STOCKHOLM ENVIRONMENT INSTITUTE

Negishi Welfare Weights: The Mathematics of Global Inequality

Elizabeth A. Stanton

May 14, 2009

Stockholm Environment Institute

Working Paper WP-US-0902

Abstract

In a global climate policy debate fraught with differing understandings of right and wrong, the importance of making transparent the ethical assumptions used in climate-economics models cannot be overestimated. Negishi weighting is a key ethical assumption at work in climate-economics models, but one that is virtually unknown to most model users. Negishi weights freeze the current distribution of income between world regions; without this constraint, IAMs that maximize global welfare would recommend an equalization of income across regions as part of their policy advice. With Negishi weights in place, these models instead recommend a course of action that would be optimal only in a world in which global income redistribution cannot and will not take place. This article describes the Negishi procedure and its origin in theoretical and applied welfare economics, and discusses the policy implications of the presentation and use of Negishi-weighted model results, as well as some alternatives to Negishi weighting in climate-economics models.

For more information about this document,
contact Elizabeth A. Stanton at liz.stanton@sei-us.org

Stockholm Environment Institute - US
11 Curtis Avenue
Somerville, MA 02144-1224, USA
www.sei-us.org and www.sei.se

1. Introduction

Climate change and global income inequality are intimately and inextricably interrelated. Regional income per capita is well correlated with past and present greenhouse gas emissions. Damages from climate change are expected to be far worse in developing countries, many of which have special geographic vulnerabilities (like tropical climates or low-lying islands) and most of which can ill afford the adaptation measures necessary to fend off climate impacts. International climate negotiations hinge on the resolution of philosophical differences between developed and developing countries on the meaning equity: What determines each country's responsibility in reducing emissions? What determines each country's responsibility in paying for emissions reductions and adaptation measures at home and abroad?

Climate-economics models play an important role in this policy debate by quantifying the value of climate damages and abatement costs under various mitigation scenarios. Many of these integrated assessment models (IAMs) offer policy advice in the form of an optimal scenario: a recommended course of action telling us how much emissions should be abated in what time periods to achieve the greatest human well-being[1] across the centuries. In a policy debate fraught with differing understandings of right and wrong, the importance of making transparent the ethical assumptions used in these models cannot be overestimated.

Like any economic models, IAMs are not value-free – many of the assumptions that go into building climate-economics models are based on moral judgments, and not on scientific facts. The choice of a discount rate is, perhaps, the best known example. The optimal course of action recommended by an IAM can only be understood in the context of the discount rate employed – the higher the discount rate the lower the value that we place on the well-being of future generations (2006).

Negishi weighting is another key ethical assumption at work in climate-economics models, but one that is virtually unknown to most model users. Negishi weights freeze the current distribution of income between world regions; without this constraint, IAMs that maximize global welfare would recommend an equalization of income across regions as part of their policy advice. With Negishi weights in place, these models instead recommend a course of action that would be optimal only in a world in which global income redistribution cannot and will not take place.

The next sections of this article describe the Negishi procedure and its origin in theoretical and applied welfare economics. The final section discusses the policy implications of the presentation and use of Negishi-weighted model results, as well as some alternatives to Negishi weighting in climate-economics models.

[1] This article follows the convention in economics of using the terms "well-being", "welfare", and "utility" interchangeably.

2. Negishi Weighting in Climate-Economics Models

In welfare optimizing climate-economics models, production results in both emissions and consumption; emissions affect the climate, causing damages that reduce production and, as a result, consumption. These models maximize the discounted present value of welfare (which grows with consumption, although at an ever-diminishing rate)[2] across all time periods. They accomplish this by choosing how much to abate emissions in each time period, where abatement costs reduce production and consumption. Examples of welfare optimizing IAMs include both global (single region) models – DICE-2007 (Nordhaus 2008), ENTICE-BR (Popp 2006), DEMETER-1CCS (Gerlagh 2006), and MIND (Edenhofer *et al.* 2006) – and regionally disaggregated models – RICE-2004 (Yang and Nordhaus 2006), FEEM-RICE (Bosetti *et al.* 2006), FUND (Tol 1999), MERGE (Manne and Richels 2004), CETA-M (Peck and Teisberg 1999), GRAPE (Kurosawa 2004), and AIM/Dynamic Global (Masui *et al.* 2006).[3]

Global welfare optimizing IAMs

Some of the IAMs that are solved by optimizing welfare approach the world as one homogenous region. A social welfare function for the world as a whole, summed across all time periods, is maximized by adjusting the shares of global output that are spent on emissions abatement – avoiding future damages – or recycled as investment in productive enterprises, thereby increasing consumption. Models chose unique abatement spending and investment values for every time period simultaneously, as if there were perfect foresight. Social welfare is measured as individual utility, which is a function of consumption, multiplied by the population. If little is spent on abatement in a given period, emissions are higher and future climate damages are higher, reducing future consumptions levels. In time periods when abatement spending is prioritized over regular investment, emissions are lower but so too is current day consumption.

Welfare optimizing IAMs maximize utility, not income or consumption. All individuals are assumed to have identical utility functions, which are modeled with diminishing marginal returns to utility from changes in consumption. Each new unit of additional consumption provides a little less utility than the last unit, so that as consumption rises, so too does utility, but at a diminishing rate. Consumption's contribution to utility depends on the individual's income, and a new dollar of consumption means more to the poor than to the rich. In a global welfare optimizing IAM, there is only one individual – the world as a whole – but there are many time periods. In these models, a new dollar of consumption contributes more to social welfare in a poorer time period than in a richer one.

---

[2] In these models, consumption's returns to welfare are always positive but diminish as we grow wealthier. Formally, the first derivative of welfare is always positive and the second is always negative. A popular, though not universal, choice defines individual welfare, arbitrarily, as the logarithm of per capita consumption or income.

[3] There are several varieties of climate-economics models: welfare optimization, general equilibrium, partial equilibrium, simulation, and cost minimization models (see Stanton *et al*. (2009)). Only the regionally disaggregated subset of welfare optimization IAMs have a built-in imperative for income equalization across regions and, therefore, only regionally disaggregated welfare optimization IAMs employ Negishi welfare weights to counteract income redistribution.

When diminishing marginal returns to utility from consumption are combined with the near universal assumption that per capita income will grow over time, the result is that in these IAMs a new dollar of consumption contributes more to social welfare in the (relatively poorer) present than it does in the (relatively richer) future. If the worst climate damages are to be averted, far-reaching emissions abatement measures must take place in the next few decades. If abatement measures do not take place, then large scale climate damages will occur closer to the end of this century and extending from there into the future. As a generalization, abatement costs reduce current consumption for our "poorer" generation while climate damages reduce future consumption for a "richer" generation. The same reduction to consumption, whether from current abatement costs or future damages, does not have the same impact on social welfare when measured according to these assumptions. If returns to utility from consumption are diminishing on the margin and income per capita is expected to grow over time, today's $1 million abatement cost is assigned a much bigger reduction to social welfare than $1 million in future climate damages.

The degree to which the marginal utility of income diminishes as income grows is one of two components to the discount rate. Frank Ramsey's (1928) widely used "prescriptive" approach to discounting distinguishes between: the rate of pure time preference, or how human society feels about costs and benefits to future generations regardless of the resources and opportunities that may exist in the future; and an income-based component – an elasticity applied to the rate of growth of real consumption – that reflects the diminishing marginal utility of consumption over time as society becomes richer.

The discount rate, r(t), combines these two elements: the rate of pure time preference, ρ, plus the product of the elasticity of marginal utility with respect to consumption per capita, η, and the growth rate of consumption per capita, g(t).[4]

1) $r(t) = \rho + \eta g(t)$

Choices about the discount rate inevitably reflect value judgments made by modelers. The selection of a value for the pure rate of time preference is determined by a modeler's ethical viewpoint, not economic theory or scientific fact. It is one thing to observe an individual's time preference over his or her own lifetime, and quite another to use this time preference as part of weighting the importance of costs and benefits to different individuals or generations (see Ackerman et al. (2009)). A pure time preference of zero would imply that (holding real incomes constant) benefits and costs to future generations are just as important as the gains and losses that we experience today. The higher the rate of pure time preference, the less we value harm to future generations from climate change and the less we value the benefits that we can confer on future generations by averting climate change.

The degree to which the marginal utility of consumption diminishes is controlled by the elasticity of marginal utility with respect to consumption per capita, η, a function of per capita consumption and the marginal utility of consumption.

[4] See also Arrow et al. (1996) and Stern (2006:Ch.2)

2) $$\eta = \frac{-\,c/u'(c)}{\delta c/\delta u'(c)}$$

Both the rate at which the marginal utility of consumption falls and the inputs to the production function that determine the projected growth in per capita income are chosen by the modeler. In most IAMs per capita income is assumed to grow at something like its long-term historical trend; of course, projecting future growth requires accurate forecasting of both global output and the change in global population over time – neither of which can be known with much confidence over the next 50 years, much less over the multi-century time-scale of many climate-economics models.

The existence of a diminishing marginal utility of consumption – such that the first dollar of consumption adds more to our welfare than the second dollar, and the second more than third – is a standard assumption in economic analysis. There is, however, no standard value, or even a range of standard values, for the degree to which utility diminishes as consumption grows. There is no standard value because conventional economic theory maintains that utility cannot be measured (the history of this theoretical notion and its obvious contradiction in applied economic models that measure utility, like IAMs, are discussed in a subsequent section).

Regionally disaggregated IAMs

In models with more than one region, the IAM maximizes the sum of regional utilities. Without Negishi weights, IAMs treat gains in consumption as having a greater contribution to welfare in a poor region than in a rich region. Because of the diminishing marginal utility of consumption, the welfare contribution of an abatement or damage cost depends on the per capita income in the region where it occurs. An unconstrained model will allocate income gains to poorer regions (where each dollar has the biggest impact on welfare) and will allocate income losses to richer regions (where each dollar has the smallest impact on welfare). The result is an equalization of incomes across all regions – a radical redistribution of world income. From scholarly articles describing two leading models:

> *For RICE:* Under the utilitarian weights, regions have different shadow prices of capitals. The social planner can improve the global welfare by moving capitals from low price [richer] places to high price [poorer] places…However, there are no conventional capital flow channels in our setting. The capital from [the richest region] "inundates" the small scale inter-connections of technological transfers. Such flows are redistributions under the pretext of technological transfers. …reasonable and correct magnitude and directions of technological transfers can only be modeled by using the Negishi social welfare weights. (Yang and Nordhaus 2006:738, 731)

> *For MERGE:* A fixed set of Negishi weights defines a so-called Negishi welfare problem, the solving of which corresponds to the maximizing of the global welfare function. MERGE updates iteratively the Negishi weights in solving sequentially the corresponding Negishi welfare problems. The steps to update the Negishi weights are performed until a pareto-optimal equilibrium solution is found. (Kypreos 2005:2723)

Many models discuss Negishi weighting in terms of equalizing the shadow price (or marginal price) of capital instead of using the more transparent language of equalizing the marginal utility of consumption. Because both terms are proportional to the labor-capital ratio, these two goals – equalizing the shadow price of capital or the marginal utility of consumption – have an identical impact on the regional contribution to social welfare of additional consumption.[5]

Modelers have viewed the tendency toward equalization of incomes across regions as a problem, where the solution is to constrain the model to view the marginal utility of income as being the same in every region (in any given time period). A set of "Negishi weights" is included in the regional utility functions such that the weighted contribution to social welfare of one dollar of additional consumption is the same in all regions. Higher weights are assigned to welfare in richer countries while welfare in poorer countries receives lower weights. This procedure obviates the IAMs' equalization of income, preventing any redistribution of income from taking place. The course of action recommended by a Negishi-weighted model takes the current income distribution among regions of the world as immutable.

In essence, climate economists using Negishi-weighted models are trying to have it both ways by embracing the diminishing marginal utility of income when it appears convenient (for intra-generational distribution) but suppressing it when it inconveniently calls for change (in intra-regional distribution). The choice to ignore equity concerns amounts to a lack of transparency in how costs and benefits to different people or generations are weighted in IAMs, along with the use of opaque technical assumptions that anyone without an advanced degree in economics would find ethically dubious.

Negishi weighting explained

Negishi (1972), writing after the Arrow-Debreu (1954) proofs of the existence and optimality of general equilibrium, offered a simpler but closely related proof. His social welfare function is a

[5] Equalizing the marginal product of capital across regions has an identical impact on regional per capita income as equalizing the marginal utility of income across regions, as suggested by a simple example with a Cobb-Douglas production function:

1) $Y = zK^{\alpha}L^{1-\alpha}$

2) $\delta Y/\delta K = \alpha Y/K = \alpha z(L/K)^{1-\alpha}$

3) $\delta U/\delta Y = (Y/L)^{1-\eta}/(1-\eta) = [z(L/K)^{-\alpha}]^{1-\eta}/(1-\eta)$

Here Y is gross income, which is a function of capital (K) and labor (L). Equation 2 is the shadow price or marginal product of capital. In Equation 3, the marginal utility (U) of income has an elasticity with respect to per capita consumption of η. Both the marginal product of capital and the marginal utility of income depend directly on (L/K). Assuming that the total factor productivity (z), the elasticity of capital (α), and the elasticity of marginal utility with respect to per capita consumption (η) are all constant, equalizing either the marginal product of capital or the marginal utility of income would require the addition of a set of weights (by region and time period) that would have the effect of equalizing the labor-capital ratio.

weighted sum of individual utilities, with constant weights that add up to one across the population.[6] Negishi demonstrated that maximizing this social welfare function also maximizes individual utility – that is, everyone is at the maximum level of utility allowed by their initial endowments – *if and only if* the weights are equal to the inverse of individuals' marginal utility of income at that maximum.

From this proof Negishi offered a procedure for constructing the optimal equilibrium point for a given set of endowments (i.e., for a given *status quo ante* distribution of income and wealth): (1) make an arbitrary starting guess at the weights; (2) calculate how far off the budget constraint each individual is; (3) adjust the weights to bring everyone closer to their budget lines; (4) repeat until the process converges. (For a detailed accounting of this procedure see Rutherford (1999) on sequential joint maximization.) The Negishi weighting procedure results in a Pareto-optimal allocation that is compatible with the given initial endowments, essentially freezing the distribution of income and suppressing any tendency for global utility calculations to call for equalizing incomes.[7] In the words of modelers using this approach,[8]

> The Negishi weights are an instrument to account for regional disparities in economic development. They equalize the marginal utility of consumption in each region for each period in order to prevent large capital flows between regions. …although …such capital flows would greatly improve social welfare, without the Negishi weights the problem of climate change would be drowned by the vastly larger problem of underdevelopment. (Keller *et al.* 2003:7)

> In order to apply this approach, Negishi weights must be determined for each of the regions. These are chosen so that each region's outlays do not exceed the value of its wealth. In equilibrium, the weight must equal each region's share of the world's wealth…The Negishi weights may be interpreted as each region's dollar voting rights in the allocation of the world's resources. (Manne 1999:393, 394)

In less technical language, if consumption is assumed to have diminishing returns to utility, the only way to achieve the Negishi result – such that a dollar has the same impact on utility regardless of the region's income per capita – is to weigh the welfare of richer regions more heavily than that of poorer regions. In climate-economics models, the Negishi procedure works like this: First, the regionally disaggregated welfare optimizing IAM is run without weights and without any possibility of trade or transfers between regions (an autarkic run). Then for each time period, a set of weights is created that when multiplied by the marginal utility of consumption gives an equal product across all regions (that is, the weights are the inverse of each region's marginal utility of consumption). These weights are then included as fixed values to each region's utility function by time period, and the model is run a second time with these fixed weights in place. In the Negishi-weighted model, every region has the same effective marginal utility of income (in any one time period), which means that a dollar gained or lost is

[6] Negishi's social welfare function is defined over individual production and consumption choices, not public goods or public choices, and hence is not comparable to Arrow's (1950) better-known discussion of the impossibility of a useful social welfare function.
[7] Pareto optimality requires a solution in which no one can be made better off without making someone else worse off. This concept is explained in greater detail below.
[8] See also Nordhaus and Yang (1996:746).

treated just the same regardless of the region. Dollars in different time periods, however, are still treated differently; not only is the future discounted, but in addition, the Negishi weights need not be constant over time.

In short, in Negishi-weighted models human welfare is more valuable in richer regions and redistribution of income among world regions is suppressed, maintaining the existing income distribution. According to model descriptions, these assumptions have a large impact on the optimal results or course of action recommended by the model. When such important modeling choices are dryly described in technical terms – equalizing the shadow price of capital – and not terms of the equity implication and the full meaning of the final results, something is lost.

The dysfunctional relationship between maximizing social welfare, diminishing marginal returns to consumption, and reducing in income inequality has a long history in the field of economics, both theoretical and applied. Only by viewing Negishi-weights in this historical context is it possible to reconcile modelers' conflicting assumptions and to understand the ethical precepts that are at work inside the black box of economic modeling.

3. Theoretical Welfare Economics: A Short Intellectual History

In conventional, or neo-classical, economics "utility" is a term that has come to mean an individual's mental state of satisfaction, with the proviso that levels of satisfaction or utility cannot be compared across individuals. It is a concept that is simultaneously too broad and too narrow. Almost anything can be seen to have and give utility, albeit with diminishing returns. At the same time, the standard theoretical treatment of the "utility" concept suffers from severe limitations. In the absence of inter-personal comparability, the utility of individuals cannot be aggregated in order to consider social welfare, nor can it be compared in order to consider distribution. While a theory of well-being that can address neither aggregate social welfare nor inequality seems of little practical or conceptual use, this modern definition of utility has nonetheless been the dominant measure of human welfare used in much of economic theory since the 1930s.

A prescription for income redistribution

Jeremy Bentham's (1970 [1789]:Ch.1) *Introduction to the Principles of Morals* was not the first, but may be the best remembered discussion of the philosophy of Utilitarianism, in which human behavior is described as motivated by pleasure and pain – their net satisfaction being "utility." Society's well-being was the sum of these utilities, such that an ethical course of action was that which led to "the greatest happiness for the greatest number." In theory, utility could be summed across individuals to determine "social welfare," but utilitarianism did not offer any practical way to actually measure either individual or societal well-being.

The most direct antecedents of today's neo-classical economists were called the Material or Marginalist Welfare School; writing roughly a century after Bentham, these theorists preserved the basic precepts of Utilitarianism, but used new mathematical tools to make their arguments. At the center of their economic theory were two related ideas: First, that the goal of individuals was to maximize utility, and, second, that utility was concave, or diminishing on the margin.[9]

Following the work of Arthur Cecil Pigou, the marginalists restricted their analysis to the necessities of life, using money as a "measuring stick." Focusing on the most material aspects of welfare led to the insight that additional income was more useful to the poor than the rich. Many of the marginalists were explicitly in favor of income redistribution because it would lead to more material wants being satisfied. As economics' most famous author of undergraduate textbooks, Paul Samuelson (1956:12), noted:

> Because Marshall, Edgeworth, Walras, Wicksell, Böhm-Bawerk, and the others thought that men were much alike and subject to interpersonally summable diminishing marginal utility, they all tended to regard existing capitalistic society as too unequal in its income distribution. They felt that only after there takes place a redistribution of the initial wealth could one regard the dollars voted in the market place as being of ethically equal weight; only then would the invisible hand of perfectly competitive markets lead to the social optimum.

In short, the economic thought of the early 20th century assumed that: 1) individual utilities could be summed together as a measure of social welfare; 2) maximizing total social welfare was viewed as a societal objective; and 3) diminishing marginal utility of income gave a strong formal justification for income redistribution. This is the "problem" that regionally disaggregated welfare optimizing IAMs – which attempt to maximize social welfare measured as the sum of regional welfares and assume a diminishing marginal utility of income – "solve" by using Negishi weights.

The Ordinalist Revolution

In 1932, British economist Lionel Robbins called for the rejection of all interpersonal comparisons of utility, arguing that cardinal measurement and interpersonal comparisons could never capture the unobservable utility or satisfaction of others; therefore, in his view, it could not be demonstrated or assumed that the marginal utility of income for the poor is greater than the marginal utility of income for the rich (Robbins 1984 [1932]). The success of Robbins' rejection of cardinal measures of utility led to the so-called "Ordinalist Revolution" in economics, and the birth of neo-classical economics as we know it today. According to ordinalist theory – well preserved in modern microeconomics – utility can only be classified according to an ordinal numbering system, not a cardinal one: that is, utility can be compared only in terms of a rank ordering of preferences and never in terms of any absolute scale.

Robbins noticed that if one were to combine the utilitarian concept of social welfare (defined as the sum of individual welfares) with another important marginalist assumption, the diminishing marginal utility of income, the logical outcome is a very subversive result: Social welfare reaches its maximum

[9] Versions of these ideas were published independently by Willam Stanley Jevons, Carl Menger, Leon Walras, and Alfred Marshall starting in the 1870s. For discussion and analysis of the marginalist revolution in economics see Ackerman (1997) and Cooter and Rappoport (1984)

when income is distributed equally across the population. If income were unequally distributed and welfare were concave in income, you could always increase social welfare by redistributing some income from the rich to the poor. Ian Little (1955:11-15) elaborated on Robbins critique arguing that individual satisfactions cannot be summed up, that satisfaction is never comparable among different individuals, and that the field of welfare economics up until that time had been – to its detriment – entirely normative.

The utilitarian definition of social welfare was gradually replaced in welfare economics by the idea of "Pareto optimality." In the concept of Pareto optimality, individual welfare is still utility, but social welfare is defined by the absence or presence of Pareto optimality (a situation in which no one can be made better off without making someone else worse off). This is a somewhat empty concept of social welfare since a wide array of situations can be Pareto optimal, and the only real opportunities for "Pareto improvements" – when someone is made better off while no one is made worse off – occur when there are unclaimed or wasted resources.[10] In *On Ethics and Economics*, Nobel Laureate Amartya Sen (1987:33-34) calls this redefinition of social well-being the narrowing of welfare economics: "In the small box to which welfare economics got confined, with Pareto optimality as the only criterion of judgement, and self-seeking behaviour as the only basis of economic choice, the scope for saying something interesting in welfare economics became exceedingly small."

A Pareto optimal outcome can be achieved on the basis of any initial allocation of income, however equal or unequal.[11] Negishi weighting accepts the current distribution of income and the Pareto optimal outcome that arises from it. One of the most important conceptual limitations of neo-classical economics is its treatment of the distribution of endowments – who owns what – as "given" (or as off-limits to economic analysis), a practice that damages this approach's effectiveness in applied analysis, whether of modern day issues like climate change and the global distribution of income, or of countless historical issues with distributional consequences. Outside of the narrow confines of mainstream economic thought, the distribution of a global atmospheric commons, or of income and wealth among the countries of the world and the residents of those countries, is very much on the table: in international negotiations regarding climate and development; in the diplomatic agendas of the nations of the Global South; and in the studies of environment and development law, policy, and philosophy (Botzen *et al.* 2008; Heyward 2007; Klinsky and Dowlatabadi 2009; Ott *et al.* 2008).

The innovations of Robbins, Little, and Pareto avoid advocacy of income redistribution by 1) rejecting the summability of individual utilities into an aggregate social welfare and 2) rejecting the objective of maximizing total social welfare in favor of the more restrictive aim of changes that leave some better off while leaving none worse off ("Pareto improvement"). Negishi's method can be thought of as a mathematical elaboration of the idea of Pareto improvement, since the solution to his system is that each region attains the maximum utility or welfare consistent with a given initial allocation of resources. Since the Negishi solution is a Pareto optimal outcome, no further Pareto improvement is possible – by this definition, social welfare cannot be improved by income redistribution.

[10] See Chapter 2 in Ackerman and Heinzerling (2004) for a more complete discussion.
[11] Note that the criterion of Pareto improvement prohibits any form of redistribution, which must necessarily leave the rich worse off.

Applied welfare economics and welfare optimizing IAMs

Pareto optimality is the core of theoretical neo-classical welfare economics. The applied economics of social welfare takes the form of cost/benefit analysis (CBA) – of which welfare optimizing climate-economics models form a subset. CBA is a common tool for making decisions about whether a project will improve social welfare (and therefore should be carried out) or will reduce social welfare (and should not be carried out). According to CBA, if the value of the future stream of costs and benefits of a project is positive, we should carry out the project, but if the value is negative we should not carry out the project. Abstracting from the vexing question of discount rates, this means that any addition to the size of the "economic pie" is good, regardless of the distribution of costs and benefits (a definition that allows changes that improve the welfare of some while diminishing that of others to qualify as social welfare improvements).

Connecting CBA back to ordinalist welfare economics takes a blind eye and a few, difficult to justify, conceptual leaps. The compensation test, introduced by Nicholas Kaldor (1939) and John Hicks (1940), is a method for determining whether or not there has been a *potential* Pareto improvement such that those who receive net benefits (the winners) could in principle compensate those who bear net costs (the losers) and still be better off. When benefits net of costs are positive, if one group gets all of the benefits but has to pay back everyone who suffers costs, the first group could potentially pay all of the losers and still have a positive benefit left for itself. Of course, this fails to bring solace to the losers unless they are compensated in practice. As Sen (2000:947) put it:

> If compensations are actually paid, then of course we do not need the comparison criterion since the actual outcome already includes the paid compensations and can be judged without reference to compensation tests…On the other hand if compensations are not paid, it is not at all clear in what sense it can be said that this is a social improvement…The compensation tests are either redundant or unconvincing.

The criterion of potential Pareto improvement (or the compensation test) runs sharply counter to that of Pareto optimality, since it embraces changes that make some better off while making others worse off. In this respect, it is very similar to Utilitarian social welfare as the sum of all individual welfares, but with one key difference: money is summed rather than utility. Thus in applied welfare economics, inter-personal comparability re-enters through the back door (income is summed), but the diminishing marginal utility of income drops out of sight (income is treated as equivalent to utility).

Negishi-weighted models have an inherent contradiction. The diminishing marginal utility of consumption is in play inter-temporally, but not inter-regionally. Any model that includes diminishing marginal returns to social welfare from consumption and maximizes the sum of multiple utility functions (whether by time period or by region) invites redistribution to equalize incomes. Across time periods, if the future is assumed to be richer than the past, this redistribution takes the form of prioritizing increases to the consumption of the present generation. Across regions, this redistribution would take the form of prioritizing increases to consumption for poorer regions. In these models, the prioritization of current consumption is permitted, but the prioritization of increased consumption in poor regions is suppressed by the Negishi weights.

4. Policy Implications and Alternatives

The presentation of Negishi-weighted results without a clear explanation of their equity implications introduces an unknown and unexamined political bias into the policy debate. IAM results are taken as impartial truth, instead of one position in the current ethical debate regarding who should pay for world-wide emissions reductions and adaptation measures (Baer *et al.* 2007; Baer *et al.* 2008).

Welfare optimizing climate-economics models ignore equity in three ways. First, welfare optimizing IAMs abstract over groups, winners, and losers to focus on a global result. Even regionally disaggregated welfare optimizing IAMs employ this logic: their decision rule maximizes the sum of regions' social welfare. Second, welfare-optimizing IAMs commonly employ a very high discount rate, which has the effect of prioritizing short-term savings (avoided abatement costs) over long-term savings (avoided climate damages). Third, welfare optimizing IAMs differ from many other forms of cost-benefit analysis in that they optimize a cardinal measure of utility, modeled with diminishing returns, instead of optimizing dollars. This means that any regionally disaggregated welfare optimizing IAM includes both a social welfare function defined as the sum of regional welfares and measured in cardinal units, and the principle of diminishing marginal returns – exactly the theoretical combination against which Robbins and Little cautioned. Negishi weights counteract the model's resultant preference for equalizing incomes across regions, thereby making the implicit assumption that utility in richer regions is more important to global social welfare than utility in poorer regions.

Regionally disaggregated welfare optimizing climate-economics models – like RICE and MERGE – use Negishi weights to prioritize the well-being of rich regions over that of poor regions. In the absence of these weights, the models would recommend equalization of income on a global scale as part of the optimal course of action to maximize social welfare. With the weights, the models return a result that has policy relevance only if decision makers agree that income redistribution is impossible, forbidden, or otherwise impractical.

Clearly, global income redistribution is neither impossible nor forbidden in reality. Richer countries' development assistance to poorer countries in the form of grants, loans, technical assistance, and debt forgiveness is income redistribution, however modest the scale. Equalizing incomes around the world may seem impractical or politically infeasible – the rich having a certain attachment to their higher incomes and greater wealth – but subjective practicality is a poor criterion for the inclusion of a policy mechanism in a climate-economics model. IAMs commonly recommend a scale of abatement measures that is beyond the horizon of today's technology, and warn of climate damages that are outside the bounds of the historical record.

IAMs that employ Negishi weights give results that can only be accurate within a circumscribed universe in which income redistribution is beyond the pale. Technical model descriptions that explain the Negishi process most often use the language of an equalization of the shadow price of capital across regions. It is not uncommon for model results to be presented devoid of any context regarding their modeling assumptions, including any explanation of the Negishi weighting process. It is extremely likely that many if not most end users of these model results – policy makers, NGOs, the public – are unaware that a model's recommended course of action is "optimal" only given the assumptions that rich countries' welfare is prioritized over that of poor countries, and income redistribution cannot take place.

There are four broad solutions to the opacity of Negishi weighting in welfare optimizing IAMs and the bias that this weighting lends to results: continue weighting but make the impact on model results transparent; discontinue weighting thereby allowing income to be redistributed; optimize money instead of utility; or use other kinds of IAMs that avoid the pitfalls of welfare optimization.

1) *Continue weighting but make the impact on model results transparent*: Negishi weighting may be useful in avoiding a result – global income redistribution – that would seem politically infeasible in the Global North. As with any modeling assumptions that have important impacts on results and policy advice (for example, the choice of the discount rate or the rate at which new technology is adopted), end users need to be made aware of the context in which a result is said to be "optimal".[12]

2) *Discontinue weighting thereby allowing income to be redistributed*: Just as Negishi weighted results (coupled with full disclosure regarding modeling assumptions) may seem more politically feasible in the Global North, unweighted results that treat income redistribution as a key component of a viable climate policy may have a greater political trenchancy in the Global South. Presenting paired model results, with and without Negishi weights, would give policy makers the most complete understanding of the interaction of climate damages, climate abatement, and the global distribution of income.

3) *Optimize money instead of utility*: Maximizing the sum of regional gross incomes, modeled without any diminishing returns, is a plausible decision rule for global climate policy; this is the strict cost-benefit analysis approach, which has become almost ubiquitous in environmental decision making. This approach eliminates diminishing returns to income gains both in richer regions and in richer time periods – a more consistent application of the theoretical principles embodied in the Negishi weighting process. If policy makers wish to consider a weighting of income gains by regional income per capita, it would be computationally simple to add a transparent mechanism for equity weighting that could be ratcheted up or down as a policy lever.

4) *Use other kinds of IAMs that avoid the pitfalls of welfare optimization*: In IAMs that do not optimize welfare, assumptions regarding the inter-regional effects of diminishing marginal utility of income are not negated by Negishi weights. In the Stern Review's PAGE model (Hope 2006) – a simulation model – for example, the elasticity of the marginal utility of income is set at 1 (such that utility is the natural logarithm of per capita income), but no radical equalization of per capita income across regions occurs because utility is not maximized. In

[12] For an example of a Negishi-weighted model using more explicit assumption regarding distribution, see Decanio (2009).

IAMs that do not optimize welfare (and, therefore, do not use Negishi welfare weights), inter-regional equity has the potential to be an active, and central, model parameter.

The solutions recommended here cover the gamut from improved transparency of the existing methodology to a wholesale rejection of regionally disaggregated welfare optimizing IAMs. Any one or a combination of these solutions would be a vast improvement to the current *status quo*. Achieving a successful climate policy will be one of the greatest challenges of the 21st century and it is already apparent that equity concerns will be at the center of global climate negotiations (Edmonds *et al.* 2008; Kok *et al.* 2008; Ott *et al.* 2008). In order to achieve political relevance, all climate-economics models must make explicit their equity assumptions.

References

Ackerman, F. (1997). Utility and Welfare I: The History of Economic Thought. Human Well-Being and Economic Goals. F. Ackerman, D. Kiron, N. R. Goodwin, J. M. Harris and K. Gallagher. Washington, D.C., Island Press

Ackerman, F., S. J. DeCanio, R. B. Howarth and K. Sheeran (2009). "Limitations of Integrated Assessment Models of Climate Change." Climactic Change: forthcoming.

Ackerman, F. and L. Heinzerling (2004). Priceless: On Knowing the Price of Everything and the Value of Nothing. New York, The New Press.

Arrow, K. J. (1950). "A Difficulty in the Concept of Social Welfare." Journal of Political Economy 58(4): 328-346.

Arrow, K. J., W. R. Cline, K.-G. Maler, M. Munasinghe, R. Squitieri and J. E. Stiglitz (1996). Chapter 4 - Intertemporal Equity, Discounting, and Economic Efficiency. Climate Change 1995 - Economic and Social Dimensions of Climate Change. Contribution of Working Group III to the Second Assessment Report of the IPCC. J. P. Bruce, H. Lee and E. F. Haites. New York, IPCC and Cambridge University Press**:** 125-144.

Arrow, K. J. and G. Debreu (1954). "Existence of an Equilibrium for a Competitive Economy." Econometrica 22(3): 265-290.

Baer, P., T. Athanasiou and S. Kartha (2007). The Right to Development in a Climate Constrained World. Berlin, Heinrich Böll Foundation

Baer, P., T. Athanasiou, S. Kartha and E. Kemp-Benedict (2008). The Greenhouse Development Rights Framework: The right to development in a climate constrained world. Berlin, Heinrich Böll Foundation.

Bentham, J. (1970 [1789]). Introduction to the Principles of Morals. London, Athlone.

Bosetti, V., C. Carraro and M. Galeotti (2006). "The Dynamics of Carbon and Energy Intensity in a Model of Endogenous Technical Change." Energy Journal Special Issue on Endogenous Technological Change and the Economics of Atmospheric Stabilisation: 191-206.

Botzen, W. J. W., J. M. Gowdy and J. C. J. M. Van Den Bergh (2008). "Cumulative CO2 emissions: shifting international responsibilities for climate debt." Climate Policy 8: 569-576.

Cooter, R. and P. Rappoport (1984). "Were the Ordinalists Wrong About Welfare Economics?" Journal of Economic Literature XXII: 507-530.

DeCanio, S. J. (2009). "The political economy of global carbon emissions reductions." Ecological Economics 68(3): 915-924.

Edenhofer, O., K. Lessmann and N. Bauer (2006). "Mitigation Strategies and Costs of Climate Protection: The Effects of ETC in the Hybrid Model MIND." Energy Journal Special Issue on Endogenous Technological Change and the Economics of Atmospheric Stabilisation: 207-222.

Edmonds, J., L. Clarke, J. Lurz and M. Wise (2008). "Stabilizing CO2 concentrations with incomplete international cooperation." Climate Policy 8: 355-376.

Gerlagh, R. (2006). "ITC in a Global Growth-Climate Model with CCS: The Value of Induced Technical Change for Climate Stabilization." Energy Journal Special Issue on Endogenous Technological Change and the Economics of Atmospheric Stabilisation: 223-240.

Heyward, M. (2007). "Equity and international climate change negotiations: a matter of perspective." Climate Policy 7: 518-534.

Hicks, J. R. (1940). " The valuation of the social income." Economica X: 105-124.
Hope, C. (2006). "The Marginal Impact of CO2 from PAGE2002: An Integrated Assessment Model Incorporating the IPCC's Five Reasons for Concern." Integrated Assessment Journal 6(1): 19-56.
Kaldor, N. (1939). "Welfare propositions of economics and interpersonal comparisons of utility." The Economic Journal 49: 549-552.
Keller, K., Z. Yang, M. Hall and D. F. Bradford (2003). "Carbon Dioxide Sequestration: When And How Much?" Working Paper 94 Center for Economic Policy Studies, Princeton University
Klinsky, S. and H. Dowlatabadi (2009). "Conceptualizations of justice in climate policy." Climate Policy 9: 88-108.
Kok, M., B. Metz, J. Verhagen and S. van Rooijen (2008). "Integrating Development and Climate Policies: National and International Benefits." Climate Policy 8: 103-118.
Kurosawa, A. (2004). "Carbon concentration target and technological choice." Energy Economics 26: 675-684.
Kypreos, S. (2005). "Modeling experience curves in MERGE (model for evaluating regional and global effects)." Energy 30(14): 2721-2737.
Little, I. (1955). A Critique of Welfare Economics. Oxford, Clarendon Press.
Manne, A. S. (1999). Greenhouse gas abatement: Toward Pareto-optimality in integrated assessment. Education in a Research University. K. J. Arrow, R. W. Cottle, B. C. Eaves and I. Olkin, Springer Netherlands.
Manne, A. S. and R. G. Richels (2004). "MERGE: An Integrated Assessment Model for Global Climate Change." http://www.standford.edu/group/MERGE/
Masui, T., T. Hanaoka, S. Hikita and M. Kainuma (2006). "Assessment of CO2 Reductions and Economic Impacts Considering Energy-Saving Investments." Energy Journal Special Issue on Endogenous Technological Change and the Economics of Atmospheric Stabilisation: 175-190.
Negishi, T. (1972). General Equilibrium Theory and International Trade. Amsterdam-London, North-Holland Publishing Company.
Nordhaus, W. D. (2008). A Question of Balance: Economic Modeling of Global Warming New Haven, Yale University Press.
Nordhaus, W. D. and Z. Yang (1996). "A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies." American Economic Review 86(4): 741-765.
Ott, H. E., W. Sterk and R. Watanabe (2008). "The Bali roadmap: new horizons for global climate policy." Climate Policy 8: 91-95.
Peck, S. C. and T. J. Teisberg (1999). "CO2 Emissions Control Agreements: Incentives for Regional Participation." Energy Journal Special Issue on the Costs of the Kyoto Protocol: A Multi-Model Evaluation: 367-390.
Popp, D. (2006). "Comparison of Climate Policies in the ENTICE-BR Model." Energy Journal Special Issue on Endogenous Technological Change and the Economics of Atmospheric Stabilisation: 163-174.
Ramsey, F. P. (1928). "A mathematical theory of saving." The Economic Journal 138(152): 543-59.
Robbins, L. (1984 [1932]). Essay on the Nature and Significance of Economic Science,. Houndmills, UK, Palgrave Macmillan.
Rutherford, T. F. (1999). Sequential joint maximization. Energy and Environmental Policy Modeling. J. Weyant. New York, Springer: 139-156.

Samuelson, P. (1956). "Social Indifference Curves." Quarterly Journal of Economics 70(1): 1-22.
Sen, A. (1987). On Ethics and Economics. New York, Blackwell.
Sen, A. (2000). "The discipline of cost-benefit analysis." The Journal of Legal Studies 29(2): 931-952.
Stanton, E. A., F. Ackerman and S. Kartha (2009). "Inside the Integrated Assessment Models: Four Issues in Climate Economics." Climate and Development 1.2.
Stern, N. (2006). The Stern Review: The Economics of Climate Change. London, HM Treasury
Tol, R. S. J. (1999). "Kyoto, Efficiency, and Cost Effectiveness: Applications of FUND." Energy Journal Special Issue on the Costs of the Kyoto Protocol: A Multi-Model Evaluation: 131-156.
Yang, Z. and W. D. Nordhaus (2006). "Magnitude and direction of technological transfers for mitigating GHG emissions." Energy Economics 28: 730-741.