

Simon Dietz, Nicoleta Anca Matei

Spaces for agreement: a theory of time-stochastic dominance and an application to climate change

**Article (Published version)
(Refereed)**

Original citation:

Dietz, Simon and Matei, Nicoleta Anca (2016) *Spaces for agreement: a theory of time-stochastic dominance and an application to climate change*. *Journal of the Association of Environmental and Resource Economists*, 3 (11). pp. 85-130. ISSN 2333-5955

DOI: [10.1086/683684](https://doi.org/10.1086/683684)

© 2016 [Association of Environmental and Resource Economists](http://www.aere.org/)

This version available at: <http://eprints.lse.ac.uk/64182/>

Available in LSE Research Online: March 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

Spaces for Agreement: A Theory of Time-Stochastic Dominance and an Application to Climate Change

Simon Dietz, Nicoleta Anca Matei

Abstract: Many investments involve both a long time horizon and risky returns. Making investment decisions thus requires assumptions about time and risk preferences. Such assumptions are frequently contested, particularly in the public sector, and there is no immediate prospect of universal agreement. Motivated by these observations, we develop a theory and method of finding “spaces for agreement.” These are combinations of classes of discount and utility function, for which one investment dominates another (or “almost” does so), so that all those whose preferences can be represented by such combinations would agree on the option to choose. The theory combines the insights of stochastic dominance and time dominance and offers a nonparametric approach to intertemporal, risky choice. We then apply the theory to climate change and show using a popular simulation model that even tough carbon emissions targets would be chosen by almost everyone, barring those with arguably “extreme” preferences.

JEL Codes: D61, H43, Q54

Keywords: Almost stochastic dominance, Climate change, Discounting, Integrated assessment, Project appraisal, Risk aversion, Stochastic dominance, Time dominance, Time-stochastic dominance

WHEN MAKING INVESTMENT DECISIONS one is frequently confronted with long time horizons and risky returns. Therefore, assumptions about time and risk preferences are important. Making such assumptions is always tricky. In the area of public project appraisal they are especially contested, because, on top of the usual

Simon Dietz (corresponding author) is at the Grantham Research Institute on Climate Change and the Environment, ESRC Centre for Climate Change Economics and Policy, and Department of Geography and Environment, London School of Economics and Political Science (s.dietz@lse.ac.uk). Nicoleta Anca Matei is at the European Commission, Joint Research Centre (JRC), Institute for Prospective Technological Studies (JRC-IPTS), Seville. The

Received October 16, 2014; Accepted August 20, 2015; Published online January 11, 2016.

JAERE, volume 3, number 1. © 2016 by The Association of Environmental and Resource Economists. All rights reserved. 2333-5955/2016/0301-0003\$10.00 <http://dx.doi.org/10.1086/683684>

challenges of estimating individual preferences, there are positions to be taken on how to aggregate individual preferences into social preferences.

A particularly good example of a long-run, risky public investment is climate-change mitigation. It comes as no surprise then that great controversy surrounds policy proposals to abate greenhouse gas emissions and that this controversy has turned in large measure on positions taken on time and risk preferences. By now the debate will be familiar to readers, so a very short summary might suffice here.

In the context of a model where social welfare is the discounted sum of individual utilities, the pioneering studies of Cline (1992) and Nordhaus (1991, 1994) staked out debating positions on pure time preference that still hold today—Cline set the utility discount rate to 0% based on so-called prescriptive ethical reasoning, while Nordhaus set it to 3% based on a more conventional descriptive analysis of market rates of investment returns.¹ More recently, the *Stern Review* (Stern 2007) set the utility discount rate to 0.1% and advocated aggressive emissions abatement, with the former assumption seemingly causing the latter result.² However, the *Stern Review* also prompted debate about the appropriate utility function, which in the standard model simultaneously represents risk preferences and preferences to smooth consumption over time. Questions have included the appropriate degree of risk/inequality aversion in an iso-elastic function (e.g., Gollier 2006; Dasgupta 2007; Stern 2008) and the appropriate function itself (Pindyck 2011; Ikefuji et al. 2013).

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 preferences legitimately exists and will endure. Given the ingredients of the debate and the current state of knowledge, “reasonable minds may differ” (Hepburn and Beckerman 2007). Why is the debate difficult to resolve? It contains normative and positive elements. There is a clear sense in which normative differences may never be completely eliminated. Positive “uncertainties” could in principle be eliminated by collecting more

views expressed are purely ours and may not in any circumstances be regarded as stating an official position of the European Commission. We would like to acknowledge very constructive comments from the editor and two anonymous referees. We have also benefited from the input of Alec Morton, Claudio Zoli, and seminar participants at the Australian Agricultural and Resource Economics Society annual conference in 2014, the Global-IQ meetings in Brussels and Rome, the Graduate Institute Geneva, Hamburg University, London School of Economics, Manchester University, the University of Paris I, and the World Congress of Environmental and Resource Economists in 2014. This research has received funding from the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement 266992—Global-IQ “Impacts Quantification of Global Changes.” The usual disclaimer applies.

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 critiques of the *Stern Review*.

empirical data from, for instance, market behavior, questionnaire surveys, or laboratory experiments, but in reality it is likely that they will also persist (as with long-standing puzzles in the economics of risk, such as the equity premium and risk-free rate).

Consequently, we are in the search for partial rather than complete orderings of choices. We want to establish a theory and method of identifying whether there exist “spaces for agreement,” that is, combinations of classes of discount and utility function, for which one investment dominates another (or “almost” does so), so that all decision makers whose preferences can be represented by such combinations would agree on the option to be chosen.

Why might this be useful? Given disagreement about appropriate time and risk preferences, our approach does not require decision makers to make a priori choices of functional form or parameter values. While this nonparametric approach could be used to inform investment choice in the private sector, its main use is more likely to be to bring renewed clarity to hotly contested choices in public policy, such as mitigation of climate change. In these areas, the debate about time and risk preferences might have become a distraction, preventing us from asking whether in fact there are some meaningful courses of action that both sides could agree to take.

The intellectual antecedents of this paper lie in the theory of Stochastic Dominance (Fishburn 1964; Hadar and Russell 1969; Hanoach and Levy 1969; Rothschild and Stiglitz 1970) and its offshoots, in particular, Almost Stochastic Dominance (Leshno and Levy 2002), Time Dominance (Bøhren and Hansen 1980; Ekern 1981), and extensions of dominance analysis to multivariate problems (Levy and Paroush 1974b; Atkinson and Bourguignon 1982; Karcher, Moyes, and Trannoy 1995).

Stochastic Dominance (SD) is a fundament of the theory of decision making under uncertainty. It is undoubtedly useful for the sort of problems we have just set out, precisely because it offers a nonparametric approach to risky choice, whereby one tests for SD relations for whole preference classes. However, the basic theory of SD is atemporal. In effect, decisions are made and payoffs obtained in the same time period. While extensions have been made to the multiperiod case (Levy 1973; Levy and Paroush 1974a), the decision maker is not permitted to prefer flows of utility in some periods of time more than in others.³ This is a serious drawback, as it is clear that most decision makers are impatient, preferring utility now to utility later on. Time preference is, by contrast, the core focus of the theory of time dominance (Bøhren and Hansen 1980; Ekern 1981), which takes the SD machinery and applies it to cash flows, that is, instead of working with cumulative distributions over the consequence space of a decision, one works with cumulative distributions over time.

3. One exception we are aware of is Scarsini (1986), who looked at a special case of utility discounting. We will clarify the relationship between his paper and ours later.

Like SD, one tests for a Time Dominance (TD) relation for whole preference classes, rather than having to pre-specify and parameterize a discount function. The drawback of TD, however, is the obverse of SD, namely, that the basic theory has been developed for certain, rather than uncertain, cash flows, and can only be extended to the latter under restrictive assumptions (see appendix A).

Another drawback of the basic theory of SD is nicely illustrated by a stylized example from Levy (2009)—try to use SD criteria to rank two prospects, one of which pays out \$0.5 with a probability of 0.01 and \$1 million with a probability of 0.99, and the other of which pays out \$1 for sure. While it would seem that virtually any investor would prefer the former, SD cannot be established.⁴ Arguably this paradox betrays the disadvantage of SD's generality—within the classes of utility function considered, there are some “extreme” (Leshno and Levy 2002) or even “pathological” (Levy 2009) utility functions, according to which the latter prospect is preferred.⁵ For this reason Leshno and Levy (2002) derived Almost Stochastic Dominance (Almost SD), according to which one compares the area between the cumulative distributions in which SD is violated with the total area between the distributions. Crucially, the ratio of the former to the latter can be given an interpretation in terms of restrictions on the class of utility functions, and if the restriction is very small, an Almost SD relation can be argued to exist.

This sets the conceptual task for the present paper, which is to unify the theories of SD and TD so that we have at our disposal a general framework for choosing between risky, intertemporal prospects, which admits the possibility of pure-time discounting and makes weak assumptions about the risk characteristics of the prospects: Time-Stochastic Dominance (TSD). In addition, we extend the notion of Almost SD to our bidimensional time-risk setup, defining Almost TSD. This provides a way to exclude extreme combinations of time and risk preferences and promises to greatly increase the practical usefulness of the framework.

We then make an empirical application of the theory to climate change, by analyzing a set of trajectories for global greenhouse gas emissions—a set of “policies”—using a stochastic version of the benchmark DICE integrated assessment model devised by

4. Where $F^n(x)$ and $G^n(x)$ are respectively the n th-order cumulative distributions of the former and latter prospects over realizations x , this is because the first nonzero values of $G^n(x) - F^n(x)$ are negative as x increases from its lower bound. However, n th-order SD requires that $G^n(x) - F^n(x) \geq 0, \forall x$, and $G^{k-1}(b) - F^{k-1}(b) \geq 0, k=1, \dots, n$, with at least one strict inequality. See appendix A for further explanation.

5. In the example used, one would be

$$u(x) = \begin{cases} x & \text{for } x \leq 1 \\ 1 & \text{for } x > 1 \end{cases}.$$

Nordhaus.⁶ Our results show the climate-change debate in a new light. Although the profile of net benefits from climate mitigation is such that “exact” TSD cannot be established, the less restrictive concept of Almost TSD allows us to show that the space for agreement on climate change is indeed large. Since Almost TSD is based on the notion of excluding extreme combinations of time and risk preferences, this result in particular lends itself to the following rather stark interpretation: only those with “extreme” preferences over time and risk would prefer not to cut carbon emissions by a large amount.

The remainder of the paper is set out as follows. In the next short section, we deal with some analytical preliminaries; in particular, we set out the classes of utility and discount function that will be of primary focus. In section 2 we establish the theory of (exact) TSD, while in section 3 we do the same for Almost TSD. Section 4 describes how we set up the DICE model, while section 5 presents our results, and section 6 concludes.

1. SPACES FOR AGREEMENT

Readers interested in quickly getting up to speed with the existing literatures on SD and TD theory are referred to the short primer in appendix A. Building on this, let us take the task at hand as being to rank two prospects X and Y , both of which yield random cash flows over time. The underlying purpose is to compare the expected discounted utilities of the prospects at $t = 0$, that is, for prospect X we compute

$$NPV_{v,u}(X) = \int_0^T v(t)E_F u[x(t)]dt = \int_0^T v(t) \left[\int_a^b u(x)f(x,t)dx \right] dt, \tag{1}$$

where x is a realization of the cash flow of prospect X , v is a discount function, and u is a utility function. Both functions v and u are assumed to be continuous and continuously differentiable at least once. We make the assumptions, characteristic in the dominance literature, that the random cash flows of X and Y are both supported on the finite interval $[a, b]$ and that each prospect pays out over a finite, continuous time horizon $[0, T]$. Therefore we can characterize a probability density function (pdf) for prospect X at time $t \in [0, T]$, $f(x, t)$, and a counterpart cumulative distribution function (cdf) with respect to realization $x \in [a, b]$ at time $t \in [0, T]$, $F^1(x, t) = \int_a^x f(s, t)ds$. Note that because utility is additively separable across time in (1), no particular assumption is required about the serial correlation of the probability distribution (Levy and Paroush 1974a).

6. The first version appeared in Nordhaus (1993b, 1993a). We build on the version in Nordhaus (2008).

Before characterizing time-stochastic dominance (TSD), we need to define classes of utility and discount functions. Our broadest class of utility function $u : [a, b] \rightarrow \mathbb{R}$ is $U_1 = \{u : u'(x) \geq 0\}$, that is, the class of utility functions, whereby utility is nondecreasing as a function of consumption, representing nothing more than (weak) nonsatiation. It is hard to imagine relevant circumstances in which the appropriate utility function would not be in U_1 . More generally, any subset m of utility functions is defined recursively as

$$U_m = \{u : u \in U_{m-1} \text{ and } (-1)^m u^m(x) \leq 0\},$$

where, among other things, m represents the number of times that $u(x)$ is differentiated. As well as U_1 , in this paper we focus on $U_2 = \{u : u \in U_1 \text{ and } u''(x) \leq 0\}$, which is the class of nondecreasing, weakly concave utility functions, ruling out risk seeking. Whether the appropriate utility function is in U_2 is a little less clear, but it is almost certainly a good description of most individual behavior, and there are few if any arguments for public policy evaluation to be based on risk seeking. Eventually we establish a theorem for TSD of an arbitrarily high order with respect to both time and risk.

Let us define a corresponding set of discount functions on the time domain, $v : [0, T] \rightarrow \mathbb{R}$. The broadest class of discount functions requires simply that at any point in time more is preferred to less, $V_0 = \{v : v(t) > 0\}$. However, V_0 is of little interest, since some positive degree of time preference is always required, however small. Therefore, without compromising the generality of our theory, let us focus our attention on the first- and second-order restrictions on V_0 :

$$V_1 = \{v : v \in V_0, \text{ and } v'(t) < 0\},$$

$$V_2 = \{v : v \in V_1, \text{ and } v''(t) > 0\}.$$

The class V_1 comprises strictly decreasing discount functions, exhibiting positive time preference, while V_2 is the class of strictly decreasing, convex discount functions, according to which impatience decreases over time. More generally, any subset n of discount functions is defined recursively as

$$V_n = \{v : v \in V_{n-1} \text{ and } (-1)^n v^n(t) > 0\}.$$

Note that V_1 and V_2 admit both exponential and hyperbolic discounting as special cases. Exponential discounting has long been the conventional approach to pure time preference, with debate focusing on the discount rate rather than the functional specification. However, arguments have been advanced for hyperbolic discounting, including that it is a more appropriate description of real individual behavior (Laib-

son 1997) and that it can result from the aggregation of heterogeneous individual preferences.⁷

Combinations of these classes of utility and discount functions constitute possible *spaces for agreement*. The combination $V_1 \times U_1$ is the largest possible space for agreement that we consider, encapsulating any decision maker whose preferences can be represented by, respectively, a strictly decreasing discount function and a nondecreasing utility function, in other words any impatient decision maker with any attitude to risk. Presumably virtually all decision makers belong to this combination of classes. By contrast $V_1 \times U_2$, for instance, encapsulates any impatient decision maker who is not risk seeking. Whether there is an actual space for agreement depends of course on whether any dominance relations can be established between projects, for the combination in question. Note that in section 3 we narrow these spaces for agreement further by placing additional restrictions on V and U with a view to excluding “extreme” combinations of time and risk preferences.

2. TIME-STOCHASTIC DOMINANCE

A further piece of notational apparatus will enable us to work in a compact, bidimensional form. Denote the integral over time of the pdf by $F_1(x, t) = \int_0^t f(x, w)dw$, while the integral over time of the cdf is

$$F_1^1(x, t) = \int_a^x F_1(s, t)ds = \int_0^t F^1(x, w)dw = \int_0^t \int_a^x f(s, w)dsdw.$$

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:

$$D_n(z, t) = \int_0^t D_{n-1}(z, w)dw$$

$$D^m(z, t) = \int_a^z D^{m-1}(s, t)ds$$

$$D_n^m(z, t) = \int_0^t D_{n-1}^m(z, w)dw = \int_a^z D_n^{m-1}(s, t)ds = \int_0^t \int_a^z D_{n-1}^{m-1}(s, w)dsdw,$$

7. Even though those individual preferences are represented by exponential discounting (see Gollier and Zeckhauser 2005).

where $i \in \{1, 2, \dots, n\}$ is the order of TD (i.e., the number of integrations with respect to time) and $j \in \{1, 2, \dots, m\}$ is the order of SD (i.e., the number of integrations with respect to the probability distribution). Note that our concept of TD relates to pure time discounting, whereas standard TD relates to discounting of consumption.

With all of our notation now set out, let us characterize TSD for various combinations of classes of U_j and V_i .

Definition 1: [Time-stochastic dominance of order n, m] For any two risky, intertemporal prospects X and Y ,

$$X >_{nTmS} Y \text{ if and only if } \Delta \equiv NPV_{v,u}(X) - NPV_{v,u}(Y) \geq 0, \text{ for all } (v, u) \in V_n \times U_m.$$

In this definition, the ordering $>_{nTmS}$ denotes pure TD of the n^{th} order, combined with SD of the m^{th} order. For example, $>_{1T1S}$, which we can shorten to $>_{1TS}$, denotes pure-time and stochastic dominance of the first order.

Proposition 1: [First-order time-stochastic dominance] $X >_{1TS} Y$ if and only if

$$D_1^1(z, t) \geq 0, \forall z \in [a, b] \text{ and } \forall t \in [0, T],$$

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

Proof: See appendix B.

Proposition 1 tells us that any impatient planner with monotonic nondecreasing preferences will prefer prospect X to prospect 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 income levels and all time periods, and is strictly larger somewhere. 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 within these classes.⁸

Consider the following stylized example, comprising discrete cash flow distributions in discrete time. The use of discrete data makes the exposition easy; moreover, it is also the form of data that would typically be encountered in practical applications; for instance, the output of the DICE climate-change model is in just this form. However, it means that we have to relate proposition 1, stated in terms of cumulative distribution functions, to a parallel theorem stated in terms of quantile distribution functions.

8. Proposition 1 is similar to theorem 3 in Scarsini (1986). However, Scarsini did not consider any other cases, that is, any other combinations of time and risk preference.

Example 1: Consider prospects X and Y , each of which comprises a cash flow over five periods of time and in four states of nature with equal probability (i.e., uniform discrete distributed):

Prospect	Probability	Time Period				
		0	1	2	3	4
X	1/4	-2	-3	2	2	1
	1/4	-1	-2	-2	3	1
	1/4	0	-2	-2	5	6
	1/4	0	0	-2	4	2
	1/4	-5	-3	2	3	7
Y	1/4	-4	-3	2	3	1
	1/4	-4	-1	-1	0	1
	1/4	-4	0	1	1	6

Instead of integration with respect to time, we simply use summation. For each additional restriction placed on the curvature of the discount function, a new round of summation of the cash flows is performed, $X_n(t) = \sum_{w=0}^t X_{n-1}(w)$. Matters on the stochastic dimension are a little more involved: we extend the quantile approach of Levy and Hanoch (1970) and Levy and Kroll (1979). 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

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

Where $H_1^{-1,r}(p, t) = F_1^{-1,r}(p, t) - G_1^{-1,r}(p, t)$, we can characterize first-order time-stochastic dominance for quantile distributions:

Proposition 2: (1TSD for quantile distributions). $X >_{1TS} Y$ if and only if

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

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

Proof: See appendix B.

It can easily be shown that proposition 2 applies to discrete data.⁹ However, when the data are discrete and are not serially independent, as is the case in all of our examples here and in our climate application later, the order of cumulation matters. The functions $F_1^{-1}(p, t)$ and $G_1^{-1}(p, t)$ are obtained by first cumulating the cash flows across time and then reordering from lowest to highest in each time period. Taking the difference between them gives us $H_1^{-1}(p, t)$. In essence, since the quantile distribution function is just the inverse of the cumulative distribution function, 1TSD requires $F_1^{-1}(p, t) - G_1^{-1}(p, t) \geq 0$, that is, the inverse of the requirement for 1TSD in terms of cumulative distributions.

In the case of example 1, computing the quantile distributions gives us:

p	Time Period				
	0	1	2	3	4
.25	3	3	1	4	4
.5	3	4	2	2	1
.75	4	3	2	3	0
1	4	4	1	4	3

Therefore, by propositions 1 and 2, $X >_{1TS} Y$.

Having established first-order TSD, we can proceed from here by placing an additional restriction on the discount function and/or on the utility function. A particularly compelling case is the assumption of impatience combined with risk aversion/neutrality— $(v, u) \in V_1 \times U_2$ —since few would be uncomfortable with the notion of excluding risk-seeking behavior a priori, especially in the public sector.

Proposition 3: [First-order time and second-order stochastic dominance] $X >_{1T2S} Y$ if and only if

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

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

Proof: See appendix B.

9. Choose an arbitrary quantile $p^*(t) \in [0, 1]$ for any t and denote $G_1^{-1}(p^*, t) = z_2(t)$ and $F_1^{-1}(p^*, t) = z_1(t)$. We need to show that $z_1(t) \geq z_2(t)$ for each t . Assume that $z_1(t) < z_2(t)$. By definition, $x_2(t)$ represents the smallest value for which equation (2) holds and for this reason $z_1(t)$ and $z_2(t)$ cannot be located on the same step of the $G_1^1(z, t)$ for any t . Therefore $G_1^1(z_1, t) < G_1^1(z_2, t)$. We have that $G_1^1(z_1, t) < G_1^1(z_2, t) = p^*(t) = F_1^1(z_1, t) < F_1^1(z_2, t)$. Thus $G_1^1(z_1, t) < F_1^1(z_1, t)$, which contradicts the initial assumption. This proves sufficiency, and necessity can be demonstrated in a very similar way.

It is evident from proposition 3 and its proof that, in line with the classical approach to SD, restricting the utility function by one degree corresponds to integrating the bidimensional probability distribution $D_1^1(z, t)$ once more with respect to the consequence space.

Example 2: Now consider two different prospects X and Y :

Prospect	Probability	Time Period				
		0	1	2	3	4
X	1/4	-4	-1	2	3	9
	1/4	-1	-3	2	2	7
	1/4	-1	-1	2	0	4
	1/4	0	0	2	2	2
Y	1/4	-5	-1	2	2	2
	1/4	-2	-3	-1	3	6
	1/4	-2	0	0	2	5
	1/4	0	0	2	1	8

In this example $H_1^{-1,1}(p, t)$ is:

p	Time Period				
	0	1	2	3	4
.25	1	1	3	3	4
.5	1	1	2	2	3
.75	1	0	2	0	2
1	0	0	0	1	-2

While in the first four time periods $H_1^{-1,1}(p, t) \geq 0$, the opposite is true when $p = 1$ in the terminal period. Therefore first-order TSD cannot be established between these two prospects. However, cumulating once more with respect to the consequence space gives $H_1^{-1,2}(p, t)$, which here is:

p	Time Period				
	0	1	2	3	4
.25	1	1	3	3	4
.5	2	2	5	5	7
.75	3	2	7	5	9
1	3	2	7	6	7

Thus from proposition 3 and by extension of proposition 2 we can say that $X >_{1T2S} Y$. What this example illustrates is that, when the violation of first-order TSD is restricted to the upper quantiles of $F_1^{-1,1}$ and $G_1^{-1,1}$, the additional restriction that $u \in U_2$, which excludes risk-seeking behavior, makes it disappear, because relatively greater weight is placed on outcomes with low income.

If we want to pursue the further case of $(v, u) \in V_2 \times U_2$, representing a risk-averse or risk-neutral planner with a decreasing and convex discount function, then integrate $D_1^2(z, t)$ once more with respect to time.

Proposition 4: [Second-order time-stochastic dominance] $X >_{2TS} Y$ if and only if

i)

$$D_1^2(z, T) \geq 0, \forall z \in [a, b],$$

ii)

$$D_2^2(z, t) \geq 0, \forall z \in [a, b] \text{ and } \forall t \in [0, T],$$

and there is at least one strict inequality.

Proof: See appendix B.

The second part of the dominance condition tells us that, in order for X to be preferred to Y by any decision maker with preferences consistent with $(v, u) \in V_2 \times U_2$, the cdf of X , integrated twice over time and once more over the consequence space, must be nowhere larger than its counterpart for Y . Additionally, first-order time and second-order stochastic dominance must hold with respect to the difference between the distributions in the terminal period T .

Example 3: Now consider another two different prospects:

Prospect	Probability	Time Period				
		0	1	2	3	4
X	1/4	-5	-2	2	1	8
	1/4	-3	-3	2	4	10
	1/4	-1	-1	-2	0	0
	1/4	0	-2	-1	2	4
Y	1/4	-5	-2	-2	5	0
	1/4	-4	-3	-2	5	2
	1/4	-2	-3	2	0	7
	1/4	0	0	2	3	9

The reader can verify that in this example the condition for first-order TSD of either X or Y is not met. Further, $H_1^{-1,2}$ is:

p	Time Period				
	0	1	2	3	4
.25	0	0	4	0	0
.5	1	1	9	0	5
.75	2	4	8	2	5
1	2	2	3	-3	1

Therefore in this case neither is the condition for first-order time and second-order stochastic dominance met. The next step is to inspect $H_2^{-1,2}$:¹⁰

p	Time Period				
	0	1	2	3	4
.25	0	0	4	4	12
.5	1	2	11	15	23
.75	2	6	14	17	29
1	2	4	7	4	5

Thus since $H_2^{-1,2} \geq 0, \forall z, t$ with mostly strict inequalities, and from above $H_1^{-1,2} \geq 0, \forall p, X \succ_{2TS} Y$.

The previous cases provide us with the machinery we require to offer a theorem for TSD that is generalized to the n th order with respect to time and the m th order with respect to risk.

Suppose that information regarding the first n derivatives of the discount function ($v \in V_n$) and the first m of the utility function ($u \in U_m$) is provided. Then:

Proposition 5: [n th-order time and m th-order stochastic dominance] X n th-order time and m th-order stochastic dominates Y if and only if

i)

$$D_{i+1}^{j+1}(b, T) \geq 0,$$

10.

$$H_2^{-1,2}(p, t) = [F_2^{-1,2}(p, t) - G_2^{-1,2}(p, t)] = \sum_{w=0}^p [F_2^{-1,1}(w, t) - G_2^{-1,1}(w, t)].$$

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)

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

with iv holding as a strong inequality over some subinterval and where $i = \{0, \dots, n-1\}$ and $j = \{1, \dots, m-1\}$.

Proof: See appendix B.

Proposition 5 gives necessary and sufficient conditions for dominance for any decision maker having time and risk preferences represented by $(v, u) \in V_n \times U_m$. Note that, for appropriate values of m and n , propositions 1 (1TSD), 3 (1T2SD), and 4 (2TSD) can be obtained as specific cases. In these cases, conditions i–iii are not always explicitly required, as the satisfaction of condition iv, which is always required, can imply that some or all of the other conditions also hold.

3. ALMOST TIME-STOCHASTIC DOMINANCE

In practice, the usefulness of what we might call “exact” dominance analysis can be limited, since even a very small violation of the conditions for dominance is sufficient to render the rules unable to order investments. As the example above showed, if a violation exists in particular at the lower bound of the domain of the cumulative distribution functions, then no amount of restrictions will make it vanish. Put another way, the downside of a flexible, nonparametric approach is that the broad classes of preference on which the dominance criteria are based include a small subset of “extreme” or “pathological” functions, whose implications for choice would be regarded by many as perverse.¹¹ Leshno and Levy (2002) recognized this problem in the context of SD and developed a theory of almost stochastic dominance (Almost

11. What is “extreme” is clearly subjective, an obvious difficulty faced by the Almost SD approach. However, Levy, Leshno, and Leibovitch (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 most risk-seeking and risk-averse participants mark out the limits, and preferences outside these limits can be considered extreme.

SD), according to which restrictions are placed on the derivatives of the utility function, so that extreme preferences are excluded (see appendix A).¹² Dominance relations between risky prospects are then characterized for “almost” all decision makers.

It is obvious that exact TSD faces the same practical constraints as exact SD. In this section, we therefore extend our theory to “Almost TSD,” excluding extreme combinations of time and risk preferences so that prospects can still be ranked. In particular, by extending the theory to our bidimensional time-risk set-up, we define and characterize almost first-order TSD and almost first-order time and second-order stochastic dominance. In doing so, the attention of the analysis shifts subtly to asking: how many preference combinations must be excluded in order to obtain a ranking? Put another way, how much smaller is the space for agreement? In general, the less that need be excluded, the better.

Let us start with Almost first-order TSD. Our basic approach is analogous to Leshno and Levy (2002) in that we measure the violation of 1TSD relative to the nonviolation of 1TSD and give the resulting, relative measure of violation meaning by linking it with a restriction on time and risk preferences. We will need two restrictions on preferences, and two corresponding violation measures.

For expositional ease, we begin with a measure of violation ε_{1T} , which is in fact the violation measure in Leshno and Levy (2002), albeit in our bidimensional framework it is for $t = T$ specifically. For every $0 < \varepsilon_{1T} < 0.5$, define the following subset of U_1 :

$$U_1(\varepsilon_{1T}) = \left\{ u \in U_1 : \frac{u'(z)}{\inf[u'(z)]} \leq \left[\frac{1}{\varepsilon_{1T}} - 1 \right], \forall z \in [a, b], t = T \right\}.$$

The class $U_1(\varepsilon_{1T})$ is the set of nondecreasing utility functions with the added restriction that the ratio between maximum and minimum marginal utility is bounded by $(1/\varepsilon_{1T}) - 1$, that is, extreme concavity/convexity is ruled out. It is easiest to see what this restriction entails in the case of $u \in U_1(\varepsilon_{1T})$, where $u''(z)$ is monotonic. Then we are restricting how much (little) marginal utility members of the class of functions associate with low income levels at the same time as restricting how little (much) marginal utility they associate with high income levels. Further narrowing the scope to the very common case of utility functions with constant elasticity of marginal

12. Tzeng et al. (2012) showed that Leshno and Levy’s theorem for almost second-order stochastic dominance is incorrect and redefined the concept. They also extended the results to higher orders.

utility, the restriction is on the absolute value of the elasticity— $|[u''(z)z]/[u'(z)]|$ —such that it cannot be large negative or large positive, and the larger is ε_{1T} the smaller $|[u''(z)z]/[u'(z)]|$ must be. Of course this is merely an illustration—the set of utility functions $U_1(\varepsilon_{1T})$ is much larger than the constant-relative-risk aversion functions alone. In the limit as ε_{1T} approaches 0.5, the only function in $U_1(\varepsilon_{1T})$ is linear utility, where $u''(z) = 0$. Conversely as ε_{1T} approaches zero, $U_1(\varepsilon_{1T})$ coincides with U_1 .

We further introduce a restriction γ_1 on the product of the marginals of the discount and utility functions, such that

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

The class $(V_1 \times U_1)(\gamma_1)$ is the set of all combinations of decreasing pure time discount function and nondecreasing utility function, with the added restriction that the ratio between the maximum and minimum products of $-v'(t) u'(z)$ is bounded by $(1/\gamma_1) - 1$. The supremum (infimum) of $-v'(t) u'(z)$ is attained when $v'(t) < 0$ is the infimum (supremum) of its set and $u'(z) \geq 0$ is the supremum (infimum) of its.¹³ Therefore, the combinations of preferences that we are excluding here will tend to comprise extreme concavity or convexity of the utility and discount functions somewhere on their respective domains.

Now define the set of realizations $z \in [a, b]$ where there is a violation of first-order TSD as S_1^1 :

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

We also explicitly define $S^{1,T}$ as the subset of S_1^1 when $t = T$, that is, the difference between the single-dimensional cumulative distributions over the consequence space at time T :

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

13. It is worth noting that when T is large and $v \in V_2$ (say $v = e^{-rt}$), $\inf[-v'(t)]$ will be attained when $t = T$, and clearly in the limit as $T \rightarrow \infty$, $\inf[-v'(t)] = 0$. Therefore finite time is required in order to guarantee the existence of $\sup[-v'(t)]/\inf[-v'(t)]$. A long time horizon does not necessarily guarantee TD, especially if a violation occurs at $t = 0$ (see Matei and Zoli [2012] for a discussion on this drawback and the dictatorship of the present in TD criteria).

Definition 2: [Almost first-order time-stochastic dominance] X dominates Y by Almost first-order time-stochastic dominance, denoted $X >_{A1TS} Y$, if and only if

i)
$$\int \int_{S_t^c} -D_1^1(z, t) dz dt \leq \gamma_1 \int_0^T \int_a^b |D_1^1(z, t)| dz dt$$
 and

ii)
$$\int_{S^{1,T}} -D_1^1(z, T) dz \leq \varepsilon_{1T} \int_a^b |D_1^1(z, T)| dz.$$

Proposition 6: [A1TSD] $X >_{A1TS} Y$ if and only if, for all $(v, u) \in (V_1 \times U_1)$ (γ_1) and $u \in U_1(\varepsilon_{1T})$,

$$NPV_{v,u}(X) \geq NPV_{v,u}(Y).$$

Proof: See appendix B.

To reiterate, the definition of almost first-order TSD contains two measures of the violation of exact first-order TSD. γ_1 measures the cumulative violation of the nonnegativity condition on D_1^1 over all t , relative to the total volume enclosed between the distributions over all t , while ε_{1T} measures the violation of the same condition at time T only, relative to the total area enclosed between the distributions at that time. All decision makers exhibiting the “nonextreme” combination of preferences expressed by the discount and utility functions $(v, u) \in (V_1 \times U_1)$ (γ_1) and $u \in U_1(\varepsilon_{1T})$ will prefer X to Y if and only if conditions i and ii in proposition 6 are satisfied.

Moving now to almost first-order time and second-order stochastic dominance, we need three restrictions on preferences and three corresponding violation measures. First, $\gamma_{1,2}$ measures the relative violation of the nonnegativity condition on D_1^2 over all t . It is equivalent to the following restriction on combined time and risk preferences:

$$(V_1 \times U_2)(\gamma_{1,2}) = \left\{ v \in V_1, u \in U_2 : \frac{v'(t)u''(z)}{\inf[v'(t)u''(z)]} \leq \left[\frac{1}{\gamma_{1,2}} - 1 \right] \right. \\ \left. \forall z \in [a, b], \forall t \in [0, T] \right\}$$

The set $(V_1 \times U_2)$ ($\gamma_{1,2}$) represents all combinations of decreasing pure time discount functions and nondecreasing, weakly concave utility functions, with the added restriction that the ratio between the maximum and minimum of $v'(t) u''(z)$ is bounded by $(1/\gamma_{1,2}) - 1$. The supremum (infimum) of $v'(t) u''(z)$ is attained when

$v'(t) < 0$ and $u''(z) \leq 0$ are the suprema (infima) of their respective sets, and these sets are defined with respect to all realizations and time periods.

Second, ε_{2T} measures the relative violation of the nonnegativity condition on D_1^2 at time T only. As per Leshno and Levy (2002), for every $0 < \varepsilon_{2T} < 0.5$,

$$U_2(\varepsilon_{2T}) = \left\{ u \in U_2 : \frac{-u''(z)}{\inf[-u''(z)]} \leq \left[\frac{1}{\varepsilon_{2T}} - 1 \right], \forall z \in [a, b], t = T \right\}.$$

Class $U_2(\varepsilon_{2T})$ is the set of nondecreasing, weakly concave utility functions with the added restriction that the ratio between maximum and minimum $u''(z)$ is bounded by $(1/\varepsilon_{2T}) - 1$. Therefore large changes in $u''(z)$ with respect to z are excluded, where only realizations at time T are considered.

Third, we need to measure a violation of the nonnegativity condition on the integral with respect to time of $D_1^2(b, t)$. We denote this λ_{1b} and it is equivalent to restricting time preferences as follows:

$$V_1(\lambda_{1b}) = \left\{ v \in V_1 : \frac{-v'(t)}{\inf[-v'(t)]} \leq v \left[\frac{1}{\lambda_{1b}} - 1 \right], z = b, \forall t \in [0, T] \right\}.$$

Function $V_1(\lambda_{1b})$ is the set of decreasing pure time discount functions with the added restriction that the ratio between maximum and minimum $v'(t)$ is bounded by $(1/\lambda_{1b}) - 1$. Hence large changes in $v''(t)$ are excluded.

Parcel out the subset of realizations S_1^2 where $D_1^2 < 0$, that is, where the condition for exact first-order time and second-order stochastic dominance is violated:

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

Further explicitly define $S^{2,T}$ as the subset of S_1^2 when $t = T$:

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

And in this case we also need to define a subset of realizations where $D_1^2(b, t) < 0$, for any t where $z = b$:

$$S_{1,b}(D_1^2) = \{z = b, t \in [0, T] : D_1^2(b, t) < 0\}.$$

Definition 3: [Almost first-order time and second-order stochastic dominance] X almost first-order time and second-order stochastically dominates Y , denoted $X >_{A1T2S} Y$ if and only if

- i)
$$\int \int_{S_1^T} -D_1^2(z, t) dz dt \leq \gamma_{1,2} \int_0^T \int_a^b |D_1^2(z, t)| dz dt,$$
- ii)
$$\int_{S^{2T}} -D_1^2(z, T) dz \leq \varepsilon_{2T} \int_a^b |D_1^2(z, T)| dz,$$
- iii)
$$\int_{S_{1b}} D_1^2(b, t) dt \leq \lambda_{1b} \int_0^T |D_1^2(b, t)| dt, \text{ and}$$
- iv)
$$D_1^2(b, T) \geq 0.$$

Proposition 7: [A1T2SD] $X >_{A1T2S} Y$ if and only if, for all $(v, u) \in (V_1 \times U_2)$ $(\gamma_{1,2})$, $u \in U_2(\varepsilon_{2T})$, and $v \in V_1(\lambda_{1b})$,

$$NPV_{v,u}(X) \geq NPV_{v,u}(Y).$$

Proof: See appendix B.

Notice that the definition of almost first-order time and second-order stochastic dominance has a similar structure to proposition 5. It contains three measures of the violation of strict dominance, as well as the requirement that $D_1^2(b, T) \geq 0$, that is, that the difference between the undiscounted mean values of projects X and Y respectively is at least zero. The proposition says that all decision makers exhibiting the “nonextreme” combination of preferences expressed by the discount and utility functions $(v, u) \in (V_1 \times U_2)$ $(\gamma_{1,2})$, $u \in U_2(\varepsilon_{2T})$, and $v \in V_1(\lambda_{1b})$ will prefer X to Y if and only if conditions i–iv are satisfied.

Example 4. Consider the following two prospects:

Prospect	Probability	Time Period				
		0	1	2	3	4
X	1/4	-5	-3	0	4	7
	1/4	0	-3	1	2	10
	1/4	0	-2	1	3	9
	1/4	0	0	0	1	1
Y	1/4	-5	-1	0	3	9
	1/4	-4	-2	-1	0	1
	1/4	-2	-3	1	1	5
	1/4	-2	-1	-1	2	1

In this example $H_1^{-1,1}$ is:

p	Time Period				
	0	1	2	3	4
.25	0	-2	-1	3	8
.5	4	3	4	3	4
.75	2	3	3	4	8
1	2	3	4	4	5

First-order TSD cannot be established between these two prospects. Moreover it can easily be shown that the occurrence of the violation in the lowest quantile of $H_1^{-1,1}$, in early time periods, means that the violation will persist despite infinitely repeated cumulation with respect to time and/or the consequence space. However, it is quite evident from the tables that X performs better than Y most of the time, so let us inspect this example within the framework of Almost TSD. Doing the necessary calculations:

A1TSD		A1T2SD		
γ_1	ε_{1T}	γ_2	ε_{2T}	γ_{1b}
.04	0	.02	0	0

The small violations reflect what is intuitively obvious from $H_1^{-1,1}(p, t)$, namely, that only a small restriction on the combination of classes of discount and utility functions is required in order for dominance to be established, since $F < G$ most of the time in most quantiles.

4. MODELING CLIMATE MITIGATION POLICIES

We now turn to our application of the theory of TSD and Almost TSD to climate-change mitigation. The question we ask is: can we make choices on emissions abatement, without having to agree on how precisely to structure and parameterize time and risk preferences in economic models of climate mitigation? 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?

To offer answers to these questions, we generate quantile data on the consumption benefits of emissions reduction policies using the DICE model. DICE essentially couples a Ramsey-Cass-Koopmans growth model to a simple climate model by generating carbon dioxide emissions as a side-effect of production and by connecting climate change back to output and welfare via a so-called damage function. The model

is described fully in Nordhaus (2008,) and so we confine our discussion here to changes that we have made.

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 randomizes eight parameters in the model so that Monte Carlo simulation can be undertaken. Table 1 lists the eight parameters, and the form and parameterization of the pdfs assigned to them.

These eight random parameters, alongside the model's remaining nonrandom parameters and initial conditions (as per Nordhaus 2008), are inputs to a Monte Carlo simulation. In particular, a Latin Hypercube sample of 1,000 runs of the model is taken. Each run solves the model for a particular, exogenous policy, which as described below is a schedule of values for the rate of control of CO₂ emissions. From this is produced a schedule of distributions of consumption per capita (where consumption per capita is equivalent to a cash flow in our theory), which is the focus of the TSD analysis. The policies themselves are in the spirit of a cost-effectiveness approach to meeting prespecified climate targets; they are obtained by choosing at $t = 0$ a trajectory of emissions controls in order to minimize expected discounted CO₂ abatement costs, subject to a constraint on the expected stock of atmospheric CO₂.

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. Their pdfs should all, to varying degrees, be interpreted as subjective. The first four parameters in table 1 play a role in determining CO₂ emissions. In one-sector growth models like DICE, CO₂ emissions are directly proportional to output, which is in turn determined in significant measure by productivity (i)¹⁴ and the stock of labor (ii). However, while CO₂ emissions are proportional to output, the proportion is usually assumed to decrease over time due to autonomous structural and technical change (iii). A further check on industrial CO₂ emissions is provided in the long run by the finite total remaining stock of fossil fuels (iv).

The fifth uncertain parameter is the price of a CO₂-abatement backstop technology. In DICE, the coefficient of the abatement cost function depends on 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. Parameter vi captures uncertainty about the carbon cycle, via the pro-

14. In particular, we randomize the initial growth rate of TFP. This is a scalar quantity that propagates through to TFP growth in future years via the structure of DICE (Nordhaus 2008).

Table 1. Uncertain Parameters for Simulation of DICE

Parameter	Units	Functional Form	Mean	Standard Deviation
(i) Initial growth rate of TFP	Per year	Normal	.0092	.004
(ii) Asymptotic global population	Millions	Normal	8,600	18,92
(iii) Rate of decarbonization	Per year	Normal	-.007	.002
(iv) Total resources of fossil fuels	Billion tons of carbon	Normal	6,000	1,200
(v) Price of back-stop technology	US\$ per ton of carbon replaced	Normal	1,170	468
(vi) Transfer coefficient in carbon cycle	Per decade	Normal	.189	.017
(vii) Climate sensitivity	°C per doubling of atmospheric CO ₂	Log-normal	1.099*	.3912*
(viii) Damage function coefficient α_3	Fraction of global output	Normal	.082	.028

* In natural logarithm space.

portion of CO₂ in the atmosphere in a particular time period, which dissolves into the upper ocean in the next period. Uncertainty about the relationship between a given stock of atmospheric CO₂ 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 CO₂. 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; Roe and Baker 2007), that the pdf 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:

$$\Omega(t) = \frac{1}{1 + \alpha_1 Y(t) + \alpha_2 Y(t)^2 + [\tilde{\alpha}_3 Y(t)]^7}, \quad (3)$$

where Ω is the proportion of output lost, Y is the increase in global mean temperature over the pre-industrial level, and α_i , $i \in \{1, 2, 3\}$ are coefficients. The term $\tilde{\alpha}_3$

is a normally distributed random coefficient (viii), so the higher-order term $[\tilde{\alpha}_3 Y(t)]^7$ 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 6% of output. By contrast, Weitzman (2012) suggests a functional form that 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.

Policies to Be Evaluated

We evaluate a set of five exogenous 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, that is, “business as usual,” or BAU. Our aims in generating this set are to achieve consistency with the modeling framework described just now, as well as a degree of representativeness of the broader policy literature on emissions reduction trajectories (e.g., Clarke et al. 2014).

Each of the five policies limits the atmospheric stock of CO_2 to a pre-specified level. This approach is very similar to many real policy discussions, which aim for a “stabilization” level of atmospheric CO_2 in the very long run. In order to render the policies consistent with the assumptions we make, we use the stochastic version of DICE itself to generate the five policy paths. BAU is the baseline scenario from Nordhaus (2008).

The control variable is the percentage reduction in industrial CO_2 emissions relative to uncontrolled emissions (i.e., not relative to BAU). Each policy path is generated by solving a stochastic optimization problem, whereby the schedule of emissions cuts is chosen to minimize abatement costs¹⁵ subject to the constraint that the mean 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 (see fig. 1). This is done under initial uncertainty about parameters i–vi, since these affect the cost of abatement and its impact on atmospheric CO_2 . As with most of the literature, we assume that the cost-effective path at $t = 0$ is adhered to, despite the resolution of all uncertainty just after $t = 0$, which should be contrasted with more complex

15. Of course, what is cost effective depends on the social objective, so for this part of the analysis we cannot avoid pre-specifying and parameterizing the social welfare and utility functions. For this purpose, we make representative choices, namely, that $\delta(t) = 1.5\%$, $\forall t$, and the coefficient of relative risk aversion is two.

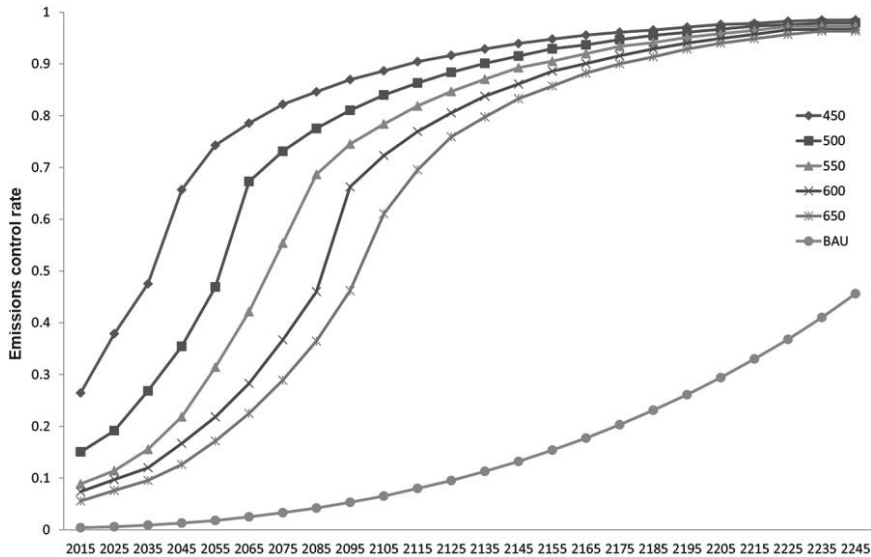


Figure 1. Abatement policies in terms of the emissions control rate. A color version of this figure is available online.

approaches that model learning and consequent revisions to the controls (e.g., Kelly and Kolstad 1999; Lemoine and Traeger 2014).

In an integrated assessment model such as DICE, and especially in running Monte Carlo simulation, solving this cost-minimization problem is a nontrivial computational challenge. We solve it using a genetic algorithm (Riskoptimizer) and with two modifications to the basic optimization problem.¹⁶ In addition, we limit the Latin Hypercube sample to 250 draws just for this task.¹⁷

16. First, we only solve for the emissions control rate from 2015 to 2245 inclusive, rather than all the way out to 2395. This considerably reduces the scope of the optimization problem, in return for making little difference to the results, since, in the standard version of DICE, the optimal emissions control rate is 100% when $t > 2245$, as the backstop abatement technology becomes the lowest cost energy technology. Our first period of emissions control is 2015, since 2005, the first period of the model, is in the past. Second, we guide the optimization by imposing the soft constraint that the emissions control rate is nondecreasing everywhere (via an exponential penalty function when the control rate decreases between any two time periods). We were able to verify that the algorithm's best solution satisfied the property of nondecreasingness in the emissions control rate, and that no solution was found which returned lower costs, where the control rate was decreasing at any point.

17. 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.

5. RESULTS

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-stochastically 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 nontrivial 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 reduction 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 2 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. With the darkest shaded areas indicating a violation of the nonnegativity condition on $H_1^{-1,1}(p, t)$, visual inspection is sufficient to establish that no abatement policy first-order time-stochastically dominates BAU, not even the most accommodating 650 parts per million (ppm) concentration limit.

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. Proposition 3 and its quantile equivalent show that this 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 exact 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 realized under the mitigation policy rather than BAU.

Unable to establish exact TSD of abatement over BAU, we now turn to analyzing Almost TSD. In particular, we look at both almost first-order TSD as set out in definition 2 and proposition 6, and almost first-order time and second-order stochastic dominance as set out in definition 3 and proposition 7. Recall that γ_k denotes the overall volume of violation of exact TSD relative to the total volume enclosed between G_t^j and F_t^j . The term ε_{kT} is the violation of exact TSD in the final time period only, while λ_{1b} is the violation of exact first-order time and second-order stochastic dominance with respect to realization b . As $\gamma_k, \varepsilon_{kT}, \lambda_{1b} \rightarrow 0.5$, the volume/area of violation accounts for half of the entire volume/area between the cumulative distributions being compared, while as $\gamma_k, \varepsilon_{kT}, \lambda_{1b} \rightarrow 0$ there is no violation.

What is striking about the results of analyzing Almost TSD in table 2 is how small the violations are. For all of the policies, in particular it is the violation of exact first-order TSD that is tiny relative to the total volume/area between the distributions. Therefore, we have a formal result showing that everyone would prefer any of

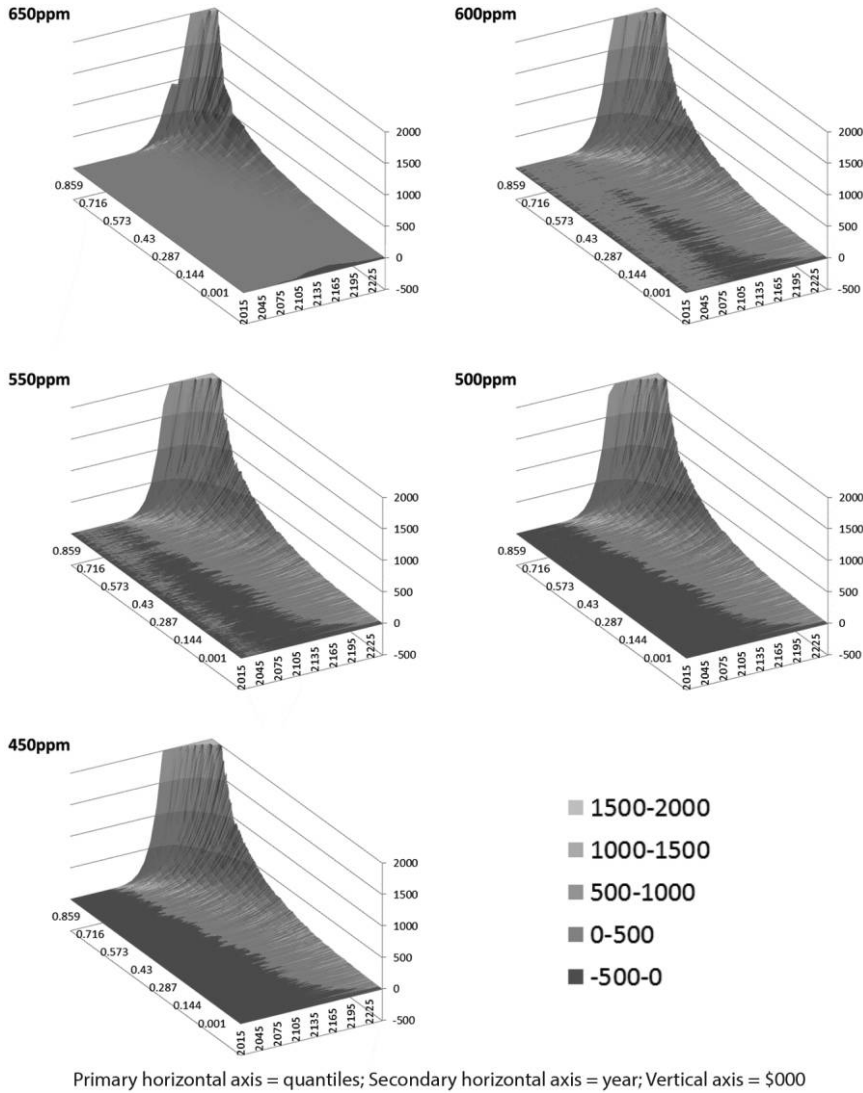
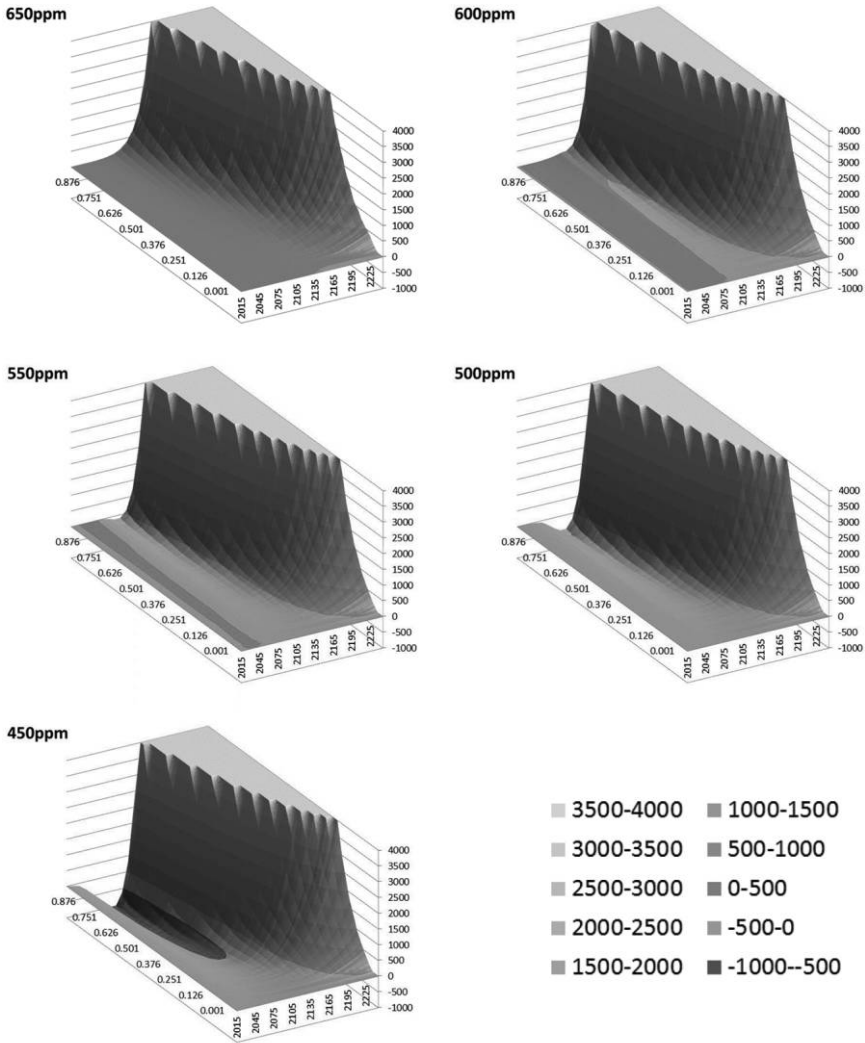


Figure 2. $H_1^{-1,1}(p, t)$ for $\overline{M}^{AT} \in \{450, 500, 550, 600, 650\}$. A color version of this figure is available online.

the abatement policies to BAU, as long as their time and risk preferences can be represented by functions in the sets $(V_1 \times U_1)$ (γ_1) and $U_1(\varepsilon_{1T})$. Moreover, we can say that those who do not prefer the abatement policies have an extreme combination of time and risk preferences. Violation of first-order time and second-order stochastic dominance is also on the whole very small, and note that the condition on



Primary horizontal axis = quantiles; Secondary horizontal axis = year; Vertical axis = \$000

Figure 3. $H_1^{-1,2}(p, t)$ for $\overline{M^{AT}} \in \{450, 500, 550, 600, 650\}$. A color version of this figure is available online.

$D_1^2(b, T)$ in definition 3—equivalently $H_1^{-1,2}(p, T) \geq 0$ —is met by all policies. The overall violation increases with the stringency of the policy.

Let us now use TSD analysis to compare the various abatement policies with each other. We know from the analysis above that exact TSD will not exist either to a first order or to a second order with respect to SD. Therefore we can proceed directly to analyzing violations. In doing so we confine our attention to the least

Table 2. Violations of Exact First-Order TSD and Exact First-Order Time and Second-Order Stochastic Dominance

CO ₂ Limit (ppm)	γ_1	ϵ_{1T}	$\gamma_{1,2}$	ϵ_{2T}	λ_{1b}
650	.00009	.00003	.00002	8E-07	0
600	.00045	.00003	.00045	2E-06	6.01E-08
550	.00092	.00003	.00231	2E-06	.00014
500	.00188	.00004	.00605	3E-06	.00086
450	.00388	.00004	.01363	4E-06	.00245

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 exact first-order TSD. The table should be read such that F_1^1 is the CO₂ 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 exact first-order TSD for $\overline{M}^{AT} = 450$ over $\overline{M}^{AT} = 650$.

Although we might have expected the violations to be relatively 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 CO₂ 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 $(V_1 \times U_1)$ (γ_1), for tough emissions reduction targets, as tough as $\overline{M}^{AT} = 450$.

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 the effect. 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

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	.00255	.00012						
550	.00351	.00011	.01054	.00034				
500	.00517	.00011	.01260	.00032	.01764	.00050		
450	.00859	.00013	.01870	.00036	.02480	.00052	.03701	.00107

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 figure 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. There are two different sources of violation of first-order TSD. The first is that, in early years, the climate is close to its initial, relatively benign state, 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 it is under BAU. The second source of violation is productivity growth, a large source of uncertainty affecting BAU consumption per capita and all that depends on it. In particular, when the realization of the random productivity-growth parameter is at its lowest, consumption per capita is also at its lowest; moreover, in these contingencies carbon emissions are very low. In these circumstances even mild emissions reductions are net costly. This latter effect is therefore isolated in figure 2 where $\overline{M^{AT}} = 650$.

On the other hand, 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 \bar{a}_3 , in (3). It is here that the driving force can be found behind the tiny violations of exact TSD in table 2, namely, the small possibility, in the second half of the modeling 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., Weitzman 2009; Dietz 2011).

6. CONCLUSIONS

In this paper, we have proposed a theory of time-stochastic dominance for ordering risky, intertemporal prospects. Our theory is built by unifying the insights of stochastic dominance (SD) and time dominance (TD). Like these earlier theories, the approach is nonparametric and allows orderings to be constructed only on the basis of partial information about preferences. But our approach generalizes the applica-

tion of simple SD to intertemporal prospects, by permitting pure temporal preferences, just as it generalizes the application of simple TD to risky prospects, by avoiding the need to make strong assumptions about the characteristics of the prospects (prospects may belong to different risk classes and cash flows may be large/nonmarginal).

Like other dominance criteria, a possible practical disadvantage of (exact) TSD is that it may not exist in the data, despite one prospect paying out more than another most of the time, in most states of nature. Various approaches can be taken to deal with this. Our choice has been to extend the notion of Almost SD pioneered by Levy and others to our bidimensional time-risk setup, giving rise to Almost TSD.

The theory can in principle be applied to any investment problem involving multiple time periods and uncertainty about payoffs; however, given the involving nature of the analysis, it might prove most useful in highly contentious public-investment decisions, where there is disagreement about appropriate rates of discount and risk aversion. A leading example might be the mitigation of climate change, so we applied the theory to this policy controversy using a stochastic version of the DICE model, in which eight key model parameters were randomized and Monte Carlo simulation was undertaken.

We were unable to establish exact TSD in the data, even when moving to second-order stochastic dominance (with first-order time dominance). However, when we analyze the related theory of Almost TSD we find that the volume/area of violation of exact TSD is generally very small indeed, so that we can say that almost all decision makers would indeed favor any of our mitigation policies over BAU, and moreover that they would favor 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 parameterized it; our approach is only nonparametric as far as preferences are concerned. Of particular note are the key roles played by uncertainty about climate sensitivity, the curvature of the damage function, and productivity growth. Our parameterization of the former two is key in producing a small violation of exact TSD, because when a high climate sensitivity combines with a high curvature on the damage function, the difference in the relevant cumulative payoff distributions becomes very large. Our parameterization 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. It will be very interesting to see what results are obtained with different integrated assessment models, or with different implementations of DICE.

Our interpretation of γ_k , ε_{kT} , and λ_{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. An interim approach that could be taken, which does rather run counter to the spirit of the TSD approach, is to suppose particular functional forms for $u(x)$ and $v(t)$, most obviously iso-elastic and exponential respectively, and to calculate the set of combinations of parameter values for which dominance holds, based on the violation measures. This is done in Pottier (2015).

APPENDIX A

A Primer on Stochastic Dominance and “Almost” Stochastic Dominance

Stochastic dominance (SD) determines the order of preference of an expected-utility maximizer between risky prospects, while requiring minimal knowledge of her utility function. Take any two risky prospects with probability distributions F and G respectively 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 stochastically 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 realization from the distribution of payoffs possible from a prospect. Moreover, it can be shown that any expected-utility maximizer with a utility function belonging to the set of nondecreasing 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 stochastically 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 maximizer with a utility function belonging to the set of all nondecreasing and (weakly) concave utility functions $U_2 = \{u : u \in U_1 \text{ and } u''(x) \leq 0\}$ would prefer F , that is, any such (weakly) risk-averse decision maker. Hence second-order SD can rank inter alia prospects with the same mean but different variances.

Nonetheless, the practical usefulness of second-order SD is still limited, as the example above illustrated. One could proceed by placing an additional restriction on the decision maker’s preferences, defining the set $U_3 = \{u : u \in U_2 \text{ and } u'''(x) \geq 0\}$ and looking 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). Where $F^3(x) = \int_a^x F^2(s)ds$ and $G^3(x) = \int_a^x G^2(s)ds$, F third-order stochastically 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. However, it can easily be verified in the example 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 in this example, because the first nonzero values of $G^1(x) - F^1(x)$ are negative as x increases from its lower bound, yet $E_F(x) > E_G(x)$. Successive rounds of integration will not make this go away.

A more fruitful route is the theory of Almost SD set out by Leshno and Levy (2002) and recently further developed by Tzeng, Huang, and Shih (2012). Almost SD places restrictions on the derivatives of the utility function with the purpose of excluding the extreme preferences that prevent exact SD from being established. Dominance relations are then characterized 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 :

$$U_1(\varepsilon_1) = \left\{ u \in U_1 : \frac{u'(x)}{\inf[u'(x)]} \leq \left[\frac{1}{\varepsilon_1} - 1 \right], \forall x \right\}$$

and

$$U_2(\varepsilon_2) = \left\{ u \in U_2 : \frac{-u''(x)}{\inf[-u''(x)]} \leq \left[\frac{1}{\varepsilon_2} - 1 \right], \forall x \right\}.$$

The class $U_1(\varepsilon_1)$ is the set of nondecreasing utility functions with the added restriction that the ratio between maximum and minimum marginal utility is bounded by $(1/\varepsilon_1) - 1$. 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 . The expression $U_2(\varepsilon_2)$ is the set of nondecreasing, 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 realizations over which exact first-order SD is violated as

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

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

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

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

Defining the set of realizations over which exact second-order SD is violated as

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

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

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

and

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

Any expected-utility maximizer with a utility function belonging to $U_2(\varepsilon_2)$ would prefer F . From these definitions of first- and second-order Almost SD 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 exact SD of the k th order is violated. The smaller is ε_k , the smaller is the relative violation.

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 cash flows of any two investments X_1 and Y_1 ,¹⁸ 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, that is, the criterion $X_1(t) \geq Y_1(t), \forall t$ is restrictive.¹⁹ Therefore, proceeding again by analogy to SD, X second-order time dominates Y if and only if

$$X_1(T) \geq Y_1(T)$$

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 set of all decreasing, convex consumption discount functions $\widehat{V}_2 = \{\widehat{v} : \widehat{v} \in \widehat{V}_1 \text{ and } \widehat{v}''(t) > 0\}$ would prefer X . This set includes both the exponential and the hyperbolic discounting.

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

19. Indeed, in the domain of deterministic cash flows 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 normalizing the cash flows to the size of the investment (Bøhren and Hansen 1980).

For the case where no discounting is applied the first constraint imposes a necessary condition, stating that the undiscounted value of the intertemporal prospect X must be higher in order to guarantee dominance. 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 n th order (see Ekern 1981).

Notice that TD applies to deterministic cash flows. It would be possible to apply the method to uncertain cash flows, if X and Y were expected cash flows and if a corresponding risk adjustment were made to $\{\hat{v}\}$. However, since any two cash flows X and Y would then be discounted using the same set of risk-adjusted rates, it would be necessary to assume that the cash flows 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 $\{\hat{v}\}$ is cash flows 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 nonmarginal).

APPENDIX B

Proof of Proposition 1

Sufficiency

We want to prove that

$$D_1^1(z, t) \geq 0 \quad \Rightarrow \quad NPV_{v,u}(X) \geq NPV_{v,u}(Y)$$

for all t and z for all $u \in U_1, v \in V_1$.

Assume that $a \leq z \leq b$. This implies that for $z \leq a, D^j(z, t) = 0$ and for $z \geq b, D^j(z, t) = 0$ for all $t \in [0, T]$. Furthermore, we assume that $D_i^j(z, 0) = 0$ for all $z \in [a, b]$. Denote by

$$\begin{aligned} \Delta &= NPV_{v,u}(X) - NPV_{v,u}(Y) = \int_0^T v(t)E_F u(x)dt - \int_0^T v(t)E_G u(y)dt \\ &= \int_0^T v(t) \left[\int_a^b -d(z, t)u(z)dz \right] dt. \end{aligned}$$

Integrating by parts with respect to z , we obtain

$$\Delta = \int_0^T v(t) \left[u(z)(-D^1(z, t)) \Big|_a^b - \int_a^b (-D^1(z, t))u'(z)dz \right] dt.$$

Since $a \leq z \leq b$, the first term in the square brackets is equal to zero. Therefore, we are left with

$$\Delta = \int_0^T \int_a^b v(t)D^1(z, t)u'(z)dzdt. \tag{B1}$$

Integrating by parts with respect to t , we have

$$\begin{aligned} \Delta &= \int_a^b \left[D_1^1(z, t)v(t) \Big|_0^T - \int_0^T D_1^1(z, t)v'(t)dt \right] u'(z)dz \\ &= \int_a^b D_1^1(z, T)v(T)u'(z)dz - \int_0^T \int_a^b D_1^1(z, t)v'(t)u'(z)dzdt, \end{aligned} \tag{B2}$$

as $D_1^1(z, 0) = 0$ for all $z \in [a, b]$. Note that by Fubini's Theorem (B2) can alternatively be obtained by integrating first with respect to t and then with respect to z . From our initial assumption we know that $D_1^1(z, t) \geq 0$. Hence $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$, for all $u \in U_1$ and $v \in V_1$.

Necessity

We have to prove that

$$\begin{aligned} NPV_{v,u}(X) \geq NPV_{v,u}(Y) &\Rightarrow D_1^1(z, t) \geq 0 \\ \text{for all } u \in U_1, v \in V_1 &\qquad \text{for all } t \text{ and } z. \end{aligned}$$

Suppose there is a violation and let (\tilde{z}, \tilde{t}) be the smallest (in the lexicographic sense) pair (z, t) such that $D_1^1(\tilde{z}, \tilde{t}) < 0$. We will show that there is then a utility function $\tilde{u} \in U_1$ and a discount function $\tilde{v} \in V_1$, for which $D_1^1(\tilde{z}, \tilde{t}) < 0$ implies that $NPV_{\tilde{v},\tilde{u}}(X) < NPV_{\tilde{v},\tilde{u}}(Y)$, thus contradicting the original assumption.

Since D_1^1 is continuous, the violation will also exist in the range $\tilde{z} \leq z \leq \tilde{z} + \varepsilon$. Define the following step function:

$$\tilde{u}(z) = \begin{cases} \tilde{z} & z < \tilde{z} \\ z & z \in [\tilde{z}, \tilde{z} + \varepsilon], \\ \tilde{z} + \varepsilon & z > \tilde{z} + \varepsilon \end{cases}$$

which can be approximated arbitrarily closely by a continuously differentiable function in U_1 (see Fishburn and Vickson 1978, 75).

Similarly, the following discount function can be defined:

$$\tilde{v}(t) = \begin{cases} 1 + pe^{-pt} & \text{if } 0 \leq t \leq \tilde{t} \\ 0 + pe^{-pt} & \tilde{t} < t \leq T \end{cases}$$

which again can be approximated arbitrarily closely by a function in V_1 (see Ekern 1981).

Substituting these functions into (B1) we obtain

$$\begin{aligned} \Delta &= \int_{\tilde{z}}^{\tilde{z}+\varepsilon} \left[\int_0^{\tilde{t}} D^1(z, t) dt + p \int_0^T e^{-pt} D^1(z, t) dt \right] dz \\ &= \int_{\tilde{z}}^{\tilde{z}+\varepsilon} \left[D_1^1(z, t) \Big|_0^{\tilde{t}} + p \int_0^T e^{-pt} D^1(z, t) dt \right] dz \\ &= \int_{\tilde{z}}^{\tilde{z}+\varepsilon} \left[D_1^1(z, \tilde{t}) + p \int_0^T e^{-pt} D^1(z, t) dt \right] dz. \end{aligned}$$

In the limit as $p \rightarrow 0$, $p \int_0^T e^{-pt} D^1(z, t) dt = 0$; therefore, for a sufficiently small p , $D_1^1(\tilde{z}, \tilde{t}) < 0$ implies that $NPV_{v,u}(X) < NPV_{v,u}(Y)$, contradicting the initial assumption and showing it is necessary that $D_1^1(\tilde{z}, \tilde{t}) \geq 0$ for all $z \in [a, b]$ and $t \in [0, T]$. QED

Proof of Proposition 2

We need to prove that the following equivalence holds:

$$H_1^{-1}(p, t) = F_1^{-1}(p, t) - G_1^{-1}(p, t) \geq 0, \quad \Leftrightarrow \quad D_1^1(z, t) = G_1^1(z, t) - F_1^1(z, t) \geq 0$$

$$\forall p \in [0, 1] \text{ and } t \in [0, T] \qquad \qquad \qquad \forall z \in [a, b] \text{ and } \forall t \in [0, T]$$

Assume first that $F_1^1(z, t) \leq G_1^1(z, t)$ for all $z \in [a, b]$ and all $t \in [0, T]$. This means that for an arbitrary $x^*(t)$ we have $F_1^1(x^*, t) = p_1^*(t) \leq G_1^1(x^*, t) = p_2^*(t)$. In this way, for given t , x^* will represent both the p_1^{*th} quantile of distribution F and the p_2^{*th} quantile of distribution G .

Since, by assumption, F and G are monotonic increasing functions of z , the quantile functions are monotonic increasing functions of $p \in [0, 1]$. Therefore, knowing that $p_1^*(t) \leq p_2^*(t)$ and due to the monotonicity of the quantile function, $G_1^{-1}(p_1^*, t) \leq G_1^{-1}(p_2^*, t)$. Remembering that $x^*(t) = G_1^{-1}(p_2^*, t) = F_1^{-1}(p_1^*, t)$, it follows that $G_1^{-1}(p_1^*, t) \leq F_1^{-1}(p_1^*, t)$.

We conclude that, for every $t \in [0, T]$, the condition $F_1^1(z, t) \leq G_1^1(z, t)$, $\forall z \in [a, b]$ implies $F_1^{-1}(p, t) \geq G_1^{-1}(p, t) \forall p$. The analogous logic can be applied to show that the reverse condition also holds, that is for a given t , $F_1^{-1}(p, t) \geq G_1^{-1}(p, t)$ will imply $F_1^1(z, t) \leq G_1^1(z, t)$.

Proof of Propositions 3 and 4

Sufficiency

Integrate the expression in (B2) once more with respect to z :

$$\begin{aligned} \Delta &= \int_a^b D_1^1(z, T)v(T)u'(z)dz - \int_0^T \int_a^b D_1^1(z, t)v'(t)u'(z)dzdt, \\ &= v(T) \left[u'(z)D_1^2(z, T) \Big|_a^b - \int_a^b u''(z)D_1^2(z, T)dz \right] - \\ &\quad - \int_0^T v'(t) \left[u'(z)D_1^2(z, t) \Big|_a^b - \int_a^b u''(z)D_1^2(z, t)dz \right] dt. \end{aligned}$$

Now, by the bounding of z , $D_1^2(a, t) = 0$ for all $t \in [0, T]$ and therefore

$$\begin{aligned} \Delta &= v(T)u'(b)D_1^2(b, T) - v(T) \int_a^b u''(z)D_1^2(z, T)dz - \\ &\quad - \int_0^T v'(t)u'(b)D_1^2(b, t)dt + \int_0^T \int_a^b u''(z)v'(t)D_1^2(z, t)dzdt. \end{aligned} \tag{B3}$$

From here we can extract the conditions for dominance with respect to $V_1 \times U_2$ presented in proposition 3. That is, $D_1^2(z, t) \geq 0$ for all $z \in [a, b]$ and all $t \in [0, T]$ is a sufficient condition for $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$ for all $\{v, u\} \in V_1 \times U_2$.

Integrating by parts once more with respect to time, we get the dominance conditions for second-order TSD for all $\{v, u\} \in V_2 \times U_2$:

$$\begin{aligned} \Delta &= u'(b)v(T)D_1^2(b, T) - \int_a^b u''(z)v(T)D_1^2(z, T)dz - u'(b)v'(t)D_2^2(b, t) \Big|_0^T + \\ &\quad + u'(b) \left[\int_0^T v''(t)D_2^2(b, t)dt \right] + \int_a^b u''(z)v'(t)D_2^2(z, t)dz \Big|_0^T - \\ &\quad - \int_a^b u''(z) \left[\int_0^T v''(t)D_2^2(z, t)dt \right] dz, \\ &= u'(b)v(T)D_1^2(b, T) - \int_a^b u''(z)v(T)D_1^2(z, T)dz - u'(b)v'(T)D_2^2(b, T) + \\ &\quad + u'(b) \left[\int_0^T v''(t)D_2^2(b, t)dt \right] + \int_a^b u''(z)v'(T)D_2^2(z, T)dz - \\ &\quad - \int_a^b u''(z) \left[\int_0^T v''(t)D_2^2(z, t)dt \right] dz. \end{aligned} \tag{B4}$$

From here it is easy to see that the following assumptions

- i) $D_1^2(z, T) \geq 0$ for all $z \in [a, b]$
- ii) $D_2^2(z, t) \geq 0$ for all $z \in [a, b]$ and all $t \in [0, T]$

imply that

$$NPV_{E_f, v} \geq NPV_{E_c, v} \text{ for all } (v, u) \in V_2 \times U_2.$$

This completes the sufficiency part of proposition 4. As in the proof of proposition 1, the same dominance conditions can also be obtained when integration is performed first with respect to time t and then with respect to z .

Necessity

We pursue a similar approach to the proof of necessity of proposition 1 for each element of (B3) and (B4) respectively. Consider a differentiable approximation of the increasing and concave, piecewise linear utility function defined by

$$\tilde{u}(z) = \begin{cases} z - \tilde{z} & \text{for } a \leq z < \tilde{z} \\ 0 & \text{for } \tilde{z} \leq z \leq b \end{cases}$$

Replicate the procedure in the proof of necessity of proposition 1, using this utility function and the previous discount function. QED

Proof of Proposition 5

Sufficiency

The proof is constructed as a simple extension of the previous analysis. Integrating by parts repeatedly, we obtain:

$$\begin{aligned} NPV_{v,u}(X) - NPV_{v,u}(Y) &= \sum_{i=0}^{n-1} (-1)^i v^i(T) \left[\sum_{j=1}^{m-1} (-1)^{j+1} u^j(b) D_{i+1}^{j+1}(b, T) \right] + \\ &+ \sum_{j=1}^{m-1} (-1)^{j+1} u^j(b) \left[\int_0^T (-1)^n v^n(t) D_n^{j+1}(b, t) dt \right] + \\ &+ \sum_{i=0}^{n-1} (-1)^i v^i(T) \left[\int_a^b (-1)^{m-1} u^m(z) D_{i+1}^m(z, T) dz \right] + \\ &+ (-1)^{m+n+1} \int_a^b \int_0^T v^n(t) u^m(z) D_n^m(z, t) dt dz. \end{aligned}$$

Remember that by assumption the odd-numbered derivatives of the utility function are positive while the even-numbered derivatives are negative. On the other hand, the odd-numbered derivatives of the discount function are negative, while the even-numbered ones are positive. The condition for dominance with respect to $V_n \times U_m$ can then be derived. Therefore, if

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

then $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$ for all $\{v, u\} \in V_n \times U_m$.

Necessity

If we assume that a violation exists and $D_n^m(z, t) \leq 0$, we can find a suitable utility function and follow a similar process to the proofs of proposition 1 and proposition 3 to show that $NPV_{v,u}(X) < NPV_{v,u}(Y)$, contradicting the initial assumption.

The utility function must satisfy the following conditions:

- i)
$$u^{k-1}(z) \text{ is a piecewise linear function, where } k = 2, \dots, m,$$
- ii)
$$u^{j-1}(b) = 0, j = 2, 3, \dots, k.$$

Furthermore, $u^m(\tilde{z}) \neq 0$ and $u^m(z) \neq 0$ for $z \in [\tilde{z}, \tilde{z} + \varepsilon]$, and the sign of u^m remains constant on $[\tilde{z}, \tilde{z} + \varepsilon]$. Following a similar process as before, one obtains the required contradiction. The same method is employed for each of the terms of the dominance condition. QED

Proof of Proposition 6

Sufficiency

We want to prove that

$$\begin{aligned} X &>_{A1TS} Y \\ \Rightarrow NPV_{v,u}(X) &\geq NPV_{v,u}(Y) \\ \forall v \in (V_1 \times U_1)(\gamma_1) \text{ and } \forall u \in U_1(\varepsilon_{1T}). \end{aligned}$$

Going back to

$$\begin{aligned} \Delta &= \int_a^b D_1^1(z, T)v(T)dz - \int_a^b \int_0^T D_1^1 v'(t)u'(z)dt dz \\ &= v(T) \int_a^b u'(z)D_1^1(z, T)dz + \int_a^b \int_0^T (-)D_1^1 v'(t)u'(z)dt dz \\ &= E + \Gamma. \end{aligned}$$

Separate the range $[a, b]$ at time T between the part $S^{1,T}$, where $D_1^1(z, T) < 0$, and the complementary part $\overline{S^{1,T}}$, where $D_1^1(z, T) \geq 0$:

$$\begin{aligned} E &= v(T) \int_a^b u'(z) [D_1^1(z, T)] dz \\ &= v(T) \int_{S^{1,T}} u'(z)D_1^1(z, T)dz + v(T) \int_{\overline{S^{1,T}}} u'(z)D_1^1(z, T)dz \geq 0. \end{aligned}$$

Note that the integral over the range $S^{1,T}$ is negative and the integral over $\overline{S^{1,T}}$ is positive. In order for $E \geq 0$, the area where $D_1^1(z, T) < 0$ must be ε_{1T} smaller than the total area enclosed between the two distributions. This restriction can be obtained from the proof of almost first-order stochastic dominance by Leshno and Levy (2002), simply by requiring that the utility function belong to the subset $U_1(\varepsilon_{1T})$, where the subscript indicates that the bounds on maximum and minimum marginal utility are established with respect to period T specifically.

Turning to Γ , separate $[a, b]$ for all t into S_t^1 , defined over ranges where $D_1^1(z, t) < 0$, and $\overline{S_t^1}$, the range over which $D_1^1(z, t) \geq 0$, so that we obtain

$$\begin{aligned} \Gamma &= \int \int_{S_t^1} [D_1^1(z, t)](-v'(t)u'(z))dz dt \\ &\quad + \int \int_{\overline{S_t^1}} [D_1^1(z, t)](-v'(t)u'(z))dz dt \geq 0. \end{aligned}$$

The first element of Γ is negative and is minimized when the product of the marginals of the discount and utility functions $[-v'(t) u'(z)]$ is maximized, while the second element is positive and minimized when $[-v'(t) u'(z)]$ is minimized. Hence denoting $\inf_{z \in [a,b] \forall t} \{-v'(t)u'(z)\} = \underline{\theta}$ and $\sup_{z \in [a,b] \forall t} \{-v'(t)u'(z)\} = \overline{\theta}$, the minimum value of Γ is

$$\Gamma^* = \overline{\theta} \int \int_{S_t^1} [D_1^1(z, t)] dz dt + \underline{\theta} \int \int_{\overline{S_t^1}} [D_1^1(z, t)] dz dt \geq 0.$$

It follows that, for a given combination of discount and utility functions, $\Gamma \geq 0$ if $\Gamma^* \geq 0$, which can be rewritten as

$$\sup[-v'(t)u'(z)] \leq \inf[-v'(t)u'(z)] \frac{\int_{J^*} D_1^1(z, t) dz dt}{\int_{S_1} D_1^1(z, t) dz dt}.$$

Let $(v, u) \in (V_1 \times U_1) (\gamma_1)$, then by definition of $(V_1 \times U_1) (\gamma_1)$, we know that

$$[-v'(t)u'(z)] \leq \sup[-v'(t)u'(z)] \leq \inf[-v'(t)u'(z)] \left[\frac{1}{\gamma_1} - 1 \right],$$

which implies $\Gamma^* \geq 0$ and therefore $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$.

Necessity

Begin by assuming the opposite of necessity, that is, that $NPV_{v,u}(X) < NPV_{v,u}(Y)$, for all functions $(v, u) \in (V_1 \times U_1) (\gamma_1)$ and for all $u \in U_1(\varepsilon_{1T})$, implies $X >_{A1TS} Y$. We will prove that this cannot be the case.

Suppose that

i)

$$\int \int_{S_1^t} [-D_1^1(z, t)] dz dt > \gamma_1 \int_0^T \int_a^b |D_1^1(z, t)| dz dt \text{ and}$$

ii)

$$\int_{S_1^T} [-D_1^1(z, T)] dz > \varepsilon_{1T} \int_a^b D_1^1(z, T) dz.$$

Let $\bar{\theta}$ and $\underline{\theta}$ be two positive real numbers such that $\gamma_1 = [\underline{\theta} / (\underline{\theta} + \bar{\theta})]$. Consider the pair of functions $(v, u) \in (V_1 \times U_1) (\gamma_1)$ and where $u \in U_1(\varepsilon_{1T})$, whose product has the following properties:

$$\begin{aligned} v'(t)u(b) &= 0, \\ v(T)u'(z) &= 0, \\ v'(t)u'(z) &= -\underline{\theta} \text{ on } \bar{S}_1^1 \text{ and} \\ v'(t)u'(z) &= -\bar{\theta} \text{ on } S_1^1. \end{aligned}$$

In other words, the product of v and u is a function proportional to

$$v(t)u(z) = zt - bt - zT + bT.$$

It follows then that

$$\begin{aligned} NPV_{v,u}(X) - NPV_{v,u}(Y) &= \bar{\theta} \int \int_{S_1^i} [D_1^1(z, t)] dz dt + \underline{\theta} \int \int_{\bar{S}_1^i} [D_1^1(z, t)] dz dt \\ &= \int \int_{S_1^i} [D_1^1(z, t)] dz dt - \frac{\underline{\theta}}{\underline{\theta} + \bar{\theta}} \int_0^T \int_a^b |D_1^1(z, t)| dz dt \\ &\geq 0, \end{aligned}$$

which contradicts the initial assumption and proves that

$$\begin{aligned} NPV_{v,u}(X) \geq NPV_{v,u}(Y) &\Rightarrow X >_{A1T2S} Y. \\ \forall (v, u) \in (V_1 \times U_1)(\gamma_1) \text{ and } \forall u \in U_1(\varepsilon_{1T}) \end{aligned}$$

QED

Proof of Proposition 7

Sufficiency

We want to prove that

$$\begin{aligned} X &>_{A1T2S} Y \\ &\Rightarrow NPV_{v,u}(X) \geq NPV_{v,u}(Y) \\ \forall (v, u) \in (V_1 \times U_2)(\gamma_{1,2}), \forall u \in U_2(\varepsilon_{2T}) \text{ and } \forall v \in V_1(\lambda_{1b}) \end{aligned}$$

Integrate the previous expression for Δ once more with respect to z , obtaining

$$\begin{aligned} \Delta &= \nu(T) \left[u'(z) D_1^2(z, T) \Big|_a^b - \int_a^b u''(z) D_1^2(z, T) dz \right] + \\ &+ \int_0^T -v'(t) \left[u'(z) D_1^2(z, t) \Big|_a^b - \int_a^b u''(z) D_1^2(z, t) dz \right] dt \geq 0 \\ &= \nu(T) u'(b) D_1^2(b, T) + u'(b) \int_0^T -v'(t) D_1^2(b, t) dt - \\ &- \nu(T) \int_a^b u''(z) D_1^2(z, T) dz + \int_0^T \int_a^b (-v'(t)) (-u''(z)) D_1^2(z, t) dz dt \geq 0 \quad (B5) \\ &= \nu(T) u'(b) D_1^2(b, T) + \Lambda + \bar{E} + \bar{\Gamma} \geq 0. \end{aligned}$$

Hence in the case of almost first-order time and second-order stochastic dominance four elements must be nonnegative. The product $\nu(T) u'(b) D_1^2(b, T)$ must simply

be nonnegative. The remaining three elements must be nonnegative overall, but can be partitioned into a region of violation and a region of nonviolation, with three respective restrictions on the relative violation.

Define the set of realizations where $D_1^2(b, t) < 0$, for any t where $z = b$ as $S_{1,b}$ and its complement as $\overline{S_{1,b}}$, so that

$$\Lambda = u'(b) \int_{S_{1,b}} (-v'(t))D_1^2(b, t)dt + u'(b) \int_{\overline{S_{1,b}}} (-v'(t))D_1^2(b, t)dt.$$

The integral over $S_{1,b}$ is negative while the integral over its complement $\overline{S_{1,b}}$ is positive. Therefore, in an analogous fashion to the proof of proposition 6, in order for $\Lambda \geq 0$ the area where $D_1^2(b, t) < 0$ must be λ_{1b} smaller than the total area enclosed between the two distributions, where the restriction is obtained by requiring that any discount function v belong to $V_1(\lambda_{1b})$.

\overline{E} is similar to E in the previous proof. By restricting the utility function to belong to the subset $U_2(\varepsilon_{2T})$, we obtain the requirement that in period T the area where $D_1^2(z, T) < 0$ cannot be larger than ε_{2T} multiplied by the total area between the two distributions.

Moving to $\overline{\Gamma}$, define an interval of violation and its complement in the usual way:

$$\overline{\Gamma} = \int \int_{S_1^c} (-v'(t))(-u''(z))D_1^2(z, t)dzdt + \int \int_{S_1} (-v'(t))(-u''(z))D_1^2(z, t)dzdt.$$

Again, following the proof of proposition 6, define $\inf_{z \in [a,b] \forall t} \{v'(t)u''(z)\} = \underline{g}$ and $\sup_{z \in [a,b] \forall t} \{v'(t)u''(z)\} = \overline{g}$, so that the minimum Ω is

$$\overline{\Gamma}^* = \underline{g} \int \int_{S_1^c} D_1^2(z, t)dzdt + \overline{g} \int \int_{S_1} D_1^2(z, t)dzdt.$$

Both elements of $\overline{\Gamma}$ are relatively larger than the corresponding elements of $\overline{\Gamma}^*$.

We are looking for a set of preferences $(V_1 \times U_2) (\gamma_{1,2})$ for which $\overline{\Gamma}^* \geq 0$, which are

$$\begin{aligned} \sup[v'(t)u''(z)] &\leq \inf[v'(t)u''(z)] \frac{\int \int_{\overline{S}} [D_1^2(z, t)] dzdt}{\int \int_{\overline{S}} [F_1^2(z, t) - G_1^2(z, t)] dzdt} \\ \sup[v'(t)u''(z)] &\leq \inf[v'(t)u''(z)] \frac{\int \int_{\overline{S}} [G_1^2(z, t) - F_1^2(z, t)] dzdt}{\int \int_{S_1} [F_1^2(z, t) - G_1^2(z, t)] dzdt} \end{aligned}$$

By letting $(v, u) \in (V_1 \times U_2) (\gamma_{1,2})$, then, by definition of $(V_1 \times U_2) (\gamma_{1,2})$, we know that

$$[v'(t)u''(z)] \leq \sup[v'(t)u''(z)] \leq \inf[v'(t)u''(z)] \left[\frac{1}{\gamma_{1,2}} - 1 \right],$$

which implies that $\Omega^* \geq 0$ holds and therefore, $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$.

Necessity

Starting from equation (B5) and using the increasing and concave utility function defined in proving necessity in proposition 3, the proof proceeds in just the same fashion as for proposition 6 and is therefore omitted. QED

REFERENCES

- Arrow, Kenneth J., William R. Cline, Karl-Göran Mäler, Mohan Munasinghe, R. Squitieri, and Joseph E. Stiglitz. 1996. Intertemporal equity, discounting, and economic efficiency. In *Climate change 1995: Economic and social dimensions of climate change; Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, ed. James P. Bruce, Hoesung Lee, and Erik F. Haites. Cambridge: Cambridge University Press, 127–44.
- Atkinson, Anthony B., and Francois Bourguignon. 1982. The comparison of multi-dimensional distributions of economic status. *Review of Economic Studies* 49 (2): 183–201.
- Bøhren, Øyvind, and Terje Hansen. 1980. Capital budgeting with unspecified discount rates. *Scandinavian Journal of Economics* 82 (1): 45–58.
- Clarke, Leon, Kejun Jiang, Keigo Akimoto, Babiker Mustafa, Geoffrey Blanford, Karen Fisher-Vanden, Jean-Charles Hourcade, Volker Krey, Elmar Kriegler, Andreas Löschel, David McCollum, Sergey Paltsev, Steven Rose, Priyadarshi R. Shukla, Massimo Tavoni, Bob van der Zwaan, and Detlef van Vuuren. 2014. Assessing transformation pathways. In *Climate change 2014: Mitigation of climate change; Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. Ottmar Edenhofer, Ramon Pichs-Madruga, Youba Sokona, Ellie Farahani, Susanne Kadner, Kristin Seyboth, Anna Adler, Ina Baum, Steffen Brunner, Patrick Eickemeier, Benjamin Kriemann, Jussi Savolainen, Steffen Schlömer, Christoph von Stechow, Timm Zwicker, and Jan C. Minx. Cambridge: Cambridge University Press.
- Cline, William R. 1992. *The economics of global warming*. Washington, DC: Peterson Institute for International Economics.
- Dasgupta, Partha. 2007. The Stern Review's economics of climate change. *National Institute Economic Review* 199 (1): 4–7.
- Dietz, Simon. 2011. High impact, low probability? An empirical analysis of risk in the economics of climate change. *Climatic Change* 103 (3): 519–41.
- Dietz, Simon, and Geir B. Asheim. 2012. Climate policy under sustainable discounted utilitarianism. *Journal of Environmental Economics and Management* 63:321–35.
- Dietz, Simon, and Cameron J. Hepburn. 2013. Benefit-cost analysis of non-marginal climate and energy projects. *Energy Economics* 40:61–71.

- Ekern, Steinar. 1981. Time dominance efficiency analysis. *Journal of Finance* 36 (5): 1023–33.
- Fishburn, Peter C. 1964. *Decision and value theory*. New York: Wiley.
- Fishburn, Peter C., and Raymond G. Vickson. 1978. Theoretical foundations of stochastic dominance. In *Stochastic dominance: An approach to decision-making under risk*, ed. George A. Whitmore and M. Chapman Findlay. Lexington, MA: Heath, 39–113.
- Gollier, Christian. 2006. An evaluation of Stern's report on the economics of climate change. IDEI Working paper 464, Institut d'Économie Industrielle (IDEI), Toulouse.
- Gollier, Christian, and Richard Zeckhauser. 2005. Aggregation of heterogeneous time preferences. *Journal of Political Economy* 113 (4): 878–96.
- Hadar, Josef, and William R. Russell. 1969. Rules for ordering uncertain prospects. *American Economic Review* 59 (1): 25–34.
- Hanoch, Giora, and Haim Levy. 1969. The efficiency analysis of choices involving risk. *Review of Economic Studies* 36 (3): 335–46.
- Hepburn, Cameron J., and Wilfred Beckerman. 2007. Ethics of the discount rate in the Stern Review on the economics of climate change. *World Economics* 8 (1): 187–210.
- Ikefuji, Masako, Roger J. A. Laeven, Jan R. Magnus, and Chris Muris. 2013. Pareto utility. *Theory and Decision* 75 (1): 43–57.
- Karcher, Thierry, Patrick Moyes, and Alain Trannoy. 1995. The stochastic dominance ordering of income distributions over time: The discounted sum of the expected utilities of incomes. In *Social choice, welfare, and ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics*, ed. William A. Barnett, Herve Moulin, Maurice Salles, and Norman J. Schofield. Cambridge: Cambridge University Press.
- Kelly, David L., and Charles D. Kolstad. 1999. Bayesian learning, growth, and pollution. *Journal of Economic Dynamics and Control* 23 (4): 491–518.
- Kimball, Miles S. 1990. Precautionary saving in the small and in the large. *Econometrica* 58 (1): 53–73.
- Laibson, David. 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112 (2): 443–78.
- Lemoine, Derek, and Christian Traeger. 2014. Watch your step: Optimal policy in a tipping climate. *American Economic Journal: Economic Policy* 6 (1): 137–66.
- Leshno, Moshe, and Haim Levy. 2002. Preferred by all and preferred by most decision makers: Almost stochastic dominance. *Management Science* 48 (8): 1074–85.
- Levy, Haim. 1973. Stochastic dominance, efficiency criteria, and efficient portfolios: The multi-period case. *American Economic Review* 63 (5): 986–94.
- Levy, Haim, and Giora Hanoch. 1970. Relative effectiveness of efficiency criteria for portfolio selection. *Journal of Financial and Quantitative Analysis* 5 (1): 63–76.
- Levy, Haim, and Yoram Kroll. 1979. Efficiency analysis with borrowing and lending: Criteria and their effectiveness. *Review of Economics and Statistics* 61 (1): 125–30.
- Levy, Haim, Moshe Leshno, and Boaz Leibovitch. 2010. Economically relevant preferences for all observed epsilon. *Annals of Operations Research* 176 (1): 153–78.
- Levy, Haim, and Jacob Paroush. 1974a. Multi-period stochastic dominance. *Management Science* 21 (4): 428–35.
- . 1974b. Toward multivariate efficiency criteria. *Journal of Economic Theory* 7 (2): 129–42.
- Levy, Moshe. 2009. Almost stochastic dominance and stocks for the long run. *European Journal of Operational Research* 194 (1): 250–57.

- Matei, Anca N., and Claudio Zoli. 2012. On finite time intertemporal dominance. Unpublished manuscript, University of Verona.
- Meehl, Gerard A., Thomas F. Stocker, William D. Collins, Pierre Friedlingstein, Arnadou Thierno Gaye, Jonathan M. Gregory, Akio Kitoh, Reto Knutti, James M. Murphy, Akira Noda, Sarah C. B. Raper, Ian G. Watterson, Andrew J. Weaver, and Zong-Ci Zhao. 2007. Global climate projections. In *Climate change 2007: The physical science basis; Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. Susan Solomon, Dahe Qin, Martin Manning, Zhenlin Chen, Melinda Marquis, Kristen B. Averyt, Henry LeRoy Tignor, and Melinda M. B. Miller. Cambridge: Cambridge University Press.
- Nordhaus, William D. 1991. To slow or not to slow: The economics of the greenhouse effect. *Economic Journal* 101 (407): 920–37.
- . 1993a. Optimal greenhouse-gas reductions and tax policy in the “DICE” model. *American Economic Review* 83 (2): 313–17.
- . 1993b. Rolling the “DICE”: An optimal transition path for controlling greenhouse gases. *Resource and Energy Economics* 15 (1): 27–50.
- . 1994. *Managing the global commons: The economics of climate change*. Cambridge, MA: MIT Press.
- . 2007. A review of the “Stern Review on the Economics of Climate Change.” *Journal of Economic Literature* 45:686–702.
- . 2008. *A question of balance: Weighing the options on global warming policies*. New Haven, CT: Yale University Press.
- Pindyck, Robert S. 2011. Fat tails, thin tails, and climate change policy. *Review of Environmental Economics and Policy* 5 (2): 258–74.
- Pottier, Antonin. 2015. Time-stochastic dominance and iso-elastic discounted utility functions. Unpublished manuscript, CERNA, MINES ParisTech.
- Roe, Gerard, and Marcia Baker. 2007. Why is climate sensitivity so unpredictable? *Science* 318 (5850): 629–32.
- Rothschild, Michael, and Joseph E. Stiglitz. 1970. Increasing risk: I. A definition. *Journal of Economic Theory* 2 (3): 225–43.
- Scarsini, Marco. 1986. Comparison of random cash flows. *IMA Journal of Management Mathematics* 1 (1): 25–32.
- Stern, Nicholas. 2007. *The economics of climate change: The Stern review*. Cambridge: Cambridge University Press.
- . 2008. The economics of climate change. *American Economic Review: Papers and Proceedings* 98 (2): 1–37.
- Tol, Richard S. J. 2012. On the uncertainty about the total economic impact of climate change. *Environmental and Resource Economics* 53 (1): 97–116.
- Tzeng, Larry Y., Rachel J. Huang, and Pai-Ta Shih. 2012. Revisiting almost second-degree stochastic dominance. *Management Science* 59 (5): 1250–54.
- Weitzman, Martin L. 2007. A review of the Stern Review on the economics of climate change. *Journal of Economic Literature* 45 (3): 703–24.
- . 2009. On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics* 91 (1): 1–19.
- . 2012. GHG targets as insurance against catastrophic climate damages. *Journal of Public Economic Theory* 14 (2): 221–44.