Centre for Climate Change Economics and Policy

Grantham Research Institute on Climate Change and the Environment

Spaces for agreement: a theory of Time-Stochastic Dominance

Simon Dietz and Anca N. Matei

October 2013

Centre for Climate Change Economics and Policy Working Paper No. 153

Grantham Research Institute on Climate Change and the Environment

Working Paper No. 137

The Grantham Foundation for the Protection of the Environment

Global Green Growth Institute

ESRC Economic & Social Research Council

Munich RE

University of Leeds

LSE The London School of Economics and Political Science

Spaces for agreement: a theory of Time-Stochastic Dominance

Simon Dietz* and Anca N. Matei†

August 20, 2013

Abstract

Many investments involve both a long time-horizon and risky returns. Making investment decisions thus requires assumptions about time and risk preferences. In the public sector in particular, such assumptions are frequently contested 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 decision-makers whose preferences can be represented by such combinations would agree on the option to be chosen. The theory is built on combining the insights of stochastic dominance on the one hand, and time dominance on the other, thus offering a non-parametric approach to inter-temporal, risky choice.

Keywords: almost stochastic dominance, discounting, project appraisal, risk aversion, stochastic dominance, time dominance, time-stochastic dominance

JEL codes: D61, H43

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

†Department of Economics, University of Verona.

We would like to thank Claudio Zoli for valuable discussions. 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". SD would also like to acknowledge the financial support of the Grantham Foundation for the Protection of the Environment, as well as the Centre for Climate Change Economics and Policy, which is funded by the UK's Economic and Social Research Council and by Munich Re. The usual disclaimer applies.

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

1 Introduction

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 challenges of estimating individual preferences, there are positions to be taken on how to aggregate individual preferences in order to construct social preferences. Some of these positions are positive in nature, some are normative.

Perhaps the best recent illustration of how assumptions about time and risk preferences are contested in project appraisal is the debate about the findings of the British Government's *Stern Review on the Economics of Climate Change* (Stern, 2007). A prominent part of this review was a cost-benefit analysis of targets for global greenhouse gas emissions. Once emitted to the atmosphere, carbon dioxide, the principal greenhouse gas, resides there for centuries. Moreover the dynamics of the climate system are such that there is a lag of many decades between abating carbon dioxide emissions and the peak pay-off from doing so. Together these features make deciding on whether to cut emissions today one of the ultimate examples of an investment with a long pay-back. At the same time, the impacts of emissions reductions are highly uncertain (ranging from ineffectual to essential for the survival of humanity – e.g. Weitzman 2009), so it is also a risky investment *par excellence*.

Assumptions about time and risk preferences were therefore going to be important, and Stern's were distinctive – within a standard, discounted utilitarian framework, the rate of pure time preference was a very low 0.1%, while he opted for a logarithmic utility function (together resulting in an unusually low social discount rate). Consequently he recommended immediate and deep cuts in global emissions, but his approach was quickly the subject of intense debate, with a number of prominent scholars arguing for different formulations of discounting and utility, in particular greater impatience and/or a greater elasticity of marginal utility of consumption.[1] Indeed, the more frequent conclusion of economic evaluation of climate-change policy has been slow and rather shallow emissions reductions (e.g. Nordhaus and Boyer 2000; Nordhaus 2008).

However, the Stern Review is merely one of the latest and most prominent manifestations of disagreement about risk and time in project appraisal (landmarks include Lind et al. 1982, and Portney and Weyant 1999). Other examples of public projects that have both long time-horizons and risky returns are radioactive waste disposal and the protection of wilderness and biological diversity.

Our starting point for this paper is that such a debate legitimately exists

[1]So many opinions have been voiced on the (de-)merits of the Stern Review that summarising them comprehensively would be a research project in itself. Hepburn and Beckerman (2007), Nordhaus (2007) and Weitzman (2007) are notable critiques of Stern's approach to pure time preference, while Dasgupta (2007), Gollier (2006) and Weitzman (2007), among others, took issue with his approach to risk aversion. Dietz et al. (2007b,a) and Dietz and Stern (2008) offered a response.

and will continue for the foreseeable future. While there is potentially a long and difficult philosophical discussion to embark upon here, we simply observe that the ingredients for the debate include normative differences and positive uncertainties, neither of which seem easy to resolve. The normative differences at hand are often rationalised in terms of the opposition between believers in a 'descriptive' approach to parameterising utility and social welfare, which relies on appropriate data from markets or other samples of representative consumer/individual behaviour, and adherents to a 'prescriptive' approach, where choosing functional form and setting parameter values is an exercise in philosophical introspection on the part of the researcher. The dichotomy is due to Arrow et al. (1996) and, since it was suggested, many justifications of both approach have been offered. The debate endures. Positive uncertainties result from the wealth of relevant but often conflicting data to inform parameterisation of utility and social welfare, including market transactions, responses to questionnaire surveys and behaviour in laboratory experiments, at different times, in different places and with respect to different goods. It may be rather easier to envisage – in principle – how these positive uncertainties could be resolved by the collection of more data, but in practice they are also likely to be long-lasting.

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 specification of time and risk preferences, our approach does not require decision-makers to make *a priori* choices of functional form or parameter values. Rather we attempt to rank alternatives based on incomplete or partial information. While this non-parametric approach could be used to inform investment choice in either the public or private sectors, our hope is that one of its main uses might be to bring renewed clarity to certain critical and hotly contested choices in public policy, such as mitigation of climate change. In these areas, it has arguably become forgotten in the debate about the formalisation of time and risk preferences, which can appear intractable, that in fact choices might be able to be made, without unanimity on parameterisation – i.e. the structure of the investment problem could be such that commentators of many shades can unite on the desirability of one course of action over another, without having that much in common. Even if this does not turn out to be true, we learn something in the process of testing for it.

The intellectual antecedents of this paper lie in the theory of Stochastic Dominance (Fishburn, 1964; Hanoch and Levy, 1969; Hadar and Russell, 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 et al., 1995).

Stochastic Dominance (SD) is an integral part 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 non-parametric approach to risky choice, whereby one tests for SD relations for whole preference classes. Yet SD sometimes faces practical limitations, nicely illustrated by a stylised 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.[2] 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.[3] For this reason Leshno and Levy (2002) derived Almost Stochastic Dominance (ASD), 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 it is very small (to be defined precisely later), an ASD relation can be argued to exist.

The basic theory of SD is a-temporal. In effect, decisions are made and pay-offs 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 have temporal preferences, that is to prefer flows of utility in some periods of time more than in others.[4] 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 cashflows, i.e. instead of working with cumulative distributions over the consequence space of a decision, one works with cumulative distributions over time. Proponents of the approach make arguments in its favour that are analogous to those made for SD – one tests for a Time Dominance (TD) relation for whole preference classes, rather than having to pre-specify and parameterise 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, cashflows, and can only be extended to the latter under restrictive assumptions. This would be done by analysing TD between expected cashflows, having made a risk adjustment to the set of discount functions under consideration. However, since all cashflows would then be discounted using the same set

[2] Where F^1 and G^1 are respectively the cumulative distributions of the former and latter prospects over realisations x, this is because the first nonzero values of $G^1(x) - F^1(x)$ are negative as x increases from its lower bound, yet $E_F(x) > E_G(x)$. n^{th}-order SD requires that $G^n(x) - F^n(x) \geq 0, \forall x$, $E_F(x) \geq E_G(x)$ and there is at least one strict inequality.

[3] In the example used, one would be $u(x) = \begin{cases} x \text{ for } x \leq 1 \\ 1 \text{ for } x > 1 \end{cases}$.

[4] 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.

of risk-adjusted rates, it would be necessary to assume that all cashflows belong to the same risk class, for example under the capital asset pricing model they would have to share the same covariance with the market portfolio. It would also be necessary to assume that any investments being compared are small (i.e. marginal), since TD assumes common *consumption* discount rates, which depend on a common growth rate.

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 way to choose between risky, inter-temporal prospects, which admits the possibility of pure-time discounting and makes weak assumptions about the risk characteristics of the prospects. As a problem in two dimensions (risk and time), our theory is related to other problems of bivariate dominance, such as bi-dimensional inequality in Atkinson and Bourguignon (1982) (income and life expectancy) and most notably changes in the distribution of income over time in Karcher et al. (1995).

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. Combinations of these classes constitute possible spaces for agreement. In Section 3 we establish the theory of (standard) Time-Stochastic Dominance, while in Section 4 we do the same for Almost Time-Stochastic Dominance. We offer several worked examples in Section 5, noting that we have applied the theory to a much more complex example of climate-change mitigation in a companion paper (Dietz and Matei, 2013). These examples are stylised, for expositional purposes, but they are intended to be a realistic reflection of practice in one important respect – they deal with discrete data. Hence applying our dominance criteria to the example data requires us also to adapt our theorems to distribution quantiles. Section 6 concludes.

2 Spaces for agreement

Let us take the task at hand as being to rank two prospects X and Y, both of which yield random cashflows over time. The underlying purpose is to compare the expected discounted utilities of the prospects at $t = 0$, i.e. for prospect X we compute

$$NPV_{v,u}(X) = \int_0^T v(t) E_F u(x,t) dt = \int_0^T v(t) \int_a^b u(x) f(x,t) dx dt,$$

where x is a realisation of the cashflow 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 cashflows of X and Y are both supported on the finite interval $[a, b]$, $-\infty < a < b < +\infty$ and that each prospect pays out over a finite, continuous time-horizon $[0, T]$. Therefore we can characterise a probability density function for prospect X at time $t \in [0,T]$, $f(x,t)$, and a counterpart cumulative distribution function with respect to realisation $x \in [a,b]$, $F^1(x,t) = \int_a^x f(s,t)ds$.

Before characterising Time-Stochastic Dominance (TSD), we need to define classes of utility and discount functions. Starting with utility functions $u : [a,b] \to \mathbb{R}$, we will focus on two specific classes:

$$\begin{aligned} U_1 &= \{u : u'(x) \geq 0\}\,, \\ U_2 &= \{u : u \in U_1 \text{ and } u''(x) \leq 0\}\,. \end{aligned}$$

As usual then, U_1 is the class of utility functions, whereby utility is non-decreasing as a function of consumption, representing nothing more than (weak) non-satiation. It is hard to imagine relevant circumstances in which the appropriate utility function would not be in U_1. U_2 is the class of non-decreasing, weakly concave utility functions, which rules 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 individual behaviour, for instance. In the literature on SD, it is common to proceed further to a third class U_3 in which $u \in U_2$ and $u'''(x) \geq 0$, which is a necessary (but insufficient) condition for a particular kind of risk aversion, decreasing absolute risk aversion. However, we will not work explicitly with U_3 in this paper, since we would have to contend with too many combinations of utility and discount functions. Nevertheless the theory is perfectly capable of handling it, and we will eventually 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, $v : [0,T] \to \mathbb{R}$. 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 typically of little interest, since some positive degree of time preference is always required (even by Stern, 2007, on the grounds that there is at least a positive risk of human extinction). Therefore, without compromising the generality of our theory, let us focus our attention on the first- and second-order restrictions on V_0:

$$\begin{aligned} V_1 &= \{v : v \in V_0, \text{ and } v'(t) < 0\} \\ V_2 &= \{v : v \in V_1, \text{ and } v''(t) > 0\}\,. \end{aligned}$$

V_1 is the class of 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. Note that V_1 admits 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 behaviour (Laibson,

1997) and that it can result from the aggregation of heterogeneous individual preferences.[5]

Combinations of utility and discount functions constitute possible *spaces for agreement*. $V_1 \times U_1$ 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 non-decreasing utility function, in other words any impatient decision-maker with any attitude to risk from seeking to averse. Presumably virtually all decision-makers belong to this combination of classes. By contrast $V_2 \times U_2$, for instance, encapsulates decision-makers whose impatience decreases over time and who are risk averse or neutral. 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 4 we narrow the space for agreement further by placing additional restrictions on $\{u\}$ and $\{v\}$ with a view to excluding 'extreme' preferences.

3 Time-Stochastic Dominance

A further piece of notational apparatus will enable us to work in a compact, bi-dimensional 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:

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

where $i \in \{1, 2, \ldots, n\}$ is the order of TD (i.e. the number of integrations with respect to time) and $j \in \{1, 2, \ldots, m\}$ is the order of SD (i.e. the number of integrations with respect to the 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 characterise TSD for various combinations of classes of U_j and V_i.

[5] Even though those individual preferences are represented by exponential discounting (see Gollier and Zeckhauser, 2005).

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

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

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

In this definition, the ordering $>_{iTjS}$ denotes pure TD of the i^{th} order, combined with SD of the j^{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 the Appendix. □

Proposition 1 tells us that any impatient planner with monotonic non-decreasing 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 wealth 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.[6]

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 behaviour *a priori*, especially in the public sector.

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

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

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

Proof. See the Appendix. □

It is evident from Proposition 2 and its proof that, in line with the classical approach to SD, restricting the utility function by one degree corresponds to integrating the bi-dimensional probability distribution $D_1^1(z, t)$ once more with respect to the consequence space.

[6] Proposition 1 is similar to Theorem 3 in Scarsini (1986). However Scarsini did not consider any other cases, i.e. any other combinations of time and risk preference.

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 impatience decreasing over time, then integrate $D_1^2(z,t)$ once more with respect to time.

Proposition 3 (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 the Appendix. □

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 pure time and second-order stochastic dominance must hold with respect to the difference between the distributions in the terminal period T.

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

Proposition 4 ($\mathbf{n}^{th}$-order Time and $\mathbf{m}^{th}$-order Stochastic Dominance). *X n^{th}-order time and m^{th}-order stochastic dominates Y if and only if*

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

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

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

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

4 Almost Time-Stochastic Dominance

In practice, the usefulness of (standard) 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 in the Introduction 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, non-parametric 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. Leshno and Levy (2002) recognised this problem in the context of SD and developed a theory of Almost Stochastic Dominance (ASD), according to which restrictions are placed on the derivatives of the utility function, so that extreme preferences are excluded.[7] Dominance relations between risky prospects are then characterised for 'almost' all decision-makers.

What is 'extreme' is clearly subjective, an obvious difficulty faced by the ASD approach. However, Levy et al. (2010) offer an illustration of how to define it using laboratory data on participant choices when faced with binary lotteries. Extreme risk preferences are marked out by establishing gambles that all participants are prepared to take. By making the conservative assumption that no participant has extreme risk preferences, the least and most risk-averse participants mark out the limits, and preferences outside these limits can be considered extreme.

It is obvious that standard TSD faces the same practical constraints as standard SD. In this section we therefore extend our theory to 'Almost TSD', characterising Almost First-order TSD and Almost First-order Time and Second-order Stochastic Dominance.

Let us start with the former. Define the set of realisations $z \in [a,b]$ for all t where there is a violation of First-order TSD as S_1^1:

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

Similarly define the set of realisations $z \in [a,b]$ at time T where there is a violation as

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

Definition 2 (**Almost First-order Time-Stochastic Dominance).** *X dominates Y by Almost First-order Time-Stochastic Dominance, denoted $X \succ_{1ATS} Y$, if and only if*

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

[7]Tzeng et al. (2012) showed that Leshno and Levy's theorem for Almost Second-order Stochastic Dominance is incorrect and re-define the concept. They also extend the results to higher orders.

Proposition 5 (A1TSD). *$X >_{A1TS} Y$ if, for all $(v,u) \in V_1(\gamma_1) \times U_1(\gamma_1)$ and $u \in U_1(\varepsilon_{1T})$,*
$NPV_{v,u}(X) \geq NPV_{v,u}(Y)$.

Proof. See the Appendix. □

The definition of Almost First-order TSD contains two measures of the violation of strict First-order TSD. γ_1 measures the cumulative violation of the non-negativity 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. The latter violation measure has the same interpretation in terms of utility as the corresponding violation measure in Leshno and Levy (2002). Adapting their theorem to our context, for every $0 < \varepsilon_{1T} < 0.5$, define the following subset of U_1:

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

$U_1(\varepsilon_{1T})$ is the set of non-decreasing utility functions with the added restriction that the ratio between maximum and minimum marginal utility is bounded by $\frac{1}{\varepsilon_{1T}} - 1$, i.e. 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 give to low wealth levels at the same time as restricting how little (much) marginal utility they give to high wealth levels. Further narrowing the scope to the very common case of utility functions with constant elasticity of marginal utility, the restriction is on the absolute value of the elasticity – $|\frac{u''(z)z}{u'(z)}|$ – such that it cannot be large negative or large positive, and the larger is ε_{1T} the smaller $|\frac{u''(z)z}{u'(z)}|$ must be. 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. Note that the bounds on $u'(z)$ are established with respect to the set of realisations when $t = T$.

γ_1 is defined in terms of the product of the marginals of the discount and utility functions as follows:

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

$V_1(\gamma_1) \times U_1(\gamma_1)$ is the set of all combinations of decreasing pure time discount function and non-decreasing utility function, with the added restriction that the ratio between the maximum and minimum products of $[-v'(t)u'(z)]$ is bounded by $\frac{1}{\gamma_1} - 1$. 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. Bounding the ratio between maximum and minimum $v'(t)$ amounts to excluding preferences exhibiting a very large change in impatience over time. Therefore the combinations of preferences that we are excluding here comprise extreme concavity or convexity of the utility and discount functions somewhere on their respective domains. Note that the bounds on $[-v'(t)u'(z)]$ are established with respect to all realisations and all time-periods.

Moving now to Almost First-order Time and Second-order Stochastic Dominance, parcel out for all t the subset of realisations S_1^2 where $D_1^2 < 0$, i.e. where the condition for strict First-order Time and Second-Order Stochastic Dominance is violated:

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

Similarly define the set of realisations $z \in [a,b]$ at time T where there is a violation as

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

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

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

Definition 3 (**Almost First-order Time and Second-order Stochastic Dominance**). *X Almost First-order Time and Second-order Stochastic Dominates Y, denoted* $X >_{A1T2S} Y$ *if and only if*

$$\begin{gathered} i)\ \int_0^T \int_{S_1^2} -D_1^2(z,t)dzdt \leq \gamma_2 \int_0^T \int_a^b \left|D_1^2(z,t)\right| dzdt, \\ ii)\ \int_{S^{2,T}} -D_1^2(z,T)dz \leq \varepsilon_{2T} \int_a^b \left|D_1^2(z,T)\right| dz, \\ iii)\ \int_{S_{1,b}} D_1^2(b,t)dt \leq \gamma_{1b} \int_0^T \left|D_1^2(b,t)\right| dt,\ \text{and} \\ iv) D_1^2(b,T) \geq 0 \end{gathered}$$

Proposition 6 (**A1T2SD**). $X >_{A1T2S} Y$ *if, for all* $(v,u) \in V_1(\gamma_2) \times U_2(\gamma_2)$, $u \in U_1(\varepsilon_{1T})$ *and* $(v,u) \in V_1(\gamma_{1b}) \times U_1(\gamma_{1b})$,
$NPV_{v,u}(X) \geq NPV_{v,u}(Y)$.

Proof. See the Appendix. □

The definition of Almost First-order Time and Second-order Stochastic Dominance contains three measures of the violation of strict dominance, as well as the requirement that $D_1^2(b,T) \geq 0$. First, γ_2 measures the relative violation of the non-negativity condition on D_1^2 over all t. It is equivalent to the following restriction on combined time and risk preferences:

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

The set $V_1(\gamma_2) \times U_2(\gamma_2)$ represents all combinations of decreasing pure time discount functions and non-decreasing, weakly concave utility functions, with the added restriction that the ratio between the maximum and minimum of $[v'(t)u''(z)]$ is bounded by $\frac{1}{\gamma_2} - 1$. 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 realisations and time-periods.

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

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

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

Third, we need to measure a violation of the non-negativity condition on the integral with respect to time of $D_1^2(b,t)$. We denote this γ_{1b}, because the restriction is on the product $[-v'(t)u'(b)]$ (see proof), therefore it has the same interpretation as γ_1, except that in this case the bounds on u' are with respect to realisation b specifically.

Propositions 5 and 6 characterise sufficient conditions for Almost TSD, rather than necessary and sufficient conditions. To see why this is so, let us dip into the proofs. First express the difference in NPV Δ between prospects X and Y in terms of the difference in their respective *cdfs*:

$$\begin{aligned} \Delta &= NPV_{v,u}(X) - NPV_{v,u}(Y) \\ &= \textstyle\int_0^T v(t) \int_a^b D^1(z,t) u'(z) dz dt \geq 0. \end{aligned}$$

Integrating with respect to time we obtain an expression in terms of D_1^1, i.e. in terms of First-order TSD:

$$\begin{aligned} \Delta &= \textstyle\int_a^b D_1^1(z,T) v(T) dz - \int_a^b \int_0^T D_1^1(z,t) v'(t) u'(z) dt dz \\ &= v(T) \textstyle\int_a^b u'(z) D_1^1(z,T) dz + \int_a^b \int_0^T (-) D_1^1(z,t) v'(t) u'(z) dt dz \geq 0. \end{aligned} \tag{1}$$

And integrating once more with respect to the consequence space we obtain an expression in terms of D_1^2, i.e. in terms of First-order Time and Second-order Stochastic Dominance:

$$\begin{aligned} \Delta &= v(T) u'(b) D_1^2(b,T) + \textstyle\int_0^T - v'(t) u'(b) D_1^2(b,t) dt - \\ &- v(T) \textstyle\int_a^b u''(z) D_1^2(z,T) dz + \int_0^T \int_a^b (-v'(t)) (-u''(z)) D_1^2(z,t) dz dt \geq 0. \end{aligned} \tag{2}$$

The proofs are built around the notion that it is sufficient for TSD that each element on the right-hand side of Equations (4) and (5) is non-negative. We can then define the maximum violation of strict dominance for each element of the equation, which is consistent with this. The great advantage of this approach is that the violation measures have meaningful interpretations in terms of discount and utility functions. However since the various elements are additive and moreover since the various restrictions are defined on different domains of $\{v(t)\}$ and $\{u(x)\}$, there is no reason why any particular element must be non-negative.

On the other hand, we can point to some informative limiting cases. In particular, in the case where

$$\frac{v(T)\int_a^b u'(z)D_1^1(z,T)dz}{\int_a^b\int_0^T v'(t)u'(z)(-)D_1^1(z,t)dtdz} \to 0,$$

it is a necessary condition for Almost First-order TSD that

$$\int_0^T\int_{S_1^1} - D_1^1(z,t)dzdt \leq \gamma_1\int_0^T\int_a^b \left|D_1^1(z,t)\right| dzdt.$$

We might find real examples that approach this limiting case, for which the time horizon for the longest cashflow is very long (e.g. many decades if not centuries) and/or for which the differences in prospects' cashflows in the terminal period are very small. More generally, the measure of violation of strict First-order TSD over all time, γ_1, is evidently an important variable that would be of primary interest in most applications.

The definition of Almost First-order Time and Second-order Stochastic Dominance has four parts. The limiting case we are considering at present will make two elements of Equation (5) vanish, but in addition we would need the difference in prospects' cashflows to be vanishingly small at the maximum wealth level.

5 Worked examples: Time-Stochastic Dominance with quantiles

In this section we present some simple, stylised examples of the TSD criteria at work. The examples are based on quantiles of discrete cashflow distributions in discrete time. This is in part for ease of exposition, but it is also because in practical applications the data to be analysed will very often be in this form, for instance it would be the typical output of a Monte Carlo simulation of a structural model. Since we have so far set out our theory with respect to continuous cumulative distributions, it is therefore an opportunity for us to re-express it in terms of quantile distributions and show that it applies just as well to discrete data.

On the time dimension, integration is simply replaced with summation. For each additional restriction placed on the curvature of the discount function, a new round of summation of the cashflows is performed, $X_n(t) = \sum_{w=0}^{t} X_{n-1}(w)$.

On the stochastic dimension, 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

$$\begin{aligned} F^{-1,1}(p,t) \quad &: \quad = \inf\{x : F^1(x,t) \geq p(t)\}, \forall t \in [0,T] \\ F^{-1,r}(p,t) \quad &: \quad = \textstyle\int_0^p F^{-1,1}(y,t)dy, \ \forall p \in [0,1]\,, \ \forall t \text{ and } r \geq 2. \end{aligned} \tag{3}$$

Proposition 7 (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 the Appendix. □

Proposition 7 characterises First-order Time-Stochastic Dominance for quantile distributions. Notice that since the quantile distribution function is just the inverse of the cumulative distribution function, 1TSD requires $F_1^{-1,1}(p,t) - G_1^{-1,1}(p,t) \geq 0$, i.e. the inverse of the requirement for 1TSD in terms of cumulative distributions.

Proposition 7 also applies to discrete data. To show this briefly, we 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 3 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.

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):

			Time	period		
Prospect	Probability	0	1	2	3	4
	1/4	-2	-3	2	2	1
X	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

$F_1^{-1,1}(p,t)$ and $G_1^{-1,1}(p,t)$ are obtained by first cumulating the cashflows across time, and then reordering from lowest to highest in each time period. Taking the difference between them gives us $H_1^{-1,1}(p,t)$:

		Time	period		
p	0	1	2	3	4
0.25	3	3	1	4	4
0.5	3	4	2	2	1
0.75	4	3	2	3	0
1	4	4	1	4	3

Therefore by Propositions 1 and 7 $X >_{1TS} Y$.

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

			Time	period		
Prospect	Probability	0	1	2	3	4
	1/4	-4	-1	2	3	9
X	1/4	-1	-3	2	2	7
	1/4	-1	-1	2	0	4
	1/4	0	0	2	2	2
	1/4	-5	-1	2	2	2
Y	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:

		Time	period		
p	0	1	2	3	4
0.25	1	1	3	3	4
0.5	1	1	2	2	3
0.75	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:

		Time	period		
p	0	1	2	3	4
0.25	1	1	3	3	4
0.5	2	2	5	5	7
0.75	3	2	7	5	9
1	3	2	7	6	7

Thus from Proposition 2 and by extension of Proposition 7 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 behaviour, makes it disappear, because relatively greater weight is placed on outcomes with low wealth.

Proposition 8 (2TSD for quantile distributions). *$X >_{2TS} Y$ if and only if*

$$H_2^{-1,2}(p,t) \geq 0, \ \forall p \in [0,1] \ \textit{and} \ t \in [0,T]$$

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

The proof is constructed as a simple extension of the previous analysis.

Example 3. Now consider another two different prospects:

			Time	period		
Prospect	Probability	0	1	2	3	4
	1/4	-5	-2	2	1	8
X	1/4	-3	-3	2	4	10
	1/4	-1	-1	-2	0	0
	1/4	0	-2	-1	2	4
	1/4	-5	-2	-2	5	0
Y	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:

		Time	period		
p	0	1	2	3	4
0.25	0	0	4	0	0
0.5	1	1	9	0	5
0.75	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 $H_2^{-1,2}$:[8]

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

p	0	Time 1	period 2	3	4
0.25	0	0	4	4	12
0.5	1	2	11	15	23
0.75	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 $H_1^{-1,2} \geq 0$, $\forall p$, $X >_{2TS} Y$.

Example 4. Finally consider the following two prospects:

Prospect	Probability	0	Time 1	period 2	3	4
	1/4	-5	-3	0	4	7
X	1/4	0	-3	1	2	10
	1/4	0	-2	1	3	9
	1/4	0	0	0	1	1
	1/4	-5	-1	0	3	9
Y	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	0	Time 1	period 2	3	4
0.25	0	-2	-1	3	8
0.5	4	3	4	3	4
0.75	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:

A1TSD		A1T2SD		
γ_1	ε_{1T}	γ_2	ε_{2T}	γ_{1b}
0.04	0	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.

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) on the one hand with those of Time Dominance (TD) on the other hand. Like these earlier theories, the approach is non-parametric and allows orderings to be constructed only on the basis of partial information about preferences. But our approach generalises the application of simple SD to intertemporal prospects, by permitting pure temporal preferences, just as it generalises 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 cashflows may be large/non-marginal).

Like other dominance criteria, a possible practical disadvantage of (standard) Time-Stochastic Dominance 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 dealing with this. Our choice has been to extend the notion of Almost SD pioneered by Levy and others, 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, we suggest that it might prove most useful in the case of some highly contentious public-investment decisions, where there is disagreement about appropriate rates of discount and risk aversion. An example might be the mitigation of climate change, and this is considered in the companion paper Dietz and Matei (2013).

References

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

Atkinson, A. B., Bourguignon, F., 1982. The comparison of multi-dimensioned distributions of economic status. The Review of Economic Studies 49 (2), 183–201.

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

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

Dietz, S., Anderson, D., Stern, N., Taylor, C., Zenghelis, D., 2007a. Right for the right reasons: a final rejoinder on the stern review. World Economics 8 (2), 229–258.

Dietz, S., Hope, C., Stern, N., Zenghelis, D., 2007b. Reflections on the stern review (1): a robust case for strong action to reduce the risks of climate change. World Economics 8 (1), 121–168.

Dietz, S., Matei, A., 2013. Is there space for agreement on climate change? a non-parametric approach to policy evaluation.

Dietz, S., Stern, N., 2008. Why economic analysis supports strong action on climate change: a response to the stern review's critics. Review of Environmental Economics and Policy 2 (1), 94–113.

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

Fishburn, P. C., 1964. Decision and Value Theory. Wiley.

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

Gollier, C., Zeckhauser, R., 2005. Aggregation of heterogeneous time preferences. Journal of Political Economy 113 (4), 878–896.

Hadar, J., Russell, W. R., 1969. Rules for ordering uncertain prospects. American Economic Review 59 (1), 25–34.

Hanoch, G., Levy, H., 1969. The efficiency analysis of choices involving risk. Review of Economic Studies 36 (3), 335–346.

Hepburn, C. J., Beckerman, W., 2007. Ethics of the discount rate in the stern review on the economics of climate change. World Economics 8 (1).

Karcher, T., Moyes, P., Trannoy, A., 1995. The stochastic dominance ordering of income distributions over time: The discounted sum of the expected utilities of incomes. In: Barnett, W. A., Moulin, H., Salles, M., Schofield, N. J. (Eds.), Social Choice, Welfare, and Ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics. Cambirgde University Press.

Laibson, D., 1997. Golden eggs and hyperbolic discounting. The Quarterly Journal of Economics 112 (2), 443–478.

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

Levy, H., 1973. Stochastic dominance, efficiency criteria, and efficient portfolios: The multi-period case. The American Economic Review 63 (5), 986–994.

Levy, H., 2006. Stochastic Dominance: Investment Decision Making under Uncertainty, 2nd Edition. Springer.

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

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

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

Levy, H., Paroush, J., 1974a. Multi-period stochastic dominance. Management Science 21 (4), 428–435.

Levy, H., Paroush, J., 1974b. Toward multivariate efficiency criteria. Journal of Economic Theory 7 (2), 129–142.

Levy, M., 2009. Almost stochastic dominance and stocks for the long run. European Journal of Operational Research 194 (1), 250–257.

Lind, R. C., Arrow, K. J., Corey, G. R., Dasgupta, P., Sen, A. K., Stauffer, T., Stiglitz, J. E., Stockfisch, J., Wilson, R., 1982. Discounting for Time and Risk in Energy Policy. Resources for the Future, Washington, DC.

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

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

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

Portney, P. R., Weyant, J. P., 1999. Discounting and Intergenerational Equity. RFF Press.

Rothschild, M., Stiglitz, J. E., 1970. Increasing risk: I. a definition. Journal of Economic Theory 2 (3), 225–243.

Scarsini, M., 1986. Comparison of random cash flows. IMA Journal of Management Mathematics 1 (1), 25–32.

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

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

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

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

A Appendix

A.1 Proof of Proposition 1

Sufficiency:

We want to prove that

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

Assume that z is bounded from below and above, $a \leq z \leq b$. This implies that for $z < a$, $F^1(x,t) = G^1(y,t) = D^1(z,t) = 0$ for all $t \in [0,T]$, while similarly for $z > b$, $D^1(z,t) = G^1(y,t) - F^1(x,t) = 1 - 1 = 0$ for all $t \in [0,T]$.

Denote by

$$\begin{aligned} \Delta &\equiv NPV_{v,u}(X) - NPV_{v,u}(Y) = \textstyle\int_0^T v(t) E_F u(x) dt - \int_0^T v(t) E_G u(y) dt \\ &= \textstyle\int_0^T v(t) \int_a^b f(x,t) u(x) dx dt - \int_0^T v(t) \int_a^b g(y,t) u(y) dy dt \\ &= \textstyle\int_0^T v(t) \int_a^b - d(z,t) u(z) dz dt. \end{aligned}$$

Integrating by parts with respect to z we obtain:

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

The first term in the square brackets is equal to zero (recall that for $z = b$, we have $D^1(b,t) = 1 - 1 = 0$ for all t and for $z = a$ we have $D^1(a,t) = 0$ for all t). Therefore, we are left with

$$\begin{aligned} \Delta &= \textstyle\int_0^T v(t) \left[-\int_a^b (-) D^1(z,t) u'(z) dz \right] dt \\ &= \textstyle\int_0^T \int_a^b v(t) D^1(z,t) u'(z) dz dt. \end{aligned} \tag{1.2}$$

Integrating by parts with respect to t we have:

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

as $D_1^1(z,0) = G_1^1(y,0) - F_1^1(x,0) = 0$ for all $z \in [a,b]$.

From our initial assumption about the bounding of z, we know that $D_1^1(z,t) \geq 0$ and $v(T) \geq 0$. Hence for all $u \in U_1$ and $v \in V_1$, $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$.

Necessity:

We have to prove that

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

Starting from Eq. 1.2, let $(\widetilde{z}, \widetilde{t})$ be the smallest (in the lexicographic sense) pair (z,t) such that $D_1^1(\widetilde{z}, \widetilde{t}) < 0$. We will show that there is a utility function

$\widetilde{u} \in U_1$ and a discount function $\widetilde{v} \in V_1$ for which our supposition implies that $NPV_{v,u}(X) < NPV_{v,u}(Y)$, thus contradicting the original assumption.

Supposing that a violation $D_1^1(\widetilde{z}, \widetilde{t}) < 0$ does exist, since D_1^1 is continuous it will also exist in the range $\widetilde{z} \leq z \leq \widetilde{z} + \varepsilon$. Define the following utility function:

$$\widetilde{u}(z) = \begin{cases} \widetilde{z} & z < \widetilde{z} \\ x & \widetilde{z} \leq z \leq \widetilde{z} + \varepsilon \\ \widetilde{z} + \varepsilon & z > \widetilde{z} + \varepsilon \end{cases}$$

noting that $\widetilde{u}(z) \notin U_1$, strictly speaking, but that it can be approximated arbitrarily closely by a function that does belong to U_1 (see Fishburn and Vickson 1978, p. 75).

Similarly define the following discount function:

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

which again is discontinuous but can be approximated arbitrarily closely by some $\widetilde{v} \in V_1$.

Substituting these functions into Equation 1.2 we obtain

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

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

A.2 Proof of Proposition 2 and Proposition 3

Sufficiency:

Starting with the expression derived in the previous proof

$$\Delta = \int_0^T v(t) \int_a^b u'(z) D^1(z,t) dz dt,$$

we continue by integrating again with respect to z:

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

Now integrating by parts with respect to time t,

$$\begin{aligned}\Delta &= u'(b)v(t)D_1^2(b,t)|_0^T - u'(b)\textstyle\int_0^T v'(t)D_1^2(b,t)dt - \\ &- \textstyle\int_a^b u''(z)v(t)D_1^2(z,t)|_0^T + \int_a^b u''(z)\int_0^T v'(t)D_1^2(z,t)dtdz.\end{aligned}$$

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

From this last expression we can extract the conditions for dominance with respect to $V_1 \times U_2$ presented in Proposition 2. 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 $\{v,u\} \in V_2 \times U_2$:

$$\begin{aligned}\Delta &= u'(b)v(T)D_1^2(b,T) - \textstyle\int_a^b u''(z)v(T)D_1^2(z,T)dz - u'(b)v'(t)D_2^2(b,t)|_0^T + \\ &+ u'(b)\textstyle\int_0^T v''(t)D_2^2(b,t)dt + \int_a^b u''(z)v'(t)D_2^2(z,t)dz|_0^T - \\ &- \textstyle\int_a^b u''(z)\int_0^T v''(t)D_2^2(z,t)dtdz.\end{aligned}$$

$$\begin{aligned}\Delta &= u'(b)v(T)D_1^2(b,T) - \textstyle\int_a^b u''(z)v(T)D_1^2(z,T)dz - u'(b)v'(T)D_2^2(b,T) + \\ &+ u'(b)\textstyle\int_0^T v''(t)D_2^2(b,t)dt + \int_a^b u''(z)v'(T)D_2^2(z,T)dz - \\ &- \textstyle\int_a^b u''(z)\int_0^T v''(t)D_2^2(z,t)dtdz.\end{aligned}$$

From here it is easy to note that the following assumptions

$$\begin{aligned} i)\ D_1^2(z,T) &\geq 0 \text{ for all } z \in [a,b] \\ ii)\ D_2^2(z,t) &\geq 0 \text{ for all } z \in [a,b] \text{ and all } t \in [0,T]\end{aligned}$$

imply that

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

This completes the sufficiency part of Proposition 3.

Necessity:

Consider the increasing and concave utility function defined by

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

and let $\widetilde{U} \in U_2$ be a suitable approximation of $\widetilde{u}$. The proofs of necessity are similar to the proofs of necessity of the previous proposition and are therefore omitted. □

A.3 Proof of Proposition 5

We want to prove that

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

Going back to

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

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 $D_1^1(z,T) \geq 0$:

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

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 $\Lambda \geq 0$, 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_1^1, defined over ranges where $D_1^1(z,t) < 0$, and $\overline{S_1^1}$, the range over which $D_1^1(z,t) \geq 0$, so that we obtain

$$\begin{aligned}\Gamma = &\int_0^T\int_{S_1^1}\left[D_1^1(z,t)\right](-v'(t)u'(z))\,dzdt\, +\\&\int_0^T\int_{\overline{S_1^1}}\left[D_1^1(z,t)\right](-v'(t)u'(z))\,dzdt \geq 0.\end{aligned}$$

The first element of Γ is negative and is minimised when the product of the marginals of the discount and utility functions $[-v'(t)u'(z)]$ is maximised, while the second element is positive and minimised when $[-v'(t)u'(z)]$ is minimised. 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

$$\Gamma^* = \overline{\theta}\int_0^T\int_{S_1^1}\left[D_1^1(z,t)\right]dzdt + \underline{\theta}\int_0^T\int_{\overline{S_1^1}}\left[D_1^1(z,t)\right]dzdt \geq 0.$$

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

$$\sup[-v'(t)u'(z)] \leq \inf[-v'(t)u'(z))] \frac{\int_0^T \int_{\overline{S_1^1}} D_1^1(z,t)dzdt}{\int_0^T \int_{S_1^1} D_1^1(z,t)dzdt}$$

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

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

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

A.4 Proof of Proposition 6

We want to prove that

$$\begin{aligned}
&X >_{A1T2S} Y \\
&\Rightarrow NPV_{v,u}(X) \geq NPV_{v,u}(Y) \\
&\forall\,(v,u) \in V_1(\gamma_2) \times U_2(\gamma_2),\ \forall u \in U_1(\varepsilon_{1T}) \text{ and } \forall\,(v,u) \in V_1(\gamma_{1b}) \times U_1(\gamma_{1b}))
\end{aligned}$$

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

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

$$\begin{aligned}
v(T)u'(b)D_1^2(b,T) + \int_0^T - v'(t)u'(b)D_1^2(b,t)dt- \\
-v(T)\int_a^b u''(z)D_1^2(z,T)dz + \int_0^T \int_a^b (-v'(t))(-u''(z))\, D_1^2(z,t)dzdt \quad &\geq \quad 0
\end{aligned}$$

$$v(T)u'(b)D_1^2(b,T) + \overline{\Gamma} + \overline{\Lambda} + \Omega \geq 0.$$

Hence in the case of Almost First-order Time and Second-order Stochastic Dominance four elements must be non-negative. $v(T)u'(b)D_1^2(b,T)$ must simply be non-negative. The remaining three elements must be non-negative overall, but can be partitioned into a region of violation and a region of non-violation, with three respective restrictions on the relative violation.

Define the set of realisations 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

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

The integral over $S_{1,b}$ is negative while the integral over its complement $\overline{S_{1,b}}$ is positive. Therefore, following the proof of Proposition 5, in order for $\overline{\Gamma} \geq 0$ the area where $D_1^2(b,t) < 0$ must be γ_{1b} smaller than the total area enclosed between the two distributions, where the restriction is obtained by requiring that any pair of discount and utility functions (v, u) belong to $V_1(\gamma_{1b}) \times U_1(\gamma_{1b})$, where the bounds on the product of the marginals of the discount and utility functions are established with respect to $z = b$ specifically.

$\overline{\Lambda}$ is similar to Λ 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 that ε_{2T} multiplied by the total area between the two distributions.

Moving to Ω, define an interval of violation and its complement in the usual way:

$$\Omega = \int_0^T \int_{S_1^2} (-v'(t)) (-u''(z)) D_1^2(z, t) dz dt + \int_0^T \int_{\overline{S_1^2}} (-v'(t)) (-u''(z)) D_1^2(z, t) dz dt.$$

Again following the proof of Proposition 5 define $\inf_{z\in[a,b]\forall t} \{v'(t)u''(z)\} = \underline{\vartheta}$ and $\sup_{z\in[a,b]\forall t} \{v'(t)u''(z)\} = \overline{\vartheta}$, so that the minimum Ω is

$$\Omega^* = \underline{\vartheta} \int_0^T \int_{S_1^2} D_1^2(z, t) dz dt + \overline{\vartheta} \int_0^T \int_{\overline{S_1^2}} D_1^2(z, t) dz dt.$$

Both elements of Ω are relatively larger than the corresponding elements of Ω^* .

We are looking for a set of preferences $V_1(\gamma_2) \times U_2(\gamma_2)$ for which $\Omega^* \geq 0$, which are

$$\sup[v'(t)u''(z)] \leq \inf[v'(t)u''(z)] \frac{\int_0^T \int_{\overline{S_1^2}} \left[D_1^2(z, t)\right] dz dt}{\int_0^T \int_{S_1^2} \left[F_1^2(z, t) - G_1^2(z, t)\right] dz dt}$$

$$\sup[v'(t)u''(z)] \leq \inf[v'(t)u''(z))] \frac{\int_0^T \int_{\overline{S_2^2}} \left[G_1^2(z, t) - F_1^2(z, t)\right] dz dt}{\int_0^T \int_{S_2^2} \left[F_1^2(z, t) - G_1^2(z, t)\right] dz dt}$$

By letting $(v, u) \in V_1(\gamma_2) \times U_2(\gamma_2)$, then, by definition of $V_1(\gamma_2) \times U_2(\gamma_2)$, we know that

$$\sup[v'(t)u''(z)] \leq \inf[v'(t)u''(z)] \left[\frac{1}{\gamma_2} - 1\right],$$

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

A.5 Proof of Proposition 7

We need to prove that the following holds:

$$H_1^{-1}(p, t) = F_1^{-1}(p, t) - G_1^{-1}(p, t) \geq 0, \quad \forall p \in [0, 1] \text{ and } t \in [0, T] \iff D_1^1(z, t) = G_1^1(z, t) - F_1^1(z, t) \geq 0 \quad \forall z \in [a, b] \text{ and } \forall t \in [0, T]$$

Given that $F_1^1(z, t) \leq G_1^1(z, t)$ is an optimal decision rule for all $(v, u) \in V_1 \times U_1$, if the above expression holds, so will $F_1^{-1}(p, t) \geq G_1^{-1}(p, t)$.

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 the reverse condition also holds, that is for a given t, $F_1^{-1}(p,t) \geq G_1^{-1}(p,t)$ will imply $F_1^1(z,t) \leq G_1^1(z,t)$. □