Is there space for agreement on climate change? A non-parametric approach to policy evaluation

Simon Dietz[*] and Anca N. Matei[†]

20th August 2013

Abstract

Economic evaluation of climate policy is notoriously dependent on assumptions about time and risk preferences, since reducing greenhouse gas emissions today has a highly uncertain pay-off, far into the future. These assumptions have always been much debated. Rather than occupy a position in this debate, we take a non-parametric approach here, based on the concept of Time-Stochastic Dominance. Using an integrated assessment model, we apply Time-Stochastic Dominance analysis to climate change, asking are there global emissions abatement targets that everyone who shares a broad class of time and risk preferences would agree to prefer? Overall we find that even tough emissions targets would be chosen by almost everyone, barring those with arguably 'extreme' preferences.

Keywords: almost stochastic dominance, climate change, discounting, integrated assessment, risk aversion, stochastic dominance, time dominance, time-stochastic dominance

JEL codes: Q54

[*]Grantham Research Institute on Climate Change and the Environment, and Department of Geography and Environment, London School of Economics and Political Science.
[†]Department of Economics, University of Verona.

Email for correspondence: s.dietz@lse.ac.uk

1 Introduction

For over three decades, economists have been evaluating the abatement of global greenhouse gas emissions, to mitigate climate change. It has become increasingly evident in that time that the value of emissions abatement depends sensitively on the social planner's time preference and attitude to risk. This makes sense, because emissions abatement is a fine example of an investment that pays off mainly in the very long run and whose pay-offs are subject to significant uncertainty. Unfortunately there is much debate about appropriate time and risk preferences, hence there is much debate about optimal climate mitigation.

Most studies − certainly most empirical studies − have been based on a model in which social welfare is the discounted sum of individual utilities. To fix ideas, let us call $\beta(t)$ the discount factor in year t. Standard practice is to set $\beta(t) = 1/\left[1 + \delta(t)\right]^t$, so it is immediately clear that over a long time horizon even small differences in the utility discount rate $\delta(t)$ can have a large effect on the valuation of future utility. Utility is assumed to be a concave function of per-capita consumption of an aggregate good, net of the effect of climate change, $u'(c) > 0$ and $u''(c) < 0$. Consumption may be uncertain, and one would take the expectation of utility in such a model. Ordinarily the utility function exhibits constant relative risk aversion, $u(c) = [1/(1 - \eta)] \cdot c(t)^{1-\eta}$, where $-\eta$ is the elasticity of marginal utility with respect to consumption, i.e. η is the coefficient of relative risk aversion.

Within this framework, there has for many years been a vigorous debate about $\delta(t)$ and η. In many respects, the pioneering studies of Cline (1992) and Nordhaus (1991; 1994) staked out debating positions that still hold today. It is well known that Cline set $\delta(t) = 0$, $\forall t$ based on so-called 'prescriptive' ethical reasoning (like e.g. Ramsey, Pigou and even Koopmans before him), while Nordhaus set $\delta(t) = 3\%$, $\forall t$ based on a more conventional 'descriptive' analysis of market rates of investment returns.[1] These contrasting positions were one of the main reasons why Cline and Nordhaus made different recommendations on emissions abatement, with Cline advocating substantially faster, deeper cuts.

This is what most remember from their debate, but it is worth pointing out that Cline's conclusions also depended on applying a somewhat informal notion of risk aversion to aggregate across a selection of his modelling 'cases'. That is, he did not treat consumption per capita as a random variable, but he did explore different combinations of parameter values in sensitivity analysis, and in drawing conclusions attempted to informally aggregate across them. Thus while it would not be until some years later that risk began to formally enter the analysis as foreseen above, its importance was already evident at this early stage.

More recently, the *Stern Review on the Economics of Climate Change* (Stern, 2007) has revived and updated 'Cline vs. Nordhaus', by setting $\delta(t) = 0.1\%$, $\forall t$ and advocating aggressive emissions abatement.[2] However, by implementing

---

[1] See Arrow et al. (1996) for a classic comparison of these two points of view, from where the labels descriptive and prescriptive hail.

[2] See Nordhaus (2007; 2008) for critique of the Stern Review.

formal Monte Carlo analysis and including in the parameter space some combinations of values implying catastrophic climate damage (Dietz et al., 2007), the Stern Review and reactions to it also stimulated a more detailed and formal debate about the role of risk and risk aversion (e.g. Pindyck, 2011; Weitzman, 2007, 2009).[3] In an intertemporal model with growth and risk, η governs not only risk aversion but also aversion to inequality in consumption over time[4], thus the dependence of the value of emissions abatement on η is complex. Across the many perspectives on the Stern Review one can find support for $\eta \in [0.5, 4]$ (e.g. Dasgupta, 2007; Gollier, 2006; Stern, 2008). Various papers have shown how sensitive the value of emissions abatement is to η, perhaps the most convincing of which is Anderson et al. (2012), which is a global sensitivity analysis using Nordhaus' DICE integrated assessment model, showing that the social cost of carbon is far more sensitive to η than it is to any other model parameter whatsoever.

While the debate as thus far described has been limited to which constant values of δ and η to choose, the forms of the discount and utility functions are also open to debate. For instance, Gerlagh and Liski (2012) investigate the implications for climate policy of hyperbolic discounting.[5] Similarly alternative utility functions have recently been proposed (Ikefuji et al., 2012; Pindyck, 2011). Millner (forthcoming) provides a recent review.

Rather than attempting to settle the debate, in this paper we embrace it. Our starting point is the supposition that debate about time and risk preference in climate economics legitimately exists. Given the ingredients of the debate and the current state of knowledge, "reasonable minds may differ" (?). Moreover, we take it as given that it always will, or at the very least it will persist long enough to cloud a sequence of important choices about global emissions faced in reality, on which economists might be able to give more one-handed advice, as President Truman might have put it. Why is the debate difficult to resolve? It contains both positive and normative elements. Even those who take a descriptive approach as described above are taking a normative position in doing so. While the positive 'uncertainties' could in principle be eliminated by collecting more empirical data from, for instance, market behaviour, questionnaire surveys or laboratory experiments, in reality it is likely that significant uncertainties will persist. Witness longstanding difficulties with, and ongoing differences in approach to, puzzles in the economics of risk such as the equity premium / risk-free rate. Moreover it is even more difficult to resolve normative differences simply by appeal to empirical information. Such differences may never be completely eliminated.

Given ongoing debate about how precisely to structure and parameterise time and risk preference in economic models of climate mitigation, the question

---

[3]The debate now extends to second-order uncertainty about probabilities, often known as 'ambiguity'. However, we abstract from this issue here; see e.g. Millner et al. (2013).

[4]In a model disaggregated into multiple regions (which Stern's was not), η also governs aversion to inequality in consumption over space.

[5]In fact Nordhaus himself was for a time of the view that $\delta(t)$ should decline over time (Nordhaus and Boyer, 2000).

we ask is; can we make choices on emissions abatement without having to agree on how exactly to represent time and risk preference? Put more formally, are there combinations of whole classes of discount and utility functions, for which it is possible to say that some abatement policies are preferred to others? These classes admit different functional forms, and, for given functional forms, often wide ranges of parameter values, so the approach we take is non-parametric. Where preference orderings over abatement policies can be constructed for certain combinations of discount and utility function, we say there is a *space for agreement*. Hence a space for agreement is a partial ordering.

The theoretical machinery for analysing spaces for agreement builds on the concepts of Stochastic Dominance (Section 2) and Time Dominance (Section 3), long-established frameworks for ordering risky prospects and cashflows over time, respectively. However, until now Stochastic Dominance and Time Dominance have been limited in their applicability to climate policy by the fact that the former does not admit pure time discounting, while the latter cannot be applied to uncertain future cashflows except under very strong assumptions. The one's strength is the other's weakness in this regard. Therefore, in a companion paper we unify Stochastic Dominance and Time Dominance to produce a theory of Time-Stochastic Dominance, which is able to handle choices between investments that are both inter-temporal and risky (Dietz and Matei, 2013). In this paper we present the theory in a stripped-down manner, highlighting the key theorems and providing cross-references to the proofs for the interested reader (Section 4).

We make an empirical application by analysing a set of trajectories for global greenhouse gas emissions – a set of 'policies' – using a stochastic version of Nordhaus' DICE model. DICE provides a theoretically coherent representation of the coupled climate-economy system and is well understood, being openly available and relatively simple. Section 5 describes how the model is set up, as well as the policies to be compared. Our version of the model was developed by Dietz and Asheim (2012) and, unlike standard DICE, incorporates uncertainty via a set of random parameters that are inputs to Monte Carlo simulation. The policies to be evaluated differ in the maximum atmospheric concentration of carbon dioxide that is permitted, i.e. each is an emissions path that maximises social welfare subject to a constraint on atmospheric CO_2.

Section 6 presents our results. It indicates that, although the profile of net benefits from climate mitigation is such that strict Time-Stochastic Dominance cannot be established, we can use the less restrictive concept of Almost Time-Stochastic Dominance to show that the space for agreement on climate change is indeed large. Section 7 completes the paper by providing a discussion.

2 A primer on Stochastic Dominance and 'Almost' Stochastic Dominance

Stochastic Dominance (hereafter SD) determines the order of preference of an expected-utility maximiser between risky prospects, while requiring minimal knowledge of her utility function. Take any two risky prospects F and G, and denote their cumulative distributions F^1 and G^1 respectively. Assuming the cumulative distributions have finite support on $[a, b]$, F is said to first-order stochastic dominate G if and only if $F^1(x) \leq G^1(x)$, $\forall x \in [a, b]$ and there is a strict inequality for at least one x, where x is a realisation from the distribution of payoffs possible from a prospect. Moreover it can be shown that any expected-utility maximiser with a utility function belonging to the set of non-decreasing utility functions $U_1 = \{u : u'(x) \geq 0\}$ would prefer F.

First-order SD does not exist if the cumulative distributions cross, which means that, while it is a powerful result in the theory of choice under uncertainty, the practical usefulness of the theorem is limited. By contrast, where $F^2(x) = \int_a^x F^1(s)ds$ and $G^2(x) = \int_a^x G^1(s)ds$, F second-order stochastic dominates G if and only if $F^2(x) \leq G^2(x)$, $\forall x \in [a, b]$ and there is a strict inequality for at least one x. It can be shown that any expected-utility maximiser with a utility function belonging to the set of all non-decreasing and (weakly) concave utility functions $U_2 = \{u : u \in U_1 \text{ and } u''(x) \leq 0\}$ would prefer F, i.e. any such (weakly) risk-averse decision-maker. Hence second-order SD can rank prospects with, for example, the same mean but different variances.

Nonetheless the practical usefulness of second-order SD is still limited. Consider the following example. Let us try to use SD criteria to rank two prospects; F pays out \$0.5 with a probability of 0.01 and \$1 million with a probability of 0.99, while G pays out \$1 for sure. While it would seem that virtually any investor would prefer F, second-order SD does not exist as $G^2(x) - F^2(x) < 0$, $x \in [0.5, 1)$. Intuitively, the reason for the violation of second-order SD is that the broad class of preferences admitted in U_2 includes risk aversion so extreme that the decision-maker effectively only cares about the 0.01 probability of sacrificing \$0.5 by taking the gamble.

From here the obvious step would be to place additional restrictions on the decision-maker's preferences. Thus we can define the set $U_3 = \{u : u \in U_2 \text{ and } u'''(x) \geq 0\}$ and look for third-order SD. Decision-makers exhibiting decreasing absolute risk aversion have preferences represented by utility functions in U_3 and such decision-makers will also exhibit 'prudence' in intertemporal savings decisions (Kimball, 1990). F third-order stochastic dominates G if and only if $F^3(x) \leq G^3(x)$, $\forall x \in [a, b]$ and $E_F(x) \geq E_G(x)$, and there is at least one strict inequality.[6] However, it can easily be verified that $G^3(x) - F^3(x) < 0$, $x \in [0.5, 1)$, yet $E_F(x) >> E_G(x)$, so third-order SD does not exist. Moreover SD cannot be established to any order, because the first non-zero values of $G^1(x) - F^1(x)$ are negative as x increases from its lower bound, yet $E_F(x) > E_G(x)$.

---

[6] $F^3(x) = \int_a^x F^2(s)ds$ and $G^3(x) = \int_a^x G^2(s)ds$.

A more fruitful route is the theory of 'Almost Stochastic Dominance' (ASD) set out by Leshno and Levy (2002) and recently further developed by Tzeng et al. (2012). ASD places restrictions on the derivatives of the utility function with the purpose of excluding the extreme preferences that prevent strict SD from being established. Dominance relations are then characterised for 'almost' all decision-makers.

For every $0 < \varepsilon_k < 0.5$, where $k = 1, 2$ corresponds to first- and second-order SD respectively, define subsets of U_k:

$$
\begin{aligned}
U_1(\varepsilon_1) &= \left\{ u \in U_1 : u'(x) \leq \inf[u'(x)] \left[\frac{1}{\varepsilon_1} - 1 \right], \forall x \right\} \text{ and} \qquad (1) \\
U_2(\varepsilon_2) &= \left\{ u \in U_2 : -u''(x) \leq \inf[-u''(x)] \left[\frac{1}{\varepsilon_2} - 1 \right], \forall x \right\}.
\end{aligned}
$$

$U_1(\varepsilon_1)$ is the set of non-decreasing utility functions with the added restriction that the ratio between maximum and minimum marginal utility is bounded by $\frac{1}{\varepsilon_1} - 1$. In the limit as ε_1 approaches 0.5, the only function in $U_1(\varepsilon_1)$ is linear utility. Conversely as ε_1 approaches zero, $U_1(\varepsilon_1)$ coincides with U_1. $U_2(\varepsilon_2)$ is the set of non-decreasing, weakly concave utility functions with an analogous restriction on the ratio between the maximum and minimum values of $u''(x)$. In the limit as ε_2 approaches 0.5, $U_2(\varepsilon_2)$ contains only linear and quadratic utility functions, while as ε_2 approaches zero, it coincides with U_2.

Defining the set of realisations over which strict first-order SD is violated as

$$
S^1(F, G) = \left\{ x \in [a, b] \; : \; G^1(x) < F^1(x) \right\},
$$

F is said to first-order almost stochastic dominate G if and only if

$$
\int_{S^1} \left[F^1(x) - G^1(x) \right] dx \leq \varepsilon_1 \cdot \int_a^b \left| \left[F^1(x) - G^1(x) \right] \right| dx.
$$

Moreover, in a similar vein to strict SD, it can be shown that any expected-utility maximiser with a utility function belonging to $U_1(\varepsilon_1)$ would prefer F.

Defining the set of realisations over which strict second-order SD is violated as

$$
S^2(F, G) = \left\{ x \in [a, b] \; : \; G^2(s) < F^2(s) \right\},
$$

F second-order almost stochastic dominates G if and only if

$$
\int_{S^2} \left[F^2(x) - G^2(x) \right] dx \leq \varepsilon_2 \cdot \int_a^b \left| \left[F^2(x) - G^2(x) \right] \right| dx \text{ and}
$$

$$
E_F(x) \geq E_G(x).
$$

Any expected-utility maximiser with a utility function belonging to $U_2(\varepsilon_2)$ would prefer F. From these definitions of first- and second-order ASD one can see that ε_k intuitively represents the proportion of the total area between

F^k and G^k in which the condition for strict SD of the kth order is violated. The smaller is ε_k, the smaller is the relative violation.

One could say that ASD fudges the issue somewhat, insofar as the clarity that strict SD brings to the ordering of prospects is partly lost. However, that is to overlook the value empirically of scrutinising the area of violation of strict SD, i.e. with ASD the focus turns to analysing the size of ε_k. Another obvious difficulty is in determining just how large ε_k can be before it can no longer be said that one prospect almost stochastic dominates another, i.e. what is an 'extreme' preference? This is clearly subjective, but Levy et al. (2010) offer an illustration of how to define it using laboratory data on participant choices when faced with binary lotteries. Extreme risk preferences are marked out by establishing gambles that all participants are prepared to take. By making the conservative assumption that no participant has extreme risk preferences, the set of non-extreme preferences is at least as large as that marked out by the least and most risk-averse participants. Preferences outside these limits can be considered extreme.

3   A primer on Time Dominance

The theory of Time Dominance (TD) builds on the SD approach to choice problems under uncertainty, and transfers it to problems of intertemporal choice (Bøhren and Hansen, 1980; Ekern, 1981). Denoting the cumulative cashflows of any two investments X_1 and Y_1,[7] X is said to first-order time dominate Y if and only if $X_1(t) \geq Y_1(t)$, $\forall t \in [0, T]$ and there is a strict inequality for some t, where T is the terminal period of the most long-lived project. Moreover it can be shown that any decision-maker with a discount function belonging to the set of all decreasing consumption discount functions $\widehat{V_1} = \{\widehat{v} \; : \; \widehat{v}'(t) < 0\}$ would prefer X. Thus if the decision-maker prefers a dollar today to a dollar tomorrow, she will prefer X if it first-order time dominates Y.

Just like SD, first-order TD has limited practical purchase, because the set of undominated investments remains large, i.e. the criterion $X_1(t) \geq Y_1(t)$, $\forall t$ is restrictive.[8] Therefore, proceeding again by analogy to SD, X second-order time dominates Y if and only if

$$X_1(T) \geq Y_1(T) \text{ and}$$

$$X_2(t) \geq Y_2(t), \; \forall t \in [0, T],$$

where $X_2(t) = \int_0^t X_1(\tau)d\tau$ and $Y_2(t) = \int_0^t Y_1(\tau)d\tau$, and there is at least one strict inequality. Any decision-maker with a discount function belonging to the

---

[7] $X_1(t) = \int_0^t x(\tau)d\tau$ and $Y_1(t) = \int_0^t y(\tau)d\tau$.

[8] Indeed, in the domain of deterministic cashflows over multiple time-periods, the requirement that $X^1(0) \geq Y^1(0)$ means that one investment cannot dominate another by a first, second or higher order, if the initial cost is higher, no matter what the later benefits are. This makes it difficult to compare investments of different sizes. However, this can be worked around by normalising the cashflows to the size of the investment (Bøhren and Hansen, 1980).

set of all decreasing, convex consumption discount functions $\widehat{V_2} = \{\widehat{v} \; : \; \widehat{v} \in \widehat{V_1}$ and $\widehat{v}''(t) > 0\}$ would prefer X. Noting how the conditions for second-order TD are obtained from their counterparts for first-order TD by integration, TD can be defined to the nth order (see Ekern, 1981).

Notice that TD applies to deterministic cashflows. It would be possible to apply the method to uncertain cashflows, if X and Y were expected cashflows and if a corresponding risk adjustment were made to $\{\widehat{v}\}$. However, since any two cashflows X and Y would then be discounted using the same set of risk-adjusted rates, it would be necessary to assume that the cashflows belong to the same risk class (Bøhren and Hansen, 1980), for example under the capital asset pricing model they would have to share the same covariance with the market portfolio. This significantly limits the reach of the method to uncertain investments. It would also be necessary to assume that any investments being compared are small (i.e. marginal), since the domain of $\{\widehat{v}\}$ is cashflows and therefore depends on a common assumed growth rate. Neither of these assumptions is likely to hold in the case of climate change (see Weitzman, 2007, for a discussion of the covariance between climate mitigation and market returns, and Dietz and Hepburn, 2013, for a discussion of whether climate mitigation is non-marginal).

This sets the scene for a theory that unifies the capacity of SD to order risky prospects with the capacity of TD to order intertemporal cashflows. The resulting theory of Time-Stochastic Dominance has the additional advantage that time and risk preferences can be disentangled and each scrutinised explicitly (whereas in applying TD, even if the assumptions just discussed would hold, both time and risk preference would be bound up in the concept of risk-adjusted discount functions).

4   Time-Stochastic Dominance and Almost Time-Stochastic Dominance

The Time-Stochastic Dominance (TSD) approach is developed formally in Dietz and Matei (2013). Here we will summarise the key concepts, referring the interested reader to the companion paper for more details, including proofs.

Take two prospects X and Y, both of which yield random cashflows over time. The underlying purpose is to compare the expected discounted utilities of the prospects at $t = 0$, i.e. for prospect X one would compute

$$NPV_{v,u}(X) = \int_0^T v(t) \cdot E_F u(x,t) dt = \int_0^T v(t) \cdot \int_a^b u(x) \cdot F^1(x,t) dx dt \quad (2)$$

where it is important to note that v is a utility or pure time discount function, rather than a consumption discount function as in the case of TD. Nonetheless we otherwise borrow the terminology developed above by analysing combinations of classes of discount and utility functions, such that $V_i \times U_j$ denotes the

combination of the ith class of pure time discount function with the jth class of utility function. For example, a natural first point of reference would be $V_1 \times U_1$, the set of all combinations of decreasing pure time discount function and non-decreasing utility function. These combinations are the basis for our notion of a *space for agreement*. In other words, we will be looking for the least restricted combination $V_i \times U_j$ such that one prospect − one climate mitigation policy − dominates another.

Where $f(x,t)$ represents the probability density function for prospect X at time t,

$$F_1^1(x,t) = \int_a^x F_1(s,t)ds = \int_0^t F^1(x,\tau)d\tau = \int_0^t \int_a^x f(s,\tau)dsd\tau.$$

Defining $d(z,t) = g(y,t) - f(x,t)$, we set

$$D_i^j(z,t) = G_i^j(y,t) - F_i^j(x,t)$$

for all $x, y, z \in [a,b]$ and all $t \in [0,T]$. Given information on the first n and m derivatives of the discount and utility functions respectively, we recursively define:

$$D_n(z,t) = \int_0^t D_{n-1}(z,\tau)d\tau$$
$$D^m(z,t) = \int_a^z D^{m-1}(s,t)ds$$
$$D_n^m(z,t) = \int_0^t D_{n-1}^m(z,\tau)d\tau = \int_a^z D_n^{m-1}(s,t)ds = \int_0^t \int_a^z D_{n-1}^{m-1}(s,\tau)dsd\tau,$$

where $i \in \{1, 2, \ldots, n\}$ is the order of pure TD (i.e. the number of integrations with respect to time) and $j \in \{1, 2, \ldots, m\}$ is the order of SD (i.e. the number of integrations with respect to the consequence space).

Definition 1 *(Time-Stochastic Dominance of order i,j). For any two risky, intertemporal prospects X and Y*

$$X >_{iTjS} Y \text{ if and only if } \Delta \equiv NPV_{v,u}(X) - NPV_{v,u}(Y) \geq 0,$$

for all $(v,u) \in V_i \times U_j$.

In this definition, the ordering $>_{iTjS}$ denotes pure TD of the ith order, combined with SD of the jth order. For example, $>_{1T1S}$, which we can shorten to $>_{1TS}$, denotes time and stochastic dominance of the first order.

Proposition 1 *(First-order Time-Stochastic Dominance). $X >_{1TS} Y$ if and only if*
$$D_1^1(z,t) \geq 0, \ \forall z \in [a,b] \ and \ \forall t \in [0,T],$$

and there is a strict inequality for some (z,t).

9

Proposition 1 tells us that X first-order time-stochastic dominates Y provided the integral over time of the cdf of Y is at least as large as the integral over time of the cdf of X, for all wealth levels and all time-periods. It maps out a space for agreement, as we can say that all decision-makers with preferences that can be represented by $V_1 \times U_1$ will rank X higher than Y, no matter what precisely is their discount function or utility function, not to mention what precisely is their rate of impatience $\delta(t)$ or CRRA η.

Having established first-order TSD, we can proceed from here either by placing an additional restriction on the discount function, or on the utility function, or on both. A particularly compelling case is $(v, u) \in V_1 \times U_2$ − since few would be uncomfortable with the notion of excluding risk-seeking behaviour *a priori*, especially in the public sector.

Proposition 2 (First-order Time and Second-order Stochastic Dominance). $X >_{1T2S} Y$ *if and only if*

$$D_1^2(z,t) \geq 0 \ \forall z \in [a,b] \ \text{and} \ \forall t \in [0,T],$$

with at least one strict inequality.

Proposition 2 delineates a space for agreement for all decision-makers who are at the same time impatient and risk averse, a subset of the set of decision-makers in Proposition 1.

It is evident from Proposition 2 that restricting the utility function by one degree corresponds to integrating $D_1^1(z,t)$ once more over the consequence space. If we want to pursue the further case of $(v, u) \in V_2 \times U_2$, representing a risk-averse planner with impatience decreasing over time, then we would integrate $D_1^2(z,t)$ once more with respect to time (see Dietz and Matei, 2013). Taking this to its logical conclusion, we can generalise TSD to the nth order with respect to time and the mth order with respect to risk.

Proposition 3 (n^{th}-order Time and m^{th}-order Stochastic Dominance). X n^{th}*-order time and* m^{th}*-order stochastic dominates* Y *if*

$$
\begin{aligned}
&i) \ D_{i+1}^{j+1}(b,T) \geq 0 \\
&ii) \ D_n^{j+1}(b,t) \geq 0, \ \forall t \in [0,T] \\
&iii) \ D_{i+1}^m(z,T) \geq 0, \forall z \in [a,b] \\
&iv) \ D_n^m(z,t) \geq 0, \forall z \in [a,b], \forall t \in [0,T]
\end{aligned}
$$

with (iv) holding as a strong inequality over some sub interval and where $i = \{0, \ldots, n-1\}$ *and* $j = \{0, \ldots, m-1\}$.

Dominance criteria have strong appeal, because they offer non-parametric rankings for entire preference classes. However, as discussed above in relation to SD, it is in the nature of their characterisation that they can fail to establish superiority of one investment over another, even if the violation of strict dominance is very small and the order of preference would seem to be common

sense. Strict TSD evidently shares this weakness. To counter it, we establish theorems for Almost TSD, extending the framework of Leshno and Levy (2002) and Tzeng et al. (2012) into our bi-dimensional time-risk set-up. In particular, we will characterise almost first-order TSD and almost first-order time and second-order stochastic dominance.

Consider the following combinations of preferences:

$$V_1(\gamma_1) \times U_1(\gamma_1) = \{v \in V_1, \, u \in U_1 \, : \, sup[-v'(t)u'(z) \leq$$
$$inf[-v'(t)u'(z)] \left[\frac{1}{\gamma_1} - 1\right], \, \forall z \in [a,b], \, \forall t \in [0,T]\} \quad (3)$$

$$V_1(\gamma_2) \times U_2(\gamma_2) = \{v \in V_1, \, u \in U_2 \, : \, sup[v'(t)u''(z) \leq$$
$$inf[v'(t)u''(z)] \left[\frac{1}{\gamma_2} - 1\right], \, \forall z \in [a,b], \, \forall t \in [0,T]\} \quad (4)$$

In words, $V_1(\gamma_1) \times U_1(\gamma_1)$ is the set of all combinations of decreasing pure time discount function and non-decreasing utility function, with the added restriction that the ratio between the maximum and minimum products of $[-v'(t) \cdot u'(z)]$ is bounded by $\frac{1}{\gamma_1} - 1$. The supremum (infimum) of $[-v'(t) \cdot u'(z)]$ is attained when $v'(t) < 0$ is the infimum (supremum) of its set and $u'(z) \geq 0$ is the supremum (infimum) of its. $V_1(\gamma_2) \times U_2(\gamma_2)$ is similarly defined, except that we are now focused on the products of $[v'(t) \cdot u''(z)]$ with respect to $\frac{1}{\gamma_2} - 1$. The supremum (infimum) of $[v'(t) \cdot u''(z)]$ is attained when $v'(t) < 0$ and $u''(z) \leq 0$ are the suprema (infima) of their respective sets.

Conceptually, bounding the ratio of $v'(t)$ amounts to restricting $v''(t)$, such that preferences exhibiting a very large change in impatience over time are excluded, for example preferences exhibiting a very rapid decrease in impatience. Similarly, bounding the ratio of $u'(z)$ or $u''(z)$ amounts to restricting $u''(z)$ or $u'''(z)$ respectively, such that extreme concavity (risk aversion) or convexity (risk seeking) of $u(z)$ is ruled out, as are large changes in prudence with respect to z. Dietz and Matei (2013) formalise this story.

Now divide the interval $[a,b]$ into two sets, for all $t \in [0,T]$. The first subset S_1^1 includes only realisations where $D_1^1 < 0$, i.e. where the condition for strict first-order TSD is violated:

$$S_1^1(D_1^1) = \left\{z \in [a,b], \, \forall t \in [0,T] : D_1^1(z,t) < 0\right\}.$$

As before, consider also decision-makers who simultaneously exhibit impatience and risk aversion/neutrality, i.e. $(v,u) \in V_1 \times U_2$. In this case we parcel out the subset of realisations S_1^2 where $D_1^2 < 0$:

$$S_1^2(D_1^2) = \left\{z \in [a,b], \, \forall t \in [0,T] : D_1^2(z,t) < 0\right\}.$$

Definition 2 *(Almost First-order Time-Stochastic Dominance).*
X almost first-order time-stochastic dominates Y, denoted $X \succeq_{A1TS} Y$, if and only if

$$\int_0^T \int_{S_1^1} \left[-D_1^1(z,t) \right] dz dt \leq \gamma_1 \int_0^T \int_a^b \mid \left[D_1^1(z,t) \right] \mid dz dt \text{ and}$$
$$\int_{S^1,T} \left[-D_1^1(z,T) \right] dz \leq \varepsilon_{1T} \int_a^b \mid \left[D_1^1(z,T) \right] \mid dz.$$

The left-hand side of the first inequality constitutes the violation γ_1 of strict first-order TSD across all time-periods. In addition, as the second inequality shows, the definition also requires the violation of strict first-order TSD to be no larger than ε_{1T} in the final time-period, where ε_{1T} is defined as in (1) but is now measured with respect to $D_1^1(z,T)$.

Proposition 4 *(A1TSD).* $X \succeq_{A1TS} Y$ *if, for all $(v,u) \in V_1(\gamma_1) \times U_1(\gamma_1)$ and $u \in U_1(\varepsilon_{1T})$,*
    $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$ *and there is at least one strict inequality.*

In defining almost first-order time and second-order stochastic dominance it is necessary to measure three violations. In addition to the violation γ_2 that prevents us from obtaining strict first-order time and second-order stochastic dominance with respect to the whole consequence space $[a,b]$ and time-horizon $[0,T]$, and the violation in the final time-period ε_{2T}, we need to measure a further violation of the non-negativity condition on the integral with respect to time of $D_1^2(b,t)$. To do this we divide the time horizon $[0,T]$ into two sets, when $z = b$. The first subset includes only realisations where $D_1^2(b,t) < 0$:

$$S_1(D_1^2) = \left\{ z = b, \ t \in [0,T] : D_1^2(t) < 0 \right\}.$$

Definition 3 *(Almost First-order Time and Second-order Stochastic Dominance).*
X almost first-order time and second-order stochastic dominates Y, denoted $X \succeq_{A1T2S} Y$ if and only if

$$i) \int_0^T \int_{S_1^2} \left[-D_1^2(z,t) \right] dz dt \leq \gamma_2 \int_0^T \int_a^b \mid \left[D_1^2(z,t) \right] \mid dz dt \text{ and}$$
$$ii) \int_{S^2} \left[-D_1^2(z,T) \right] dz \leq \varepsilon_{2T} \int_a^b \mid \left[D_1^2(z,T) \right] \mid dz, \text{ and}$$
$$iii) \int_{S_1} \left[D_1^2(b,t) \right] dt \leq \gamma_{1b} \int_0^T \mid \left[D_1^2(b,t) \right] \mid dt, \text{ and}$$
$$iv) D_1^2(b,T) \geq 0.$$

Note that the restriction $\int_{S_1} \left[D_1^2(b,t) \right] dt / \int_0^T \mid \left[D_1^2(b,t) \right] \mid dt \leq \gamma_{1b}$ requires $(v,u) \in V_1(\gamma_1) \times U_1(\gamma_1)$ for all t when $z = b$, which we hence write as $(v,u) \in V_1(\gamma_{1b}) \times U_1(\gamma_{1b})$, while there is also a specific requirement that $D_1^2(b,T) \geq 0$.

Proposition 5 *(A1T2SD).* $X \succeq_{A1T2S} Y$ *if, for all $(v,u) \in V_1(\gamma_2) \times U_2(\gamma_2)$, $V_1(\gamma_{1b}) \times U_1(\gamma_{1b})$, and $u \in U_2(\varepsilon_{2T})$,*
    $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$ *and there is at least one strict inequality.*

5 Modelling approach

5.1 A stochastic version of DICE

Standard versions of DICE are deterministic, with fixed parameters. This is a poor fit to the problem of evaluating climate policy, however, because risk is a central element. Therefore we use a stochastic version of DICE, developed by Dietz and Asheim (2012). This version randomises eight parameters in the model so that Monte Carlo simulation can be undertaken. Table 1 lists the eight parameters, and the form and parameterisation of the probability density functions assigned to them. The equations of the model can be found in the Appendix to Nordhaus (2008). Unless otherwise stated here, no changes are made to original model.

The eight random parameters were originally selected by Nordhaus (2008), based on his broader assessment of which of all the model's parameters had the largest impact on the value of policies. The first four parameters in Table 1 play a role in determining CO_2 emissions. In one-sector growth models like DICE, CO_2 emissions are directly proportional to output, which is in turn determined in significant measure by productivity (i) and the stock of labour (ii). However, while CO_2 emissions are proportional to output, the proportion is usually assumed to decrease over time due to structural change away from CO_2-intensive production activities (e.g. from manufacturing to services), and to decreases in CO_2 intensity in any given activity. In DICE, this is achieved by virtue of a variable representing the autonomous ratio of emissions/output (i.e. in the absence of policy-driven emissions control), which decreases over time as a function of a rate-of-decarbonisation parameter (iii). A further check on industrial CO_2 emissions is provided in the long run by the finite total remaining stock of fossil fuels (iv).

[TABLE 1 HERE]

The fifth uncertain parameter is the price of a CO_2-abatement backstop technology, capable of completely eliminating CO_2 emissions. Such a technology could most plausibly be a zero-emissions energy technology such as solar or geothermal power.[9] The initial backstop price (v) is very high (mean = US\$1170/tC), but it declines over time. In DICE, the coefficient of the abatement cost function is a function of the backstop price, hence we obtain abatement cost uncertainty as a result of backstop price uncertainty.

The sixth and seventh parameters in Table 1 capture important uncertainties in climate science. On the one hand parameter (vi) captures uncertainty about the carbon cycle via the proportion of CO_2 in the atmosphere in a particular time-period, which dissolves into the upper ocean in the next period. On the other hand uncertainty about the relationship between a given stock of

---

[9]It could also be a geo-engineering technology such as artificial trees to sequester atmospheric CO_2, except that DICE has exogenous emissions of CO_2 from land use.

atmospheric CO_2 and temperature is captured by specifying a random climate-sensitivity parameter (vii). The climate sensitivity is the increase in global mean temperature, in equilibrium, that results from a doubling of the atmospheric stock of CO_2. In simple climate models like DICE's, it is critical in determining how fast and how far the planet is forecast to warm in response to emissions. There is by now much evidence, derived from a variety of approaches (see Meehl et al., 2007, and Roe and Baker, 2007), that the probability density function for the climate sensitivity has a positive skew.

The eighth and final uncertain parameter is one element of the damage function linking temperature and utility-equivalent losses in output. In Dietz and Asheim's (2012) version of DICE, the damage function has the following form:

$$\Omega(t) = \frac{1}{1 + \alpha_1 \Upsilon(t) + \alpha_2 \Upsilon(t)^2 + [\tilde{\alpha}_3 \Upsilon(t)]^7}, \tag{5}$$

where Ω is the proportion of output lost, Υ is the increase in global mean temperature over the pre-industrial level, and α_i, $i \in \{1, 2, 3\}$ are coefficients. $\tilde{\alpha}_3$ is a normally distributed random coefficient (viii), so the higher-order term $[\tilde{\alpha}_3 \Upsilon(t)]^7$ captures the uncertain prospect that significant warming of the planet could be accompanied by a very steep increase in damages. That such a possibility exists has been the subject of recent controversy, with the approaches of Nordhaus (2008) and Weitzman (2012) marking out opposing stances. The controversy exists, because there is essentially no empirical evidence to support calibration of the damage function at high temperatures (Dietz, 2011; Tol, 2012); instead there are simply beliefs. In standard DICE, $\alpha_3 = 0$, thus there is no higher-order effect and 5°C warming, as a benchmark for a large temperature increase, results in a loss of utility equivalent to 6% of output. By contrast Weitzman (2012) suggests a functional form which can be approximated by $\alpha_3 = 0.166$. Here $\tilde{\alpha}_3$ is calibrated such that the Nordhaus and Weitzman positions represent minus/plus three standard deviations respectively, and at the mean 5°C warming results in a loss of utility equivalent to around 7% of output. Thus the mean value of function (5) remains fairly conservative.

Random parameters (i)-(viii), alongside the model's remaining non-random parameters and initial conditions (as per Nordhaus, 2008), are inputs to a Monte Carlo simulation. In particular, a Latin Hypercube sample of 1000 runs of the model is taken. Each run solves the model for a particular policy, which as described below is a schedule of values for the rate of control of CO_2 emissions. From this is produced a schedule of distributions of consumption per capita (where consumption per capita is equivalent to a cashflow in our theory), which is the focus of the Time-Stochastic Dominance analysis.

5.2 Policies to be evaluated

We evaluate a set of five, representative policies governing the rate of control of CO_2 emissions, plus a sixth path representing a forecast of emissions in the absence of policy-driven controls, i.e. 'business as usual'. These policies are

exogenous, since it makes no sense to optimise within the framework of Time-Stochastic Dominance, the whole point of which is to construct partial orderings given disagreement over the precise form and parameterisation of the discount and utility functions.

Each of the five policies limits the atmospheric stock of CO_2 to a pre-specified level. This approach maps on to real policy discussions, which often aim to derive emissions targets based on a 'stabilisation' level for the atmospheric concentration of CO_2 in the very long run. In order to try, as far as possible, to render the policies consistent with the assumptions we make, we use the stochastic version of DICE itself to generate the five policy paths, rather than relying on someone else's work with other assumptions and another model. The sixth path, 'business-as-usual' or BAU, is the baseline scenario from Nordhaus (2008).

The control variable is the emissions control rate $\mu(t) \in [0, 1]$, which is linked to production and emissions via the following equation:

$$E_{IND}(t) = \sigma(t)[1 - \mu(t)]Y(t),$$

where $E_{IND}(t)$ is industrial CO_2 emissions, $\sigma(t)$ is the ratio of uncontrolled industrial emissions to output and $Y(t)$ is production gross of abatement costs and climate-change damages.

Each policy path is generated by solving a stochastic optimisation problem, whereby the schedule $\{\mu(t)\}, \forall t$ is chosen to minimise abatement costs[10] subject to the constraint that the atmospheric stock of CO_2, $M^{AT}(t) \leq \overline{M^{AT}}$, where $\overline{M^{AT}} \in \{450, 500, 550, 600, 650\}$ and where the units are parts per million volume. This is done under uncertainty about parameters (i)-(vi), since these affect the cost of abatement and its impact on atmospheric CO_2. The constraint on the stock of atmospheric CO_2 is in terms of its mean value. Notice that such a policy objective does not necessarily entail $M^{AT}(t) = \overline{M^{AT}}$. In our policies, $M^{AT}(t) < \overline{M^{AT}}$ if it is cost-effective to do so. By contrast, much of the literature on stabilisation of atmospheric CO_2 requires the policy to satisfy this equality in the long run.

In an integrated assessment model such as DICE, and especially in running Monte Carlo simulation, solving this cost-minimisation problem is a non-trivial computational challenge. We solve it using a genetic algorithm (Riskoptimizer) and with two modifications to the basic optimisation problem.[11] In addition,

---

[10]Of course, what is cost-effective depends on the social objective, so for this part of the analysis we cannot avoid pre-specifying and parameterising the social welfare and utility functions. For this purpose, social welfare is given, as in standard DICE, by (??), where $\delta(t) = 1.5\%, \forall t$, and utility is CRRA, where $\eta = 2$.

[11]First, we only solve $\mu(t)$ from 2015 to 2245 inclusive, rather than all the way out to 2395. This considerably reduces the scope of the optimisation problem, in return for making little difference to the results, since, in the standard version of DICE, $\mu(t) = 1$ when $t > 2245$ (i.e. abatement yields high benefits relative to costs in the far-off future). Our first period of emissions control is 2015, since 2005, the first period of the model, is in the past. Second, we guide the optimisation by imposing the soft constraint that $\mu(t)$ is non-decreasing everywhere (via an exponential penalty function when $\mu(t)$ decreases between any two time-periods). Otherwise, the algorithm struggles to find a path towards the global optimum. As a soft constraint, the penalty does not enter the welfare evaluation. We were able to verify that

we limit the Latin Hypercube Sample to 250 draws.[12]

[FIGURE 1 HERE]

5.3 Time-Stochastic Dominance with quantiles

The DICE model output, for any policy setting, is in the form of $N = 1000$ discrete time series of consumption per capita, each with a discrete value in every time period. Each time series has an equal probability of $1/N$.

The TD algorithm simply involves repeated summation of cashflows. For each additional restriction on the curvature of the discount function, a new round of repeated summation is performed. Therefore, when $v \in V_1$, $X_1(t) = \sum_{\tau=0}^{t} x(\tau)$, while recursively $X_n(t) = \sum_{\tau=0}^{t} X_{n-1}(\tau)$. We first cumulate across time in this way.

The SD algorithm is based on comparing, for first- and second-order SD, the quantile functions of the distributions considered, a methodology developed by Levy and Hanoch (1970) and Levy and Kroll (1979) for uniform discrete distributions. Take X to be an integrable random variable with, for each $t \in [0, T]$, a *cdf* $F^1(x,t)$ and an *r-quantile* function $F^{-1,r}(p,t)$, the latter of which is recursively defined as

$$F^{-1,1}(p,t) := \inf\{x : F^1(x,t) \geq p(t)\}, \forall t \in [0,T]$$
$$F^{-1,2}(p,t) := \int_0^p F^{-1,1}(y,t)dy, \ \forall p \in [0,1] , \ \forall t \ and \ r \geq 2$$

Proposition 6 (1TSD for quantile distributions). $X >_{1TS} Y$ *if and only if*

$$H_1^{-1,1}(p,t) = F_1^{-1,1}(p,t) - G_1^{-1,1}(p,t) \geq 0, \ \forall p \in [0,1] \ and \ t \in [0,T]$$

and there is a strict inequality for some (p,t).

Proposition 6 characterises First-order Time-Stochastic Dominance for quantile distributions. Notice that since the quantile distribution function is just the inverse of the cumulative distribution function, 1TSD requires $F_1^{-1,1}(p,t) - G_1^{-1,1}(p,t) \geq 0$, i.e. the inverse of the requirement for 1TSD in terms of cumulative distributions.

Proposition 7 (1T2SD for quantile distributions). $X >_{1T2S} Y$ *if and only if*
$$H_2^{-1,2}(p,t) \geq 0, \ \forall p \in [0,1] \ and \ t \in [0,T]$$

and there is a strict inequality for some (p,t).

---

the algorithm's best solution satisfied the property of non-decreasingness in $\mu(t)$, and that no solution was found which returned lower costs, where $\mu(t)$ was decreasing at any point.

    [12]In order to ensure comparability with the results of the Time-Stochastic Dominance analysis, the smaller sample is calibrated on the sample statistics of the larger sample.

6   Results

6.1   Time-Stochastic Dominance analysis

We carry out the TSD analysis in two parts. In the first part we examine whether any of the abatement policies time-stochastic dominates BAU. That is to ask, can we use the analysis to establish that there is a space for agreement on acting to reduce greenhouse gas emissions by some non-trivial amount? This would already be of considerable help in understanding the scope of the debate about climate mitigation. In the second part we use the framework to compare the emissions reductions policies themselves – can we further use the framework to discriminate between the set of policies, so that we end up with a relatively clear idea of the policy that would be preferred?

Recall from Propositions 1 and 6 that first-order TSD requires $H_1^{-1,1}(p,t) \geq 0$, $\forall z, t$, with at least one strict inequality. Figure 2 plots $H_1^{-1,1}(p,t)$ when $\overline{M^{AT}} \in \{450, 500, 550, 600, 650\}$ is compared with BAU. Hence visual inspection is sufficient to establish that no abatement policy first-order time-stochastic dominates BAU, not even the most accommodating 650ppm concentration limit.

[FIGURE 2 HERE]

Although first-order TSD cannot be established between abatement and BAU, it could still be that one or more of the policies is preferred to BAU according to first-order time and *second*-order stochastic dominance. Propositions 2 and 7 showed that first-order time and second-order stochastic dominance requires $H_1^{-1,2}(p,t) \geq 0$, $\forall z, t$, with at least one strict inequality. Figure 3 plots $H_1^{-1,2}$ when each abatement policy is compared with BAU. Again, it is straightforward to see that the condition for strict first-order time and second-order stochastic dominance is not satisfied for any of the policies. This is because, for all policies, there exists a time-period in which the lowest level of consumption per capita is realised under the mitigation policy rather than BAU.

[FIGURE 3 HERE]

Unable to establish strict TSD of abatement over BAU, we now turn to analysing almost TSD. In particular, we look at both almost first-order TSD as set out in Definition 2 and almost first-order time and second-order stochastic dominance as set out in Definition 3. Recall that γ_k denotes the overall volume of violation of strict TSD relative to the total volume enclosed between G_i^j and F_i^j. ε_{kT} is the violation of strict TSD in the final time-period only, while γ_{1b} is the violation of strict first-order time and second-order stochastic dominance with respect to realisation b. As $\gamma_k, \varepsilon_{kT}, \gamma_{1b} \to 0.5$, the volume/area of violation accounts for half of the entire volume/area between the cumulative distributions being compared, while as $\gamma_k, \varepsilon_{kT}, \gamma_{1b} \to 0$ there is no violation.

What is striking about the results of analysing almost TSD in Table 2 is how small the violations are. For all of the policies, especially the violation of strict

first-order TSD is tiny relative to the total volume/area, as measured by γ_1 and ε_{1T}. Therefore we have a formal result showing that everyone would prefer any of the abatement policies to BAU, as long as their time and risk preferences can be represented by functions in the set $V_1(\gamma_1) \times U_1(\gamma_1)$. Moreover we can say that those who do not prefer the abatement policies have an extreme combination of time and risk preferences. Violation of strict first-order time and second-order stochastic dominance is also on the whole very small, and note that the condition on $D_1^2(b, T)$ in Definition 3 – equivalently $H_1^{-1,2}(p, T) \geq 0$ – is met by all policies.

[TABLE 2 HERE]

Notice that the overall violation increases with the stringency of the policy. Indeed it does so at an increasing rate, so that when $\overline{M^{AT}} = 450$ it is no longer entirely clear, given the inherent ambiguity in the concept γ_k, whether we should think of it as being favoured by almost every decision maker in $V_1 \times U_2$. Rather, for this policy especially, we must recognise that a non-trivial subset of decision-makers would not prefer to implement emissions reductions over BAU.

Let us now use TSD analysis to compare the various abatement policies with each other. We know from the analysis above that strict TSD will not exist either to a first order or in the case of first-order time and second-order stochastic dominance. Therefore we can proceed directly to analysing violations. In doing so we confine our attention to the least restrictive first-order TSD, given the wealth of pairwise comparisons that could potentially be made. Table 3 presents the results, in terms of violations of strict first-order TSD. The table should be read such that F_1^1 is the CO_2 limit in the first column and G_1^1 is the limit in the top row. So, for example, $\gamma_1 = 0.00859$ is the violation of strict first-order TSD for $\overline{M^{AT}} = 450$ over $\overline{M^{AT}} = 650$.

Although we might have expected the violations to be in the main large, since the abatement policy controls are much more similar to each other than they are to BAU – and they do tend to be higher than in the comparison with BAU – in fact they are all relatively small in absolute terms, such that for any pair of policies the lower CO_2 limit in the pair is almost dominant. Therefore we can go further and say that there exists a broad space for agreement, represented by everyone whose preferences are in the set $V_1(\gamma_1) \times U_1(\gamma_1)$, for tough emissions reduction targets, as tough as $\overline{M^{AT}} = 450$.

[TABLE 3 HERE]

6.2   How DICE yields these results

The topography of the panels in Figure 2 tells us much about the effect of emissions abatement on consumption per capita in DICE, how this effect is related to time and the nature of the uncertainty about it. In this century we can see it is often the case that $H_1^{-1,1} < 0$, but the surface appears flat as there is little difference between the cumulative distributions. In the next

century, however, the surface rises to a peak at high quantiles, revealing that the mitigation policies can yield much higher consumption per capita than BAU, albeit there is much uncertainty about whether this will eventuate and there is only a low probability associated with it. Comparing the policies, we can see that it is more likely that $H_1^{-1,1} < 0$, the more stringent is the limit on the atmospheric stock of CO_2. However, what 2 does not show, due to truncating the vertical axes in order to obtain a better resolution on the boundary between $H_1^{-1,1} < 0$ and $H_1^{-1,1} \geq 0$, is that conversely the peak difference in consumption per capita is higher, the more stringent is the concentration limit.

What lies behind these patterns? In fact, Figure 2 can be seen as a new expression of a well known story about the economics of climate mitigation. In early years, the climate is close to its initial, relatively benign conditions, yet significant investment is required in emissions abatement. This makes it rather likely that consumption per capita will initially be lower under a mitigation policy than under BAU. How much lower depends in the main on uncertainty about the cost of mitigation, and this in turn depends in the main on backstop-price uncertainty in our version of DICE. However, in later years the BAU atmospheric stock of CO_2 is high, so the possibility opens up that emissions abatement will deliver higher consumption per capita. How much higher depends in the main on how much damage is caused by high atmospheric CO_2 and therefore how much damage can be avoided by emissions abatement. In our version of DICE this is highly uncertain − much more so than the cost of emissions abatement − and depends principally on the climate sensitivity and the damage function coefficient $\tilde{\alpha}_3$ in (5). It is here that the driving force can be found behind the tiny violations of strict TSD in Table 2, namely the small possibility, in the second half of the modelling horizon, that the mitigation policies will deliver much higher consumption per capita than business as usual. This is consistent with the observation in previous, related research that the tails of the distribution are critical in determining the benefits of emissions abatement (e.g. Dietz, 2011; Weitzman, 2009).

In addition, productivity growth is a large source of uncertainty throughout, which affects BAU consumption per capita and all that depends on it. When productivity grows quickly, consumption per capita does likewise, as do CO_2 emissions, all else being equal. This increases the share of output that must be diverted to meeting a given limit on atmospheric CO_2, but at the same time it increases the global mean temperature, climate damage and the benefits of emissions abatement, again all else being equal. In our version of DICE, low productivity and output is associated with the lowest realisations of consumption per capita. Since in these states of nature there is little benefit to emissions reductions, it is low productivity that is pivotal in generating the violation of strict TSD in the first place.

19

7 Conclusions

In this paper we ask, is there space for agreement on climate change, in the specific sense of asking, are there climate policies that everyone who shares a broad class of time and risk preferences would prefer, first to business as usual and second to other policies? To find out we applied a new theory of Time-Stochastic Dominance, which enables time and risk preferences to be disentangled and dominance relations to be tested amongst options based only on partial information about the decision-maker or social planner's preferences on both dimensions. Our application was based on a stochastic version of the DICE model, in which eight key model parameters were randomised and Monte Carlo simulation was undertaken.

We were unable to establish strict TSD in the data, even when moving to second-order stochastic dominance (with first-order time dominance). However, when we analyse the related theory of Almost TSD we find that the volume of violation of strict TSD is generally very small indeed, so that we can say that almost all decision-makers would indeed favour any of our mitigation policies over BAU, and moreover that they would favour tougher mitigation policies over slacker alternatives. So the space for agreement is large in this regard.

Clearly our empirical results depend on the structure of the DICE model and how we have parameterised it. Of particular note are the key roles played by uncertainty about climate sensitivity, the curvature of the damage function, and productivity growth. Our parameterisation of the former two is key in producing a small violation of strict TSD, because when a high climate sensitivity combines with a high curvature on the damage function, $D_i^j(z,t)$ explodes. Our parameterisation of initial TFP growth, specifically our assumption via an unbounded normal distribution that it could be very low or even negative over long periods, is conversely key in producing a violation in the first place.

Our interpretation of γ_k, ε_{kT} and γ_{1b} in the application of Almost TSD is also open to debate, given the nature of the concept. Research on almost dominance relations is still at a relatively early stage, so we lack data on the basis of which we can say with high confidence that some preferences are extreme, while others are not. Nonetheless our violations are for the most part so small that we are somewhat immune to this criticism.

References

Anderson, B., Borgonovo, E., Galeotti, M., Roson, R., 2012. Uncertainty in integrated assessment modelling: can global sensitivity analysis be of help?

Arrow, K. J., Cline, W., Maler, K., Munasinghe, M., Squitieri, R., Stiglitz, J., 1996. Intertemporal equity, discounting, and economic efficiency. In: IPCC (Ed.), Climate Change 1995: Economic and Social Dimensions of Climate Change. Cambridge University Press, pp. 127–144.

Bøhren, Ø., Hansen, T., 1980. Capital budgeting with unspecified discount rates. Scandinavian Journal of Economics, 45–58.

Cline, W. R., 1992. The Economics of Global Warming. Peterson Institute.

Dasgupta, P., 2007. The stern review's economics of climate change. National Institute Economic Review 199 (1), 4–7.

Dietz, S., 2011. High impact, low probability? an empirical analysis of risk in the economics of climate change. Climatic Change 103 (3), 519–541.

Dietz, S., Asheim, G., 2012. Climate policy under sustainable discounted utilitarianism. Journal of Environmental Economics and Management 63, 321–335.

Dietz, S., Hepburn, C. J., 2013. Benefit-cost analysis of non-marginal climate and energy projects. Energy Economics 40, 61–71.

Dietz, S., Hope, C., Patmore, N., 2007. Some economics of dangerous climate change: Reflections on the stern review. Global Environmental Change 17 (3), 311–325.

Dietz, S., Matei, A. N., 2013. Spaces for agreement: a theory of time-stochastic dominance.

Ekern, S., 1981. Time dominance efficiency analysis. Journal of Finance 36 (5), 1023–1033.

Gerlagh, R., Liski, M., 2012. Carbon prices for the next thousand years.

Gollier, C., 2006. An evaluation of stern's report on the economics of climate change.

Ikefuji, M., Laeven, R. J., Magnus, J. R., Muris, C., 2012. Pareto utility. Theory and Decision, 1–15.

Kimball, M. S., 1990. Precautionary saving in the small and in the large. Econometrica, 53–73.

Leshno, M., Levy, H., 2002. Preferred by all and preferred by most decision makers: Almost stochastic dominance. Management Science 48 (8), 1074–1085.

Levy, H., Hanoch, G., 1970. Relative effectiveness of efficiency criteria for portfolio selection. Journal of Financial and Quantitative Analysis, 63–76.

Levy, H., Kroll, Y., 1979. Efficiency analysis with borrowing and lending: criteria and their effectiveness. Review of Economics and Statistics 61 (1), 125–130.

Levy, H., Leshno, M., Leibovitch, B., 2010. Economically relevant preferences for all observed epsilon. Annals of Operations Research 176 (1), 153–178.

Meehl, G., Stocker, T., Collins, W., Friedlingstein, A., Gaye, A., Gregory, J., Kitoh, A., Knutti, R., Murphy, J., Noda, A., Raper, S., Watterson, I., Weaver, A., Zhao, Z.-C., 2007. Global climate projections. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., Miller, H. (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Millner, A., forthcoming. On welfare frameworks and catastrophic climate risks. Journal of Environmental Economics and Management.

Millner, A., Dietz, S., Heal, G. M., 2013. Scientific ambiguity and climate policy. Environmental and Resource Economics 55 (1), 21–46.

Nordhaus, W. D., 1991. To slow or not to slow: the economics of the greenhouse effect. Economic Journal 101 (407), 920–937.

Nordhaus, W. D., 1994. Managing the Global Commons: the Economics of Climate Change. MIT Press, Cambridge, MA.

Nordhaus, W. D., 2007. A review of the "stern review on the economics of climate change". Journal of Economic Literature, 686–702.

Nordhaus, W. D., 2008. A Question of Balance: Weighing the Options on Global Warming Policies. Yale University Press.

Nordhaus, W. D., Boyer, J., 2000. Warming the World: Economic Models of Global Warming. MIT Press (MA).

Pindyck, R. S., 2011. Fat tails, thin tails, and climate change policy. Review of Environmental Economics and Policy 5 (2), 258–274.

Roe, G., Baker, M., 2007. Why is climate sensitivity so unpredictable? Science 318 (5850), 629–632.

Stern, N., 2007. The Economics of Climate Change: the Stern Review. Cambridge University Press.

Stern, N., 2008. The economics of climate change. American Economic Review: Papers and Proceedings 98 (2), 1–37.

Tol, R. S. J., 2012. On the uncertainty about the total economic impact of climate change. Environmental and Resource Economics 53 (1), 97–116.

Tzeng, L. Y., Huang, R. J., Shih, P.-T., 2012. Revisiting almost second-degree stochastic dominance. Management Science.

Weitzman, M., 2007. A review of the Stern Review on the economics of climate change. Journal of Economic Literature 45 (3), 703–724.

Weitzman, M., 2009. On modeling and interpreting the economics of catastrophic climate change. Review of Economics and Statistics 91 (1), 1–19.

Weitzman, M., 2012. Ghg targets as insurance against catastrophic climate damages. Journal of Public Economic Theory 14 (2), 221–244.

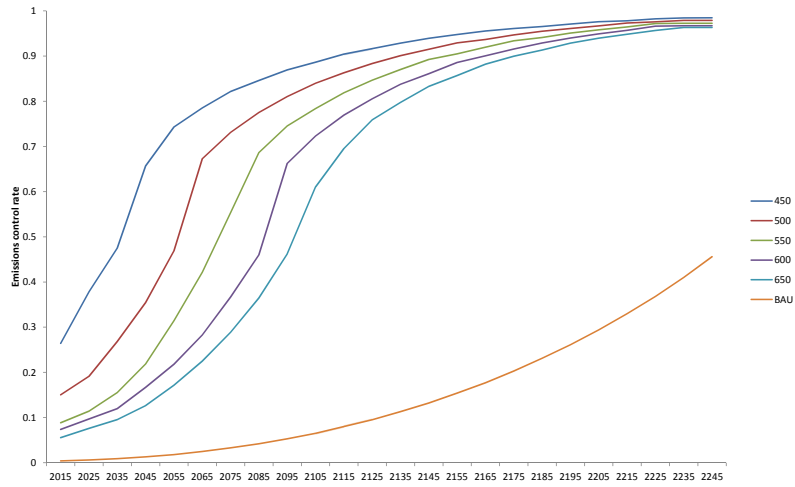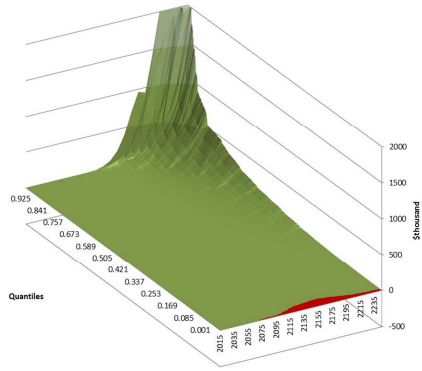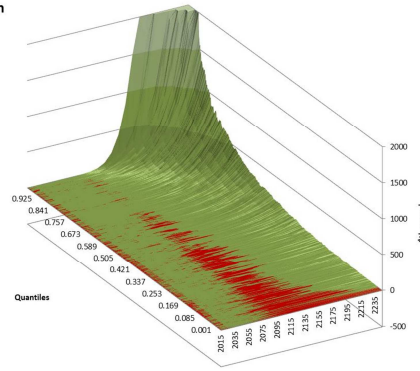Figure 1: Abatement policies in terms of the emissions control rate.

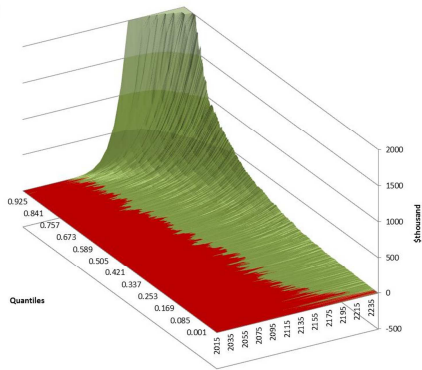Figure 2: $H_1^{-1,1}(p,t)$ for $\overline{M^{AT}} \in \{450,\ 500,\ 550,\ 600,\ 650\}$.
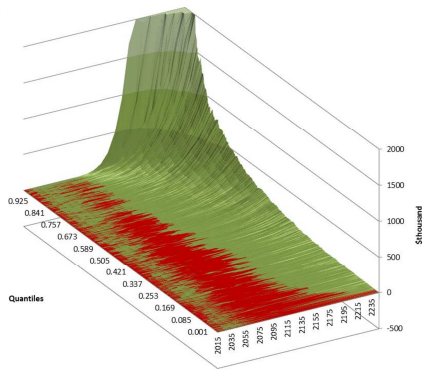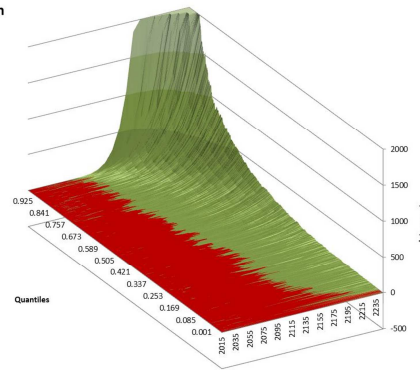
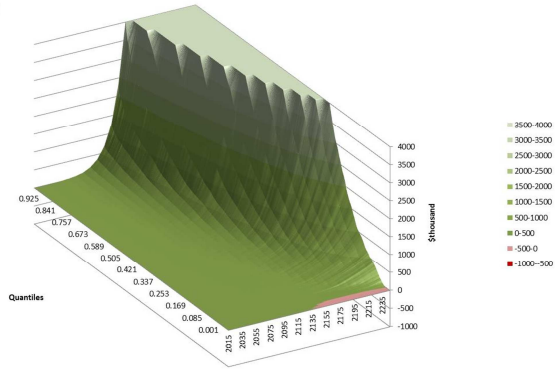Figure 3: $H_1^{-1,2}(p,t)$ for $\overline{M^{AT}} \in \{450, 500, 550, 600, 650\}$.
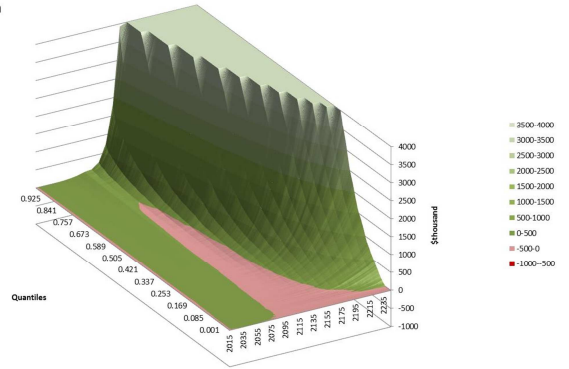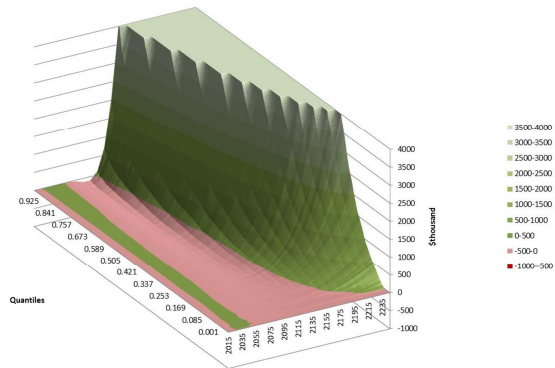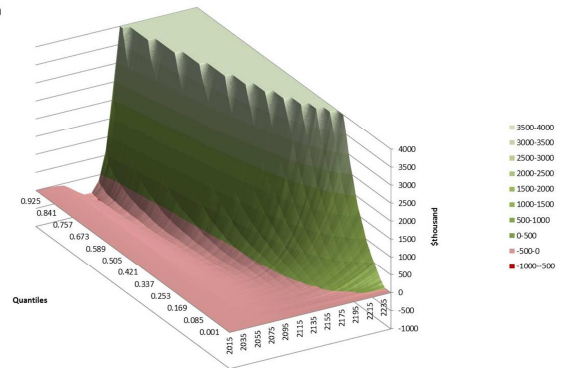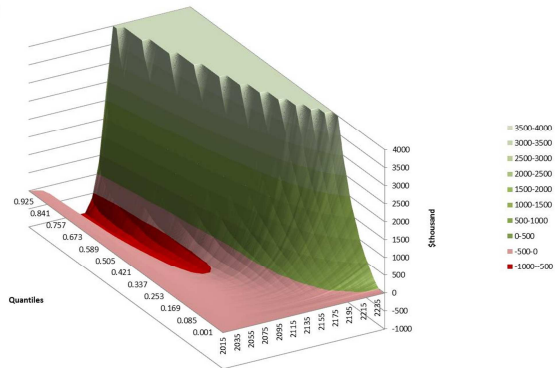
Table 1: Uncertain parameters for simulation of DICE.

Parameter	Units	Functional form	Mean	Standard deviation
(i) Initial growth rate of TFP	Per year	Normal	0.0092	0.004
(ii) Asymptotic global population	Millions	Normal	8600	1892
(iii) Rate of decarbonisation	Per year	Normal	-0.007	0.002
(iv) Total resources of fossil fuels	Billion tons of carbon	Normal	6000	1200
(v) Price of back-stop technology	US$ per ton of carbon replaced	Normal	1170	468
(vi) Transfer coefficient in carbon cycle	Per decade	Normal	0.189	0.017
(vii) Climate sensitivity	$^{\circ}$C per doubling of atmospheric CO_2	Log-normal	1.099*	0.3912*
(viii) Damage function coefficient α_3	Fraction of global output	Normal	0.082	0.028

*In natural logarithm space.

Table 2: Violations of strict first-order TSD and strict first-order time and second-order stochastic dominance.

CO_2 limit (ppm)	γ_1	ε_{1T}	γ_2	ε_{2T}	γ_{1b}
650	0.00009	0.00003	0.00002	8E-07	0
600	0.00045	0.00003	0.00045	2E-06	6.01E-08
550	0.00092	0.00003	0.00231	2E-06	0.00014
500	0.00188	0.00004	0.00605	3E-06	0.00086
450	0.00388	0.00004	0.01363	4E-06	0.00245

Table 3: First-order TSD analysis of abatement policies against each other.

CO_2 limit (ppm)	650		600		550		500	
	γ_1	ε_{1T}	γ_1	ε_{1T}	γ_1	ε_{1T}	γ_1	ε_{1T}
600	0.00255	0.00012						
550	0.00351	0.00011	0.01054	0.00034				
500	0.00517	0.00011	0.01260	0.00032	0.01764	0.00050		
450	0.00859	0.00013	0.01870	0.00036	0.02480	0.00052	0.03701	0.00107