



Published in final edited form as:

Nat Clim Chang. 2021 October ; 11(10): 827–833. doi:10.1038/s41558-021-01130-6.

Utilitarian benchmarks for emissions and pledges promote equity, climate and development

Mark B. Budolfson^{1,2,✉}, David Anthoff³, Francis Dennig^{2,4}, Frank Errickson^{2,5}, Kevin Kuruc⁶, Dean Spears^{2,7,8,9,✉}, Navroz K. Dubash¹⁰

¹Department of Environmental and Occupational Health and Justice, Rutgers School of Public Health, Center for Population-Level Bioethics, and Department of Philosophy, Rutgers University, New Brunswick, NJ, USA.

²Climate Futures Initiative, School of Public and International Affairs, Princeton University, Princeton, NJ, USA.

³Energy and Resources Group, University of California at Berkeley, Berkeley, CA, USA.

⁴Social Sciences (Economics), Yale-NUS College, Singapore, Singapore.

⁵Center for Policy Research on Energy and the Environment, School of Public and International Affairs, Princeton University, Princeton, NJ, USA.

⁶Department of Economics, University of Oklahoma, Norman, OK, USA.

⁷Department of Economics, University of Texas at Austin, Austin, TX, USA.

⁸Economics and Planning Unit, Indian Statistical Institute, New Delhi, India.

⁹IZA Institute of Labor Economics, Bonn, Germany.

¹⁰Centre for Policy Research, New Delhi, India.

Reprints and permissions information is available at www.nature.com/reprints.

✉Correspondence and requests for materials should be addressed to Mark B. Budolfson or Dean Spears. mark.budolfson@rutgers.edu; dspears@utexas.edu.

Author contributions

All authors contributed equally to this work. M.B.B., D.A., F.D., N.K.D., F.E. and D.S. designed the research. F.D., F.E. and K.K. led the computer modelling. M.B.B., F.E., K.K. and D.S. designed the figures and F.E., K.K. and D.S. constructed the figures. M.B.B., N.K.D. and D.S. wrote the main paper and M.B.B. wrote the Supplementary Information. All authors discussed the results and implications and commented on the manuscript.

Code availability

All model code used to generate results and create figures for this article is archived⁵⁴ and freely available at <https://github.com/Environment-Research/Utilitarianism>.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41558-021-01130-6>.

Competing interests

N.K.D. is a member of the committee advising the Government of India on analysis of low carbon trajectories. All other authors declare no competing interests.

Extended data is available for this paper at <https://doi.org/10.1038/s41558-021-01130-6>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41558-021-01130-6>.

Abstract

Tools are needed to benchmark carbon emissions and pledges against criteria of equity and fairness. However, standard economic approaches, which use a transparent optimization framework, ignore equity. Models that do include equity benchmarks exist, but often use opaque methodologies. Here we propose a utilitarian benchmark computed in a transparent optimization framework, which could usefully inform the equity benchmark debate. Implementing the utilitarian benchmark, which we see as ethically minimal and conceptually parsimonious, in two leading climate–economy models allows for calculation of the optimal allocation of future emissions. We compare this optimum with historical emissions and initial nationally determined contributions. Compared with cost minimization, utilitarian optimization features better outcomes for human development, equity and the climate. Peak temperature is lower under utilitarianism because it reduces the human development cost of global mitigation. Utilitarianism therefore is a promising inclusion to a set of benchmarks for future explorations of climate equity.

What tools should be used to benchmark and evaluate actual and pledged carbon emissions against criteria of equity and fairness^{1–4}? In the 2015 Paris Agreement, each country agreed to each prepare and communicate nationally determined contributions (NDCs) to be achieved through domestic mitigation efforts, and to subsequently update these pledges over a five-year cycle⁵. While there is growing consensus that global emissions reductions and countries' initial pledges are inadequate⁶, there is far less consensus about which countries should enhance their mitigation efforts and by how much. In large part, this is because there is considerable disagreement about the benchmarks by which each country's pledge should be judged to be consistent with fairness and equity. Indeed, the Intergovernmental Panel on Climate Change (IPCC) summarized the breadth of equity frameworks for allocating emissions, illustrating the disparities in approach¹. In response, a series of recent analyses apply these frameworks to the task of benchmarking national emissions and pledges. The result is a wide range of possible equity benchmarks, which lead to a diverse set of conclusions: that particular national emissions (actual or pledged) are equitable according to some benchmarks and inequitable according to others^{7–13} (<https://climateactiontracker.org/>). Over two decades of negotiations, national governments failed to agree on a top-down allocation of emissions in part because of disagreements over equity, so it is no surprise that such disagreements persist in alternative equity benchmarks. One response to disagreement over equity has been to produce new benchmarks by averaging widely disparate results across conceptually distinct alternative criteria, at the cost of transparency and clarity¹⁴. In such cases, readers can reasonably wonder whether the results reflect a coherent, defensible approach to equity, or merely a mechanical attempt at compromise¹⁵, because the average is shaped by the portfolio of equity approaches chosen, and especially by the extremes.

Our approach is different. We add to equity debates a single benchmark derived from the simple ethical theory of utilitarianism^{16,17}. Although utilitarianism has been criticized by both ethicists and economists^{10,15,18–22}, we suggest that using utilitarianism yields insights in equity benchmark debates because it is ethically minimal and conceptually parsimonious. Utilitarianism is ethically minimal because it requires agreeing only that each person's interests count equally, that policy should promote wellbeing and that a unit of foregone consumption harms the poor more than the rich. In the context of climate debates,

this is minimal, in part, because it does not seek to account for the past and perhaps other factors relevant to justice, which, for many, is an important component of climate equity^{10,15,19,20}. It is also conceptually parsimonious because, unlike other benchmarks, it is well understood in the ethics literature and therefore a transparent benchmark and does not require interpretation or the construction of composite and therefore difficult-to-understand indices^{23,24}.

In addition, our purpose here is to contrast utilitarianism with the standard approach based on monetary cost: as Adler²⁵ summarizes, “cost-benefit analysis is now the dominant policy-analysis methodology in governmental practice.” In moving beyond a focus on monetary cost, we demonstrate the simplest way in which a climate–economy model can use a social welfare approach that promotes wellbeing while weighing each person’s interests equally. Future research could apply our methods to further, more complex benchmarks for social welfare, some of which may be even more ethically compelling for some purposes²⁵. As we show below, even the ethically minimal benchmark of utilitarianism yields very different results to cost-minimization approaches, illustrating the value of explicit attention to equity benchmarks. We implement a utilitarian benchmark by making a simple, transparent modification to an existing, high-profile climate–economy model. We use the model to calculate the distribution of mitigation costs and climate damages between nations that maximizes overall global human development (which we summarize as ‘wellbeing’), weighing the interests of all persons equally, but taking into account that a dollar sacrificed by the poor subtracts more wellbeing than a dollar sacrificed by the rich. In particular, the model calculates an allocation to different regions of future emissions that would be optimal according to a utilitarian objective. Where we refer to the ‘utilitarian-optimal’ benchmarks or allocation below, we mean the time path of emissions for each world region that maximizes this utilitarian objective. Because this objective could be maximized by any optimizing climate–economy model, our approach is independent of the particular models in which we choose to implement it, as we demonstrate by implementing our approach in two different models from the literature^{26,27} and finding qualitatively complementary (although, of course, quantitatively distinct) results. In short, we offer such a utilitarian-optimal allocation of emission shares as an alternative tool to benchmark national emissions and pledges that offers promising advantages for future equity discussions and research.

Our method produces benchmark paths for emissions shares over time. In our main result, we compare these paths with actual emissions in 2019 (ref.²⁸). We also compare the benchmark paths with pledges, here operationalized as initial NDCs²⁹. As the Paris Agreement states, each country’s plan should take into account its “common but differentiated responsibilities and respective capabilities, in light of different national circumstances,” so that outcomes are equitable³⁰. Our approach and results offer an informative alternative methodology that can join the ongoing debate on assessing the equity of emissions shares and pledges.

Utilitarianism among other climate equity benchmarks

Our use of utilitarianism as a benchmark identifies a transparent, simple to model, alternative specification of climate equity. We recognize that utilitarianism will strike many readers as inadequate as a full analysis of equity, and perhaps rightly so. Utilitarianism, like any forward-looking approach, does not explicitly internalize responsibility for past emissions as a basis for future actions. Yet, it does implicitly take historical responsibility partly into account to the extent that the past shapes current capacities. This happens because historical responsibility for emissions is correlated with present-day capacity to mitigate (Extended Data Fig. 1)³¹. Similarly, our approach also does not entail any ‘latent grandfathering’ of historically high emissions: because utilitarianism is forward looking, past emissions per se do not restrict the model’s search for the allocation that will maximize future wellbeing. We do not defend the utilitarian benchmark as uniquely correct. Rather, we suggest the application of utilitarianism yields insight precisely because the key normative assumptions of utilitarianism applied to climate equity are difficult to reject, and therefore might reasonably be considered ethically minimal. These assumptions are that the consequences of climate policy for human development matter; that each person should be weighted equally in a normative evaluation of equity; and that more human development is lost when a dollar of consumption is foregone by a poorer person than a richer person.

All candidate benchmarks for equity have limitations, including utilitarianism. Despite the limitation detailed above (namely, a forward-looking perspective) utilitarianism can be an informative, transparent focal point for discussion among parties who disagree about which substantive principle of equity should guide climate policy^{18–20,23,32–34}. As we detail in Supplementary Table 1 and Extended Data Fig. 1, the utilitarian benchmark captures a core value of each of the IPCC equity categories¹, while avoiding the implications that result from other benchmarking principles that rely on one category alone as a full analysis of equity. In addition, by using optimization tools common in climate economics, the calculations behind the utilitarian benchmark are easy to summarize for policy audiences and can be straightforwardly performed in any multi-region climate optimization model with very minor modifications.

The utilitarian benchmark contrasts with past solutions of integrated assessment models (IAMs) that recommend a uniform carbon price. Two typical IAM experiments that lead to uniform carbon price policies are cost-effectiveness analysis and standard cost–benefit analyses that do not consider equity. Cost-effective solutions are found by minimizing total global abatement cost subject to a cumulative emissions constraint^{35–37}. A standard cost–benefit analysis calculates the optimal emissions trajectory that maximizes net benefits (benefits of mitigation minus the cost of reducing emissions) while ignoring equity. Such cost–benefit solutions are also cost-effective, that is, they achieve their emission reduction goal in a cost-minimizing way^{26,27,38–41}. Both IAM solutions lead to a uniform carbon price, but neither considers equity. To compute our cost-minimization benchmark transparently, we use a third approach: we optimize a utilitarian social welfare function, while imposing a constraint at each timestep that the carbon price must be equal across regions. This constraint implies that at each point in time, the emission reductions achieved are realized at minimum cost. Note that this approach is not dynamically cost minimizing,

because intertemporal allocation decisions are not driven by an interest rate. We use this as our base case benchmark because it amounts to the minimal change from the utilitarian benchmark that leads to a uniform carbon price⁴².

Using cost minimization as a benchmark for equity ignores global inequality. Cost-minimization solutions equate the marginal dollar cost of abatement across nations, thereby ignoring the fact that a dollar of foregone consumption due to mitigation cost sacrificed by a poor person subtracts more wellbeing than a dollar sacrificed by the rich. The utilitarian objective we use instead equates the marginal wellbeing costs of abatement by assigning different carbon prices to different nations. Therefore, we propose a utilitarian benchmark with different regional prices as an alternative that succeeds in including minimal standards for equity and common but differentiated responsibilities, instead of setting aside equity via uniform carbon price modelling.

Theoretical economics supports our utilitarian approach, in which carbon prices differ across regions⁴³. Chichilnisky and Heal¹⁶ prove that, in light of global inequality, there are many policies that are Pareto optimal and have different regional carbon prices. Sandmo⁴⁴ shows that in a standard optimal taxation framework one should differentiate taxes on environmental externalities if there are economic inequalities. Recent work has implemented this insight in theoretical⁴⁵ and computational models^{17,46}. Our method is novel in two ways. First, unlike other approaches to equity that are meaningful only in specific models, our method can be simply and transparently implemented in any leading multi-region optimization model. We demonstrate this by applying it to two standard models, the Regional Integrated Climate-Economy (RICE) model and Climate Framework for Uncertainty, Negotiation, and Distribution (FUND) model. Second, our method does not consider the possibility of large redistributive transfers that would be motivated by a more general concern for equity^{7,47,48}; instead, it reflects a simple, transparent standard for equity in the limited context of national emissions. In principle, a cost-minimization approach could be part of a policy package that would be superior, according to utilitarianism, to even our utilitarian-optimal benchmark; however, because this would involve very large international redistributive transfers, our model ignores this theoretical possibility as a matter of realism, as transfers of anything near this magnitude are not under discussion as part of any climate-policy package.

We implement both utilitarianism and the cost-minimization optimum in the widely used RICE model developed by William Nordhaus. Yet, the principles behind our normative arguments apply against all uniform carbon price benchmarks versus our utilitarian benchmark. Although limitations of the RICE model have been documented in the literature, RICE is widely studied and therefore suitable for such an illustration.

Our use of the RICE model assumes the same relationship between wellbeing and dollars of consumption that is assumed by Nordhaus: the parameter that controls the marginal utility of consumption is set to 1.5 (that is, a dollar forgone by a rich person is equivalent to 2.8 dollars lost by a person who is half as rich; see equation (1) in the Supplementary Information). We use a lower rate of discounting the future (0.8% annually), which we choose to deliver a peak temperature of 2 °C above pre-industrial levels. The 2 °C

assumption does not drive our results, nor does our choice to use the RICE model. Sensitivity studies in the Supplementary Information show that alternative discount rates or use of an alternative IAM (FUND) may have an effect on the timing of reductions in the utilitarian outcome, but neither substantively changes the relative reductions across regions which are our focus (Supplementary Tables 2–5, Extended Data Fig. 2 and Supplementary Fig. 1). For example, one of our sensitivity checks is to set the marginal utility of consumption equal to 1 and time preference equal to 1.5% (a common alternative in the literature); we show that this results in emissions shares comparable with our main results (Supplementary Table 3).

Utilitarian benchmark for national emissions and pledges

We compare a scenario that optimizes emission reductions using a single global carbon price (which implies cost minimization) with our utilitarian approach that optimizes by allowing carbon prices (and hence sacrifices) to vary between regions in the way that is utilitarian-optimal. Emissions shares and totals and evaluated actual and pledged (NDC) emissions under these two approaches are plotted in Fig. 1.

The optimal allocations of (industrial) CO₂ emissions in the United States and the European Union (EU) are much lower under the utilitarian optimum (with variable regional prices), and thus these regions' current NDCs are evaluated to be far from adequate by the utilitarian approach (Fig. 1a,b). In particular, although every region's actual and pledged emissions are above the utilitarian benchmark, and collectively are far above the limits required to keep global warming below 2 °C above pre-industrial levels, the emissions of the United States and EU are particularly far above. This conclusion contrasts with estimates and studies which indicate that the NDCs of developed regions such as the EU and United States are normatively adequate and equitable and that the NDCs of India and African nations are less adequate¹⁴. These results highlight that a cost-minimization framework produces large, policy-relevant differences in emission shares relative to the utilitarian approach, which is problematic because cost minimization ignores equity, as Kartha et al.¹⁵ have emphasized.

The allocations of CO₂ emissions within time periods and over the entire future are contrasted in Fig. 1f–i. A comparison of the distributions in Fig. 1g and 1i shows that the future emissions shares of high-income regions are nearly eliminated, with the difference allocated to developing regions such as Africa, India and low-income Asia. Comparing Fig. 1f and 1h shows that it is not merely the stock of future emissions that differs, but also the timing: under the utilitarian optimum, for example, Africa could continue to produce substantial emissions into the twenty-second century.

Our core qualitative findings, that utilitarianism permits poorer regions more space to develop and that there are climate benefits to the distribution of emissions recommended by utilitarianism, are robust to a variety of parametric and structural modifications to the model. Supplementary Tables 2 and 3 include alternative assumptions about time preference and inequality aversion, respectively. Supplementary Table 4 includes modifications to: the climate damage function (column 2), a 'cost-of-adjustment' penalty to slow optimal reductions in emissions (column 3) and alternative but commonly used parameter

combinations (columns 4 and 5). We perform a multi-model comparison by replacing the underlying IAM entirely, from RICE to FUND^{49,50} (presented in Extended Data Fig. 2). To be sure, these modifications yield quantitatively distinct results, but the broad consequences of utilitarianism rather than cost minimization for equity, temperature and development remain. In other words, whether overall decarbonization should be slower or faster or just as RICE or FUND recommends, the same relationship between the recommendations of utilitarianism and cost minimization emerges. Such robustness illustrates that a utilitarian objective is a simple, transparent alternative approach to equity that could be incorporated into many existing models.

Advantages for equity, climate and development

An important concern is that adequate mitigation policy may prevent currently developing countries from having the emissions budget needed for human development^{51–53}. We respond to this concern by comparing human development (using consumption as a proxy) in the three poorest regions of the world over the next half century under the cost-minimization and utilitarian approaches (Fig. 2a,b).

Development is slowed under the cost-minimization approach: the three poorest regions in the world are worse off relative to a scenario in which it is assumed that there is no near-term carbon price (business as usual; BAU). In contrast, all three of these regions experience development beyond BAU under the utilitarian benchmark. So, Fig. 2b shows that the poorest regions are net beneficiaries of optimal utilitarian policy, whereas Fig. 2a shows that they suffer decades of net losses under cost minimization. Fig. 2c,d shows the rate at which regions of the world decarbonize their economies over this time frame under the two policies, which drives these differences in human development.

These differences in near- and intermediate-term development translate into important consequences for the poorest people in the longer run, as Fig. 3 shows. Utilitarianism shifts much of the mitigation burden from poorer to richer regions, allowing emerging economies more headroom for development.

A utilitarian optimum, rather than a cost-minimization optimum, would also have important advantages for the climate, as Fig. 4 shows: utilitarianism recommends faster emissions reductions globally. Faster global decarbonization is justifiable under a utilitarian optimum that allocates the reductions to richer nations and temporarily spares the poorer nations. In contrast, the cost-minimization allocation is constrained by design to place the same-magnitude carbon price everywhere in the world, and therefore has costs that are welfare-inferior at the same level of global mitigation.

Relative to the cost-minimization approach, the optimal utilitarian approach would lower peak temperature while also allowing more room for developing country emissions into the twenty-second century. Because cost minimization forces all regions to have the same carbon price, which raises the human development cost of mitigation, it yields a larger peak temperature. So, utilitarian benchmarks have important climate, equity and development advantages.

Some readers may find the rapid decarbonization of richer countries to be an unrealistic aspect of the utilitarian benchmark, implemented by the RICE model. Here we emphasize that the utilitarian benchmark is an optimal allocation, but need not be the actual optimal outcome which occurs. In other words, the utilitarian benchmark could be interpreted as a baseline for the allocation of emission shares (for example, via permits for emissions), from which international emissions trading could provide further gains to wellbeing while allowing rich high-emitting nations the ability to ratchet down emissions along the most technologically and economically realistic pathway. The opportunity for further gains from trading arises because in any situation with different regional prices the same emissions level can be achieved in a Pareto-improving way by allowing a region with a higher price to pay a region with a lower price for a share in the latter's emission share (that is, an emissions trading scheme that would allocate initial permits according to the utilitarian benchmark). Thus, the small allocation of emissions rights to rich nations in the utilitarian approach need not imply that rich nations will not be allowed to emit beyond those levels, but merely that utilitarian-benchmarked equity requires that they should pay poorer nations for that privilege. In practice, whether policymakers adopt such a trading scheme depends on how the utilitarian benchmark and others are used in equity debates and politics.

A final observation is that our utilitarian approach takes seriously the fact that rich countries are not making large economic transfers to poor countries. In principle, any cost-minimization approach to benchmark national emissions and pledges could be interpreted to implicitly (and unrealistically) assume that such large international transfers are happening. Hypothetically, such a policy package could indeed be utilitarian-optimal if it combined cost minimization with large-scale international redistribution, but only as part of such a package. Because large transfers of the required magnitude are not even under discussion as part of any climate-policy package, we assume that they do not happen in our modelling here. In the absence of large progressive transfers between nations, it would be a mistake to evaluate the shares of global emissions as if such transfers were actually happening outside of the model.

In sum, the utilitarian approach outlined here is methodologically transparent and implementable in many existing modelling frameworks; recommends better outcomes for human development, equity and the climate; and more accurately reflects the policy reality of independently determined national pledges and emissions outcomes (without large international transfers). These results are particularly useful given that, among the competing benchmarks for climate equity, utilitarianism can be considered parsimonious: ethically and conceptually minimal in requiring each person's interests to count equally, emphasizing enhancements in wellbeing, and agreement that a unit less of consumption hurts the poor more than the rich.

Of course, we offer only a benchmark: no model of the utilitarian outcome will in fact be implemented. Indeed, no single benchmark for equity or mitigation ambition will be sufficient to solve the challenge of how to differentiate countries' common responsibility to address climate change under the United Nations Framework Convention on Climate Change. And yet, our transparent and simple methodology can inform equity debates. Modellers now routinely investigate the consequences of varying assumptions such as the social discount rate in robustness checks that otherwise leave models unchanged. In this

way, future modellers can investigate the consequences of optimizing a utilitarian objective, as an informative alternative to limiting their analyses to cost minimization. This could become a standard sensitivity test in the literature.

Methods

Described here are the model and data underlying the computational exercises archived⁵⁴ and freely available at: <https://github.com/Environment-Research/Utilitarianism>. All results are therefore fully reproducible; we run all simulations on the Mimi computing platform using the Julia programming language.

The RICE model was developed by William Nordhaus and analyses the tradeoffs between investing in climate mitigation, which incurs a cost relatively soon, and climate damages, which incur costs in the more distant future²⁶. RICE is the regional counterpart to the global aggregate Dynamic Integrated Climate-Economy (DICE) model, which is one of the three leading cost–benefit models used by researchers and governments for regulatory analysis, including to estimate the social cost of carbon⁵⁵. Here we describe key aspects of the standard RICE2010 model, which has been described in more detail elsewhere^{26,56}.

Briefly, RICE is a regionalized optimization model that includes an economic component and a geophysical (climate) component that are linked. RICE divides the world into 12 regions, some of which are single countries while others are groups of countries. Each region has a distinct endowment of economic inputs including capital, labour and technology, which together produce that region’s gross output via a Cobb–Douglas production function. Carbon emissions are a function of gross output and an exogenously determined, region-specific carbon-intensity pathway. These carbon emissions can be reduced (mitigated) at a cost to gross output via control policies that are selected via a carbon price. Any remaining carbon emissions are incorporated into the climate module where they influence global temperature and, ultimately, the future economy through climate-related damages. Future climate change affects regions differently, with poorer regions generally more vulnerable to climate damages. The model is solved by maximizing the weighted sum of discounted global wellbeing, where wellbeing is a concave function of consumption. We describe the RICE model equations in more detail below.

While it is often stated that optimal global climate policy requires global harmonization of marginal abatement costs, this is only the case if distributional issues are ignored or if lump-sum transfers are made between countries. Chichilnisky and Heal¹⁶ have shown in a quite general theoretical model that if there are global inequalities and the absence of, for example, corrective lump-sum transfers between countries, then a policy in which different regions face different carbon prices may be superior to one with a single global carbon price. Still, most IAMs assume away distributional issues and calculate climate policy that assumes a single global carbon price.

However, in RICE it is possible to remove the constraint that the carbon price must be globally uniform, and instead instruct the model to vary the carbon price in each region while maintaining the same global objective (equation (1) below). This is the method we

follow here (for similar methods and results in multiple models see prior related work)^{17,46}. The resulting optimum with varying regional prices can be conceptualized as a ‘second best’ policy in the sense familiar from public economics, optimized to a situation in which massive uncorrected global income inequality is known to loom uncorrected in the background: equalizing the marginal dollar cost of abatement (as existing models require) does not take the diminishing marginal utility of consumption properly into account in such a context, where it is known that a dollar of abatement has a much larger utility cost in a poor region than in a rich region. In addition to the prior literature we join that has focused on equity, other important recent papers have assessed and contrasted NDCs with a different focus: for example, on comparability and the Sustainable Development Goals⁵⁷ or on transparency and coordination⁵⁸.

In this way, varying regional prices allows richer regions to contribute much more mitigation effort than poorer regions in a way that is optimized to a utilitarian objective, thus permitting poorer regions to continue developing. This maximizes the utilitarian improvements that are possible (as in Chichilnisky and Heal’s proof¹⁶) over the optimum that involves a single uniform global carbon price. As a result, we call the optimum with varying prices the ‘utilitarian optimum’, in contrast to the uniform carbon price optimum that implies cost minimization.

Equations and model description.

In more detail, our version of RICE follows RICE2010²⁶ and most of the literature in using a discounted and separable constant-elasticity objective function with (total utilitarian) population weights:

$$W(c_{it}) = \sum_{it} \omega_{it} \frac{L_{it}}{(1 + \rho)^t} \frac{c_{it}^{1-\eta}}{1-\eta} \quad (1)$$

where W denotes social welfare, L population, c per capita consumption, ρ the rate of pure time preference and η inequality aversion (that is, the consumption elasticity of marginal utility). We set ρ at 0.8% to deliver a 2 °C optimal temperature path. η is set to 1.5 which implies that a dollar forgone by a rich person is equivalent to 2.8 dollars lost by a person who is half as rich. The subscripts i and t are the region and time indices, respectively.

If the weights that appear in this equation, ω , are proportional to the inverse of the marginal utility of consumption then they are called time-varying Negishi weights. Time-varying Negishi weights are used in many climate–economy models, including early versions of RICE, where they were introduced to impose constraints on capital flows. (The first version of RICE was implemented like a computational general equilibrium model, in which there would be capital flows until the marginal utilities of consumption are equated across regions⁵⁹.) RICE2010 does not require weights for this purpose (because regions are autarkic), but they are still used in Nordhaus’s version of RICE2010 so that the maximization-as-market-simulation principle holds³⁸.

The objective function we use sets $\omega = 1$ for all i and t and thus does not use Negishi weights. We use uniform weights for a number of reasons. First, we do not suppose that our results represent a market simulation because we do not expect the mitigation rates that we compute to emerge as the result of an unregulated market. But more importantly for the purposes of this paper, Negishi weights distort time preferences⁶⁰ and the inter-regional trade-off⁶¹ in ways that are opaque and difficult to justify, both descriptively and normatively. We have explained this change in more detail in a previous publication⁴².

Principally, we run the model without Negishi weights to investigate the policy implications of a utilitarian objective; such an objective is ruled out by Negishi weights, which do not weigh the interests of each person equally. Because regions are autarkic in RICE, our use of uniform weights does not result in progressive income redistribution between regions; instead, the only distributional choice within the model is how to distribute future emissions via different regional carbon prices. In this way, we model climate policy relative to suboptimal background facts about global income inequality that are not counterfactually assumed to be successfully addressed by any direct global-income-redistribution scheme; but, in all of our modelling, we still set aside the issue of whether there should be general progressive income redistribution, as we do not assume or permit any such general redistribution in our modelling.

While we call equation (1) without weights a utilitarian objective, it can also be interpreted as a prioritarian objective^{62–65}, because equation (1) relies on a diminishing social welfare of consumption parameter η , which could be interpreted in the mainstream utilitarian way, but could also be interpreted as implicitly capturing the additional distinctive prioritarian idea that there is also diminishing social welfare of utility. (See Adler et al.³² for a study that further explicitly represents and distinguishes prioritarian parameters.)

The utilitarian optimum that we report involves large differences in carbon prices between regions; at the same time, we assume that regions are autarkic. The rationale for this begins with the observation that no model in the IAM literature fully endogenizes the movement of capital, technology or labour that would result from its policies. For example, the standard RICE model takes population growth to be exogenous, even though migration is well known to respond to economic incentives. We follow RICE by assuming that all factors do not endogenously relocalize globally. Such a modelling approach can be interpreted as equivalent to assuming the existence of border taxes or other controls that prevent such endogenous relocalization. In particular, differential prices can cause a competitiveness differential that could lead to relocation of energy intensive industries. The large literature on ‘carbon leakage’ looks at this issue and at policy proposals, such as border tax adjustments, to counteract the effect. The broad conclusion of this literature is that there are two channels for leakage: competitiveness differences due to carbon price differences and fossil fuel price level reductions due to decreased global demand. The consensus is that the second, price level, effect is the dominating one^{66–69}. This has two important consequences for our utilitarian benchmark. First, the larger, level, effect does not apply in a proposal in which all countries commit to emissions caps, because the overall global emission is thereby capped. The competitiveness channel persists, but it is small in size and can be

adequately dealt with by an implementation of border tax adjustments, such as that proposed in Flannery et al.⁷⁰.

The main climate and economic dynamics of our version of RICE follow RICE2010. In our version of RICE, the world is composed of the same twelve macro-region economies as in RICE2010 (some of which are in fact countries, such as China, India, the United States, Russia and Japan). And, as in RICE2010, given pre-damage and pre-mitigation cost gross output Q_{it} in region i at time t (determined by a Cobb–Douglas production function), the post-damage and post-mitigation cost net output Y_{it} is

$$Y_{it} = \left(\frac{1 - \Lambda_{it}}{1 + D_{it}} \right) Q_{it} \quad (2)$$

where D_{it} is regional damage and Λ_{it} is regional mitigation cost. For regional population L_{it} and savings rate S_{it} investment I_{it} (which adds to the stock of capital) is

$$I_{it} = S_{it} Y_{it} \quad (3)$$

and average per capita consumption (for region i at time t) is

$$\bar{c}_{it} = \frac{1 - S_{it}}{L_{it}} Y_{it}. \quad (4)$$

The function determining the damage term D_{it} is a quadratic function of temperature rise above pre-industrial levels (T), with the severity and shape of damages governed by linear and quadratic coefficients α_1 and α_2 :

$$D_{it} = \alpha_1 T_i + \alpha_2 T_i^2. \quad (5)$$

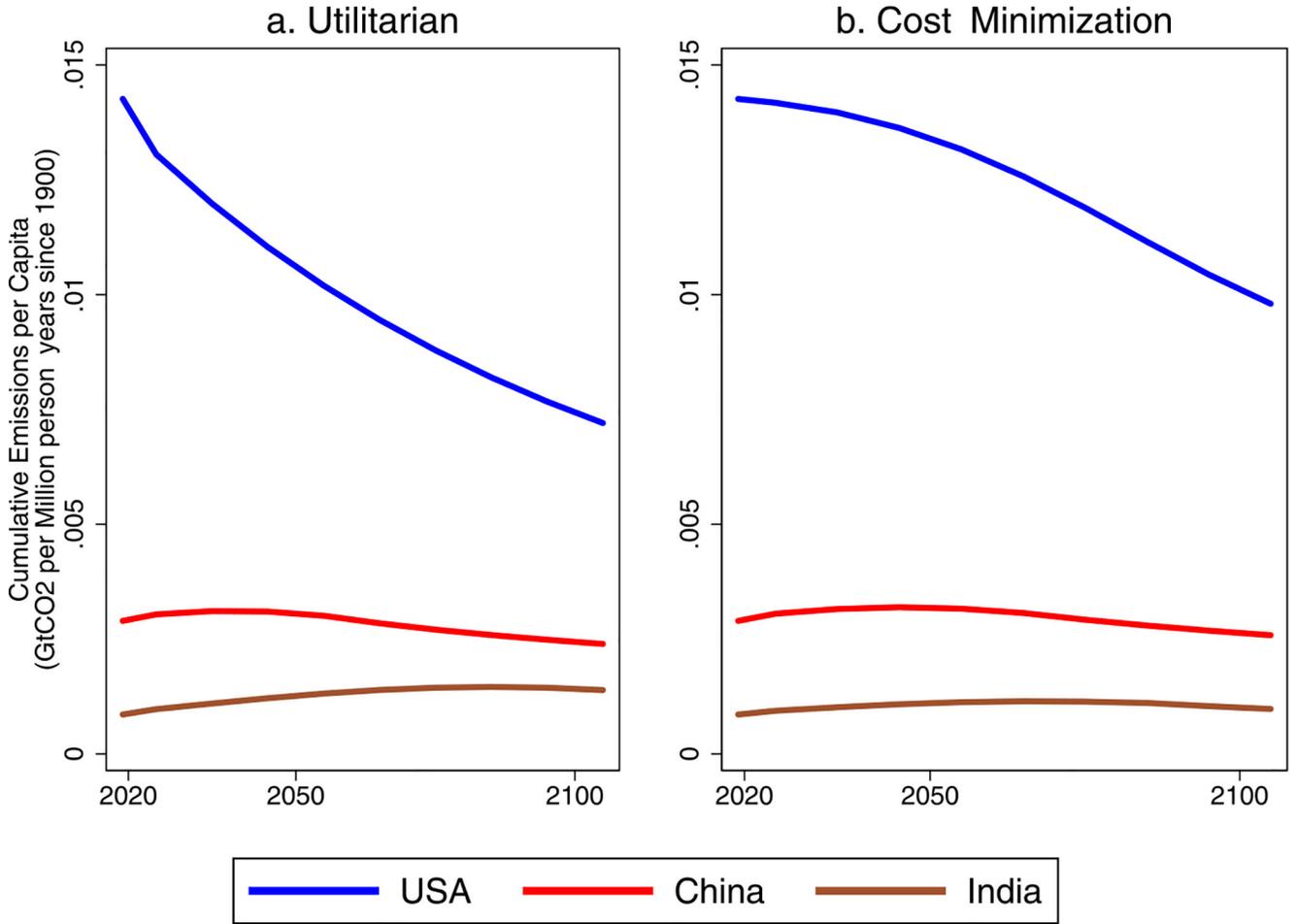
Equation (2) links the economic and geophysical (climate) components; industrial carbon emissions can be reduced (mitigated) at a cost to gross output (Λ) via control policies that are selected via a global carbon price. In the utilitarian optimum, different regions are allowed to have different carbon prices at each time point, whereas in the cost-minimization optimum regions are constrained to all have the same carbon price at each time point. Any remaining carbon emissions are incorporated into the climate module where they influence global temperature and, ultimately, the future economy through climate-related damages (D) in equation (5). Future climate change affects regions differently, with poorer regions generally more vulnerable to climate damages. All of this influences the consumption of heterogeneous future people via equation (4), which influences the sum of wellbeing in equation (1). The model is solved by maximizing equation (1).

Our version of RICE involves two other modifications beyond those described above. First, we updated the population projections to those of the United Nations 2017 medium variant⁴². Second, in RICE2010 the social objective is not clearly distinguished from the preferences of infinitely lived representative agents. In our implementation, we decouple the behaviour of the private sector from the social welfare evaluation, so that when we change a parameter in the social welfare function (discount rate or inequality aversion), this triggers no change in the private preferences of the individual agents about intertemporal allocation and savings. Savings in our version of RICE are not determined by the social welfare function but within a slightly simplified utility function with very similar parameters to ours, following prior work^{42,71}. This results in a fixed savings rate of 25.8%. The environment we study has slower depreciation of capital than the environment in which we derive the savings rate, so we have checked the sensitivity of our results with lower uniform savings rates (down to 20.8%). These results are available upon request, but do not meaningfully change the results besides leaving all nations slightly less wealthy in the far future.

Our version of RICE results in lower peak temperature at the cost-minimization optimum than RICE2010 at its cost-minimization optimum due to the modifications detailed in this section, where this change is due primarily to our modification to the objective function (that is, the fact that we do not use Negishi weights in equation (1)). (See Supplementary Table 5 below for demonstration that these changes do not drive our results.)

The preceding paragraphs summarize the key features of RICE relevant to the analysis in this paper. Additional technical discussion of RICE2010 is available in Nordhaus²⁶.

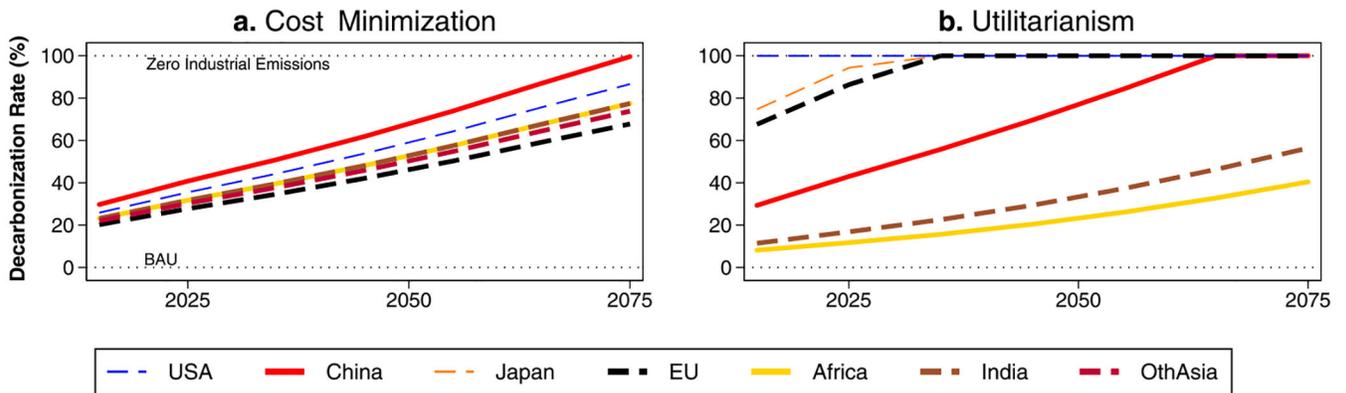
Extended Data



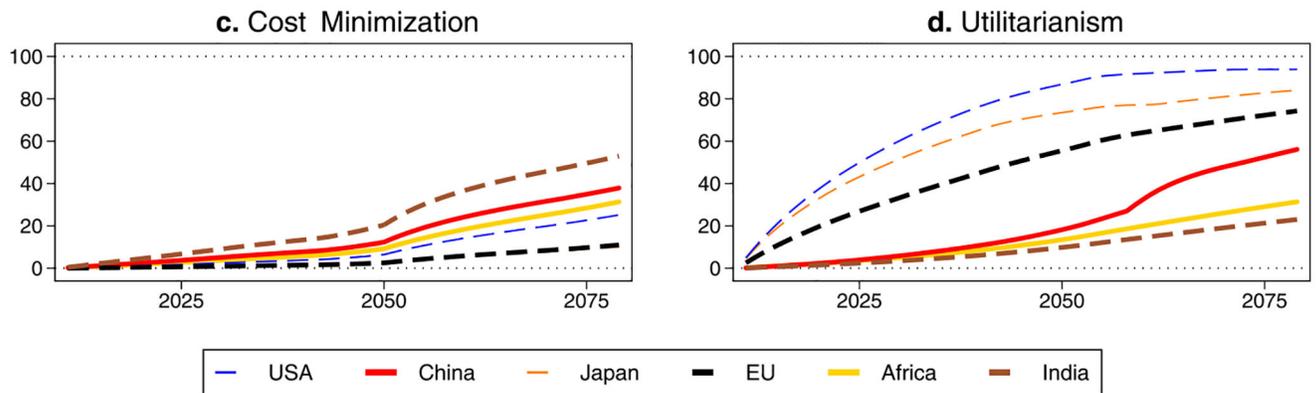
Extended Data Fig. 1 | Cumulative emissions per capita converge over the 21st century under utilitarianism.

The vertical axis plots, for each year, (cumulative emissions by region since 1900 up to that year) divided by (the total population, measured as person-years, lived in that region since 1900 up to that year). The high levels for USA in 2020 are the result of far higher emissions through the 20th century than the poorer regions plotted. Only countries are presented for which we have adequate population and emissions data going back to 1900.

RICE



FUND

**Extended Data Fig. 2 |. Multi-Model robustness.**

Implementing the utilitarian method in the FUND model. The top row replicates Fig. 2c–d for RICE, while the bottom row displays the analogous results in the FUND model, which is known to have substantively different assumptions and structure.⁵⁰ (See FUND documentation for details of the FUND model.⁵¹) The main results of the paper—that regional emission allocations are heavily tilted towards developing countries in the utilitarian optimum—also hold in the FUND model. (See also Anthoff 2011.¹⁸) We optimize FUND through 2300 with the same discounting parameters and utilitarian objective function used to generate our main results with RICE. The FUND results assume that regional carbon taxes can go no higher than \$5000/ton CO₂ and remain constant after 2200.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Research by D.S. is supported by the National Institute of Child Health and Human Development (NICHD) grants K01HD098313 and P2CHD042849. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability

All data used in our version of the model is archived⁵⁴ and freely available at <https://github.com/Environment-Research/Utilitarianism>.

References

1. Clarke L et al. in *Climate Change 2014: Mitigation of Climate Change* (eds Edenhofer O et al) Ch. 6 (Cambridge Univ. Press, 2014).
2. Adoption of the Paris Agreement FCCC/CP/2015/L.9/Rev.1 (UNFCCC, 2015); <https://unfccc.int/resource/docs/2015/cop21/eng/109r01.pdf>
3. Pauw P, Mbeva K & van Asselt H Subtle differentiation of countries' responsibilities under the Paris Agreement. *Palgrave Commun.* 5, 86 (2019).
4. Winkler H et al. Countries start to explain how their climate contributions are fair: more rigour needed. *Int. Environ. Agreem* 18, 99–115 (2018).
5. Nationally Determined Contributions (NDCs) (UNFCCC, 2020); <https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement/nationally-determined-contributions-ndcs>
6. Höhne N et al. Emissions: world has four times the work or one-third of the time. *Nature* 579, 25–28 (2020). [PubMed: 32132686]
7. Ackerman F, Stanton EA & Bueno R CRED: a new model of climate and development. *Ecol. Econ* 85, 166–176 (2013).
8. CSO Equity Review Fair Shares: A Civil Society Equity Review of INDCs. figshare 10.6084/m9.figshare.5917399 (2018).
9. CSO Equity Review Equity and the Ambition Ratchet: Towards a Meaningful 2018 Facilitative Dialogue. figshare 10.6084/m9.figshare.5917408 (2018).
10. Holz C, Kartha S & Athanasiou T Fairly sharing 1.5: national fair shares of a 1.5 °C-compliant global mitigation effort. *Int. Environ. Agreem* 18, 117–134 (2018).
11. Sheriff G Burden sharing under the Paris Climate Agreement. *J. Assoc. Environ. Resour. Econ* 6, 275–318 (2019).
12. van den Berg NJ et al. Implications of various effort-sharing approaches for national carbon budgets and emissions pathways. *Clim. Change* 162, 1805–1822 (2020).
13. Zimm C & Nakicenovic N What are the implications of the Paris Agreement for inequality? *Clim. Policy* 20, 458–467 (2020).
14. Robiou du Point Y et al. Equitable mitigation to achieve the Paris Agreement goals. *Nat. Clim. Change* 7, 38–43 (2017).
15. Kartha S et al. Cascading biases against poorer countries. *Nat. Clim. Change* 8, 348–349 (2018).
16. Chichilnisky G & Heal G Who should abate carbon emissions? *Econ. Lett* 44, 443–449 (1994).
17. Anthoff D *Optimal Global Dynamic Carbon Abatement* (Univ. California Berkeley, 2011).
18. Caney S *Cosmopolitan justice, responsibility, and global climate change*. *Leiden J. Int. Law* 18, 747–775 (2005).
19. Shue H *Climate Justice: Vulnerability and Protection* (Oxford Univ. Press, 2014).
20. Gardiner S, Caney S, Jamieson D & Shue H (eds) *Climate Ethics: Essential Readings* (Oxford Univ. Press, 2010).
21. Bykvist K *Utilitarianism: A Guide for the Perplexed* (Continuum, 2009).
22. Kagan S *Normative Ethics* (Routledge, 1998).
23. Singer P in *Climate Ethics: Essential Readings* (eds. Gardiner S et al.) 181–199 (Oxford Univ. Press, 2010).
24. Dooley K et al. Ethical choices behind quantifications of fair contributions under the Paris Agreement. *Nat. Clim. Change* 11, 300–305 (2021).
25. Adler M *Measuring Social Welfare: An Introduction* (Oxford Univ. Press, 2019).
26. Nordhaus WD Economic aspects of global warming in a post-Copenhagen environment. *Proc. Natl Acad. Sci. USA* 107, 11721–11726 (2010). [PubMed: 20547856]

27. Tol RSJ The damage costs of climate change towards a dynamic representation. *Ecol. Econ* 19, 67–90 (1996).
28. Global Carbon Atlas—CO₂ Emissions (Future Earth; accessed October 2020); <http://www.globalcarbonatlas.org/en/CO2-emissions>
29. Robiou du Pont Y A Multidimensional Equity and Warming Assessments of National Climate Pledges (Paris Equity Check, accessed 10 September 2018); <http://paris-equity-check.org/multi-equity-map.html>
30. Pauw WP & Klein RJT Beyond ambition: increasing the transparency, coherence and implementability of nationally determined contributions. *Clim. Policy* 20, 405–414 (2020).
31. Wei T et al. Developed and developing world responsibilities for historical climate change and CO₂ mitigation. *Proc. Natl Acad. Sci. USA* 109, 12911–12915 (2012). [PubMed: 22826257]
32. Adler M et al. Priority for the worse-off and the social cost of carbon. *Nat. Clim. Change* 7, 443–449 (2017).
33. Fleurbaey M & Zuber S Discounting, beyond utilitarianism. *Economics* 9, 1–52 (2015).
34. Broome J *Climate Matters: Ethics in a Warming World* (Norton, 2012).
35. Weyant J Some contributions of integrated assessment models of global climate change. *Rev. Environ. Econ. Policy* 11, 115–137 (2017).
36. Bosetti V & Frankel J Politically feasible emissions targets to attain 460 ppm CO₂ concentrations. *Rev. Environ. Econ. Policy* 6, 86–109 (2012).
37. Edenhofer O, Lessmann K, Kemfert C, Grubb M & Kohler J Induced technological change: exploring its implications for the economics of atmospheric stabilization: synthesis report from the Innovation Modeling Comparison Project. *Energy J.* 27, 57–107 (2006).
38. Nordhaus WD & Sztorc P *DICE 2013R: Introduction and User's Manual* (Yale Univ., 2013).
39. Hope CW The marginal impacts of CO₂, CH₄ and SF₆ emissions. *Clim. Policy* 6, 537–544 (2006).
40. Anthoff D *Equity and Climate Change: Applications of FUND* (Max Planck Institute, 2009).
41. Stern N *Stern Review: The Economics of Climate Change* (HM Treasury, 2006).
42. Dennig F, Budolfson MB, Fleurbaey M, Siebert A & Socolow RH Inequality, climate impacts on the future poor, and carbon prices. *Proc. Natl Acad. Sci. USA* 112, 15827–15832 (2015). [PubMed: 26644560]
43. Bataille C, Guivarch C, Hallegatte S, Rogelj J & Waisman H Carbon prices across countries. *Nat. Clim. Change* 8, 648–650 (2018).
44. Sandmo A Global public economics: Public goods and externalities. *Public Econ.* 18–19, 3–21 (2007).
45. Sheeran KA Who should abate carbon emissions? A note. *Environ. Resour. Econ* 35, 89–98 (2006).
46. Budolfson M & Dennig F in *Handbook on the Economics of Climate Change* (eds. Chichilnisky G et al.) 224–238 (Edward Elgar Press, 2020).
47. Rawls J A *Theory of Justice* (Harvard Univ. Press, 1971).
48. Scanlon TM *What We Owe to Each Other* (Harvard Univ. Press, 2000).
49. Diaz D & Moore F Quantifying the economic risks of climate change. *Nat. Clim. Change* 7, 774–782 (2017).
50. Anthoff D & Tol RSJ Fund Documentation (2016); <https://fundjl.readthedocs.io/en/latest/index.html>
51. Lamb WF & Rao ND Human development in a climate-constrained world: what the past says about the future. *Glob. Environ. Change* 33, 14–22 (2015).
52. Rao ND & Baer P “Decent living” emissions: a conceptual framework. *Sustainability* 4, 656–681 (2012).
53. Winkler H et al. *Equitable Access to Sustainable Development: Contribution to the Body of Scientific Knowledge* (BASIC Expert Group, 2011).
54. Budolfson M et al. Utilitarian benchmarks for emissions and pledges promote equity, climate and development. *Zenodo* 10.5281/zenodo.4618473 (2021).

55. Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866 (Interagency Working Group on the Social Cost of Carbon, United States Government, 2016).
56. Nordhaus WD & Boyer J Warming the World: Economic Models of Global Warming (MIT Press, 2000).
57. Iyer G et al. Implications of sustainable development considerations for comparability across nationally determined contributions. *Nat. Clim. Change* 8, 124–129 (2018).
58. Aldy J et al. Economic tools to promote transparency and comparability in the Paris Agreement. *Nat. Clim. Change* 6, 1000–1004 (2016).
59. Nordhaus WD & Yang Z A regional dynamic general-equilibrium model of alternative climate-change strategies. *Am. Econ. Rev* 86, 741–765 (1996).
60. Dennig F & Emmerling J A Note on Optima with Negishi Weights (Princeton Univ., 2017).
61. Stanton EA Negishi welfare weights in integrated assessment models: the mathematics of global inequality. *Clim. Change* 107, 417–432 (2011).
62. Broome J Weighing Goods: Equality, Uncertainty and Time (Blackwell, 2003).
63. Broome J Equality versus priority: a useful distinction. *Econ. Philos* 31, 219–228 (2015).
64. Adler M Future generations: a prioritarian view. *George Wash. Law Rev* 77, 1478–1519 (2009).
65. Voorhoeve A & Fleurbaey M Priority or equality for possible people? *Ethics* 126, 929–954 (2016).
66. Monjon S & Quirion P A border adjustment for the EU ETS: reconciling WTO rules and capacity to tackle carbon leakage. *Clim. Policy* 11, 1212–1225 (2011).
67. Lockwood B & Whalley J Carbon-motivated border tax adjustments: old wine in green bottles? *World Econ.* 33, 810–819 (2010).
68. Aldy JE & Pizer WA The competitiveness impacts of climate change mitigation policies. *J. Assoc. Environ. Resour. Econ* 2, 565–595 (2015).
69. Felder S & Rutherford TF Unilateral CO₂ reductions and carbon leakage: the consequences of international trade in oil and basic materials. *J. Environ. Econ. Manage* 25, 162–176 (1993).
70. Flannery B, Hillman J, Mares JW & Porterfield M Framework Proposal for a US Upstream Greenhouse Gas Tax with WTO-Complaint Border Adjustments (Resources for the Future, 2018).
71. Golosov M, Hassler J, Krusell P & Tsyvinski A Optimal taxes on fossil fuel in general equilibrium. *Econometrica* 82, 41–88 (2014).

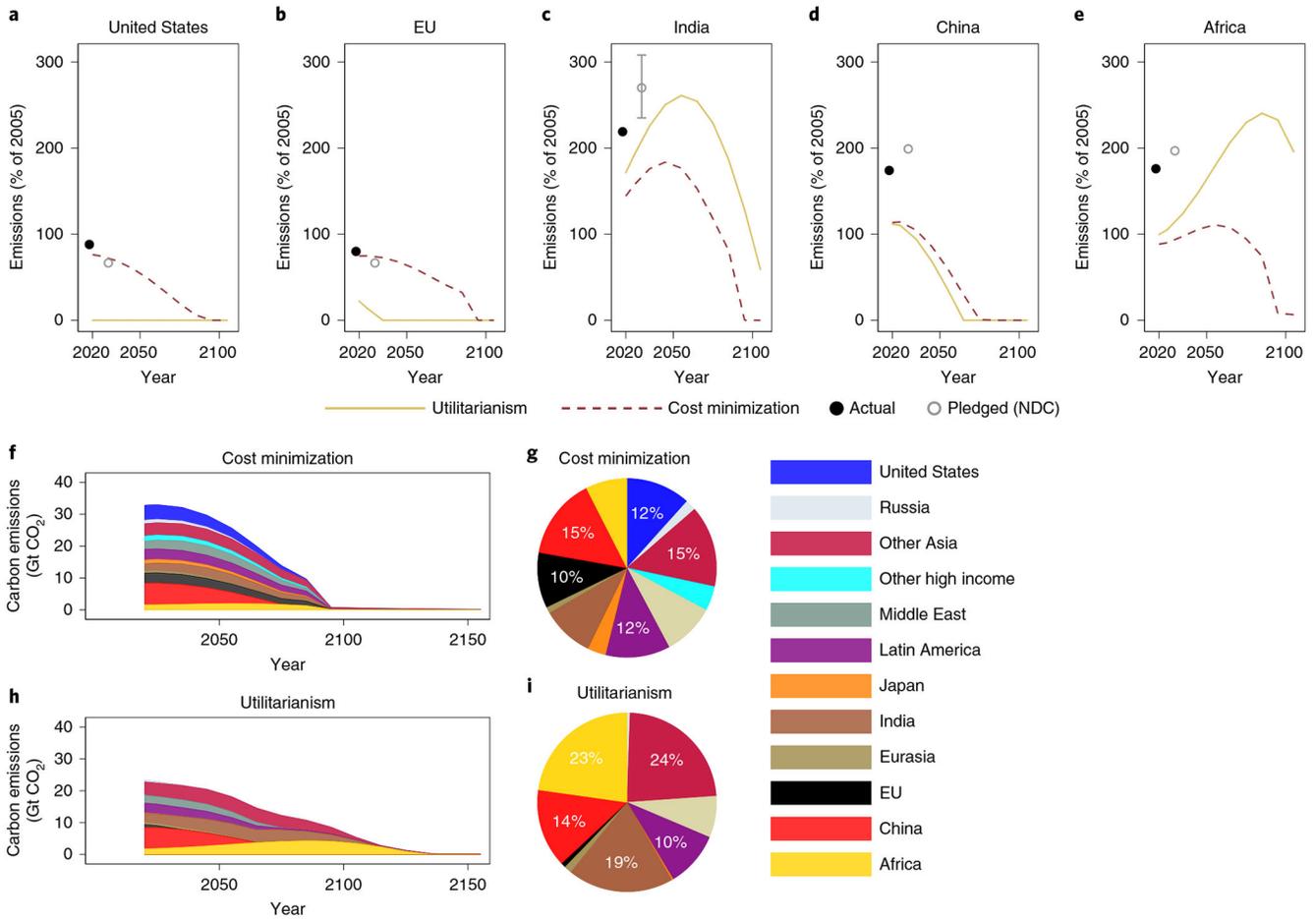


Fig. 1 |. Evaluating NDCs with the utilitarian benchmark.

Regional shares of CO₂ emissions and resulting evaluation of actual (2019) and pledged (NDC) emissions are plotted under two policies: a cost-minimization optimum with uniform global carbon prices (cost minimization) and the optimal utilitarian regime with carbon prices that can vary between regions (utilitarianism). **a–i**, Comparisons of actual and pledged emissions for selected regions with utilitarian benchmarks (**a–e**), future global CO₂ emissions with regional decomposition (**f–h**) and regional shares of all future global CO₂ emissions (**g–i**) for these two policies. Actual emissions are from Global Carbon Atlas of the Global Carbon Project²⁸. NDCs are from du Pont et al.¹⁴ and the accompanying external data visualization tool²⁹, where bars represent high versus low NDCs.

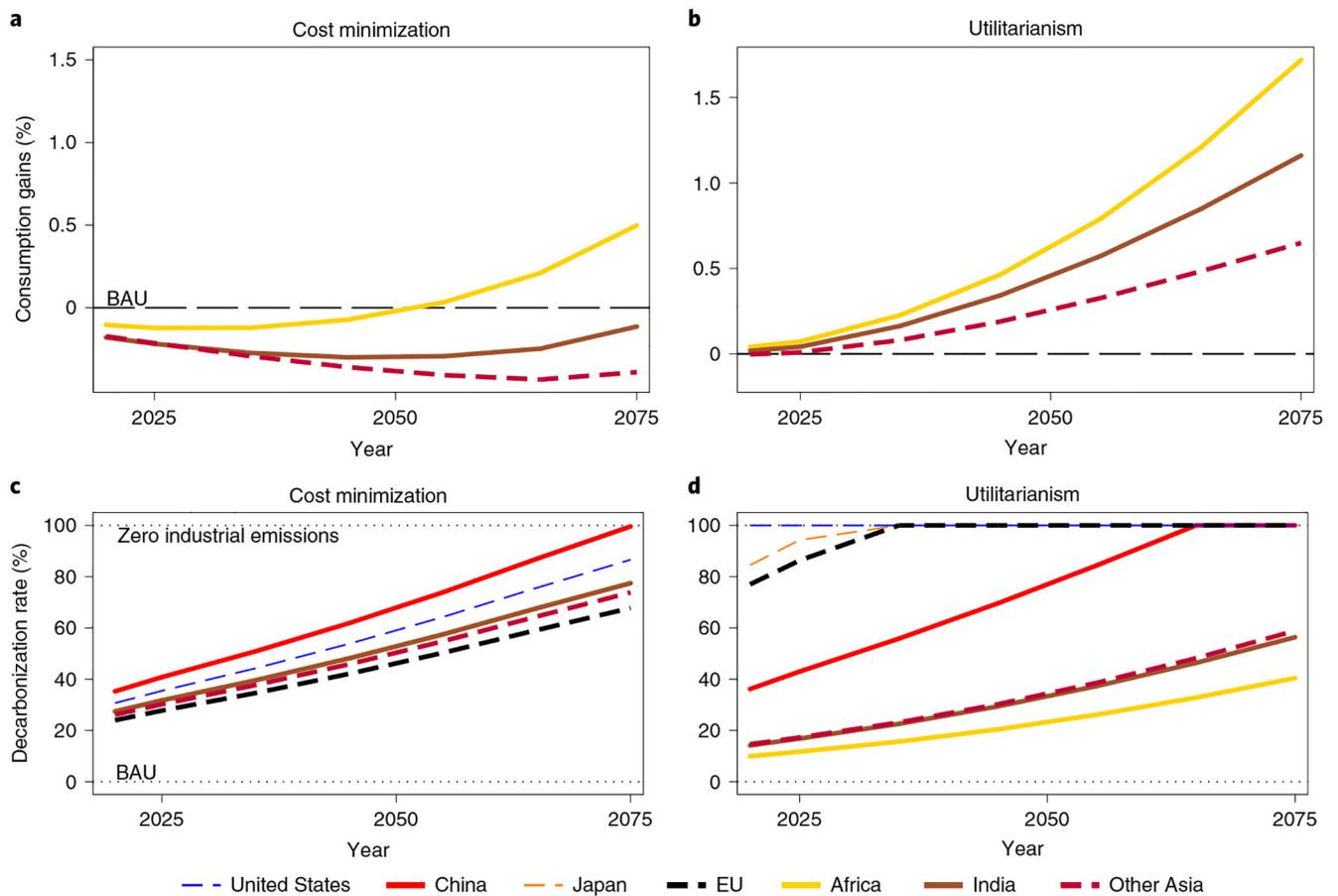


Fig. 2 |. Human development and equity advantages of the utilitarian policy.

a,b, Percent change in per capita consumption over time versus a BAU scenario in which there is no carbon price into the future for the cost-minimization policy (**a**) versus the optimal utilitarian policy (**b**). **c,d**, Percent reduction in CO₂ versus BAU (decarbonization rate) under the cost-minimization policy (**c**) versus the optimal utilitarian policy (**d**).

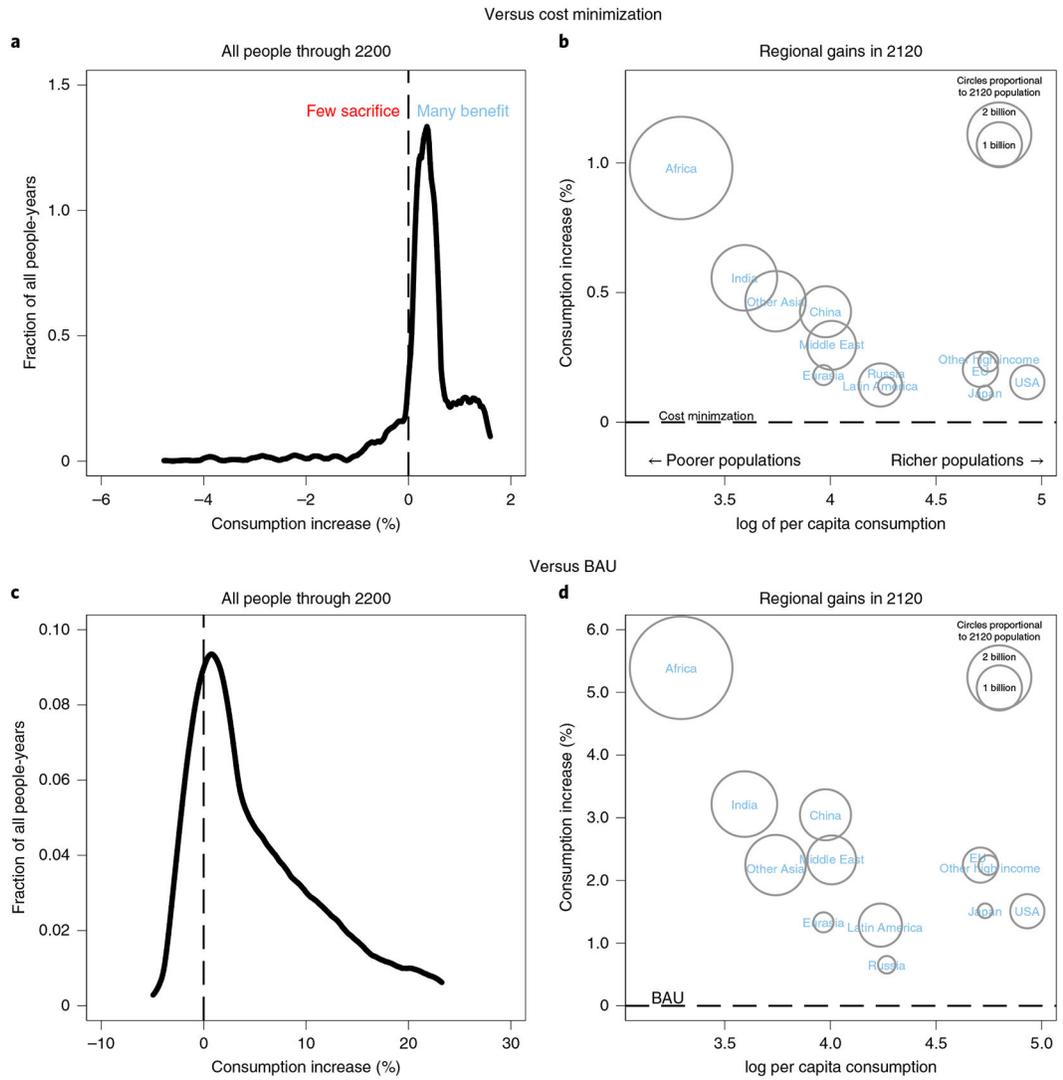


Fig. 3 |. Gains from adopting the utilitarian versus the cost-minimization optimal policy. The top row compares the utilitarian policy and cost minimization. **a**, The distribution of gains (or losses) for all persons on the planet through 2200 (that is, a density function of welfare gains for each person-year lived). The y units indicate the relative number of people with some level of consumption gains; a y value of 1, for example, means there are twice as many people with those gains than another point with a y value of 0.5. **b**, A sample year (2120) demonstrates cross-sectionally of which regions are the largest beneficiaries (selecting other years reveals a similar correlation between gains and per capita consumption). The bottom row shows the analogous plots (welfare gains for each person-year lived, **c**; regional gains, **d**) for utilitarianism versus a BAU scenario in which there is no carbon price into the future.

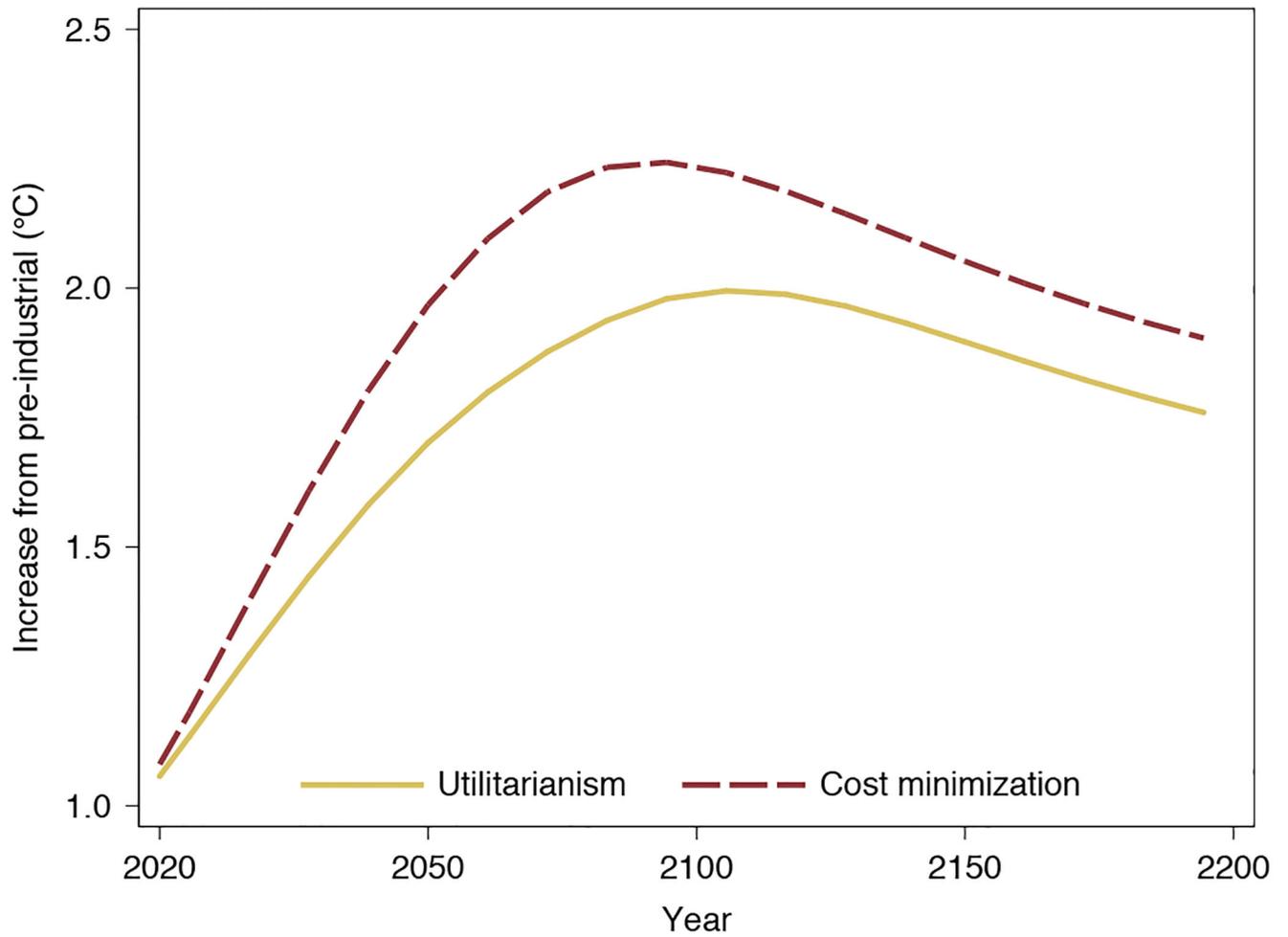


Fig. 4 |. Climate advantages of the utilitarian policy.

Comparison of global average temperature increase in each year over pre-industrial levels in RICE. The solid line reflects the optimal temperature path under the utilitarian policy; the dashed line plots the optimal path under the cost-minimization optimum.