University of Sussex
Business, Management & Economics

Economics Department Working Paper Series

No. 38-2012

Climate Policy with Bentham-Rawls Preferences

Richard S. J. Tol

Department of Economics

University of Sussex

R.Tol@sussex.ac.uk

and

Institute for Environmental Studies

Department of Spatial Economics

Vrije Universiteit

Amsterdam, The Netherlands

Abstract: A Bentham-Rawls welfare function is the weighted sum of the net present welfare (Bentham) and the welfare of the worst-off generation (Rawls). If utility is non-decreasing over time, optimal climate policy is more stringent in the near-term under Bentham preferences than under Bentham-Rawls preferences. If utility is decreasing, Bentham-Rawls abatement is higher. If there is a chance of decreasing utility, Bentham-Rawls optimal climate policy is probably less stringent than Bentham policy.

JEL Classification: Q54

Key Words: climate policy; social cost of carbon; Bentham-Rawls preferences

1. Introduction

Optimal climate policy is derived from an intertemporal welfare function. The pure rate of time preference used is hotly debated (Arrow et al. 1996;Nordhaus 2007;Stern and Taylor 2007). (Alvarez-Cuadrado and Van Long 2009) derive a welfare function that uses conventional discounting but is potentially sensitive to the very long run nonetheless. I here investigate the implications of their proposal for climate policy.

An intertemporal welfare function cannot simultaneously satisfy the two conditions of Strong Pareto and anonymity: either one prefers a situation in which one generation is better off and none worse off (Strong Pareto), or one is sensitive to a re-ordering of generations (anonymity) (van Liedekerke and Lauwers 1997). Because generations arrive in order, it is natural to violate anonymity, and discounted utilitarianism is the default choice (Koopmans 1960;Koopmans 1966;Koopmans 1967). (Asheim and Mitra 2010;Zuber and Asheim 2012) define welfare functions that satisfies anonymity (and hence violate Pareto).

(Chichilnisky 1996) replaces anonymity with weaker axioms of non-dictatorship and independence. (Alvarez-Cuadrado and Van Long 2009) further refine this, dropping independence and thus expanding the model space with optimal solutions. They derive the Bentham-Rawls utility function, which is the weighted sum of conventional net present welfare and the welfare of the worst-off generation $\underline{U}$:

(1) $$W = (1-\vartheta)\int_t U\big(C(t)\big)\mathrm{d}t + \vartheta \underline{U}(\underline{C})$$

where U is instantaneous utility, C is consumption and ϑ is the Rawls weight, the weight attached to the welfare of the poorest generation. If ϑ=0, Equation (1) reverts to the standard net present welfare, inspired by (Bentham 1789). If ϑ=1, Equation (1) describes a maximin problem inspired by (Rawls 1972). I refer to Equation (1) as Bentham-Rawls preferences, to its first component as Bentham preferences, and to its second component as Rawls preferences. Derivatives etc are referred to in the same way.

In Section 2, I analytically explore the implications for climate policy. I add numbers in Section 3. Section 4 concludes.

2. Analytical results

2.1. Preliminaries

Optimal emission reduction follows from equating the marginal costs of emission reduction to its marginal benefits. To a first approximation, the marginal costs are instantaneous:

(2) $$MC = \begin{cases} \dfrac{\partial U(0)}{\partial C(0)}\dfrac{\partial C(0)}{\partial E(0)} \text{ if } \underline{U} = U(0) \\ (1-\vartheta)\dfrac{\partial U(0)}{\partial C(0)}\dfrac{\partial C(0)}{\partial E(0)} \text{ if } \underline{U} \neq U(0) \end{cases}$$

where E are emissions.

The marginal benefits are a stream of future benefits:

(3) $$MB = (1-\vartheta)\int_t \frac{\partial U(t)}{\partial C(t)}\frac{\partial C(t)}{\partial T(t)}\frac{\partial T(t)}{\partial M(t)}\frac{\partial M(t)}{\partial E(0)}\mathrm{d}t + \vartheta \frac{\partial \underline{U}(t^*)}{\partial C(t^*)}\frac{\partial C(t^*)}{\partial T(t^*)}\frac{\partial T(t^*)}{\partial M(t^*)}\frac{\partial M(t^*)}{\partial E(0)}$$

where T is temperature (or any other indicator of climate change), M is the atmospheric concentration of greenhouse gas emissions, and t^* is the time at which instantaneous utility is minimum. Equation (3) defines the Bentham-Rawls social cost of carbon, the weighted sum of the conventional, Bentham social cost of carbon and the Rawls social cost of carbon.

There are two distinct cases. In the first, utility is increasing, or at least non-decreasing. The first generation, us, is the poorest, $t^*=0$. In the second case, utility increases at first but falls in the future because of climate change.[1] Some future generation is the poorest, $t^*>0$. I discuss both cases in turn.

2.2. *Utility is non-decreasing*

If utility is non-decreasing over time, $MC_{\vartheta>0}=MC_{\vartheta=0}$. That is, marginal abatement costs are the same whether one uses Bentham welfare or Bentham-Rawls welfare.

Climate change is a slow process. Because of the large heat capacity of the ocean, warming is very gradual. Therefore, to a first approximation:

(4) $$\frac{\partial T(0)}{\partial M(0)}\frac{\partial M(0)}{\partial E(0)} = 0$$

This is implies that $MB_{\vartheta>0}=(1-\vartheta)MB_{\vartheta=0}$. Thus, in the optimum,

(5) $$MC^*_{\vartheta>0} - MC^*_{\vartheta=0} = -\vartheta MB_{\vartheta=0}$$

That is, if utility is non-decreasing, Bentham-Rawls optimal emission control is below Bentham optimal emission control.

If the emission reduction cost function is quadratic, a not unreasonable assumption, the Bentham-Rawls optimal emission control rate is $(1-\vartheta)$ times the Bentham optimal control rate.

2.3. *Utility is decreasing*

If utility is decreasing over time, $MC_{\vartheta>0}= (1-\vartheta)MC_{\vartheta=0}$.That is, marginal abatement costs are lower under Bentham-Rawls welfare than under Bentham welfare. The abatement costs are partly ignored because they do not fall on the poorest generation.

In this case, the marginal benefits $MB_{\vartheta>0}=(1-\vartheta)MB_{\vartheta=0}+\vartheta\underline{MB}$.

Therefore, in the respective optima:

(6) $$MC^*_{\vartheta>0} - MC^*_{\vartheta=0} = \frac{\vartheta}{1-\vartheta}\underline{MB}$$

That is, if the worst-off generation faces negative impacts of climate change, at the margin, then Bentham-Rawls optimal control is more stringent than Bentham optimal emission control.

2.4. *Utility may be decreasing*

[1] If utility would fall for some other reason than climate change, we should focus on that problem instead.

If we assume that there is a probability p that utility is decreasing over time, then, combining conditions (5) and (6) and reworking, we find that for ϑ>0 and p<1emission control is more stringent under Bentham-Rawls welfare than under Bentham welfare if:

(7) $$\frac{p}{1-p}\frac{1}{1-\vartheta}\frac{\underline{MB}}{MB_{\vartheta=0}} > 1$$

This result I s intuitive. The Rawls part of Bentham-Rawls welfare has a greater effect on optimal emission control if:

1. there is a greater chance of decreasing utility;
2. a greater weight is placed on the utility of the worst-off generation; or
3. the worst-off generation is hit harder, at the margin.

Depending on the parameters chosen, Bentham-Rawls welfare may thus lead to more or less stringent emission reduction. However, the probability of decreasing utility due to climate change is small; and the weight placed on the plight of the worst-off generation is likely to be small too. That implies that the product of the first two terms of Equation (7) is much smaller than one. In other words, the marginal damage to the worst-off generation has to be much larger than the net present value of the marginal damage to all generations. If not, Bentham-Rawls welfare implies less stringent climate policy. In the next section, I further explore Equation (7), adding estimates of the numbers.

3. Numerical results

Equation (7) has four elements: (1) the probability of decreasing utility p, (2) the Rawls weight ϑ, (3) the Rawls social cost of carbon $\underline{MB}$ and (4) the Bentham social cost of carbon $MB_{\vartheta=0}$. Table 1 shows selected results of a recent meta-analysis of estimates of the Bentham social cost of carbon (Tol 2012). There is substantial uncertainty about the estimates. The pure rate of time preference is a crucial parameter. For a 3% PRTP, the mean Bentham social cost of carbon is \$25/tC. For a 0% PRTP, it is almost 12 times as large: \$296/tC.

The Rawls social cost of carbon is defined as:

(8) $$\underline{MB} = \frac{\partial \underline{U}(t^*)}{\partial C(t^*)}\frac{\partial C(t^*)}{\partial T(t^*)}\frac{\partial T(t^*)}{\partial M(t^*)}\frac{\partial M(t^*)}{\partial E(0)}$$

The climate is a complex dynamic system. A general analytical solution to Equation (8) cannot be found. However, a solution can be had if we assume that the worst-off generation is placed in the distant future, long after emissions have ceased and the atmosphere is again in equilibrium with the ocean. Under those assumptions,

(9) $$\lim_{t\to\infty, E(t)\downarrow 0} M(t) = \alpha \int_{s=0}^{t} E(s)\mathrm{d}s$$

where α is proportional to the fraction of carbon dioxide that stays in the atmosphere forever. α is also the first partial derivative of the equilibrium concentration to current emissions, that is, the fourth term in Equation (8). About 13% of carbon dioxide emissions does not degrade in the atmosphere (Maier-Reimer and Hasselmann 1987). Emissions of 2.13 10^9 tonne of carbon correspond to 1 part per million (by volume) concentration.[2]

[2] http://cdiac.ornl.gov/pns/convert.html

Equilibrium warming is proportional to the logarithm of the concentration:

(10) $$\lim_{t\to\infty,E(t)\downarrow 0} T(t) = \beta \ln M(t) \Rightarrow \frac{\partial T(t)}{\partial M(t)} = \frac{\beta}{M(t)}$$

Figure 1 shows estimates of reserves and resources (proven, probable and possible) of fossil fuels (WEC 2010), converted into the atmospheric concentration of carbon dioxide. The current concentration is 380 ppm, but if emissions would stop now, the concentration would fall to 300 ppm, 20 ppm above the pre-industrial concentration. If only the reserves of oil and gas (and coal) are burned, the permanent concentration will increase to 320 (360) ppm. If all resources are burned too, the permanent concentration will rise to 510 ppm. The peak concentrations would be much higher: 500 (oil and gas reserves), 800 (all reserves) and 2000 (all resources) ppm.

Figure 2 shows estimates of the impact of climate change, and two fitted functions: $I=\kappa T+\lambda T^2$ as proposed by (Tol 2009) and $I=\kappa' T^2+\lambda' T^6$, as proposed by (Weitzman 2012). The second term in Equation (8) is thus $\kappa+2\lambda T$ or $2\kappa' T+6\lambda' T^5$.

Assuming a CRRA utility function, the first component of Equation (8) is $C^{-\eta}$ where η is the rate of risk aversion and C is per capita consumption in the long term, say $350 per person per year.

Finally, the Rawls social cost of carbon in Equation (8) is measured in utils per tonne of carbon. I convert this to dollar per tonne of carbon by multiplication with the inverse of marginal utility in 2010 (Anthoff and Tol 2010), using the world average income of $7,000 per person per year. I further multiply with the number of people in that year ($6.7\ 10^9$), as the Rawls social cost of carbon in Equation (8) is in fact conceptualized as the impact on the worst-off person.

Table 2 shows estimates of the Rawls social cost of carbon for a variety of assumptions. The results make intuitive sense: more severe climate change and higher risk aversion imply a higher Rawls social cost of carbon. The Rawls social cost of carbon is rather low, perhaps even negative for the Tol impact function. It is much higher for the Weitzman impact function. The Rawls social cost of carbon is always less than 10 times the Bentham social cost of carbon (Table 1).

Returning to Equation (7), if the third term on the left-hand side is less than 10, then Rawls-Bentham preferences lead to lower emission abatement unless the probability of decreasing utility is fairly high. Figure 3 shows the relationship. If the Rawls social cost of carbon is as high as the maximum of Table 2 and the Bentham social cost of carbon is as low as the minimum of Table 1, then for a Rawls weight of 5%, Bentham-Rawls abatement is lower than Bentham abatement unless the probability of decreasing utility is greater than 9%. The break-even probability decreases if the Rawls weight increases. The break-even probability increases if the Rawls social cost of carbon decreases (cf. the dashed line in Figure 3).

(Fankhauser and Tol 2005) explore the impact of climate change on economic growth in a number of growth models. They find that it is unlikely for climate change to overwhelm growth. Figure 2 shows that most impact estimates of climate change are relatively small – particularly considering that a 3°C warming would take decades or longer to materialize. Furthermore, direct impacts on welfare, which do not affect growth, form a large share of these estimates.

Therefore, even if we stack the cards in favour of high Bentham-Rawls abatement, it is likely that Benthan-Rawls abatement is lower than Bentham abatement.

4. Conclusion

(Alvarez-Cuadrado and Van Long 2009) introduced the Bentham-Rawls intertemporal welfare function that would better reflect long-term concerns than standard, Bentham net present welfare. I apply this to climate policy. If utility is non-decreasing over time, Bentham-Rawls preferences discount the impact of climate change as it falls on future, better-off generations but not the impact of emission reduction; optimal control is lower as a result. If utility is decreasing over time, Bentham-Rawls preferences imply stricter climate policy. If there is a chance of decreasing utility, the result is ambiguous. However, reasonable parameterizations indicate that in all likelihood Bentham-Rawls preferences call for less emission abatement than Bentham preferences.

There are a number of caveats. I used a static, steady-state model of carbon cycle and climate. A dynamic model would be more appropriate – utility may grow then shrink then grow again – but current integrated assessment models all work with relatively short time horizons and do not properly represent the impact of climate change on development. The treatment of population and uncertainty are ad hoc, and this may bias the result. I assumed that the first partial derivative of the chance of decreasing utility to current emissions is negligibly small. Bentham-Rawls preferences put a premium on avoiding decreasing utility. This would require a realistically parameterized model of the impact of climate change on economic growth and shrink. These matters are all postponed to future research.

References

Alvarez-Cuadrado, F. and N.Van Long (2009), 'A mixed Bentham-Rawls criterion for intergenerational equity: Theory and implications', *Journal of Environmental Economics and Management*, **58**, (2), pp. 154-168.

Anthoff, D. and R.S.J.Tol (2010), 'On international equity weights and national decision making on climate change', *Journal of Environmental Economics and Management*, **60**, (1), pp. 14-20.

Arrow, K.J., W.R.Cline, K.-G.Maeler, M.Munasinghe, R.Squitieri, and J.E.Stiglitz (1996), 'Intertemporal Equity, Discounting, and Economic Efficiency', in *Climate Change 1995: Economic and Social Dimensions -- Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, J.P. Bruce, H. Lee, and E.F. Haites (eds.), Cambridge University Press, Cambridge, pp. 125-144.

Asheim, G.B. and T.Mitra (2010), 'Sustainability and discounted utilitarianism in models of economic growth', *Mathematical Social Sciences*, **59**, (2), pp. 148-169.

Bentham, J. (1789), *An Introduction to the Principles and Morals of Legislation* T. Payne, London.

Chichilnisky, G. (1996), 'An Axiomatic Approach to Sustainable Development', *Social Choice and Welfare*, **13**, (2), 219-248.

Fankhauser, S. and R.S.J.Tol (2005), 'On climate change and economic growth', *Resource and Energy Economics*, **27**, (1), 1-17.

Koopmans, T.C. (1960), 'Stationary Ordinal Utility and Impatience', *Econometrica*, **28**, (2), pp. 287-309.

Koopmans, T.C. (1966), 'On the concept of optimal economic growth', in *The Econometric Approach to Development Planning*, North Holland, Amsterdam.

Koopmans, T.C. (1967), 'Objectives, Constraints, and Outcomes in Optimal Growth Models', *Econometrica*, **35**, 1-15.

Maier-Reimer, E. and K.Hasselmann (1987), 'Transport and Storage of Carbon Dioxide in the Ocean: An Inorganic Ocean Circulation Carbon Cycle Model', *Climate Dynamics*, **2**, 63-90.

Nordhaus, W.D. (2007), 'Critical Assumptions in the Stern Review on Climate Change', *Science*, **317**, 201-202.

Rawls, J. (1972), *A Theory of Justice* Cambridge University Press, Cambridge.

Stern, N.H. and C.Taylor (2007), 'Climate Change: Risks, Ethics and the Stern Review', *Science*, **317**, (5835), 203-204.

Tol, R.S.J. (2009), 'The Economic Effects of Climate Change', *Journal of Economic Perspectives*, **23**, (2), 29-51.

Tol, R.S.J. (2012), *Targets for Global Climate Policy: An Overview*, Working Paper **3712** ,Department of Economics, University of Sussex, Falmer.

van Liedekerke, L. and L.Lauwers (1997), 'Sacrificing the patrol: Utilitarianism, future generations and infinity', *Economics and Philosophy*, **13**, pp. 159-174.

WEC (2010), *2010 Survey of Energy Resources* ,World Energy Council, London.

Weitzman, M.L. (2012), 'GHG Targets as Insurance Against Catastrophic Climate Damages', *Journal of Public Economic Theory*, **14**, (2), pp. 221-244.

Zuber, S. and G.B.Asheim (2012), 'Justifying social discounting: The rank-discounted utilitarian approach', *Journal of Economic Theory*, **147**, (4), pp. 1572-1601.

Table 1. Selected characteristics of the Bentham social cost of carbon for alternative pure rates of time preference.

	3%	1%	0%
Mean	25 $/tC	105 $/tC	296 $/tC
Mode	19 $/tC	55 $/tC	144 $/tC
Median	23 $/tC	83 $/tC	247 $/tC
Standard deviation	22 $/tC	128 $/tC	309 $/tC

Table 2. Rawls social cost of carbon under various assumptions on the impact function, the curvature of the utility function, the atmospheric concentration experienced by the worst-off generation, and the climate sensitivity.

RSCC ($/tC)		Parameters		
for impact function		η	M	CS
Tol	Weitzman	-	ppm	K/2xCO_2
0.1	20.1	1.5	800	3.0
0.0	4.5	1.0	800	3.0
0.7	90.1	2.0	800	3.0
-0.5	1.2	1.5	500	3.0
0.4	190.9	1.5	2000	3.0
-0.2	0.2	1.5	800	1.5
1.0	235.1	1.5	800	4.5

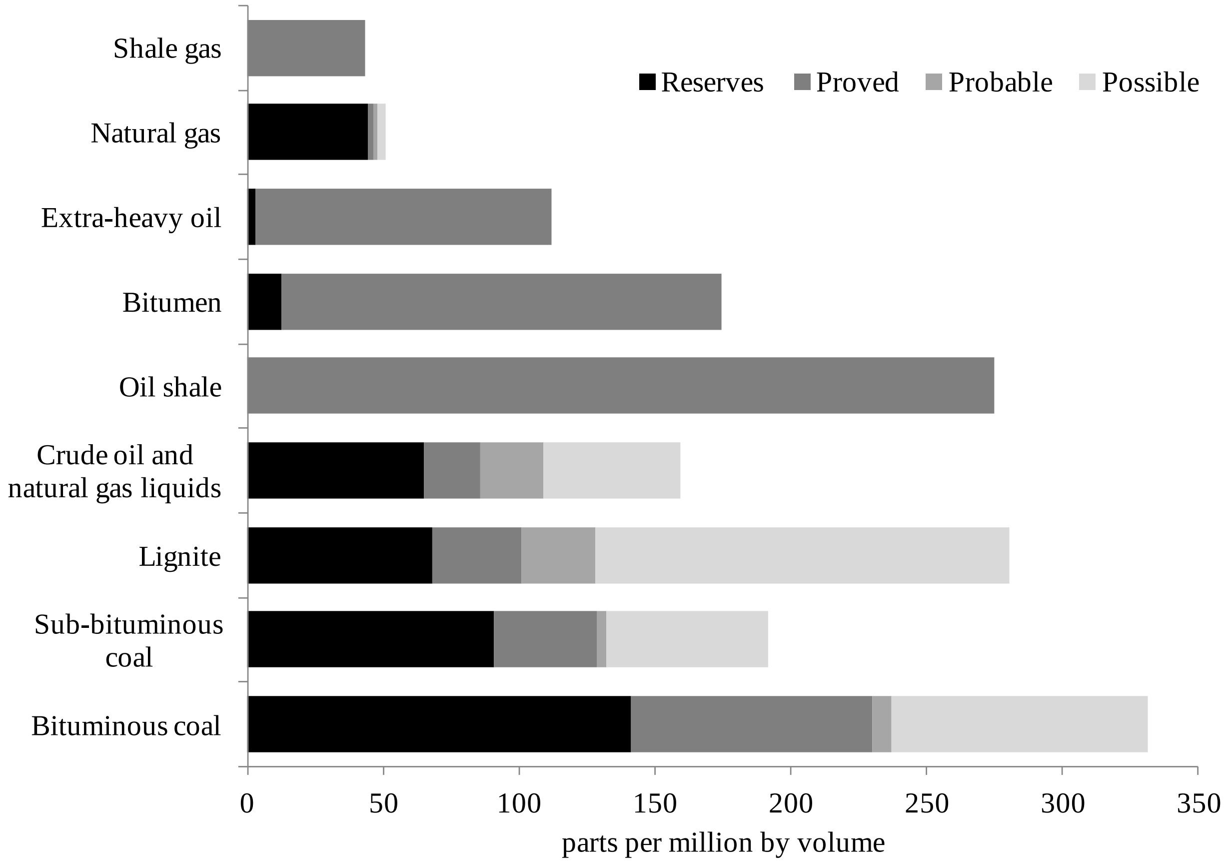

Figure 1. The atmospheric concentration of carbon dioxide corresponding to conventional and unconventional reserves and recoverable resources (proved, probable and possible) of coal, oil and gas.

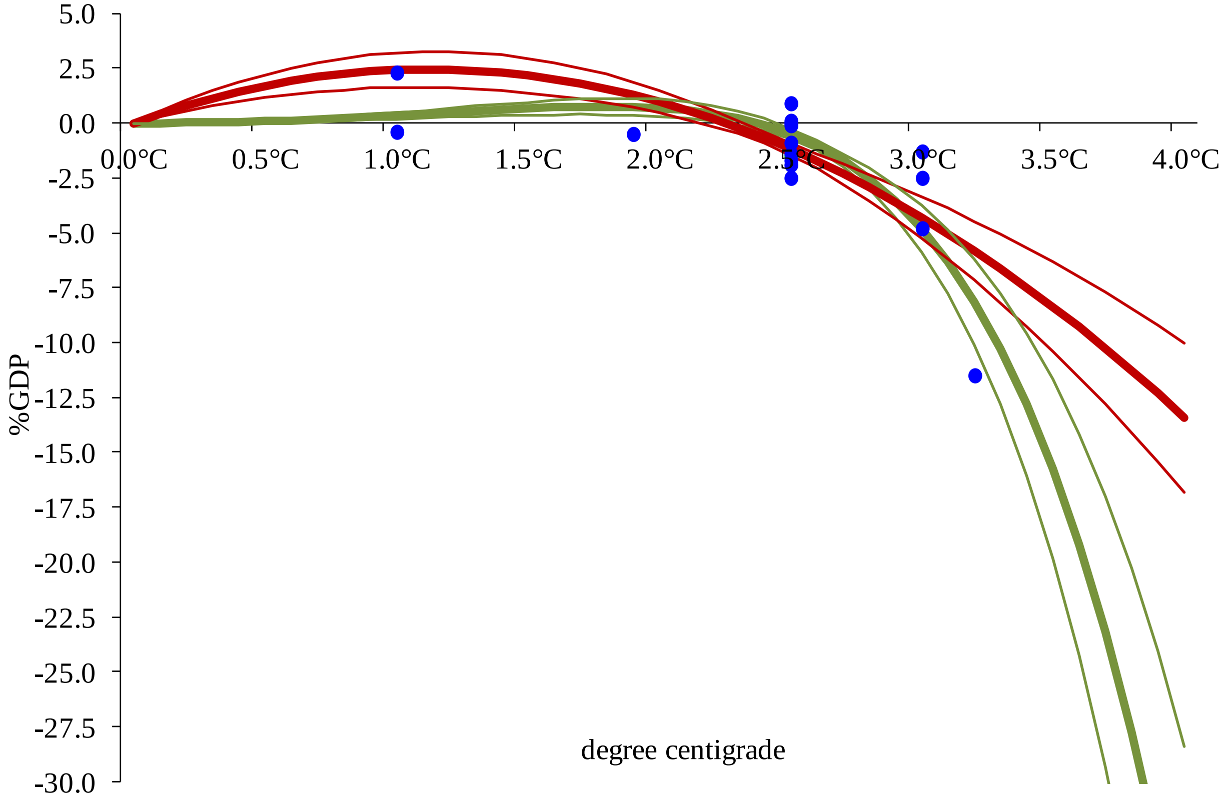

Figure 2. Estimates of the global economic impact of climate change (blue dots) and two fitted functions: I=4.33(1.49)T-1.92(0.56)T^2 (red line) and I=0.348(0.166)T^2-0.0109(0.0025)T^6 (green line); the thin lines demarcate the 95% confidence interval based on the bootstrapped standard deviation.

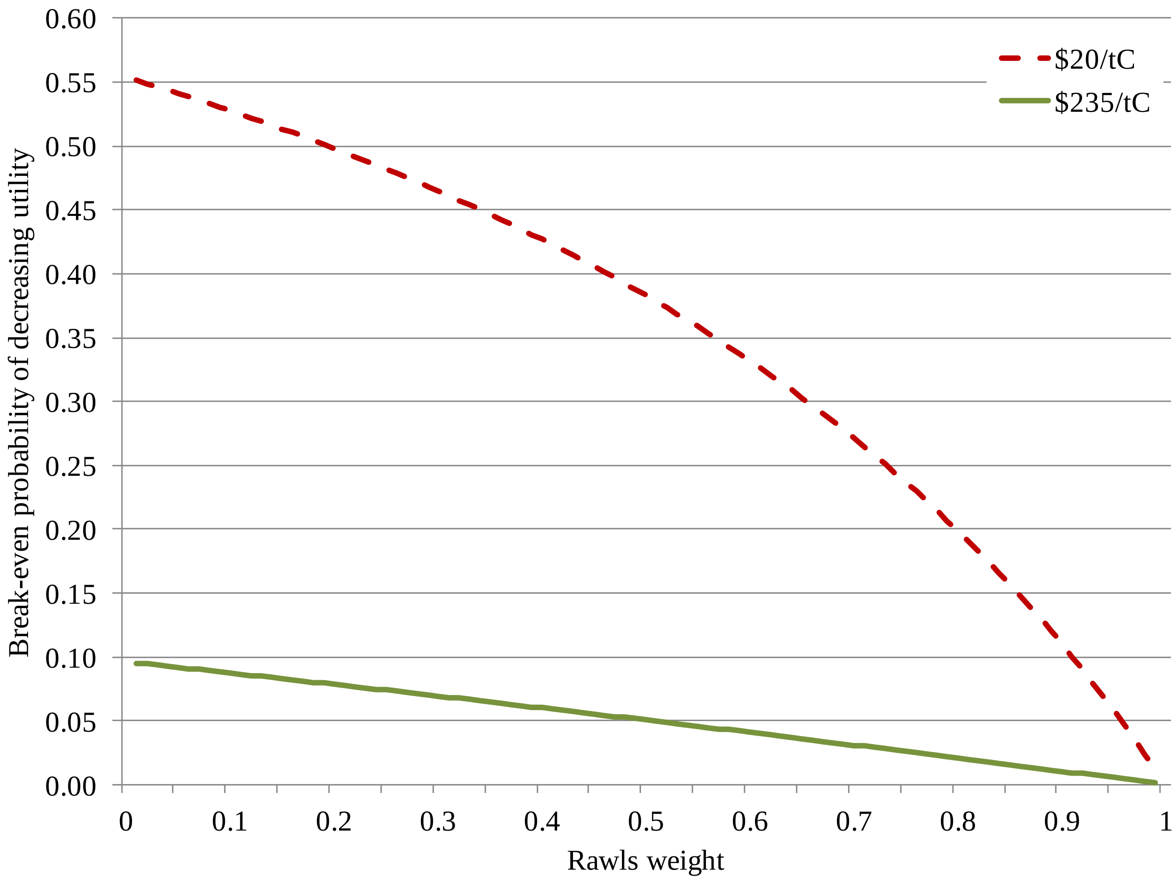

Figure 3. The break-even probability of decreasing utility as a function of the Rawls weight for two alternative estimates of the Rawls social cost of carbon.