Discounting and Intragenerational Equity*

Johannes Emmerling†
Toulouse School of Economics

December 19, 2011

Abstract

We study the optimal consumption discount rate taking into account income inequality within a generation between countries. We show that if the dispersion of income decreases over time, the global consumption discount rate should under certain conditions be lower than in the case without inequality. With actual growth predictions used in the context of climate change, we find that the discount rate for short horizons should be almost twice as high than without considering inequality, but decreasing over the time horizon. Moreover, I show that these results also hold qualitatively when we extend the setting allowing inequality, risk, and fluctuation aversion to differ. In particular, empirical evidence suggests that inequality aversion should be lower than risk aversion and fluctuation aversion.

Keywords: Discounting, Inequality, Climate Change, Convergence

JEL Classification: D63, H43, Q54

---

*I would like to thank Christian Gollier for very useful suggestions and general encouragement. I also would like to thank Geir Asheim, Ben Groom, Richard Tol, and participants of the ASSET Meeting 2010, Alicante, the ESOP Workshop "Climate Change and Distribution" in Oslo, and the Exeter Sustainability Forum 2011 for very helpful comments. The usual caveat applies.

†Corresponding Address: Toulouse School of Economics, 21 Allée de Brienne, 31000 Toulouse, France; Phone: +33689467927, E-Mail: mail@johannes-emmerling.de.

1 Introduction

The role of the discount rate has received great attention in recent years with climate change being one of the main driver of this renewed interest. The large time horizons in the discussion about costs and benefits in the context of climate change which make the discount rate of great importance for the policy recommendations derived. Since the publication of the Stern Review in 2006, the debate amongst economists has been further intensified concluding that the discount rate is the most important single factor in determining the magnitude of e.g., the optimal price of carbon. Stern et al. (2006) assumed a rather low discount rate of 1.4 per cent leading to an optimal Social Cost of Carbon (SCC) above 300 dollars per ton of carbon. Many economists argued that this result was mainly driven by the rather low discount rate. Nordhaus (2007) for instance shows that using three to four per cent, the value is reduced to around $30.

That is, the issue of inter-generational equity, or, how we compare today and future generations' well-being is pivotal. If one considers only aggregate world per consumption, as it is done e.g. in the Stern report or Nordhaus' DICE model, this is sufficient. Facing real world's income inequality between countries on the other hand, it seems likely that the evaluation of future versus present consumption differs for different countries. It seems therefore fruitful to consider the inter- and intragenerational distribution jointly in the sense of Schelling (1995). Our paper can therefore be seen as a more formal analysis of his point.

The fundamental theory of discounting shows that the discount rate should reflect the opportunity cost of an investment today with future payoffs. The idea being that any 'project' should be compared to other potential projects.

Basically there are two kinds of alternatives to compare it with: a private investment project or delaying consumption and consuming in the future. If one assumes a given project is likely to crowd out private investment, its return should exceed the so called marginal social opportunity cost of capital (SOC) which is related to the returns of private investment. If one the other hand one sees a given project as basically implying reducing today's consumption in order to increase it in the future, the relevant criterion would be the marginal social rate of time preference (SRTP) which reflects consumers' preference for present consumption. In a perfect economy without distortionary taxation, externalities, informational problems etc., the two should coincide. Moreover, from an international perspective with perfect capital mobility, they should be the same in every country. In the real world, the two will typically differ and there is an ongoing debate which approach should be used in order to compute optimal discount rate–see for instance Zhuang et al. (2007). We follow the consumption-based approach prevailing in the recent debate about climate change.

The classical model dating back to Ramsey (1928) states that the optimal discount rate depends on the pure rate of time preference (PRTP) and today's and tomorrow's marginal utility.Apart from the PRTP, the curvature of the utility on the one hand and the growth of consumption over time on the other are therefore

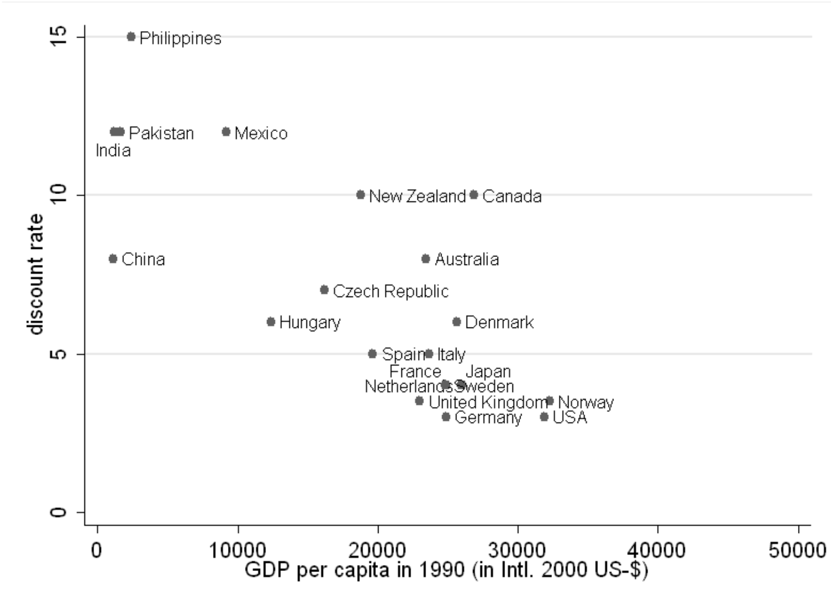

Figure 1: Officially used discount rates and GDP per capita

pivotal for the magnitude of the discount rate. This result can be illustrated using the Ramsey-Keynes condition.

In recent years, there have been several extensions of the classical model in particular looking at the role of uncertainty as in Weitzman (1998) and Gollier (2002). Recent works include the extension of the nature of uncertainty (Weitzman (2009) and Gollier (2008)), ambiguity Gierlinger and Gollier (2008) and the differentiation between different goods (Sterner and Persson (2007) and Gollier (2010b)). Moreover, heterogeneity of time preferences has been discussed based on different degrees of impatience, e.g., see Gollier and Zeckhauser (2005) and Jouini et al. (2010). Note however that in this case, a system of perfect risk sharing and perfect capital markets is assumed. In particular the assumption of risk sharing between agents allows different degrees of time preference to be insured against each other. This is crucially different from the situation considered in this paper where risk sharing and perfect capital markets are absent, a case which seems more realistic in our setting. In particular, this implies that discount rates will be different for different countries.

If one looks at the social discount rates actually applied in different countries, there is indeed a large variability depending on the method used to determine the discount rate but also economic factors such as expected growth. As it can be seen from Figure 1[1], discount rates applied in public decision making vary widely between countries. In particular, poorer countries typically apply considerably higher discount rates than richer countries reflecting to some extent the more frequently used SOC method, but also different growth prospects.

In other words, the question of how much someone is ready to give up today in order to in order to improve the well-being of future generations will be answered

[1]Sources: Spackman (2001), Zhuang et al. (2007) and the CAIT database

differently in richer or poorer countries. As an example, think of a society composed only by one Chinese and one American citizen where the American is nine times richer. Imagine that in the distant future, the income relation will be reversed, that is there is no growth of World GDP. Obviously, the Chinese would not be willing to sacrifice much of his present income knowing to be much poorer in the future while it would be the opposite for the American. Considering only average income, which remains constant, the discount rate would be just equal to the pure rate of time preference. In this particular case, this holds also when considering inequality between countries since inequality is the same at both points in time. However, if inequality changes over time, this is not true anymore. For instance, if income would be equally distributed in the future, the discount rate should be higher than the PRTP since welfare in an inequality averse and prudent society (exhibiting decreasing and convex marginal utility) is higher in the future.

The role of regional heterogeneity, in particular of income inequality, has been pointed out already in Stern et al. (2006, 32). From a development perspective, UNDP (2007) and Conceiçao and Zhang (2010) recently stressed the importance of global inequalities on the discount rate. Whereas the role of uncertainty for the discount rate has been widely studied, less is known about the effect of income differences. Azar and Sterner (1996) is a notable exception that includes discounting and inequality concerns using a stylized two-region model.

This paper provides a theoretical discussion on the optimal discount rate integrating intragenerational inequality considerations. We isolate the effect of inequality and its evolution over time on the socially optimal rate of discount which should be used for projects or initiatives on the regional or global level. We characterize the socially optimal discount rate in a world with wealth inequality and identify the difference to the standard case. We then use various scenarios that have been used in the literature on climate change in order to calibrate the model and derive the magnitude and term structure of the optimal discount rate with wealth inequality. The results suggest a higher but decreasing discount rate as compared to the case without inequality. It is noteworthy that our analysis is analytically similar to the analysis of uncertainty in the same context. It is thus based on the observation that the concepts of inequality and uncertainty can be formalized in an analytically equivalent way as in Atkinson (1970).

2 The inequality adjusted discount rate

The basic equation in consumption-based derivations of the optimal discount rate is the so-called Ramsey-Keynes Condition. One simple way of deriving it is by looking at a marginal project which costs one unit of consumption today and yields $e^{r_t t}$ units of consumption at date t with certainty. If intertemporal welfare is not altered by the project, r_t is the consumption discount rate. Using a CRRA utility function and under certainty, one can derive the basic equation $r_t = \delta + \eta g_c$ where δ is the rate of pure time preference, η the degree of risk aversion and g_c the consumption growth rate.

Now consider a project that it implemented on the global level and the costs and benefits accrue to each country proportional to its population size. We assume that the project in question will affect every individual in a given country equally in that it affects average per-capita consumption. This is certainly a strong assumption and is used here mainly to separate the issue of time preferences from distributional aspects due to different (spatial) distributions of costs and benefits. Moreover, if we were to consider the case of a country with different regions, the assumption could be less problematic.

We consider n countries (or regions but we will refer to them as countries throughout the paper) and denote by c_{it} the per-capita consumption in country i at date t while we do not consider inequality within an given country for the moment. That is, we consider what Bourguignon et al. (2006) call the *international distribution of income* as population weighted inequality between countries' per capita incomes, which today accounts for around two thirds of total world inequality. We denote by π_{it} the share of country i's population of the world population P_t.

Using a standard increasing and concave utility function $U(c)$ and exponential utility discounting with the pure rate of time preference δ–both being the same in each country–, we can express discounted expected utility in country i at time t as $e^{-\delta t}U(c_{it})$.

We use a Social Welfare Function (SWF) of the Benthamite/Utilitarian type weighting each individual equally, that is, we sum each individual's utility to obtain global welfare as $W_t = \sum_{i=1}^{n} \pi_{it} P_t U(c_{it})$. A more general social welfare function $W(U(c_{1t}), .., U(c_{nt}))$ could be used and for specifications typically used, would tend to favor poorer countries even more then the Utilitarian one. Still this specification captures inequality aversion due to the decreasing marginal utility of consumption. Using a Constant Relative Risk Aversion (CRRA) utility specification with a degree of relative risk aversion of $\eta = 2$, a reduction by 10% of consumption of one Chinese is considered equivalent to a reduction by 49% by one American who is on average 8.5 times (on PPP basis) richer. One case of particular interest is logarithmic utility ($\eta = 1$) for which *relative* changes in consumption are evaluated independently of the level of consumption.

The cost of the given project in terms of social welfare can then be expressed as the sum of marginal utilities today as $\sum_{i=1}^{n} \pi_{it} P_t U'(c_{i0})$ while the benefits at date t can be accordingly written as $e^{r_t t} E\left[\sum_{i=1}^{n} \pi_{it} P_t U'(c_{it}) e^{-\delta t}\right]$.

In the following, we assume most of the time that the population does not change over time in order to get a more tractable solution. In a world with a single representative agent, this does not alter the results given that one considers implicitly per capita consumption. In our case with different sub-populations on the other hand, this is no longer true. Differences in population growth rates lead to changing relative populations which affects average world per-capita consumption.

In general, the project would be socially desirable if

$$\sum_{i=1}^{n} \pi_i U'(c_{i0}) \leq e^{r_t t} E\left[\sum_{i=1}^{n} \pi_i U'(c_{it}) e^{-\delta t}\right]$$

and we can derive the socially optimal discount rate for the time horizon t as the rate of return which would make the social planner just indifferent between accepting or rejecting this project. In this case this condition holds with equality and it can be solved for r_t[2] as

$$r_t = \delta - \frac{1}{t} \ln \frac{\sum_{i=1}^n \pi_i E[U'(c_{it})]}{\sum_{i=1}^n \pi_i U'(c_{i0})}. \tag{1}$$

If we think of the weighted sums in this expression as expectations, this version of the Ramsey-Keynes condition looks somewhat similar to the equivalent cases when considering uncertainty, there is a substantial difference: while uncertainty affects only the future, inequality is considered both at present and in the future. Take one example where today's income were distributed equally, i.e. $c_{i0} = c_0 \; \forall i$. In this case we can rewrite (1) in the spirit of a model with parametric uncertainty as in Gollier (2008) where on top of uncertainty about future growth, its expected value is also uncertain. If we take different growth rates in different countries for this particular case, this will induce a lower discount rate over the case where the aggregate growth rate is such that aggregate consumption will be the same. This case is obviously very unrealistic and not the part we are interesting in.

We can derive a different interpretation of the general condition (1) by introducing a weighting scheme based on marginal utility. Denoting by $\bar{\pi}_i$ the weight of country i based on today's marginal utility, i.e., $\bar{\pi}_i \equiv \frac{\pi_i U'(c_{i0})}{\sum_{j=1}^n \pi_j U'(c_{jo})}$[3], we can rewrite it as

$$r_t = \delta - \frac{1}{t} \ln \sum\nolimits_{i=1}^n \bar{\pi}_i \frac{EU'(c_{it})}{U'(c_{i0})} \tag{2}$$

which shows that each country's contribution to the global discount rate is weighted by today's population-weighted marginal utility. The weights π_i are replaced by inequality-adjusted weights $\bar{\pi}_i$. These weights in a way are similar to equity weights in evaluation impacts from climate change as proposed by Fankhauser et al. (1997). In their case, however, the weights are applied to the actual damage estimates while here they apply for the discount rate due to changes in the income distribution.

To simplify notation, in the following we denote by E_t the expectation operator in a given country over time while by E_π we refer to the average of its argument weighted by population, so that we can write expected average marginal utility as $E_t E_\pi U'(c_{it})$. Since the expectation (or average) operator is linear, we can interchange its order, that is, uncertainty can be interpreted as additional inequality in the future due to the analogy between risk and future inequality. What matters is only the distribution of future marginal utility–be it due to uncertainty or inequality between countries (veil of ignorance). Moreover, for $U''' > 0$ we have that $E_\pi E_t U'(c_{it}) > E_\pi U'(E_t c_{it})$ or that uncertainty implies a lower discount rate–as in the standard case without inequality.

[2]By $r_t$we refer throughout this paper to the average discount rate as opposed to the marginal discount rate which is equivalent to the instantaneous change of the discount factor.

[3]A similar weighting scheme was found in Gollier and Weitzman (2010) in the context of uncertain discount or interest rates. The important difference is that in their case, the different initial income levels are endogenously determined based on the realization of the true interest rate while in our case they are exogenously given by the current income distribution.

If we allow for population growth, this would affect the results only in changing relative weights as discussed before. In this case we would have time-dependent weights as $\bar{\pi}^*_{it} \equiv \bar{\pi}_i \frac{\pi_{it}}{\pi_{i0}}$. Note that while as stated above global population growth does not enter the picture, a *relative* change in the geographical distribution of the population shifts the weights towards countries with faster population growth.

3 Comparison with the case without inequality

In order to analyze the effect distributional considerations have on the optimal discount rate, one needs to specify a reference case to compare the results with. In particular, one needs to compare (2) with the appropriate version without the consideration of intragenerational inequality.

Stern et al. (2006, 48) and Conceiçao and Zhang (2010) discuss a simple case where they compare two different future income distributions. In this case, one can easily conclude using the results of Atkinson (1970) that a more unequal future income distribution in the sense of Lorenz domination[4] will imply a lower discount rate if and only if $U''' > 0$. Convex marginal utility decreases in this case due to the SSD deterioration in the distribution of future income and thus implies a lower discount rate.

This result, however, does not admit the conclusion that increasing inequalities imply a lower discount rate. The reason is that if one considers inequality for the choice of the discount rate, the representative agent approach as such is not valid. Not only is the future income distribution affecting the magnitude of the discount rate, but also today's income distribution. The appropriate comparison would be between the discount rate with inequality against the discount rate using only average GDP per capita, i.e., the representative agent case.

Therefore, global GDP and its growth rate have to be used to calculate the global discount rate assuming only average GDP per capita is considered. We write $\tilde{r}_t$ for the optimal discount rate considering only average world per-capita consumption $E_\pi c_{it}$. In this case, the optimal discount rate is defined as

$$\tilde{r}_t = \delta - \frac{1}{t} \ln \frac{U'(E_\pi c_{it})}{U'(E_\pi c_{io})} \tag{3}$$

whereby we abstract from uncertainty following the discussion above. That is, comparing this condition with the discount rate r_t when inequality is considered as given by (1), it becomes clear that

$$r_t < \tilde{r}_t \Longleftrightarrow \frac{U'(E_\pi c_{it})}{U'(E_\pi c_{i0})} < \frac{E_\pi[U'(c_{it})]}{E_\pi U'(c_{i0})}. \tag{4}$$

Inequality affects both today's and tomorrow's marginal utility as it increases it relative to the no-inequality case. The overall effect is thus a priori not clear. The reason is that the effect of inequality can be offset by the growth effect.

[4] where the new Lorenz curve lies everywhere below the old one. In the context of uncertainty this is equivalent to the notion of second-order stochastic dominance (SSD).

If there is no global growth ($E_\pi c_{it} = E_\pi c_{i0}$), one can easily show that increasing inequality in the sense of adding mean-preserving spreads implies always a lower discount rate. The same holds for the case of the CARA utility function $U(c) = A^{-1}e^{-Ac}$.

A more relevant case is the case of a CRRA utility function $U(c) = (1-\eta)^{-1}c^{1-\eta}$, ubiquitous in applied work. Moreover, we consider inequality through an aggregate inequality measure since individual data is hardly available at least for projections. All of the widely used classes of inequality indices can be traced back to a particular welfare function specification and thus by choosing one of them one implicitly takes on a welfarist judgment. Given that our welfare specification is Utilitarian and using the CRRA utility function as mentioned, the family of Atkinson (1970)'s inequality indices is the appropriate concept of measuring inequality since they are derived precisely in this framework.

The Atkinson index of inequality in any region r with per-capita consumption of c_{rt} is defined as

$$I_{rt}(\eta) = 1 - \frac{c_{rt}^{ede}}{c_{rt}} \text{ where } c_{rt}^{ede} = \left(\sum_{c=1}^{C_r} \pi_{rct} c_{rct}^{1-\eta}\right)^{\frac{1}{1-\eta}} \tag{5}$$

and where η denotes the degree of inequality aversion between countries. It can be interpreted as the percentage of per-capita income that provided the same total welfare than the actual income distribution. The equally distributed equivalent level of consumption c_{rt}^{ede} introduced above can hence be expressed as $c_{rt}^{ede} = c_{rt}(1-I_{rt}(\eta))$.

Using this measure of inequality and using the CRRA utility specification also for welfare evaluations over time, we obtain the following result:

Proposition 1. *With a CRRA utility function and for small degrees of inequality, whenever economic convergence leads to a decrease of inequality in the sense that the according Atkinson measure of inequality $I(\eta)$ decreases, the optimal discount rate is higher than without considering inequality.*

For an arbitrary degree of inequality, a decrease in inequality according to the Atkinson measure $I_t(\eta+1)$ is a necessary and sufficient condition for a lower discount rate.

Proof. Using a second-order Taylor expansion of $E[c_t^{-\eta}]/(Ec_t)^{-\eta}$ around the mean of c_t yields $1 + \frac{1}{2}\eta(1+\eta)\frac{Var(c_t)}{(Ec_t)^2}$. Therefore, condition (4) depends only on the change of the coefficient of variation over time. Similarly we can develop Atkinson's inequality measure $I_t(\eta)$ the same way which is also a monotone function of the coefficient of variation. Thus for small degrees of inequalities, a decrease of the coefficient of variation implies a decrease in $I(\eta)$ as well as $r_t > \tilde{r}_t$. For the second part, we only need to write $I_t(\eta+1)$ explicitly and after some reformulations, the equivalence to (4) is immediate. □

It is important to note that there is no unambiguous link between the evolvement of inequality and the discount rate. The reason lies in difference in distribution of marginal utility, which is used to compute the discount rate, on the one hand, and of utility itself, which determines the degree of inequality at a given point in time. Even for the restrictive class of the CRRA utility function, the equivalence is only valid for small degrees of inequality.[5] The exact equivalence for the Atkinson index $I(\eta+1)$ must be seen with caution since it does only hold when the social planner has a degree of inequality aversion that is equal to the degree of risk aversion or the inverse of the intertemporal elasticity of substitution plus one.

While Gollier (2010a) obtains a less ambiguous result, his findings are due to a particular way of measuring the degree of changes of inequalities over time. In particular, he applies the notion of decreasing inequality or more convergence as a reduction of concordance between the vector of consumption and the growth factor showing that a reduction of concordance asks for a power utility function always for a higher discount rate. However, a reduction in concordance is not equivalent with a reduction in inequality.[6]

The case of a general utility function is even more complex. The reason is that even without changes in the global income distribution, the effect of considering income inequality for the discount rate is ambiguous. As it is shown in Gollier (2010a), no general sufficient condition can be found to determine the sign of the effect on the discount rate. In the following, we denote by $\overline{c}_t$ average world income at date t and by $\varepsilon_t = (\varepsilon_{1t}, .., \varepsilon_{nt})$ the zero mean income spread at date t so that we have $c_{it} = \overline{c}_t + \varepsilon_{it}$.

Now we can rewrite (4) by adding one to both sides as

$$\frac{E_\pi U'(\overline{c}_t + \varepsilon_{it}) - U'(\overline{c}_t)}{U'(\overline{c}_t)} > \frac{E_\pi U'(\overline{c}_{i0} + \varepsilon_{i0}) - U'(\overline{c}_0)}{U'(\overline{c}_0)}. \tag{6}$$

That is, whether inequality asks for a lower or higher discount rate depends on how much inequality increases average marginal utility relative to marginal utility at date 0 and date t. In general, the conditions for this inequality to hold are threefold: Firstly, the numerators are positive if $U''' > 0$ due to the Jensen inequality. Secondly, the denominators depend on the relative per capita world income over time. Thirdly, in order to compare the numerators, this depends on whether the function $-U'(c)$ (which is increasing and concave) would prefer a risk at lower or higher incomes. Given that the attitude towards downside risks depends on the third derivative of the utility function, in this case the relevant characteristics would be U''''. In general, these three effects can work in either directions so that the overall results will be ambiguous.

For a comparably low level of income inequality, we can write the (4) using second-order Taylor expansions around $\overline{c}_t$ and $\overline{c}_0$ respectively. This yields an approximation of the relative changes of marginal utility due to inequality that can be

[5] As we show in section five, however, this result can be extended to any degrees of inequalities if we impose the very common distributional assumption of a lognormal income distribution.

[6] Note that this discrepancy is related to the debate about $\sigma-$ versus $\beta-$convergence: A decrease in the relatedness between today's income and growth ($\beta-$ convergence) is not sufficient for reducing income dispersion ($\sigma-$convergence), see, e.g., Young et al. (2008).

written as

$$r_t < \tilde{r}_t \Longleftrightarrow Var[\varepsilon_t]\frac{U'''(\bar{c}_t)}{U'(\bar{c}_t)} > Var[\varepsilon_0]\frac{U'''(\bar{c}_0)}{U'(\bar{c}_0)}. \quad (7)$$

The term $U'''(c)/U'(c)$ has been used in context of uncertainty as a measure of downside risk aversion, see Menezes et al. (1980), Keenan and Snow (2002), and Crainich and Eeckhoudt (2008)[7]. It is sometimes referred to as the index of Absolute Downside Risk Aversion (ADRA).

In the context of inequality, Aversion to Downside Inequality[8] (ADI) is the equivalent concept to downside risk aversion (Davies and Hoy, 1995). It is widely accepted as a normative axiom for inequality comparisons implying that a spread in the income distribution is considered less severe if it occurs at an already higher income level.

If we accept the axiom of aversion to downside inequality, i.e., assume $U''' > 0$, we can conclude that the discount rate is lower than without inequality if and only if the downside inequality premium $Var[c] \cdot ADI$ is higher in the future than it is today. Put it differently, the condition states that the optimal discount rate should be lower if and only if the variance increases over time faster than the absolute measure of downside risk aversion decreases.

A priori, the question whether downside risk aversion is increasing or decreasing in wealth is not clear. However, in the context uncertainty, we can use the concept of standard risk aversion in Kimball (1993) together with the results of Wang and Menezes (2002) to argue that decreasing (absolute) downside risk aversion seems to be the plausible case. With decreasing ADRA, however, this implies that with non-negative global growth and inequality in the sense of the variance of income decreasing over time, the interest rate must be higher than if inequality were not considered. For the particular example of the CRRA utility function for which $ADI_t = \eta(1+\eta)c_t^{-2}$ is decreasing, the relevant condition can be further simplified and implies that the discount rate should be higher, i.e., $r_t > \tilde{r}_t$, if and only if the coefficient of variation $StdDev(c_t)/\bar{c}_t$ decreases over time as we showed before.

4 Constant growth rates

While it is not possible to get an easy-to-interpret general solution of (2), we can look at the case where the discount rate in each country i would be constant over time denoted by g_i. In this case, measures of inequality such as the Gini coefficient or other distribution-based concepts are typically changing non-monotonically over time. For instance, in the long run, the Gini coefficient in a world with different but constant growth rates, will always increase ultimately.

[7]The latter show that the amount of money a decision maker is willing to accept at the best state (income w) for reallocating a pure risk ε from this state to a state where income is Δw lower can be approximated as $0.5Var(\varepsilon)\Delta w\frac{U'''(w)}{U'(w)}$, which looks qualitatively similar to equation (7).

[8]A concept also referred to as the 'principle of diminishing transfers' (Kolm, 1976) or 'transfer sensitivity' (Shorrocks and Foster, 1987).

Together with a CRRA utility function, we can write the ratio of future and today's marginal utility in country i as $(e^{g_i t})^{-\eta}$ and get a simplified form for the optimal discount rate. Moreover, in order to analyze the effect of inequality, we need to derive the global discount rate $\tilde{r}_t$ based on per capita income only for which we introduce a different weighting scheme. The individual factors will be weighted by the level of today's consumption denoted by $\alpha_i \equiv \frac{\pi_i c_{io}}{\sum_{j=1}^{n} \pi_j c_{jo}}$. This is intuitive since the growth rate of a country with higher initial consumption has a larger effect on average consumption growth. We will use again the notation of $E_{\bar{\pi}}$ and E_α for the average of its argument taken using the weights $\bar{\pi}_i$ and α_i respectively while E_t denotes the expectation within a given country about the future. With this notation, we can compare the two cases (2) and (3) for constant growth rates as

$$r_t = \delta - \frac{1}{t} \ln E_{\bar{\pi}} \left[E_t \left[\left(e^{g_i t} \right)^{-\eta} \right] \right] \text{ vs. } \tilde{r}_t = \delta - \frac{1}{t} \ln E_t \left[\left(E_\alpha \left[e^{g_i t} \right] \right)^{-\eta} \right]. \tag{8}$$

Comparing the two conditions one can distinguish two effects of inequality on the socially optimal discount rate. The first effect is related to the different weights used in these two equations: without distributional considerations the weights are just the initial consumption levels α_i since a given growth rate will be weighted higher if occurring in a country with high initial consumption as it has an larger effect for the World GDP growth rate. With equity considerations on the other hand, the weights are based on today's marginal utility implying larger weights for poorer countries. The second effect stems from Jensen's Inequality and is qualitatively equivalent to the role of uncertainty about future growth: the average under equity considerations is taken over marginal utility while it is the marginal utility of the average consumption in the second case.

Proposition 2. *For constant country-specific growth rates, the optimal discount rate is always non-increasing over the time horizon and converges to the discount rate applicable in the country with the lowest growth rate.*

If inequality increases over time in the sense that the covariance between today's consumption and its growth rate is positive, i.e., $Cov(c_{i0}, g_i) > 0$, the optimal discount rate is always lower when income inequality is taken into account.

Proof. Computing the derivative of r_t with respect to time and using Jensen's inequality one can prove that the optimal global discount rate is always non-increasing over the time horizon and moreover tends to the lowest rate applicable in any country using a reasoning similar as in Weitzman (1998) or Gollier (2007).[9]

For the second part of the proposition, having that the covariance between today's consumption level and the growth rate is positive implies that $Cov(c_{i0}, e^{g_i t}) > Cov(c_{i0}^{-\eta}, e^{g_i t})$. After several reformulations, one can show that this is sufficient for ensuring that $E_{\bar{\pi}}[e^{g_i t}] < E_\alpha[e^{g_i t}]$. Due to the fact that marginal utility is convex for CRRA utility, this implies further the left part of the following series of

[9]By a similar argument, once can show that the average growth rate and hence the discount rate without considering inequality $\tilde{r}_t$ is increasing over the time horizon and converges to the discount rate applicable in the fastest growing country as t tends to infinity.

inequalities whereas the right one follows from the first part of the proposition: $E_{\bar{\pi}}\left[(e^{g_i t})^{-\eta}\right] > (E_{\bar{\pi}}[e^{g_i t}])^{-\eta} > (E_\alpha[e^{g_i t}])^{-\eta}$. From the two conditions in (8) it becomes now clear that in this case, the discount rate with the consideration of inequality needs to be lower than the one considering only average world income. □

These results are comparable with the models about uncertainty about the discount rate as in (Weitzman, 1998, 2001). The important difference is the weighting scheme which now has a very different interpretation: in the Weitzman case of an uncertain interest rate, initial consumption is random since it is endogenously determined once the interest rate is learned, see Gollier and Weitzman (2010) whereas here it reflects actual income differences. Therefore, the weights of different discount factors is based on the exogenous income distribution whereas in the Gollier and Weitzman case they are determined endogenously.

The results from the case of constant country-specific growth rate can easily be extended to allow for uncertainty and also relative changes in population. In particularly we look at the case where per capita consumption in each country follows a geometric Brownian Motion, i.e., $\frac{dc_i}{c_i} = g_i dt + \sigma_i dW_{it}$. For this case we get for the optimal discount rate

$$r_t = \delta - \frac{1}{t} \ln \sum_{i=1}^{n} \bar{\pi}_i e^{-\eta g_i t + .5\eta(\eta+1)\sigma_i^2 t}. \tag{9}$$

That is, only the drift and variance of each country's consumption path matter for the discount rate while any possible correlation between the individual Wiener processes $E[dW_{it}dW_{jt}]$ does not affect the results at all. This is due to the fact that in an Expected Utility framework using an Utilitarian Social Welfare Function one considers basically the (weighted) sum of marginal utility and hence the cross-moments do not matter.

If the relative population sizes change over time, this needs to be taken into account for the global discount rate. Denoting by n_i the population growth rate in country i we have that $\pi_{it} = \pi_{i0} e^{(n_i - \overline{n_t})t}$ where we write $\overline{n}_t$ for the average global population growth rate over the horizon t, i.e., $\overline{n}_t \equiv \frac{1}{t} ln\left(\frac{P_t}{P_0}\right)$ which is not constant over time. For the optimal discount rate we get in this case using again the fact that we consider only the relative change in population and not global population growth the following formula:

$$r_t = \delta - \frac{1}{t} \ln \sum_{i=1}^{n} \bar{\pi}_i e^{-\eta g_i t + .5\eta(\eta+1)\sigma_i^2 t + (n_i - \overline{n_t})t} \tag{10}$$

Finally, we apply these results and calibrate a widely used regional scenario, namely the IS92a scenario, similar to the SRES A2 baseline scenario used in the Third Assessment Report of the IPCC. They distinguish nine regions whose characteristics are summarized in the following table.

	Pop.	GDP p.c.	Growth	π_i	$\bar{\pi}_i^{\eta=1}$	$\bar{\pi}_i^{\eta=2}$
World	6,205	4,309	1.7 - 2.7%	1.00	1.00	1.00
United States of America	269	30,545	1.4%	0.04	0.00	0.00
OECD Europe / Canada	457	13792	1.5%	0.07	0.01	0.00
OECD Pacific	159	20,504	1.6%	0.03	0.00	0.00
Russia / Eastern Europe	453	8,350	1.6%	0.07	0.01	0.00
China / Central Planned Asia	1,403	661	3.4%	0.23	0.34	0.37
Middle East	175	6,098	1.9%	0.03	0.00	0.00
Africa	870	754	2.3%	0.14	0.18	0.18
Latin America	529	2,234	2.2%	0.09	0.04	0.01
South and South-East Asia	1,890	711	2.9%	0.30	0.42	0.43

Regional values of the IS92a scenario for 2000 (GDP in USD of 1990)

Taking the expected per capita growth rates over the two centuries until 2200, one can compute the optimal discount rate. For instance, for values of $\eta = 1$ and $\delta = 0.1\%$, we get for the discount rate reflecting only aggregate world consumption a value of $\tilde{r}_t = 1.8\%$ for short horizons that increases slowly reflecting World GDP growth. If we compute the respective equity weighted value, we find that for short horizons $r_t = 3.0\%$ which decreases to 2.7% over 200 years (see Figure 3). Note that this scenario predicts for instance, that the World-Gini coefficient declines from .69 in 2000 to .44 in 2100 and to .30 in 2200 ($Corr(c_{i0}, g_i) = -0.74$). This strong convergence makes future generations better of which in turn implies that the optimal discount rate should be higher than if inequality were not considered.

A related approach of computing the discount rate relies on using the speed of convergence directly in terms of a weighted convergence regression as in Gollier (2010a) of the form[10]

$$g_i = \alpha + \beta log(\frac{c_{i0}}{\overline{c_{i0}}}) + \varepsilon_i \tag{11}$$

for which we find using a first-order Taylor approximation for a relatively low convergence speed, that the discount rate can be approximated as

$$r_t \simeq \delta + \alpha \cdot \eta - \beta log \frac{E_\pi c_{io}^{-\eta}}{\overline{c_{i0}}^{-\eta}} \tag{12}$$

where the last term indicates that convergence ($\beta < 0$) implies a higher discount rate. Moreover this effect is more important whenever initial inequality is rather large as this increases the log-term. For the present calibration and the regression results $\alpha = 0.0207$, $\beta = -0.0049$, $R^2 = 0.79$, this leads to a constant discount rate of $r_t = 2.89\%$ which approximates the exact term structure fairly well.

For higher values of η the pattern remains the same in that the discount rate increases in both cases while the equity-weighted value remains at around twice the value without it. This is due to the fact that increasing η leads to a one-to-one increase in the discount rate due to the growth effect. Moreover, it shifts the weight

[10]Dividing initial income by its mean guarantees that average growth without inequality or convergence is just equal to α.

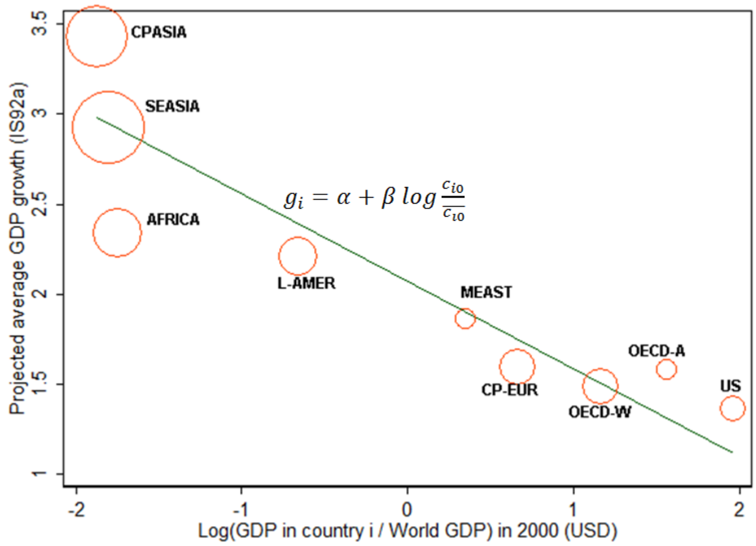

Figure 2: IS92a scenario and convergence

even more towards the poorer countries. As the poorest regions are in this scenario assumed to grow with the highest growth rates, the overall effect unambiguously increases the equity-weighted discount rate as η rises.

That is, equity weighting implies an almost twice as large discount rate and this effect is rather persistent as r_t decreases only slowly. The main reason is the relative high weight attributed to poorer countries' higher expected growth rates given that higher growth rates in poorer countries are over-weighted. This scenario predicts for instance, that the World-Gini coefficient declines from .69 in 2000 to .44 in 2100 and to .30 in 2200 ($Corr(c_{i0}, g_i) = -0.74$).

5 Risk aversion, inequity aversion, and resistance to intertemporal substitution

Until now, the function U captured the degrees of inequality aversion over three dimensions: time, individuals, and stats of nature. The curvature of this function thus constrains these three different attitudes to coincide, e.g., for the isoelastic case, expressed by the parameter η. This assumption has been criticized as restrictive as the three dimensions are qualitatively different. Moreover, experimental studies found different degrees of inequality aversion for the different dimensions, see Atkinson et al. (2009) or Carlsson et al. (2005). In this chapter we derive the discount rate in the spirit of recursive expected utility where the three parameters are allowed to differ.

In the context of discounting, risk aversion and resistance to intertemporal substitution have been disentangled in Gollier (2002) or Traeger (2009). An application to climate change is Ha-Duong and Treich (2004). With regard to risk aversion and inequality aversion, Carlsson et al. (2005) and Tol (2010) argue that the two prefer-

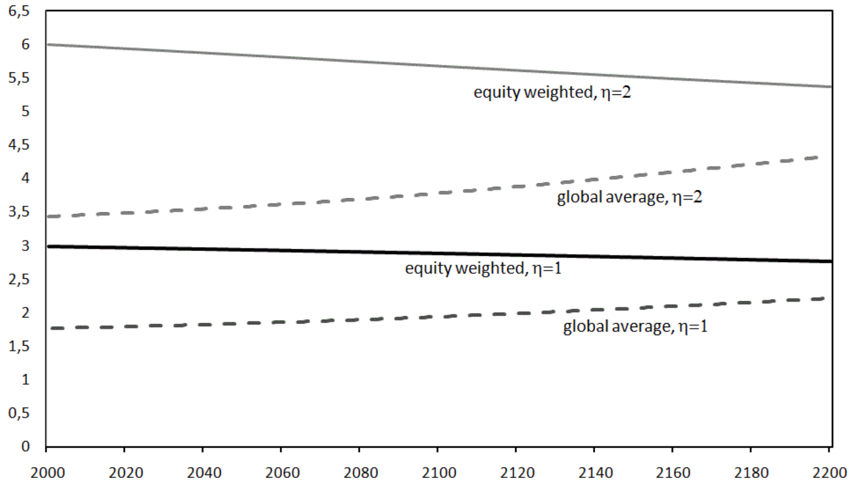

Figure 3: Optimal discount rates from the IS92a scenario

ences should not be confounded. A separation of the three different preferences of the social planner however does not exist to our knowledge.

In the following we therefore aim at providing a model, which allows to derive analytical results for the effect of the three different parameters. Our approach is in the spirit of the recursive model by Kreps and Porteus (1978) and Epstein and Zin (1989, 1991). However, the different nature of inequality as compared to uncertainty implies a somewhat different interpretation.

Making the three dimensions separable requires a threefold nested welfare function. This implies one would a three dimensional random variable for consumption as well. However, we consider consumption in the future to be characterized by a one-dimensional random variable c_{it} at any given point in time. The fundamental reason is that uncertainty and inequality can otherwise not be separated since global inequality in the future depends on the fluctuations of growth between countries. Moreover, when it comes to intra-regional or even inequality within countries, individual data are hardly available. Therefore, we will make use of parametric estimates of the income distribution within a region or country.

More precisely, we assume a continuous bivariate distribution of today's income c_{i0} and the growth factor x_{it} as $(c_{i0}, x_{it}) \sim F_t$ where the distribution depends on t. Note that $x_{it} = e^{g_{it}t}$ where g_{it} denotes the average growth rate if initial consumption was given by c_{i0} ("country i") until date t in a given state of the world. Note that despite using a continuous distribution, we use the sub-indices i in order to distinguish the inequality dimension between individuals from uncertainty. The bivariate distribution thus captures both inequality and uncertainty. Moreover, the degree of dependence between the two random variables c_{i0} and x_{it} measures the changes of inequality over time. Income at date t in a country with initial income c_{i0} is thus given by the random variable $c_{i0} \cdot x_{it}$. That is, inequality of future income can be seen as the conditional distribution of $c_{i0} \cdot x_{it}$ for any state of the world. Uncertainty, on the other hand, is captured by the conditional distribution

$x_{it} \mid c_{i0}$. This distribution can be seen as the distribution of the growth factor of any given initial level of income c_{i0}. For the case of a degenerate distribution of c_{i0} and assuming for instance that consumption follows a geometric Brownian motion, this distribution would be log-normal.

Now in order to derive welfare at date 0, we have to employ three steps from the distribution of future consumption c_{it}. First, consider a concave von Neumann-Morgenstern Utility function v applied in any country for evaluating the uncertain consumption level c_{it}in country i denoted by $v(c_{it} \mid c_{i0})$. Given that the consumption level of this country today is given by c_{i0}, we can write the conditional certainty equivalent of this random consumption level c_{it}^{CE} which solves

$$v(c_{it}^{CE}) = E_t[v(c_{it}) \mid c_{i0}]. \tag{13}$$

This certainty equivalent captures the degree of uncertainty in any given country i of future consumption.

Secondly, the social planner is assumed to be inequality averse which can be represented by a Social Welfare Function. Inequality aversion between countries or individuals is then captured by a function combining individual utilities and the degree of inequality aversion expressed by the utility function in each country. For isoelastic specifications of the two functions, one can derive the degree of consumption inequity aversion (Tol, 2010)[11] of the individual utility function $g(c)$ which we will denote γ later on. Now let the argument of utility in each country be the aforementioned certainty equivalent level of consumption where uncertainty is already accounted for. That is, current social welfare at time t can be written as $W_t = E_\pi g(c_{it}^{CE})$ where as in the precious section we denote by E_π that the population average of its argument is taken.

In a similar fashion as with the certainty equivalent above, we can use of the concept of the representative or 'equally distributed equivalent´ level of consumption or income (c_t^{EDE}) as proposed by Atkinson (1970): aggregated welfare is equivalent to welfare of the hypothetical value of c_t^{EDE} if it were to be equally distributed. For date zero where individual consumption levels are know, this implies that we have $c_0^{EDE} = g^{-1}(E_\pi [g(c_{i0})])$. At any future period, however, individual consumption levels are random and therefore we consider the uncertainty adjusted certainty equivalent level for each country as calculated above. The value of c_t^{EDE} at any date t can therefore be written as

$$g\left(c_t^{EDE}\right) = E_\pi \left[g(c_{it}^{CE})\right]. \tag{14}$$

[11] Tol (2010) shows that inequity aversion can be characterized by either using a non-Utilitarian SWF or a concave utility function. In the case of both functions being isoelastic functions, the relative degree of inequity aversion over consumption can be obtained as the sum of the individual inequality aversion parameters minus their product. A degree of relative inequality aversion of one can thus be obtained, e.g., with a Bernoulli-Nash SWF and linear utility or a Utilitarian SWF with a logarithmic utility function. Global welfare can thus be rewritten as the average of a function of individual consumption that is isoelastic with a modified degree of relative inequality aversion. This degree of inequality aversion is what we use throughout this paper.

Finally, we maintain time-separability and exponential (continuous) discounting where utility at any point in time is described by the function $u(\cdot)$. We can then write discounted social welfare for consumption at two points in time, 0 and t, as

$$W = u(c_0^{EDE}) + e^{-\delta t} u(c_t^{EDE}). \tag{15}$$

Note that the argument of u depends on the representative income c_t^{EDE} which in the case of date t itself is a function of the certainty equivalent of consumption for each initial income c_{i0}. Moreover, if the function g is linear, the social planner is inequality neutral between individuals or countries. In this case, welfare collapses to the standard recursive utility model evaluation consumption levels of $E_\pi c_{i0}$ today and $x_{it} E_\pi c_{i0}$ at date t. If on the other hand$v$ is assumed to be linear, the social planner is risk neutral and the certainty equivalent in each country is just the expected value, i.e., $c_{it}^{CE} = c_{i0} E_t[x_{it} \mid c_{i0}]$. Typically, however, one would assume concave functions for the three functions. As outlined in the last chapter and in Gollier (2010a), inequality, since it affects both today's and tomorrow's marginal utility, is rather intractable in order to derive comparative statics results. Using an even more general model, this is even more demanding.

Therefore, we use isoelastic specifications as in Epstein and Zin (1989) for the three functions f, g, and u in order to get analytical results. We denote the elasticity of marginal utility in the three dimensions by φ (v, states of nature), γ (g, space), and η (u, time).

The overall welfare at date zero of consumption at two points in time can then given as the following threefold nested expression

$$W = \left[\left(E_\pi \left[c_{i0}^{1-\gamma} \right] \right)^{\frac{1-\eta}{1-\gamma}} + e^{-\delta t} \left(E_\pi \left[\left(E_t \left[c_{it}^{1-\varphi} \mid c_{i0} \right] \right)^{\frac{1-\gamma}{1-\varphi}} \right] \right)^{\frac{1-\eta}{1-\gamma}} \right]^{\frac{1}{1-\eta}} \tag{16}$$

where η denotes the inverse of the inter-temporal elasticity of substitution (IES), the parameter γ captures the degree of inequality aversion between countries and φ measures the degree of relative risk aversion (RRA). The structural form of (16) is additively separable and uses exponential utility discounting which ensures dynamic consistency.

We can now use this formula and derive the optimal consumption discount rate as in the previous chapter. Again, the rate of discount is given by comparing one marginal unit consumption today against one marginal unit in the future. In our case, we have to add to this definition that the unit of consumption is accruing on a per-capita basis today and in the future. Moreover, the marginal unit from the assumed project is as before certain and hence independent of the states of the world considered for the consumption path. Using this further precision of the definition of the optimal social discount rate, we can write it in general as

$$- E_\pi \frac{\partial W}{\partial c_{i0}} + e^{-r_t t} E_\pi E_t \frac{\partial W}{\partial c_{it}} = 0. \tag{17}$$

Using the expression for welfare at date 0 as given by (16), it becomes clear that for any admissible distribution F_t of $\{c_{i0}, x_{it}\}$ the results are not tractable, in particular due to the inequality aversion. In order to be able to compute the discount rate analytically, we specify the bivariate distribution. In particular, we look at the case where consumption per capita is log-normally distributed today and at any date t. The log-normal distribution has proven to be a very good approximation of the world income distribution and has been the preferred distribution e.g. in Atkinson and Brandolini (2010). Moreover, with regard to the income distribution of the growth factor x_{it} for any given initial income c_{i0}, we also assume it to be lognormal distributed. One way of obtaining this result would be if consumption were to follow a geometric Brownian Motion or that instantaneous growth rates are normally distributed.

Having specified the marginals of the distribution F_t, we need to make make some assumption about the dependence structure of the two random variables. Here, we restrict the bivariate distribution between today's income distribution and the growth factor distribution to be a bivariate lognormal distribution. While there are an infinite number of dependence structures with log-normal marginal distributions possible, we take the simplest case. In any case, the dependence need to be specified based on convergence predictions, which are not very detailed anyhow. Still this is an important assumption which for the moment we have to impose.

Using this assumption and some meaningful choices of the parameters and time-dependence of this distribution, F_t can be expressed as follows:

$$\begin{pmatrix} c_{i0} \\ x_{it} \end{pmatrix} \sim LN\left(\begin{pmatrix} \mu_0 \\ \mu_g t \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_g\sqrt{t} \\ \rho\sigma_0\sigma_g\sqrt{t} & \sigma_g^2 t \end{pmatrix}\right) \tag{18}$$

From this specification we can derive several useful distributions. The current income distribution is given as $c_{i0} \sim LN(\mu_0, \sigma_0^2)$. At any given point in time time t, we have $c_{it} \sim LN(\mu_0 + \mu_g t, \sigma_0^2 + \sigma_g^2 t + 2\rho\sigma_0\sigma_g\sqrt{t})$. Most importantly, we can derive the conditional distribution of the growth factor in a country with initial level c_{i0}, which is given by

$$x_{it} \mid c_{i0} \sim LN(\mu_g t + \frac{\rho\sigma_g\sqrt{t}}{\sigma_0}(lnc_{i0} - \mu_0), \sigma_g^2(1-\rho^2)t) \tag{19}$$

implying that the growth rates for any value c_{i0} are normally distributed and have the same volatility. Using this distributional assumption together with the isoelastic functional forms, we can analytically derive the optimal discount rate. After some tedious computations, we finally get the optimal consumption discount rate as a function of the set of parameters as

$$r_t = \delta + \eta\overline{g}_t - \varphi(\eta+1)\frac{\sigma_g^2}{2}(1-\rho^2) - \gamma(\eta+1)\left(\frac{\rho\sigma_0\sigma_g}{\sqrt{t}} + \frac{\rho^2\sigma_g^2}{2}\right) \tag{20}$$

where $\overline{g}_t$ denotes the growth rate of unconditional *average* consumption between 0 and t.

The first two terms are the standard pure rate of time preference and the growth effect of the Ramsey formula. Uncertainty enters the picture through the negative term $\varphi(\eta+1)\frac{\sigma_g^2}{2}(1-\rho^2)$ which is similar to Gollier (2002) and depends on the variance of the growth rate, the degree of relative prudence (which for our specification is given by the expression $\varphi(1+\frac{1}{\eta})$, see Kimball and Weil (2009) and the inverse of the intertemporal elasticity of substitution η. The lognormal assumption implies that volatility of growth is equal and independent of the initial income level. Moreover, it is constant for all maturities as in the simple case of one geometric Brownian Motion, so that the term structure would be flat.

The changes in inequality between countries, or individuals, is captured by the fourth term of (20). The coefficients are similar to the ones of the uncertainty effect, depending on the elasticity of intertemporal substitutions and a measure of relative prudence with respect to changes in inequality while the last term in brackets needs more interpretation.

In order to decompose the variance of log income in the future into inequality and uncertainty, we can use the law of total variance and using the properties of the bivariate log-normal distribution. Both terms can be computed as $E_\pi\left[Var(lnc_{it} \mid c_{i0})\right] = \sigma_g^2(1-\rho^2)t$ and $Var(E_t\left[lnc_{it} \mid c_{i0}\right]) = (\sigma_0 + \rho\sigma_g\sqrt{t})^2$. These two terms capture uncertainty and inequality in of the future distribution of consumption. The former term gives the expected variance of future consumption due to the uncertain growth rate from which we can derive the volatility of the growth rate given above. The latter stands for the variance of the expected value of log consumption in the future. Comparing this term with today's variance, we obtain $Var(E_t\left[lnc_{it} \mid c_{i0}\right]) - Var(lnc_{i0}) = 2\rho\sigma_0\sigma_g\sqrt{t} + \rho^2\sigma_g^2 t$, the change in the variance of log-income between countries or individuals. Half of its annualized value gives exactly the term in the inequality term of (20) and we can thus write the discount rate the following way:

$$r_t = \delta + \eta\bar{g}_t - \varphi(\eta+1)\frac{E_\pi\left[Var(lnc_{it} \mid c_{i0})\right]}{2t} - \gamma(\eta+1)\frac{Var(E_t\left[lnc_{it} \mid c_{i0}\right]) - Var(lnc_{i0})}{2t} \tag{21}$$

This formula looks very similar to the generalized Ramsey rule, see Gollier (2011). The uncertainty term depends on the average volatility of the growth rate while the inequity term depends on the change of the variance of log income, a frequently used inequality measure, see Sala-I-Martin (2002). On a side note, unless the change in the variance of log income is linear over time, the term structure will not be constant.

We can express this formula in a more accessible manner based on measures of average volatility of growth σ_g and inequality as measured by the Atkinson index $I_t(\gamma)$ as above. This finally gives us the optimal discount rate as

$$r_t = \delta + \eta\bar{g}_t - \varphi(\eta+1)\frac{\sigma_g}{2} - (\eta+1)log\left(\frac{1-I_t(\gamma)}{1-I_0(\gamma)}\right)^{1/t} \tag{22}$$

and can be calibrated using data on growth, its volatility, and aggregated indices of inequality over time.[12]

[12]A similar formula could be derived if the original welfare specification were replaced by a

From this equation, we can get an interpretation along the lines of prudence in the recursive utility framework: abstracting from uncertainty, this result suggests that the discount rate can be easily computed from the standard discount rate by substituting marginal utility of per-capita consumption $\overline{c}_t^{-\eta}$ by average marginal utility which can be expressed as $\overline{c}_t^{-\eta}(1 - I_t(\gamma))^{-(\eta+1)}$. Per capita marginal utility at any date t is thus multiplied by a factor larger than one that increases both in the degree of inequality aversion γ and η. In particular, using the definition of $I_t(\gamma)$, average marginal utility can be written as $\left(\overline{c}_t e^{\frac{\gamma}{\eta}(\eta+1)\frac{Var(E_t[lnc_{it}|c_{i0}])}{2}}\right)^{-\eta}$ where the coefficient for the variance term $\frac{\gamma}{\eta}(\eta+1)$ represents the degree of relative prudence of the social planner using our preference specification, see Kimball and Weil (2009). Similar to the context of uncertainty as in Kimball (1990), prudence matters for the direction and magnitude of the inequality effect. In the context of inequality, however, it is not the absolute size of the uncertainty but rather the relative degree of inequality today and in the future that determines the effect on the discount rate.

6 Calibration and Results

In order to get an idea of the magnitudes of each effect, we need to calibrate the three effects and use some sensible values for the triple of parameters (φ, γ, η). But which values for both parameters are reasonable? For the context of the discount rate, Barsky et al. (1997), Carlsson et al. (2005) and Atkinson et al. (2009) estimated individual values of η, γ, and φ elicited from experiments. Their results show a great variability between individuals. The study by Atkinson et al. (2009) is probably the most relevant to our problem since they focused including with their sample on the long-term issue of climate change. Their obtained median values of $\eta \sim 9$, $\gamma \sim 2-3$ and $\varphi \sim 3-5$ however need to be taken with care for the normative discount rate.

Firstly, the estimated degree of intertemporal elasticity of substitution would imply a way larger discount rate through the growth effect. In the literature on discounting and climate change based on macroeconomic calibrations, η has virtually never be considered to be smaller than one or larger than two (Just to name a few suggested values: Weitzman: 2, Dasgupta: 2-3, Nordhaus: 2, Stern: 1, Cline: 1.5). Secondly, the degree of inequity aversion seem on the upper end of values suggested earlier on the basis of actual policy decisions. For instance, Evans (2005) finds $\eta \sim 1.4$ based on income tax profiles in OECD countries while Clarkson and Deyes (2002) suggest a value of γ between 0.5 and 1.5. For inequality indices for the U.S., the U.S. Census Bureau (2010) publishes inequality indices using $\gamma = \{0.25, 0.5, 0.75\}$ and suggests one as maximum value. Also compare Atkinson and Brandolini (2010) who consider values as low as 0.2 and as high as 2.5 as defensible while Tol (2010) estimates γ to be around 0.7 based on development aid by OECD countries. Moreover, as pointed out by Tol (2005), the richer countries do not reveal

welfare function as proposed in Sen and Foster (1997). The authors introduced a welfare function where instead of considering per capita level of consumption $\overline{c}$, the Gini coefficient G or other inequality measures are used to compute a new measure of welfare as $\overline{c}(1 - G)$.

as much concern for the poor as is implied by the equity weights using higher values of γ.

From these results, a value below unity, take 0.5 as a point estimate, seems to be what might be a good compromise Finally, with regard to the degree of relative risk aversion, the estimates in this range seem plausible with regard to several empirical observations. While the effect of uncertainty is relatively small even for $\varphi = 4$, the growth effect dominates the discount rate and the choice of η is crucial. Moreover, the increasing inequality implies a lower discount rate since future generations are characterized not only by higher average income and more uncertainty, but also by a higher degree of income inequality. Since this effect is scaled by the degree of inequality aversion this parameter is also important. Now given the comparably small values for γ, the separation of the three preferences in particular allows to use, say, Nordhaus' $\eta = 2$ without having the inequality effect swamp away the growth effect due to inequality and the lack of flexibility of the model.

From our perspective and based on this discussion, the combination ($\eta = 2$, $\varphi = 4$, $\gamma = 0.5$) would be our preferred parametrization and we can finally calculate the actual discount rates.

First, we look at the frequently considered case of the United States. Regarding the growth of average or per capita income, we use estimates from Kocherlakota (1996), who estimates $g_t = 1.8\%$ and an annualized volatility of the growth rate of 3.6%.

The more difficult task is to estimate the changes in inequality over time. From the discussion about different concepts of convergence, the concept that should be employed in our case is the one focusing on the actual distribution of income, i.e., whether inequality increases or decreases over time ($\sigma-$convergence), see Young et al. (2008). Based on estimates over the period of 1967-2009, the U.S. Census Bureau (2010) estimates that the Gini coefficient increased by 17.9%. For the changes in the income distribution, we need the variance of log-income, one of the set income inequality measures typically used. If it is not reported, however, we can recover it e.g., from the Gini coefficient based on the lognormal assumption by inverting the formula $G = 2\Phi(\frac{\sigma}{\sqrt{2}}) - 1$. Using their data, we obtain the variance of log income for the last forty years as depicted in Figure 4.

Inequality within the United States has been increasing over the last decades and the trend can be quite well approximated linearly. The slope of this trend is estimated at 0.0049 via OLS. Moreover, the linearity implies that the term structure will be flat. Finally, it is noteworthy to study the long-term implications of the assumed evolvement of inequality. Using this estimate, the Gini coefficient of the U.S., which rose from 0.39 in 1970 to 0.45 in 2009 is expected to further increase to a value of 0.55 in 2100. This increase does not seem to unreasonable given the expected large overall growth and which is in line with the general property of inequality indices to remain rather stable over time.

Using this calibrated values, we can compute the optimal consumption discount rate as

$$r_t = \delta + \eta \cdot 1.8\% - \varphi(\eta + 1) \cdot 0.065\% - \gamma(\eta + 1) \cdot 0.25\%.$$

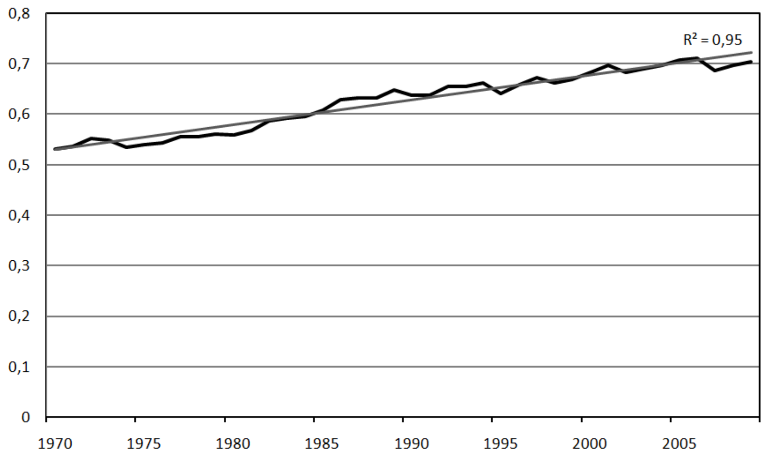

Figure 4: Variance of log income in the U.S.

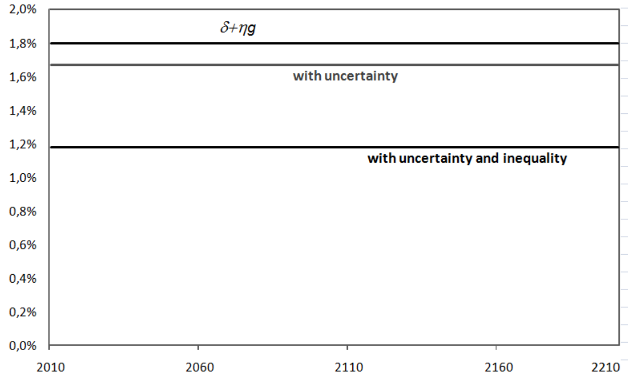

Figure 5: Optimal Social Discount rate for the U.S. ($\eta = 1, \varphi = 1, \gamma = 1$)

The growth and uncertainty effect are the same as in previous models. In particular, the uncertainty effect is negligible given the assumption about the independence of growth rates over time. The inequality effect on the other hand is much more important. First, it is negative thus reduces the discount rate since inequality is expected to increase in the future. Second, even though the estimated slow increase of inequality does not sound unrealistic, its magnitude is significant. Moreover, as argued before, the term structure is flat due to the linear increase in the variance of log income together with constant growth and volatility. Finally, the discount rate depends crucially on the choice of the three parameters for the different dimensions of inequity aversion. For a benchmark case of unity for all three parameters, the optimal discount rate (maintaining $\delta = 0\%$) is depicted in Figure 5, which shows the relative magnitude of the three effects.

For our preferred parametrization ($\eta = 2$, $\varphi = 4$, $\gamma = 0.5$), we finally obtain a value for the optimal discount rate for the U.S. economy of 2.5%.

As pointed out above, the results depend crucially on the regional aggregation. In the following, we thus look at the global rate of discount. As outlined earlier, on the global level, inequalities in the sense used in this paper have been reduced, compare,

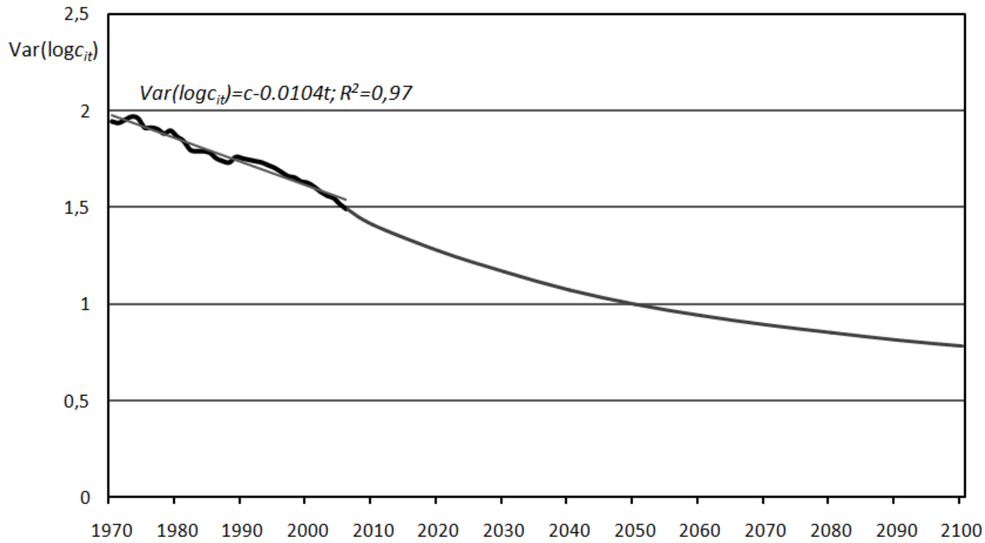

Figure 6: Global Inequality from 1970 to 2100

e.g., Pinkovskiy and Sala-I-Martin (2009) who found significant $\sigma-$convergence for the world as a whole. Based on their estimates, we compute the variance of log income which is depicted in Figure 6. The convergence is clearly visible. Moreover, according to projections used in the context of climate change, see Calzadilla (2010), this pattern is expected to continue. For instance, for Nordhaus (2010)' RICE model which uses GDP projections based on the SRES A2 scenario (similar as the one presented above), the Figure shows how convergence is expected to continue.

Based on this prediction, we can compute the optimal discount rate with Nordhaus' specification of a pure rate of time preference of $\delta = 1.5\%$ and elasticity of marginal utility of $\eta = 1.5$.[13] The resulting discount rate is depicted for three values of γ in Figure 7. As noted in the last section, the term structure is decreasing.[14] Now the disentanglemenf of η and γ has a significant effect of the magnitude. For $\gamma = \eta$, the discount rate almost doubles whereas for a moderate value of $\gamma = 0.5$ the effect is much lower and adds around one percentage point to $\tilde{r}_t$.

7 Conclusion

This paper discusses the role of a global discount rate when taking into account income inequality. First we derived the appropriate weighting scheme for the country-specific discount rates using an Utilitarian Social Welfare Function. We find that each country's contribution to the expectations about the future needs to be weighted by today's marginal utility giving poorer countries a higher weight. Moreover, the

[13] In DICE there is no explicit uncertainty so the discount rate consists only in the classical Ramsey rule and the inequality term.

[14] In RICE, the growth rates of all regions are slowly declining over time which explains the decreasing term structure of $\tilde{r}_t$ without inequality considerations.

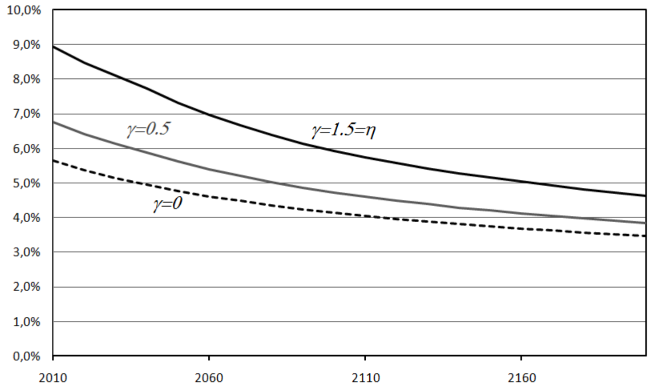

Figure 7: Inequality adjusted discount rate for the DICE 2010 model

evolvement of inequality over time is pivotal for the term structure of the discount rate. An expected deterioration of income inequality between countries over time implies under some fairly general conditions a smaller discount rate than if inequality remained constant or increased. Analytically, what matters is the relative effect of inequality on average marginal utility today and in the future. The so-called degree of aversion to downside inequality of the social planner turned out to be together with the variance of income the decisive characteristic. If the dispersion of income evaluated as a downside inequality at the mean income increases over time, the discount rate should be lower than in the case without inequality and vice versa.

A calibration of the results shows that considering inequality would imply a higher discount rate using assumptions typically used in the context of climate change. In particular we find use the SRES A2 scenario, which projects on average 1.6% GDP growth and a relatively strong speed of convergence. Using a typical specification ($\delta = 0.1\%, \eta = 1$) this would imply a global discount rate of 1.7% whereas the optimal equity-weighted discount rate yields 3.0% for short time horizons and decreases to 2.7% over two centuries. Taking into account income inequality thus could justify the use of a discount rate almost twice as large.

We extend this model allowing for different preferences of inequality aversion across time, states of nature, and space. In particular, it seems appropriate to use a lower degree of inequality aversion than resistance to intertemporal substitution and risk aversion. Otherwise, revealed social preferences in other policy areas such as development aid would be contradictory to the specification used. This is even more important since the discount rate is very sensitive to the degree of inequality aversion. Consequently, taking a more considerate value of $\gamma = 0.5$ reduces the effect of inequality.

These results suggest that using an initially higher but decreasing discount rate could be rationalized when considering inequality also within and not only between generations. In the context of climate change, this would ceteris paribus ask for a lower optimal social cost of carbon arising from the consideration of poorer countries' time preference. This provides a theoretical foundation for the results of Hope (2008) who finds that appropriate equity weighting in the context of climate change can including reduce rather than increases the optimal social cost of carbon due to the effect on the discount rate.

References

Atkinson, Anthony B., "On the measurement of inequality," *Journal of Economic Theory*, 1970, *2* (3), 244–263.

_ **and Andrea Brandolini**, "On Analyzing the World Distribution of Income," *World Bank Econ Rev*, January 2010, *24* (1), 1–37.

Atkinson, Giles D., Simon Dietz, Jennifer Helgeson, Cameron Hepburn, and Håkon Sælen, "Siblings, not triplets: social preferences for risk, inequality and time in discounting climate change," *Economics - The Open Access, Open Assessment E-Journal*, 2009, *3* (26), 1–29.

Azar, Christian and Thomas Sterner, "Discounting and distributional considerations in the context of global warming," *Ecological Economics*, November 1996, *19* (2), 169–184.

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics*, May 1997, *112* (2), 537–579.

Bourguignon, Francois, Victoria Levin, and David Rosenblatt, "Global redistribution of income," Technical Report, The World Bank July 2006.

Bureau, The U.S. Census, "Income, poverty, and health insurance coverage in the United States: 2009," Technical Report 2010.

Calzadilla, Alvaro, "Global Income Distribution and Poverty: Implications from the IPCC SRES Scenarios," Kiel Working Papers 1664, Kiel Institute for the World Economy November 2010.

Carlsson, Fredrik, Dinky Daruvala, and Olof Johansson-Stenman, "Are People Inequality-Averse, or Just Risk-Averse?," *Economica*, 2005, *72* (287), 375–396.

Clarkson, Richard and Kathryn Deyes, "Estimating the social cost of carbon emissions," Technical Report 140, HM Treasury 2002.

Conceiçao, Pedro and Yanchun Zhang, "Discounting in the context of climate change economics: the policy implications of uncertainty and global asymmetries," *Environmental Economics and Policy Studies*, June 2010, *12* (1-2), 31–57.

Crainich, David and Louis Eeckhoudt, "On the intensity of downside risk aversion," *Journal of Risk and Uncertainty*, June 2008, *36* (3), 267–276.

Davies, James and Michael Hoy, "Making Inequality Comparisons When Lorenz Curves Intersect," *The American Economic Review*, September 1995, *85* (4), 980–986.

Epstein, Larry G. and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, July 1989, *57* (4), 937–969.

_ **and** _ , "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *The Journal of Political Economy*, April 1991, *99* (2), 263–286.

Evans, David J, "The Elasticity of Marginal Utility of Consumption: Estimates for 20 OECD Countries*," *Fiscal Studies*, June 2005, *26* (2), 197–224.

Fankhauser, Samuel, Richard S.J. Tol, and David W. Pearce, "The Aggregation of Climate Change Damages: a Welfare Theoretic Approach," *Environmental and Resource Economics*, October 1997, *10* (3), 249–266.

Gierlinger, Johannes and Christian Gollier, "Socially Efficient Discounting under Ambiguity Aversion," Technical Report 561, IDEI, Toulouse, Toulouse November 2008.

Gollier, Christian, "Discounting an uncertain future," *Journal of Public Economics*, August 2002, *85* (2), 149–166.

_ , "The consumption-based determinants of the term structure of discount rates," *Mathematics and Financial Economics*, July 2007, *1* (2), 81–101.

_ , "Discounting with fat-tailed economic growth," *Journal of Risk and Uncertainty*, December 2008, *37* (2), 171–186.

_ , "Discounting, inequalities and economic convergence," Technical Report 3262, CESifo, Munich 2010.

_ , "Ecological discounting," *Journal of Economic Theory*, 2010, *145* (2), 812–829.

_ , *Pricing the future: The economics of discounting and sustainable development*, Princeton University Press, 2011.

_ **and Martin L Weitzman**, "How should the distant future be discounted when discount rates are uncertain?," *Economics Letters*, 2010, *107* (3), 350–353.

_ **and Richard Zeckhauser**, "Aggregation of Heterogeneous Time Preferences," *Journal of Political Economy*, 2005, *113* (4), 878–896.

Ha-Duong, Minh and Nicolas Treich, "Risk Aversion, Intergenerational Equity and Climate Change," *Environmental and Resource Economics*, June 2004, *28* (2), 195–207.

Hope, Chris, "Discount rates, equity weights and the social cost of carbon," *Energy Economics*, May 2008, *30* (3), 1011–1019.

Jouini, Elyès, Jean-Michel Marin, and Clotilde Napp, "Discounting and divergence of opinion," *Journal of Economic Theory*, 2010, *145* (2), 830–859.

Keenan, Donald C. and Arthur Snow, "Greater Downside Risk Aversion," *Journal of Risk and Uncertainty*, 2002, *24* (3), 267–277.

Kimball, Miles S., "Precautionary Saving in the Small and in the Large," *Econometrica*, January 1990, *58* (1), 53–73.

Kimball, Miles S, "Standard Risk Aversion," *Econometrica*, 1993, *61* (3), 589–611.

_ **and Philippe Weil**, "Precautionary Saving and Consumption Smoothing across Time and Possibilities," *Journal of Money, Credit and Banking*, March 2009, *41* (2-3), 245–284.

Kocherlakota, Narayana R., "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 1996, *34* (1), 42–71.

Kolm, Serge-Christophe, "Unequal inequalities. I," *Journal of Economic Theory*, June 1976, *12* (3), 416–442.

Kreps, David M. and Evan L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, January 1978, *46* (1), 185–200.

Menezes, C., C. Geiss, and J. Tressler, "Increasing Downside Risk," *American Economic Review*, 1980, *70* (5), 921–932.

Nordhaus, William D., "A Review of the Stern Review on the Economics of Climate Change," *Journal of Economic Literature*, 2007, *45* (3), 686–702.

_ , "Economic aspects of global warming in a post-Copenhagen environment," *Proceedings of the National Academy of Sciences*, June 2010.

Pinkovskiy, Maxim and Xavier Sala-I-Martin, "Parametric Estimations of the World Distribution of Income," Technical Report w15433 October 2009.

Ramsey, F. P., "A Mathematical Theory of Saving," *The Economic Journal*, December 1928, *38* (152), 543–559.

Sala-I-Martin, Xavier, "The Disturbing "Rise" of Global Income Inequality," Technical Report 8904, National Bureau of Economic Research April 2002.

Schelling, Thomas C, "Intergenerational discounting," *Energy Policy*, April 1995, *23* (4-5), 395–401.

Sen, Amartya Kumar and James Eric Foster, *On economic inequality*, Oxford University Press, 1997.

Shorrocks, Anthony F. and James E. Foster, "Transfer Sensitive Inequality Measures," *The Review of Economic Studies*, July 1987, *54* (3), 485–497.

Spackman, Michael, "Public lnvestment and Discounting in European Union Member States," *OECD Journal of Budgeting*, 2001, *1* (2), 213–260.

Stern, Nicholas et al., *Stern Review: The Economics of Climate Change*, London: HM Treasury, 2006.

Sterner, Thomas and U. Martin Persson, "An Even Sterner Review: Introducing Relative Prices into the Discounting Debate," Technical Report, Resources For the Future July 2007.

Tol, Richard S.J., "The marginal damage costs of carbon dioxide emissions: an assessment of the uncertainties," *Energy Policy*, November 2005, *33* (16), 2064–2074.

_ , "International inequity aversion and the social cost of carbon," *Climate Change Economics (CCE)*, 2010, *1* (01), 21–32.

Traeger, Christian P., "Recent Developments in the Intertemporal Modeling of Uncertainty," *Annual Review of Resource Economics*, October 2009, *1* (1), 261–286.

UNDP, "Human Development Report 2007/2008. Fighting climate change: Human solidarity in a divided world," *Human Development*, 2007.

Wang, X. Henry and Carmen Menezes, "The Precautionary Premium and the Risk-Downside Risk Tradeoff," Technical Report 0204a 2002.

Weitzman, Martin L., "Why the Far-Distant Future Should Be Discounted at Its Lowest Possible Rate,," *Journal of Environmental Economics and Management*, November 1998, *36* (3), 201–208.

_ , "Gamma Discounting," *The American Economic Review*, March 2001, *91* (1), 260–271.

Weitzman, Martin L, "On Modeling and Interpreting the Economics of Catastrophic Climate Change," *Review of Economics and Statistics*, February 2009, *91* (1), 1–19.

Young, Andrew T, Matthew J Higgins, and Daniel Levy, "Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data," *Journal of Money, Credit and Banking*, August 2008, *40* (5), 1083–1093.

Zhuang, Juzhong, Zhihong Liang, Tun Lin, and Franklin De Guzman, "Theory and Practice in the Choice of Social Discount Rate for Cost-Benefit Analysis: A Survey," *ERD Working Paper (ADB)*, 2007, *No. 94*.