Z be the set of feasible allocations.

For each i, the indifference relation $\sim^i$ and strict preference relation $\succ^i$ are defined in the standard way.

An allocation $x = (x^1, \ldots x^I)$ is *Pareto indifferent* to an allocation $\hat{x} = (\hat{x}^1, \ldots \hat{x}^I)$ if and only if $x^i \sim^i \hat{x}^i$ for all I. x *Pareto dominates* $\hat{x}$ if and only if $x^i \succeq^i \hat{x}^i$ for all i and x is not Pareto indifferent to $\hat{x}$. A feasible allocation is *Pareto efficient* if and only if no feasible allocation Pareto dominates it. Otherwise it is *Pareto inefficient*.

Basic assumptions:

Y is closed and each X^i is closed.

Each $\succeq^i$ is complete, transitive and reflexive.

Each $\succeq^i$ is continuous. That is: for all $x^i \in X^i$ the sets $\{\hat{x}^i | \hat{x}^i \succeq^i x^i\}$ and $\{\hat{x}^i | x^i \succeq^i \hat{x}^i\}$ are closed.

Each $\succeq^i$ is locally nonsatiated. That is: for all $x^i \in X^i$, every neighbourhood of x^i contains an $\hat{x}^i$ such that $\hat{x}^i \succ^i x^i$.

The set Z of feasible allocations is bounded.

Interpretation. The model has its standard intertemporal interpretation. Each good is dated. In the application to climate change, e_t is minus the emission of greenhouse gas at the tth date. In the first part of this appendix, I assume there is a fixed population of people. Each person has fixed dates of birth and death, but these dates differ between people. Unless people have altruistic preferences about goods consumed after they die, the private goods in each person's consumption set are all dated to dates when she is alive.

A general equilibrium model like this models the economy by means of a single intertemporal equilibrium, which ascribes prices to all goods at all times. This is not realistic. There are in practice few markets at the present time for future goods. Financial markets dealing in debt are a sort of proxy, but only a rough one. The most serious problem is that the people who are supposed to trade in these markets do not all exist at the same time. Future prices appear in practice in the form of present people's expectations of future prices, which depend on their expectations of future people's preferences. These expectations may be mistaken.

First lesson

No theorem 1. No general theorem supports the first lesson. Several theorems state sufficient

conditions for a state to be efficient, but none that I know states sufficient conditions for a state to be inefficient. Even if it contains externalities, a state might be Pareto efficient because of special features of the economy. For example, a negative externality might come in discrete lumps. The marginal lump emitted may be a genuine externality, in that removing it would have a positive external benefit. But the private cost of removing this lump may exceed the external benefit of doing so. If so, the state might be Pareto efficient despite the externality. For this reason, I expressed the first lesson with the qualification 'except in some unusual circumstances'.

However, I can informally describe one necessary condition for a state with externalities to be inefficient. If a state containing an externality is Pareto inefficient, that means a Pareto improvement is possible. To make the Pareto improvement will normally require some agent to change its production or consumption in a way that is inherently costly, and to be compensated for the cost by a transfer of goods from other agents. So compensating transfers must be possible if Pareto inefficiency is to result from an externality.

In the case of climate change, present people can change their consumption and production methods for the sake of future people, but (unless they have altruistic preferences for future goods) this will be a Pareto improvement only if they can be compensated for doing so by a transfer from future people. They will have to receive present goods, which they can consume and benefit from; a transfer of future goods in the form of debts would not be good enough unless they can be traded for present goods. Future people can compensate present people with present goods, through the mechanism of overlapping lives described in section 2.1.

An intergenerational externality such as climate change creates Pareto inefficiency between generations only in so far as intergenerational transfers are possible through overlapping lives. If there could be no such transfers, intergenerational externalities would not lead to Pareto inefficiency. This constraint is implicit in the structure of the model.

Second lesson

Theorem 2. Given the basic assumptions, if a feasible allocation is Pareto inefficient, there is a feasible Pareto efficient allocation that Pareto dominates it.

Proof of theorem 2. The set Z of feasible allocations is closed because Y is closed and each X^i is closed. It is bounded by assumption.

Take a feasible allocation $\hat{x} = (\hat{x}^1, \ldots \hat{x}^I)$ that is Pareto inefficient. Take the set of allocations $\{(x^1, \ldots x^I) | (\forall i)(x^i \succeq^i \hat{x}^i)\}$ that Pareto dominate or are Pareto indifferent to $\hat{x}$. This is a closed set because the individual sets $\{(x^i) | x^i \succeq^i \hat{x}^i\}$ are closed. Let Z' be the intersection of Z with this set. Z' is closed because it is the intersection of closed sets. It is bounded because Z is bounded.

By Debreu's (1954) representation theorem, each person i's preferences can be represented by a continuous real-valued utility function u^i. Each allocation x in Z' has a total of utilities $U(x) = \sum_i u^i(x^i)$. U is a continuous function on the closed and bounded set Z'. It therefore attains its maximum on this set.

Take an allocation x' in Z' that has maximum total of utility. x' is feasible because it is in Z. Any feasible allocation that Pareto dominates x' Pareto dominates $\hat{x}$ and is therefore in Z'. But no allocation in Z' Pareto dominates x' because any Pareto dominating allocation would have a greater total of utility. x' is therefore Pareto efficient. x' is consequently not Pareto indifferent to $\hat{x}$ because if either of a pair of Pareto indifferent allocations is efficient, so is the other, and $\hat{x}$ is not efficient. Since x' is in the set of allocations that either Pareto dominate or are Pareto indifferent to $\hat{x}$, x' Pareto dominates $\hat{x}$.

In sum, x' is a feasible Pareto efficient allocation that Pareto dominates $\hat{x}$, as the theorem requires. END OF PROOF

Third lesson

Theorem 3. Suppose a feasible allocation $x = (x^1, \dots x^I) = (e, g^1, \dots e, g^I)$ is such that there is a price vector $r = (r_1, \dots r_N)$ for private goods, and for each i there is a price vector $q^i = (q_1^i, \dots q_T^i)$ for public goods with the following properties:

$p.(e, \sum_i (g^i - w^i)) \geq p.y$ for any vector $y \in Y$, where $p = (\sum_i q^i, r)$. (So the production sector is profit-maximizing at p.)

For all i, $x^i \succeq^i \hat{x}^i$ for any $\hat{x}^i$ such that $p^i.\hat{x}^i \leq p^i.x^i$, where $p^i = (q^i, r)$. (So x^i is preference maximizing at p^i, given a budget constraint.)

Then, given the basic assumptions, x is Pareto efficient.

Interpretation. This is the so-called first theorem of welfare economics applied to an economy with public goods. Note that in theorem 3 the price of each private good is the same for each person and for producers, but each public good has different prices for different people. The price to producers of a public good is the sum of its prices for each person.

In this theorem people are not to be interpreted as actually buying public goods. Their preference-maximizing over public goods is purely notional. The preference-maximizing condition specifies that each person would be willing to buy just the prevailing amount of public goods were she required to pay for them at her own prices $(q_1^i, \dots q_T^i)$. These prices indicate her willingness to pay for public goods at the margin. For greenhouse gas, these prices measure the marginal external costs imposed on the person by emissions.

On the other hand, the production system in the theorem, which produces both public and private goods, is supposed to be genuinely profit-maximizing at the given prices. Producers are supposed genuinely to pay the prices of their inputs. For public goods, these prices are the sum of the private willingnesses to pay. In practice they would need to be imposed by the government.

The conditions of the theorem express the condition that producers' external costs are internalized. It is Paul Samuelson's (1954) condition for Pareto efficiency with public goods.

Proof of theorem 3. Suppose the allocation x is not efficient. Then there is a feasible allocation $\hat{x} = (\hat{x}^1, \dots \hat{x}^I) = (\hat{e}, \hat{g}^1, \dots \hat{e}, \hat{g}^I)$ that Pareto dominates x.

Because $\hat{x}$ Pareto dominates x, $\hat{x}^i \succ^i x^i$ for some i. Since x^i is preference maximizing at p^i, it follows that $p^i.\hat{x}^i > p^i.x^i$ for this i.

Because $\hat{x}$ Pareto dominates x, $\hat{x}^i \succeq^i x^i$ for all i. Suppose $p^i.\hat{x}^i < p^i.x^i$. Then there is a neighbourhood of $\hat{x}^i$ such that for every point $\vec{x}^i$ in this neighbourhood $p^i.\vec{x}^i < p^i.x^i$. By local nonsatiation, one of these points is strictly preferred to $\hat{x}^i$, which means it is strictly preferred to x^i. This contradicts that x^i is preference maximizing at p^i. So $p^i.\hat{x}^i \geq p^i.x^i$ for all i.

Because $p^i.\hat{x}^i \geq p^i.x^i$ for all i and $p^i.\hat{x}^i > p^i.x^i$ for some i, $\sum_i p^i.\hat{x}^i > \sum_i p^i.x^i$. That is, $\sum_i (q^i.\hat{e} + r.\hat{g}^i) > \sum_i (q^i.e + r.g^i)$. That is, $\sum_i q^i.\hat{e} + r.\sum_i \hat{g}^i > \sum_i q^i.e + r.\sum_i g^i$.

However, because x is profit-maximizing at p, $p.(e, \sum_i (g^i - w^i)) \geq p.(\hat{e}, \sum_i (\hat{g}^i - w^i))$, so $p.(e, \sum_i g^i) \geq p.(\hat{e}, \sum_i \hat{g}^i)$. That is $\sum_i q^i.e + r.\sum_i g^i \geq \sum_i q^i.\hat{e} + r.\sum_i \hat{g}^i$.

The supposition that x is not efficient therefore implies a contradiction. So x is efficient. END OF PROOF

Fourth lesson

Theorem 4. Suppose a feasible allocation $x = (x^1, \dots x^I) = (e, g^1, \dots e, g^I)$ is Pareto efficient. Assume also that the production set Y is convex, and that for each i the set $\{\hat{x}^i | \hat{x}^i \succeq^i x^i\}$ is convex. Then, under the basic assumptions, there are initial endowments of private goods $w =$

$(w^1, \ldots w^J)$, a price vector $r = (r_1, \ldots r_N)$ for private goods, and for each i a price vector $q^i = (q_1{}^i, \ldots q_T{}^i)$ for public goods with the following properties:

$p.(e, \sum_i (g^i{-}w^i)) \geq p.y$ for any vector $y \in Y$, where $p = (\sum_i q^i, r)$. (So the production sector is profit-maximizing at p.)

For all i, $x^i \succeq^i \hat{x}^i$ for any $\hat{x}^i$ such that $p^i.\hat{x}^i \leq p^i.x^i$, where $p^i = (q^i, r)$. (So x^i is preference-maximizing at p^i given a budget constraint.)

Proof of theorem 4. A proof appears in Foley (1979: 68–9). Some slight modifications are required since Foley uses slightly different assumptions.

Interpretation. This is a version of the second theorem of welfare economics for an economy with public goods. It is the converse of the first theorem. It shows that, provided the production set and upper-contour sets for preferences are convex, a Pareto efficient allocation can be supported by a price system in which the price of a public good is the sum of people's willingnesses to pay for it. If the Pareto efficient allocation is to be a market equilibrium the price of public goods will have to be established by government taxes, and initial endowments $(w^1, \ldots w^J)$ will need to be distributed appropriately by lump-sum taxes and subsidies.

Constrained efficiency

I now drop the assumption that the population of people is fixed. As before, each possible state contains a number of people living for various periods that overlap. There is one group of people, the present people, who are the same in every state, but there are also future people who exist in some states and not in others. Both the identities and numbers of future people may vary between states.

There is a set (which need not be finite) of *possible people*. For convenience, I shall assume that each possible person has fixed preferences. (If we wanted to allow for varying preferences, we could treat a possible person as a possible person–preference pair.) Each possible person i has a consumption set X^i, a preference relation $\succeq^i$ on X^i, and an initial endowment w^i. There is an aggregate production set Y.

The *present population* is a finite nonempty subset Π of the set of possible people. It has J members indexed by $i = 1, \ldots J$. A *population* is a union $\Pi \cup \Phi$ of Π with some other, disjoint, finite subset Φ of possible people, who are the *future population*. $\Pi \cup \Phi$ has I members indexed by $I = 1, \ldots I$.

Take a particular population $\Pi \cup \Phi$. An *allocation* for this population is a vector $x = (x^1, \ldots x^I) = (e, g^1, \ldots e, g^I)$ so that each person's consumption of public goods is the same. An allocation is *quasi-feasible* if and only if $x^i \in X^i$ for all $i \in \Pi \cup \Phi$, and $(e, \sum_{i \in \Pi \cup \Phi} (g^i{-}w^i)) \in Y$. Owing to the nonidentity effect, not all quasi-feasible allocations are causally compatible with the population's being $\Pi \cup \Phi$. An allocation is *feasible* if and only if it is quasi-feasible and causally compatible with its population.

A quasi-feasible allocation for $\Pi \cup \Phi$ is *Pareto quasi-efficient* if and only if no quasi-efficient allocation for $\Pi \cup \Phi$ Pareto dominates it.

A *present allocation* for the present population Π is a vector $x^\Pi = (x^1, \ldots x^J) = (e, g^1, \ldots e, g^J)$. A present allocation is *feasible* if and only if $x^i \in X^i$ for all $i \in \Pi$, and $(e, \sum_{i \in \Pi} (g^i{-}w^i)) \in Y$. The *future production* of a feasible present allocation x^Π together with a production vector $(e, g) \in Y$ is the vector $(e, g - \sum_{i \in \Pi} (g^i - w^i))$. The *constraint set* of a feasible present allocation x^Π together with a production vector (e, g) is a set of $T+N$ dimensional vectors y such that $y \geq (e, g - \sum_{i \in \Pi} (g^i - w^i))$. (Notation for vectors: $y \geq z$ means $y_s \geq z_s$ for each component s.) A quasi-feasible allocation x for a population $\Pi \cup \Phi$ includes a feasible present allocation for its present population Π consisting of its first J components, and it implies a

production vector $(e, \sum_{i \in \Pi \cup \Phi} (g^i - w^i))$. So future production and a constraint set can be ascribed to any quasi-feasible allocation x. I use the notation $f(x)$ for the future production of an allocation x.

The constraint set of an allocation x is *crude* if and only if it is the set $\{y \mid y \geq f(x)\}$.

Given a population $\Pi \cup \Phi$, the constraint set of an allocation $x = (x^1, \ldots x^J) = (e, g^1, \ldots e, g^J)$ is *preference-based* if and only if it is the set $\{(\hat{e}, \bar{g}) \mid (\exists (\hat{g}^{J+1}, \ldots \hat{g}^I))(\bar{g} = \sum_{i \in \Phi} (\hat{g}^i - w^i) \&$ $(\forall i \in \Phi)((\hat{e}, \hat{g}^i) \in X^i \& (\hat{e}, \hat{g}^i) \succeq^i x^i))\}$. This is the Scitovszky set of future productions that are sufficient to give every future person in Φ a consumption that she does not disprefer to her consumption in x.

A feasible allocation $x = (x^1, \ldots x^J)$ for $\Pi \cup \Phi$ is *constrained efficient* if and only if there is no feasible present allocation $\hat{x}^\Pi = (\hat{x}^1, \ldots \hat{x}^J)$ and production $(\hat{e}, \hat{g}) \in Y$ such that, for all $i \in \Pi$, $\hat{x}^i \succeq^i x^i$, for some $i \in \Pi$, $\hat{x}^i \succ^i x^i$ and future production is a member of the constraint set of x.

We might adopt:

> *Existence assumption*. For any given feasible present allocation $x^\Pi = (x^1, \ldots x^J)$ and production $(e, g) \in Y$, there is a feasible allocation x that has this present allocation and production.

Owing to the nonidentity effect, this is not implied by the rest of the model, but it seems plausible. Under the existence assumption, efficiency can be defined more neatly as follows:

> An allocation $\hat{x}$ *constrained dominates* an allocation x if and only if for all $i \in \Pi$, $\hat{x}^i \succeq^i x^i$, for some $i \in \Pi$, $\hat{x}^i \succ^i x^i$, and future production $f(\hat{x})$ is a member of the constraint set of x. A feasible allocation is *constrained efficient* if and only if there is no feasible allocation that constrained dominates it.

I use the existence assumption only in generalizing theorem 2 below.

For two of the theorems below, I need to assume that the production set allows present desirable consumption to be substituted for future desirable consumption as an output. To save space, I shall do this in an ad hoc manner, as follows:

> *Substitution assumption*. If $x = (x^1, \ldots x^J)$ and $\bar{x}$ are two quasi-feasible allocations for the population $\Pi \cup \Phi$ and $\bar{x}$ Pareto dominates x, there is a quasi-feasible allocation $\hat{x} = (\hat{x}^1, \ldots \hat{x}^J)$ that Pareto dominates x and such that $\hat{x}^i \succ^i x^i$ for some $i \in \Pi$.

Constrained efficiency: theorems

Theorem 2′. Let constraint sets be crude. Under the basic assumptions and the existence assumption, if a feasible allocation is constrained inefficient, there is a feasible allocation that constrained dominates it and is constrained efficient.

Proof. Let $\hat{x}$ be a feasible allocation that is constrained inefficient with future production $f(\hat{x})$. Let $\hat{x}^\Pi$ be the present allocation consisting of the first J components of $\hat{x}$. Let Z^Π be the set of feasible present allocations that can have future production at least as great as $f(\hat{x})$. That is, $Z^\Pi = \{x \mid (\forall i \in \Pi)(x^i \in X^i) \& (\exists (e, g) \in Y)((e, g - \sum_{i \in \Pi} (g^i - w^i)) \geq f(\hat{x}))\}$. Z^Π is closed because Y is closed and each X^i is closed. It is bounded because the set of feasible allocations is bounded. Let $Z^{\Pi\prime}$ be the intersection of Z^Π and the set $\{x^\Pi \mid (\forall i \in \Pi)(x^i \succeq^i \hat{x}^i)\}$ of present allocations that are not dispreferred to $\hat{x}^\Pi$ by any present person. $Z^{\Pi\prime}$ is bounded and closed, being the intersection of closed sets.

The proof then proceeds like the proof of theorem 2 by maximizing the total utility of present people on the closed and bounded set $Z^{\Pi\prime}$, while keeping future production constant. END OF PROOF.

I cannot prove a version of theorem 2 for constraint sets other than crude ones. As in the proof above, there is a bounded set on which the total utility of present people might be maximized, but this set may not be closed.

Theorem 5. Under the basic assumptions and the substitution assumption, given a population $\Pi\cup\Phi$, a quasi-feasible allocation is Pareto quasi-efficient if and only if it is constrained efficient with a preference-based constraint set.

Proof. First, take a quasi-feasible allocation $x = (x^1, \ldots x^I)$ that is not constrained efficient with a preference-based constraint set. I prove that it is not Pareto quasi-efficient. There is a feasible present allocation $\hat{x}^\Pi = (\hat{x}^1, \ldots \hat{x}^J)$ and production $(\hat{e}, \hat{g}) \in Y$ such that, (i) for all $i \in \Pi$, $\hat{x}^i \succeq^i x^i$, (ii) for some $i \in \Pi$, $\hat{x}^i \succ^i x^i$ and (iii) future production $(\hat{e}, \bar{g})$ is a member of the preference-based constraint set of x. (iii) implies that $(\hat{e}, \bar{g})$ is sufficient to give every future person in Φ a consumption that she does not disprefer to her consumption in x. That is: there is a vector of private goods $(\hat{g}^{J+1}, \ldots \hat{g}^I)$ such that $\sum_{i \in \Phi} (\hat{g}^i - w^i) = \bar{g} = \hat{g} - \sum_{i \in \Pi} (\hat{g}^i - w^i)$ and, for all $i \in \Phi$, $(\hat{e}, \hat{g}^i) \in X^i$ and $(\hat{e}, \hat{g}^i) \succeq^i x^i$. Take the allocation $\hat{x} = (\hat{x}^1, \ldots \hat{x}^I)$ made up of the present allocation $\hat{x}^\Pi$ joined to the vector $(\hat{x}^{J+1}, \ldots \hat{x}^I) = (\hat{e}, \hat{g}^{J+1}, \ldots \hat{e}, \hat{g}^I)$. Then for all $i \in \Pi\cup\Phi$, $\hat{x}^i \succeq x^i$ because of (i) above and because, for all $i \in \Phi$, $(\hat{e}, \hat{g}^i) \succeq^i x^i$. For some $i \in \Pi\cup\Phi$, $\hat{x}^i \succ x^i$ because of (ii) above. Finally, $\hat{x}$ is quasi-feasible because, for all $i \in \Pi\cup\Phi$, $\hat{x}^i \in X^i$ and because total consumption $(\hat{e}, \sum_{i \in \Pi\cup\Phi} (\hat{g}^i - w^i)) = (\hat{e}, \sum_{i \in \Pi} (\hat{g}^i - w^i) + \sum_{i \in \Phi} (\hat{g}^i - w^i)) = (\hat{e}, \hat{g})$, which is a member of Y. In sum, $\hat{x}$ is quasi-feasible and Pareto dominates x. x is therefore not Pareto quasi-efficient.

Next take a quasi-feasible allocation $x = (x^1, \ldots x^I)$ that is not Pareto quasi-efficient. I prove it is not constrained efficient with a preference-based constraint set. There is a quasi-feasible allocation $\bar{x}$ that Pareto dominates x. By the substitution assumption, there is a quasi-feasible allocation $\hat{x} = (\hat{x}^1, \ldots \hat{x}^I)$ that Pareto dominates x and such that $\hat{x}^i \succ^i x^i$ for some $i \in \Pi$. Also, $\hat{x}^i \succeq^i x^i$ for all $i \in \Pi$. The allocation $\hat{x}$ implies a feasible present allocation $\hat{x}^\Pi = (\hat{x}^1, \ldots \hat{x}^J)$ and a future production $(\hat{e}, \sum_{i \in \Phi} (\hat{g}^i - w^i))$. We already know that, for some $i \in \Pi$, $\hat{x}^i \succ^i x^i$ and that, for all $i \in \Pi$, $\hat{x}^i \succeq^i x^i$ because $\hat{x}$ Pareto dominates x. The vector $(\hat{g}^{J+1}, \ldots \hat{g}^I)$ satisfies the conditions that ensure future production $(\hat{e}, \sum_{i \in \Phi} (\hat{g}^i - w^i))$ is a member of the preference-based constraint set of x, because for all $i \in \Phi$, $(\hat{e}, \hat{g}^i) \in X^i$ & $(\hat{e}, \hat{g}^i) \succeq^i x^i$. In sum, $\hat{x}^\Pi$ and production $(\hat{e}, \hat{g}) \in Y$ are such that, for all $i \in \Pi$, $\hat{x}^i \succeq^i x^i$, for some $i \in \Pi$, $\hat{x}^i \succ^i x^i$ and future production is a member of the constraint set of x. It follows that x is not constrained efficient with a preference-based constraint set. END OF PROOF.

Theorems 3 and 4 are true if 'Pareto quasi-efficient' is substituted for 'Pareto efficient' in the statements of the theorems. The proofs are exactly the same except that in the proofs (but not the statements of the theorems) 'quasi-feasible' must be substituted for 'feasible'.

Therefore, in view of theorem 5, theorems 3 and 4 are true if 'constrained efficient' is substituted for 'Pareto efficient', so long as the constraint set is preference-based. This specifies how far the third and fourth lessons extend to the case of variable population.

References
Debreu, G. 1954. Representation of a preference ordering by a numerical function. In *Decision Processes*, ed. R. M. Thrall, C. H. Coombs and R. L. Davis, 159–65. Wiley.
Foley, D. 1970. 'Lindahl's solution and the core of an economy with public goods', *Econometrica*, 38, pp. 66–72.
Samuelson, P. 1954. The pure theory of public expenditure. *Review of Economics and Statistics* 36: 387–9.