WORKING PAPER

8/2012

(Revised February 2016)

Avoiding path dependence of distributional weights

Lessons from climate change economic assessment

Disa Thureson

Economics

ISSN 1403-0586

https://www.oru.se/institutioner/handelshogskolan/forskning/working-papers/
Örebro University School of Business
701 82 Örebro
SWEDEN

Abstract

In some cost benefit analysis (CBA) applications, such as those used for the valuation of climate change damage, distributional weights are used to account for diminishing utility of marginal income. This is usually done by means of intra-temporal distributional weights, which are combined with discounting to account for inter-temporal equity and efficiency. Here, I show that this approach might introduce some inconsistencies in terms of path dependence. In short, this inconsistency means that regional economic growth is double counted. This is because income weighting is performed both through the discount rate and through the distributional weights such that growth shows up twice in the weighting process. Using the PAGE2002 model, it is found that the inconsistency problem in the original model erases the influence of distributional weights on the social cost of carbon dioxide ($SCCO_2$) compared to a standard CBA approach. The alternative approaches proposed here yield about 20%–40% higher values of $SCCO_2$ than the old approach. While this has been briefly commented on in previous work, it has not yet been more thoroughly analyzed nor communicated to the broader community of climate policy and economic analysts who are not deeply interested in the specifications of the climate impact assessment models.

Keywords: Distributional weights; Equity weights; Discounting; Cost benefit analysis; Marginal utility; Integrated assessment model; PAGE2002; Social cost of carbon; Climate change

JEL codes: C69; H23; H43; Q54

1 INTRODUCTION

In some cost benefit analysis (CBA) applications, such as those used for the valuation of climate change damages, weights are used to account for the diminishing utility of marginal income. This is usually done by means of intra-temporal distributional weights, which are combined with discounting to account for both inter-temporal equity and efficiency. Previously, two of the leading models used for valuation of climate change mitigation suffered from a serious specification problem with respect to these issues, which I will theoretically show in this essay.

As a numerical illustration, I have used the PAGE2002 model, which is an integrated assessment model (IAM) constructed for CBA of climate change, e.g., for the estimation of $SCCO_2$. In this model, equity is accounted for in the following manner. First, distributional weights are applied for each region based on the mean per capita gross domestic product (GDP) in each time period, and these are applied independently of other time periods so that the inter-temporal equity can be ignored at this stage. Inter-temporal equity is instead accounted for in the next step by discounting using the Ramsey rule. This is done for each region separately based on each region's expected economic growth path. Anthoff et al (2009) noted that the FUND model (another IAM) previously used a method similar to PAGE2002 and that this method was incorrect. They introduced a new method to correct for the previous problems, but they did not show why the method would leads to inconsistencies. In my study I show why an inconsistency problem may arise for a more generalized model, and then I derive a general solution to the problem. The same theoretical solution has been found independently in a working paper by Richard Tol, (2015). However, in that work, focus was only at the new method, and not at all the previous inconsistency problem. Also, in the present analysis I use a more general utility function as a base, from which I derive the results, hence, the common results from Tol (2015) are based on a plausible special case of this more general specification. In addition, I use another model for the numerical results; the PAGE2002 model, instead of the FUND model. Moreover, in a recent, peer-reviewed article by Tol et al., the base case specification in the FUND model, (see Waldhoff et al., 2014), is still flawed, even though in a slightly different way than before. Although denoted as using no equity weights, in practice this approach implies regressive distributional weights, as I will show in the numerical part of the present paper.

The analysis presented here consists of two sections. In the theoretical section, I set up a general model to show that a path dependence problem exists and how it comes about, and I also offer a solution to the problem. The general requirement for an inconsistency problem to occur is when intra-temporal distributional weights are applied along with discounting, based on region-specific growth-based discounts rates, in two separate steps. In the numerical section of the analysis, I first provide a very simple example to show intuitively how the basic features of the model laid out in the theoretical section work. I then use the PAGE2002 model to estimate the magnitude of the inconsistency problem.

It is estimated that the inconsistency problem in the original PAGE2002 model eliminates the influence of distributional weights on the $SCCO_2$ compared with a standard CBA approach where equal distributional weights and no region-specific discount rates are used. The proposed alternative methods would result in about 20%–40% higher $SCCO_2$ values than the original model.

This paper continues with the background (*Section 2*), where some basic foundations upon which the paper is built are summarized. Previous literature is briefly reviewed along with a short introduction to the PAGE2002 model. The paper then continues with the theoretical part (*Section 3*) and the numerical part (*Section 4*). Finally, some concluding remarks are provided in *Section 5*.

2 BACKGROUND

2.1 Social welfare calculations

In CBAs, costs and benefits are aggregated. It is often assumed that the total welfare in society can be expressed as a function of all individuals' utilities, which in turn are functions of their consumption:

$$W = W(U_1(c_1), U_2(c_2), \dots, U_n(c_n)), \tag{1}$$

where W is the total welfare in society, U_i is the utility of an individual (or group[1]) i, and c_i is the consumption of individual (or group) i. Consumption is, in turn, a function of individual income:

$$c_i = c_i(y_i). \tag{2}$$

Then eq. (1) can be rewritten as:

$$W = W\left(\widehat{U}_1(y_1), \widehat{U}_2(y_2), \dots, \widehat{U}_n(y_n)\right), \tag{3}$$

where $\widehat{U}_i(y_i) = U_i\left(c_i(y_i)\right)$. This means that the marginal change in total welfare in society from an increase in income for individual i can be expressed as:

$$\beta_i = \frac{\partial W}{\partial y_i} = \frac{\partial W}{\partial U_i} \cdot \frac{\partial U_i}{\partial y_i}. \tag{4}$$

Here, $\frac{\partial W}{\partial U_i}$ refers to the social welfare weight that society assigns to each individual, and $\frac{\partial U_i}{\partial y_i}$ is the marginal utility of consumption. The combined factor β_i will be referred to here as the distributional weight. In standard CBAs, the

[1] Note that i can also denote different groups in society with different mean incomes. This study will focus on the case where i denotes different regions of the world.

distributional weights are usually set to 1, due to the Kaldor–Hicks criterion[2], so that:

$$W = \sum y_i \, . \tag{5}$$

For simplicity, this study will rely on the assumption that all individuals have the same utility function and the same individual weight in the welfare function, $\frac{\partial W}{\partial U_i} = 1$, so that $\beta_i = \frac{\partial W}{\partial y_i} = \frac{\partial U_i}{\partial y_i}$.[3]

2.2 Utility function

The utility function of an individual is usually defined in the positive quadrant and as being concave, reflecting the rule of diminishing marginal utility. There are many possible functional forms of utility functions, but in practice the class of hyperbolic absolute risk aversion (HARA) functions is generally used because of the desire for tractability. Within this class, specifically, CRRA (constant relative risk aversion, see below), CARA (constant absolute risk aversion), and quadratic utility all exhibit HARA and are often used. PAGE2002 uses CRRA utility, and therefore this is the specific utility function that will be considered in this article. The CRRA utility function can then be defined as:

$$U(y_i) = \begin{cases} A \cdot \frac{{y_i}^{1-\eta}}{1-\eta} & \text{for } \eta \neq 1 \\ A \cdot \ln y_i & \text{for } \eta = 1 \end{cases} \tag{6}$$

where A is a positive constant and η is the positive elasticity of marginal utility of consumption, EMUC.

The marginal utility from income is then:

$$\beta_i = \frac{\partial U_i}{\partial y_i} = \frac{A}{{y_i}^{\eta}}. \tag{7}$$

Pearce and Nash (1981) set A to be the mean income raised to η in order to normalize the weight on individuals' mean income to 1. Following this procedure, eq. (7) becomes:

$$\beta_i = \frac{\partial U_i}{\partial y_i} = \frac{\bar{y}^{\eta}}{{y_i}^{\eta}}, \tag{8}$$

[2] The Kaldor–Hicks criterion is a measure of economic efficiency that captures some of the intuitive appeal of Pareto efficiency, but it is less stringent and hence applicable to more circumstances. Under Kaldor–Hicks efficiency, an outcome is considered more efficient if a Pareto optimal outcome can be reached by arranging sufficient compensation from those that are made better off to those that are made worse off so that all would end up no worse off than before. In line with the Kaldor-Hicks concept, Harberger (1978) argued not to use distributional weights in the CBA of a specific project, if it possible to find a more efficient redistribution mechanism outside the project.

[3] For a more thorough introduction to welfare calculations and CBA, see Boadway (2006).

where $\bar{y}$ is the mean income. Eq. (5) can also be seen as a special case of eq. (8) where A is 1 and η is 0, reflecting no concavity.

2.3 Discounting

Comparisons of costs and benefits over time require a discount rate, which determines the weight placed on costs and benefits occurring at different times. The welfare in each time period of one representative agent is generally assumed to be:

$$W_t = U\big(c(t)\big) \cdot e^{-\delta t}, \tag{9}$$

(for a continuous specification of time) where δ is the pure rate of time preference (PRTP). With a discrete formulation of time, the formulation is instead:

$$W_t = U\big(c(t)\big) \cdot (1+\delta)^{-t}. \tag{10}$$

It is important to remember that eq. (9) and eq. (10) are not equivalent. In the case of discrete time periods in combination with long time horizons, as in the case of climate change CBAs, the results of the two approaches will typically be a little bit different given the same value δ.[4]
When the analysis is performed on a representative agent basis, it is also common to include the utility function in the discount rate so that eq. (9) transforms into:

$$W_t = c(t) \cdot e^{-rt}, \tag{11}$$

where r is the discount rate, which is generally assumed to depend on the PRTP and on expectations of future incomes. If we believe that future generations will be richer, it is perfectly rational and ethical to place less weight on their marginal income due to diminishing marginal utility. As a result, the discount rate will depend strongly on the expected future economic growth and to what extent we weight benefits and damages dependent on the receivers' income (i.e. the value of η). This study pays special attention to the Ramsey rule for discounting:

$$r = \delta + \eta \cdot g, \tag{12}$$

where r is the discount rate, δ is the PRTP, η is the elasticity of marginal utility, and g is the growth rate in income per capita. The Ramsey rule of discounting can be seen as a simple and effective way to approximate eq. (9) through eq. (11) in the special case of CRRA utility. If the PRTP is set low, the main part of the discount rate will be due to the second term in the Ramsey formula, which accounts for intergenerational equity.

[4] Different calibrations of δ would ultimately be preferred for the two cases of eq. (9) and eq. (10), but the general uncertainty about δ is much higher than this difference would turn out to be.

In this study, it will be assumed that the utility of each region of the world can be represented by a representative agent following eq. (10)[5] and with the same PRTP for all regions so that the total welfare in society can be stated as:

$$W = \sum_t \left(\sum_i U(y_{i,t}) \cdot (1+\delta)^{-t} \right). \tag{13}$$

When both distributional weights and discounting are used, the most common specification of a marginal change in total welfare in CBA practice is:

$$\Delta W = \sum_t \sum_i (\beta_i \cdot d_t \cdot \Delta y_{i,t}), \tag{14}$$

where $d_t = d_t(r, t)$ is the discount factor. In this study, I will call the combined weight from the distributional weight and the discounting on individual (or group) i in period t the total weight:

$$w_{i,t} = \beta_i \cdot d_t. \tag{15}$$

2.4 Climate change

The $SCCO_2$ is the cost of the damage caused by emitting one additional unit of carbon dioxide (CO_2), typically expressed in US dollars per metric ton of CO_2. Ethical considerations about the time preference and distributional weights, of which there is no consensus today, have a great influence on the results.

Since 1982, when the first estimate of the $SCCO_2$ was made (Nordhaus, 1982), hundreds of estimates have been produced by researchers using different IAMs. The essential linkages in all models are from emissions to atmospheric concentration, from concentrations to temperature change, and from temperature change to damages. There has been great progress in the sophistication and comprehensiveness of these models, but there are still shortcomings to tackle and omitted factors to include.

There are many special circumstances that distinguish estimates of the cost of climate change from many other effects considered in standard CBAs. In particular, the Kaldor–Hicks criterion might not be applicable in the case of climate change because there is no functioning authority handling global distributional issues that can enforce optimal global compensation.[6] This provides a basic rationale for the use of distributional weights.

When intra-temporal distributional weights are applied using the formula of marginal utility, this is usually done by dividing the world into several regions. The share of $SCCO_2$ originating from the different regions is then weighted using

[5] For simplicity, the rest of this article will stick to the discrete formulation of time, but all results also hold for the continuous formulation in eq. (9) as long as welfare is evaluated in discrete time periods. For example, the FUND model uses eq. (9), but it still has discrete time periods of one-year intervals.

[6] Alternatively, climate change is not one marginal project among others, so the argument that gains and losses even out in the long run does not apply. Also, Harberger's argument does not apply because there is no institution that can enforce global optimal redistribution.

the mean income in each region. Intra-regional equity is usually ignored. For a thorough discussion (and analysis) of equity weighting in the context of the social cost of greenhouse gases, see Antoff et al (2009).[7]

The development of this area of research has been characterized by deep differences and controversies. The Stern Review (Stern, 2006) triggered much debate by deriving much higher $SCCO_2$ estimates than the mainstream literature at that time, largely due to a different view on the ethics involved in intergenerational trade-offs in combination with a more sophisticated treatment of uncertainty (Dietz et al. 2007). It was argued that the only ethical rationale for the PRTP is the annual risk that the human race will go extinct independently of climate change. This led to a PRTP of 0.1% per year. (Stern, 2006, chs. 2 and 6). In the empirical part of this study, I present results both with the PAGE2002 base case discount rates and with the Stern discount rates.

2.5 The PAGE2002 model

PAGE2002 is a Monte Carlo simulation-based IAM consisting of 86 stochastic parameters describing both climate variables and economic damages. Other parameters are treated deterministically, including population growth, economic growth, and total emissions for each region and time period. Most parameter values are taken directly from the Intergovernmental Panel on Climate Change (IPCC) Third Assessment Report (TAR) (IPCC 2001).
In PAGE2002, income weighting of regions (the world is divided into eight different regions in the model) and generations is performed in an ad hoc way following the method used in Eyre et al. (1999). The model uses the Ramsey rule for discounting and interregional distributional weighting in two separate steps, but with the same value for η. Intraregional distributional weighting is performed for each time period separately using eq. (7). The utility weight for region i, $\beta_{i,T}$ compared to other regions during the same time, T, is:

$$\beta_{i,T} = \left(\frac{\bar{y}_T}{y_{i,T}}\right)^{\eta}, \qquad (16)$$

[7] Negishi weights are often used when it comes to climate change CBAs (see Stanton (2011) for a thorough discussion of this issue), which is a completely different approach compared to the one given by eq. (9). Instead of basing β_i solely on the utility function, β_i is based on money voting power (in short, Negishi welfare represents a market-based rather than a normative model). “The Negishi weighting procedure results in a Pareto-optimal allocation that is compatible with the given initial endowments” (Stanton, 2011, p. 423). This is done by inclusion of regressive weights $\frac{\partial W}{\partial U_i}$ into β_i in order to negate the progressive effect of $\frac{\partial U_i}{\partial y_i}$. The procedure for calculating $\frac{\partial W}{\partial U_i}$ is mathematically rather complicated (see Nordhaus & Yang, 1996). The rationale for Negishi weights is to suppress the redistribution of income in each period. Models that are not just interested in the $SCCO_2$ in a “business as usual” or a “most likely” scenario, but seek to calculate an optimal path, would recommend an equalization of income across regions as part of their policy advice if Negishi weights were not used. The evaluation of Negeshi weights is mostly beyond the scope of this study and will only be mentioned briefly.

where $\bar{y}$ is the mean per capita income in all regions. Discounting is made for each region separately. Because the regions' specific growth rates vary, each region gets an individual discount rate for each time period. From eq. (12) we get the time and location-specific discount rates:

$$r_{i,t} = \delta + \eta \cdot g_{i,t}, \tag{17}$$

where g is the expected per capita growth in GDP approximated by the expected GDP growth minus the regional population growth.
The discount factor is defined for each time period as:

$$d_{i,T} = \prod_{t=T_0}^{T} \left[\frac{1}{1+r_{i,t}}\right]. \tag{18}$$

The total weight from eq. (15) is then:

$$w_{i,T} = \beta_{i,T} \cdot d_{i,T}. \tag{19}$$

Note that the difference from eq. (15) is that β differs between time periods and that d differs between regions. It is also important to note that these weights are based on predetermined growth rates that are assumed to be independent of climate change. For a more thorough specification of the PAGE2002 model, see Hope (2006).

3 THEORY

3.1 The model – a general case

I will now show that the type of weighting implied by eq. (19) involves inconsistencies that are problematic in a long-run CBA. In this section, the model is generalized so that it holds for any utility function with decreasing marginal utility. The weights are then:

$$\beta_{i,T} = \frac{\frac{\partial U_{i,T}}{\partial y_{i,T}}}{\frac{\partial \hat{U}_T}{\partial \hat{y}_T}} = \frac{U'_{i,T}}{\hat{U}'_T}, \tag{20}$$

where $\hat{U}'_T$ is some normalizing measure based on the income distribution for each time period, for example, the marginal utility of mean income as in eq. (8) or of the mean income in a reference region. In addition, the time and region-specific[8] interest rate is generalized to be any increasing function of the time and region-specific growth rates of per capita income:

$$r_{i,t} = r_{i,t}(g_{i,t}). \tag{21}$$

[8] Note that in this example i denotes a region, but the same analogy could be made for individuals (or income groups) with individual discount rates based on increases in their individual income as mentioned in section 2.1.

3.2 The inconsistency

From eq. (19), the geographic utility weight for region i compared to other regions during the same time T is:

$$w_{i,T} = \beta_{i,T} \cdot d_{i,T} = \frac{U'_{i,T}}{\bar{U}'_T} \cdot \prod_{t=T_0}^{T} \left[\frac{1}{1+r_{i,t}(g_{i,t})}\right]. \tag{22}$$

The relative weights between two different regions (h, i) in the same time period are then:

$$\frac{w_{h,T}}{w_{i,T}} = \frac{\frac{U'_{h,T}}{\bar{U}'_T} \cdot \prod_{T_0=0}^{T}\left[\frac{1}{1+r_{h,t}(g_{h,t})}\right]}{\frac{U'_{i,T}}{\bar{U}'_T} \cdot \prod_{T_0=0}^{T}\left[\frac{1}{1+r_{i,t}(g_{i,t})}\right]},$$

which yields:

$$\frac{w_{h,T}}{w_{i,T}} = \frac{U'_{h,T}}{U'_{i,T}} \cdot \prod_{T_0=0}^{T} \left[\frac{1+r_{i,t}(g_{i,t})}{1+r_{h,t}(g_{h,t})}\right]. \tag{23}$$

Because this is the relative weight of two regions in the same time period, we would not expect that discounting would show up in the expression. The reason why it nonetheless shows up (the second factor of eq. (23)) is that different regions have different discount rates. Regions with high economic growth between T_0 and T will also have higher discount rates during this period. If no intra-temporal regional weights were used, this would not yield a problem (as long as initial income differences were accounted for). However, with inter-temporal weights, the differences in economic growth are already adequately accounted for through the first factor in eq. (23). This means that also taking growth into account through the discount rate yields a double counting of the differences in economic growth.

The result is that the relative weights are dependent not just on the expected income at T but also on previous economic growth in the different regions (the ex-ante expected growth leading to that income), which means that the weights are path dependent. I call this inconsistency the *regional weighting inconsistency* and refer to the second factor of eq. (23) as the *regional weighting inconsistency factor*. If the same discount rate was used in all regions, this problem would disappear.

If we instead compare the weights for the same region in two different time periods (T_1, T_2), we would get:

$$\frac{w_{i,T_1}}{w_{i,T_2}} = \frac{\frac{U'_{i,T_1}}{\bar{U}'_{T_1}} \cdot \prod_{t=T_0}^{T_1}\left[1+r_{i,t}(g_{i,t})\right]}{\frac{U'_{i,T_2}}{\bar{U}'_{T_2}} \cdot \prod_{t=T_0}^{T_2}\left[1+r_{i,t}(g_{i,t})\right]},$$

which yields:

$$\frac{w_{i,T_1}}{w_{i,T_2}} = \frac{\frac{U'_{i,T_1}}{\bar{U}'_{T_1}}}{\frac{U'_{i,T_2}}{\bar{U}'_{T_2}}} \cdot \prod_{t=T_1}^{T_2}\left[1+r_{i,t}\left(g_{i,t}\right)\right] . \tag{24}$$

Because this is the relative weight of the same region between two different time periods, we would not expect intra-temporal equity weights (the first factor of eq. (24)) to show up alongside discounting (the second factor of eq. (23)). The first factor is equal to one only by coincidence, which means that the resulting discount rate is generally different from the intended one, the second factor of eq. (24), based on eq. (18). It follows that if one region is growing faster (slower) than other regions, the discount rate for that region will be overestimated (underestimated). Again we see that economic growth is double counted, both through intra-temporal weights and through discounting. I call this inconsistency the *discount inconsistency* and refer to the first factor of eq. (24) as *the discount inconsistency factor*. This is not due to the fact that the discount rate is different in different regions, but to the fact that income weighting is performed in two different steps (discounting (inter-temporal) and regional (intra-temporal)). This means that if no intra-temporal weights were used, this problem would disappear.

In the special case of the PAGE2002 utility function (CRRA) in eq. (16) and the interest rate (Ramsey rule of discounting) in eq. (17), the equations (22), (23), and (24) become the following:

$$w_{i,T} = \frac{U'_{i,T}}{\bar{U}'_{T}} \cdot \prod_{t=0}^{T}\left[\frac{1}{1+r_{i,t}\left(g_{i,t}\right)}\right] = \left(\frac{\bar{y}_T}{y_{i,T}}\right)^{\eta} \cdot \prod_{t=0}^{T}\left[\frac{1}{1+\delta+\eta\cdot g_{i,t}}\right], \tag{25}$$

(25)

$$\frac{w_{h,T}}{w_{i,T}} = \frac{U'_{h,T}}{U'_{i,T}} \cdot \prod_{t=0}^{T}\left[\frac{1+r_{i,t}\left(g_{i,t}\right)}{1+r_{h,t}\left(g_{h,t}\right)}\right] = \left(\frac{y_{i,T}}{y_{h,T}}\right)^{\eta} \cdot \prod_{t=0}^{T}\left[\frac{1+\delta+\eta\cdot g_{i,t}}{1+\delta+\eta\cdot g_{h,t}}\right], \tag{26}$$

where the second factor is the *regional weighting inconsistency factor*, and

$$\frac{w_{i,T1}}{w_{i,T2}} = \frac{\frac{U'_{i,T_1}}{\bar{U}'_{T_1}}}{\frac{U'_{i,T_2}}{\bar{U}'_{T_2}}} \cdot \prod_{t=T_1}^{T_2}\left[1+r_{i,t}\left(g_{i,t}\right)\right] = \left(\frac{\frac{\bar{y}_{T_1}}{y_{i,T_1}}}{\frac{\bar{y}_{T_2}}{y_{i,T_2}}}\right)^{\eta} \cdot \prod_{t=T_1}^{T_2}\left[1+\delta+\eta\cdot g_{i,t}\right], \tag{27}$$

where the first factor is the *discount inconsistency factor*.

3.3 Two alternative approaches

Subsection 3.2 suggests that region-specific discount rates should not be used in combination with regional distributional weights. If regional distributional weights are not used, then region-specific discount rates can be used instead of regional intra-temporal weights if the differences in initial incomes are compensated for. The specification could, for example, be:

$$w_{i,T} = \frac{U'_{i,T_0}}{\hat{U}'_{T_0}} \cdot \prod_{t=T_0}^{T} \left[\frac{1}{1+r_{i,t}(g_{i,t})}\right]. \quad (28)$$

I call this specification *compensated discounting* (the term refers to regional discounting that is compensated for the initial difference in incomes), which leads to the following equation replacing eq. (23):

$$\frac{w_{h,T}}{w_{i,T}} = \frac{U'_{h,T_0}}{U'_{i,T_0}} \cdot \prod_{t=T_0}^{T} \left[\frac{1+r_{i,t}(g_{i,t})}{1+r_{h,t}(g_{h,t})}\right] \quad . \quad (29)$$

The relative weights depend on income differences where the first factor takes initial income differences into account and the second factor takes differences in growth rates into account.

The following equation now replaces eq. (24):

$$\frac{w_{i,T_1}}{w_{i,T_2}} = \prod_{t=T_1}^{T_2} \left[1 + r_{i,t}(g_{i,t})\right]. \quad (30)$$

The first factor in eq. (22) has now disappeared, leaving us only with the region-specific discount rate, so the expression now works as intended. The *Compensated discounting* in eq. (28) thus solves the inconsistency problems laid out in eq. (21) and eq. (22). But it is important to keep in mind that this method is only as good as the discount rate formula is in mimicking the assumed utility function in each specific case. In the case of the CRRA utility and the Ramsey rule of discounting, there is an approximation error in this transformation (see *Technical annex*).

I will now derive a more explicit solution to the problem that does not take the indirect way of discounting to account for intergenerational equity. In this solution the total weights only depend on actual income in different regions during different time periods and PRTP, and thus they are implicitly independent of the growth leading to that particular income. The assumption is that eq. (12) holds, which gives:

$$\frac{\partial W}{\partial y_{i,t}} = U'(y_{i,t}) \cdot (1+\delta)^{-t}. \quad (31)$$

The weight is then normalized to 1 for some income measure in the first time period, $\hat{y}_{T_0}$. This normalization gives the following total weights[9]:

$$w_{i,T} = \frac{U'_{i,T}}{\hat{U}'_{T_0}} \cdot (1+\delta)^{-T}. \quad (32)$$

This formulation can in turn be divided into one income component

[9] This approach is a generalized version of the approach employed in Anthoff et al (2009).

$$\beta_{i,T} = \frac{U'_{i,T}}{\widehat{U}'_{T_0}}, \tag{33}$$

and one PRTP discount component

$$d_{i,T} = (1+\delta)^{-T}. \tag{34}$$

Note that $\beta_{i,T}$ now not only represents the weight between different regions in the same time period, but it also incorporates the share of discounting that is due to income increases (or decreases). This means that growth should not be incorporated in the discount rate in eq. (17). Instead, the discount rate is specified as $r_{i,t} = \delta$.

I call this new specification of the discounting and distributional weights *ideal weighting*. Eq. (32) leads to the following equation replacing eq. (17):

$$\frac{w_{h,T}}{w_{i,T}} = \frac{U'_{h,T}}{U'_{i,T}}. \tag{35}$$

This expression only depends on actual income differences between the two regions. The following equation now replaces eq. (23):

$$\frac{w_{i,T_1}}{w_{i,T_2}} = \frac{U'_{i,T_1}}{U'_{i,T_2}} \cdot (1+\delta)^{T_1-T_2}. \tag{36}$$

The first factor now accounts for growth-based discounting and the second factor accounts for the PRTP. In other words, discounting still works as the Ramsey rule of discounting, but now the marginal incomes of different income groups are weighted in the same way no matter if the income differences are due to different places of birth or different times of birth.

The two approaches of *compensated discounting* in eq. (28) and *ideal weighting* in eq. (32) are virtually the same if $r_{i,t}(g_{i,t})$ works as a good approximation of the utility function. However, using the discount rate to account for decreasing marginal utility will always involve a discretization of the utility function (for a finite number of time periods) and hence always involve some degree of approximation error given that we believe the utility function to be continuous (see *Technical annex* for an example).

In the specific case of the PAGE2002 model, the *ideal weighting* specification is:

$$w_{i,T} = \frac{U'_{i,T}}{\widehat{U}'_{T_0}} = \left(\frac{\frac{1}{y_{i,T}}}{\frac{1}{\bar{y}_{T_0}}}\right)^{\eta} = \left(\frac{\bar{y}_{T_0}}{y_{i,T}}\right)^{\eta} \cdot (1+\delta)^{T_0-T}. \tag{37}$$

To conclude the analytical part of this paper, I have shown that, given the model outlined in sub-section 3.1, economic growth will be double counted when total weights are constructed. Growth will be taken into account both through the intra-temporal regional weights and through discounting. This will lead to inconsistencies regarding the income-based weights, making them not only dependent on income but also dependent on the path leading to that income. The numerical implications of these inconsistencies will be revealed in the next section.

4 NUMERICAL RESULTS

In this section, some implications of *Section 3* will be revealed through two numerical examples. The first one (subsection 4.1) is a simple hypothetical example designed make the results of *Section 3* intuitively easy to follow. In 4.2 the implications of a realistic model, the PAGE2002 model, are presented.

4.1 A simple numerical example

For some basic intuition, I set out a very simple example based on eqs. (25), (26), and (27). For simplicity, an elasticity of marginal utility (η) = 1 and PRTP (δ) = 0 are assumed. Imagine that there are only two regions in the world, *Region A* and *Region B*, where both are the same size but *Region A* is richer in the base year (T_0) but *Region B* has faster economic growth (there is no population growth). Table 1 shows the incomes for the two regions for a base year (T_0) and a year in far in the future (T_1). The mean income ($\bar{y}_T$) for each year is also calculated.

$y_{i,T}$	T_0	T_1
Region *A*	150	150
Region *B*	50	150
$\bar{y}_T$	**100**	**150**

Table 1: Income distribution in two different time periods for the two fictive regions.

From eq. (20) we get:

$$w_{i,T_0} = \frac{\bar{y}_{T_0}}{y_{i,T_0}} \cdot 1, \tag{38}$$

$$w_{i,T_1} = \frac{\bar{y}_{T_1}}{y_{i,T_1}} \cdot \frac{y_{i,T_0}}{y_{i,T_1}}, \tag{39}$$

where the first factor in each expression is the regional distributional weight and the second factor is the discount factor. The weights are given in Table 2.

Weights	β_{i,T_0}	β_{i,T_1}	d_{i,T_0}	d_{i,T_1}	w_{i,T_0}	w_{i,T_1}
Region *A*	0.67	1	1	1	0.67	1
Region *B*	2	1	1	0.33	2	0.33

Table 2: Resulting weights. The first pair of columns denotes regional distributional weights, the second pair of columns denotes the discount factors, and the last pair of columns denotes the total weights.

We can see that the weight in *Region B* in T_1 is strongly underestimated in comparison to *Region A*. This modeling problem is even more serious because the impacts of global warming typically occur in the far future and especially in regions corresponding to *Region B* in the example. There is, therefore, a risk that this way of performing distributional weighting is underestimating the impacts of global warming.

4.2 Analysis in PAGE2002

The effect of the original PAGE2002 weighting (as stated in subsection 2.3) is compared to four other procedures for discounting and distributional weighting. First a reference procedure similar to a standard national CBA is tested (the *No regions* procedure). In this procedure the world is not divided into regions, and there are no regional discounting or intra-temporal weights. Instead, the Ramsey rule of discounting is used with world mean growth rates for each time period. Second, a procedure is tested where no intra-temporal weights are used but region-specific discount rates are used is tested (the *No equity* procedure). This approach will yield regressive total weights as long as the fast-growing regions are poorer than the slow-growing regions in absolute terms (poor countries catching up). This approach is not recommended and is only tested here for the sake of comparison. This comparison is relevant since this is the base case approach in the FUND model, according to Waldhoff et al. (2015).

Last, the two approaches proposed in this study, as stated in subsection 3.3, are tested. All five procedures are run both for the base case parameter values in PAGE2002 for η and δ and for the parameter values used in the Stern review (Stern, 2006). Table 3 shows the input parameter values for the two schemes, and the resulting $SCCO_2$ values are shown in Table 4.

Discount schemes	Base case	Stern
Pure rate of time preference, δ	(0.1%; 1%; 2%)*	0.1%
Elasticity of marginal consumption, η	(0.5; 1; 2)*	1

Table 3: Two different parameter schemes. *Denotes a triangular probability density function with (min; mode; max), where min is the minimum of the values where the density is positive, mode is the value where the density is maximum, and max is the supremum of the values where the density is positive. This is applied in the base case model. In the Stern scheme, δ and η are treated as deterministic.

Weights	PAGE	No regions	No equity	Compensated	Ideal
Base case	28	27	21	38	34
Stern	73	73	57	101	93

Table 4: Mean $SCCO_2$ (\$/tCO_2 in 2010) for five different kinds of weight schemes (the different columns). PAGE denotes the original procedure in PAGE2002. *No regions* denotes a procedure similar to a usual national CBA where equal distributional weights are used and no regionally differentiated discount rates are used. Instead, the Ramsey rule of discounting is used with world mean growth rates for each time period. *No equity* denotes a procedure with region-specific discount rates, but with no intra-temporal distributional weights. *Compensated* denotes a procedure with region-specific discount rates, but where initial income differences have been compensated for, in line with eq. (21). *Ideal* denotes total weights that are only dependent on income and the PRTP, in line with eq. (26). The default time horizon for the calculations in the model is the year 2200, and all values are in year 2000 real prices. Unit: \$/tCO_2.

Although the original PAGE2002 weighting procedure is much more complicated, it yields almost identical results as a procedure with no regional division (similar to a standard national CBA procedure). By coincidence, the extra weight given to poorer regions in PAGE2002 is exactly offset by the lower weight for fast-growing regions. This result will not hold in general though. The results from the *No equity* procedure is that $SCCO_2$ is decreased, which is no big surprise, because this is a regressive approach putting even less weight on poorer, fast-growing economies than a standard no weights approach. Note that this is the base case approach in the FUND model, according to Waldhoff et al. (2015).

In contrast, the two procedures proposed in this study are progressive and yield higher values of $SCCO_2$ than the original PAGE2002 approach and the standard CBA approach (about 20%–40% higher). It is somewhat troublesome that the results from *Compensated* are about 10% higher than those from *Ideal* because the aim of the first method is to approximate the latter. This is because the Ramsey rule of discounting yields a fair approximation of elasticity of marginal utility for shorter time periods, but it might work worse for very long time horizons depending on the discount rate (higher discount rates yield larger approximation errors[10]). From this point of view, it is recommended that *Compensated* is used only with caution and that *Ideal* is preferable.

[10] With mean parameter values (g = 2%, η = 1.17, δ = 1.03%) in PAGE2002, the discount rate is about 3%, which yields an approximation error of about 3% at year 100 (the middle of the 200-year interval). The maximum discount rate is about 10% (g = 4%, η = 2, δ = 2%), which yields an error of about 30% at year 100.

5 CONCLUDING REMARKS

In this paper I have identified inconsistencies that result from the combination of intra-temporal distributional weighting and discounting when the discount rates are region specific and dependent on growth (and the regional weights are period specific). Of the three leading models developed for $SCCO_2$ calculations – DICE (or RICE[11]), FUND, and PAGE – the latter two use distributional weights. Until recently (when Anthoff et al 2009 changed the specification of FUND), both of these models suffered from the inconsistency problems laid out in the present study. Also, even though this original problem has been corrected, the otherwise sophisticated FUND model currently uses an even more regressive approach as base case specification (Waldhoff et al. (2015)); the same as the *No equity* approach in the present paper. In this article I have shown that this approach yields about 40% lower values than *Ideal* weighting in the PAGE2002 model while there is a much larger effect in the FUND model; in Tol (2015) equity weighted global $SCCO_2$ estimates are 3-17 times larger than the no equity weight approach (denoted *Simple sum* in Tables 2-4). It is not clear why the sensitivity in this respect is so much larger in the FUND model compared to the PAGE model, but could be due to a more detailed representation of regions, possibly in combination with different PRTP and EMUC values.

To conclude, if the objective of a policy is to maximize total global utility over generations, it is recommended that $SCCO_2$ estimates from base case specifications in the FUND model are not used, since those estimates are not consistent with such an approach. Instead, it is recommended that only equity weighted estimates from the FUND model are used in such policy evaluation.

I consider the proposed approach to adequately solve the identified inconsistency problem so that the need for further research on this topic is limited. On the other hand, there might be other methods that work as well, and it is possible that the results can be even more generalized to hold true in more situations. For example, an even more explicit method would be to use normalized utility functions directly in the CBAs instead of using weights on marginal incomes.[12] With marginal changes, however, this would yield the same answer as *Ideal* weighting. In any case, there will be a need for a continued broader discussion about equity and about how to use distributional weights in light of climate change. Stanton's article from 2011 was a welcomed contribution to such a discussion.

In an even broader context, this article brings a new issue into the CBA literature, which to my knowledge has not been recognized yet. The issue of combining discounting and distributional weighting correctly has the potential to be important in all long-term decisions where distributional weighting is present.

[11] RICE uses Negishi weights, see Nordhaus & Yang (1996)

[12] PAGE09 uses this approach

REFERENCES

Anthoff, D., Hepburn, C. & Tol, R. S. J. (2009), "Equity weighting and the marginal damage costs of climate change", Ecological Economics, Vol. 68, Pp. 836 – 849

Boadway, R. (2006), "Principles of Cost-Benefit Analysis", *Public Policy Review*, Vol.2, No.1.

Dietz, S., Anderson, D., Stern, N., Taylor, C. & Zenghelis, D. (2007), "Right for the Right Reasons: A final rejoinder on the Stern Review", *World Economics*, Vol. 8, No. 2

Eyre, N., Downing, T., Hoekstra, R. & Rennings, K. , 1999. Global warming damages. Final Report of the ExternE Global Warming Sub-task. DGXII, EC, Brussels. Available at: http://www.externe.info/

Harberger, A. C. (1978), "On the Use of Distributional Weights in Social Cost-Benefit Analysis", *Journal of Political Economy*, Vol. 86, No. 2, 87-120.

Hope, C. (2006), "The Marginal Impact of CO_2 from PAGE2002: An Integrated Assessment Model Incorporating the IPCC's Five Reasons for Concern", *The Integrated Assessment Journal*, Vol. 6, Iss. 1, Pp. 19–56

Hope, C. (2011), *The Social Cost of CO2 from the PAGE09 Model*, Discussion paper, available at: http://www.economics-ejournal.org/economics/discussionpapers/2011-39

Nordhaus, W.D. (1982), "How Fast Should We Graze the Global Commons?", *The American Economic Review*, Vol. 72, No. 2, Pp. 242-246.

Nordhaus, W.D., Yang, Z. (1996), "A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies", *The American Economic Review*, Vol. 86, No. 4, Pp. 741-765

Pearce, D., Nash, C. (1981), "The social appraisal of Projects – a Text in Cost-Benefit Analysis", London: Macmillan Press

Stanton, E. (2011), "Negishi welfare weights in integrated assessment models: the mathematics of global inequality", *Climatic Change*, Vol 107, Pp. 417–432

Stern, N. (2006), The Economics of Climate Change: The Stern Review. Cambridge, UK, Cambridge University Press. Available at: http://www.hm-treasury.gov.uk/stern_review_report.htm

Tol, R.S. J. (2011), *Modified Ramsey Discounting for Climate Change*, ESRI Working Paper No. 368

Tol, R.S. J. (2013), *Modified Ramsey Discounting for Climate Change*, University of Sussex Working Paper Series, No. 63-2013

Tol, R.S. J. (2015), *Modified Ramsey Discounting for Climate Change*, CESIFO Working Paper No. 5588