

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 Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council and has five inter-linked research programmes:

1. Developing climate science and economics
2. Climate change governance for a new global deal
3. Adaptation to climate change and human development
4. Governments, markets and climate change mitigation
5. The Munich Re Programme - Evaluating the economics of climate risks and opportunities in the insurance sector

More information about the Centre for Climate Change Economics and Policy can be found at: <http://www.cccep.ac.uk>.

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training in climate change and the environment. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute, and has five research programmes:

1. Global response strategies
2. Green growth
3. Practical aspects of climate policy
4. Adaptation and development
5. Resource security

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: <http://www.lse.ac.uk/grantham>.

This working paper is intended to stimulate discussion within the research community and among users of research, and its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper represent those of the author(s) and do not necessarily represent those of the host institutions or funders.

Spaces for agreement: a theory of Time-Stochastic Dominance

Simon Dietz* and Nicoleta Anca Matei†

October 16, 2014

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 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, ESRC Centre for Climate Change Economics and Policy, and Department of Geography and Environment, London School of Economics and Political Science.

†European Commission, Joint Research Centre (JRC), Institute for Prospective Technological Studies (JRC-IPTS), Seville.

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

We would like to acknowledge the input of Alec Morton, Claudio Zoli, and seminar participants at WCERE 2014, AARES 2014, Hamburg University, LSE, Manchester University and at the Global-IQ meetings in Rome and Brussels. 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.

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.

A particularly good example of a long-run, risky public investment is mitigation of climate change by reducing emissions of greenhouse gases. Carbon dioxide, the principal greenhouse gas, has a residence time in the atmosphere stretching into 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*. It comes as perhaps no surprise then that great controversy surrounds policy proposals to abate emissions, and that this controversy has turned in large measure on positions taken on time and risk preferences. Compare, for example, the British Government's *Stern Review on the Economics of Climate Change* (Stern, 2007), which advocated immediate and deep cuts in global emissions based on a distinctive combination of low pure-time discounting and moderate risk aversion, with critiques of it by Nordhaus (2007) and Weitzman (2007) among others. However, the Stern Review is merely one of the latest and most prominent manifestations of disagreement about time and risk in project appraisal (landmarks include Lind et al. 1982, and Portney and Weyant 1999). Other examples of public projects that have especially long time-horizons and risky returns are radioactive waste disposal and the protection of wilderness and biological diversity. Matters may usually be easier in the private sector, where the cost of capital dictates time preference and the ability to hedge risk supports risk-neutral choices, but incompleteness in financial markets can undermine these conditions and motivate a desire both to smooth cashflows over time and to avoid risk, for example in research and development.

Our starting point for this paper is that such a debate legitimately exists and will continue for the foreseeable future. The ingredients for the debate include normative differences and positive uncertainties, neither of which seems easy to resolve, certainly in the context of public investments. The normative differences at hand are often framed as between believers in a 'descriptive' approach to time and risk, which relies on appropriate data from markets or other samples of representative 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 specification of time and risk preferences, 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 this is likely to take a long time.

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. While this non-parametric approach could be used to inform investment choice in either the public or private sectors, one of its main uses is likely to be to bring renewed clarity to certain critical and hotly contested choices in public policy, such as mitigation of climate change. In these areas, the focus on the debate about time and risk preferences has arguably come at the expense of asking whether in fact there are some courses of action that both sides could agree to take. In turn, a key question is whether these courses of action are non-trivial.

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 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 non-parametric approach to risky choice, whereby one tests for SD relations for whole preference classes. However, 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.¹ 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 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

¹One 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.

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.

Another drawback of the basic theory of SD is 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 the restriction is very small (to be defined precisely later), an ASD 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, inter-temporal prospects, which admits the possibility of pure-time discounting and makes weak assumptions about the risk characteristics of the prospects. In addition, we extend the notion of Almost SD to our bi-dimensional time-risk setup, defining Almost Time-Stochastic Dominance. This provides a way to exclude extreme combinations of time and risk preferences.

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 a simple but realistic practical application 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). 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

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

²Where 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.

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

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 $u : [a, b] \rightarrow \mathbb{R}$, 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 most 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, $v : [0, T] \rightarrow \mathbb{R}$. 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 :

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

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. 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 behaviour (Laibson, 1997) and that it can result from the aggregation of

heterogeneous individual preferences.⁴

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 rate of impatience is non-increasing 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 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. Note also that these classes of time and risk preferences could alternatively be defined via a composite, multi-attribute function $Q(x, t) = v(t) * u(x)$, on which we could directly impose equivalent restrictions.⁵ We prefer to keep them separate, however, in order to make the independent role of time and risk preferences, and TD and SD procedures respectively, more explicit.

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:

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

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 characterise TSD for various combinations of classes of U_j and V_i .

⁴Even though those individual preferences are represented by exponential discounting (see Gollier and Zeckhauser, 2005).

⁵As is commonly done in multivariate dominance analysis, in particular when dealing with utility as a function of two or more goods (Atkinson and Bourguignon, 1982).

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

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

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

Proof. See Appendix 1. □

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

Consider the following stylised example, comprising discrete cashflow distributions in discrete time. The use of discrete data makes the exposition easy, but it is also the form of data that would typically be encountered in practical applications; for instance such would be the output of a Monte Carlo simulation of a structural model.

Example 1. Consider prospects X and Y , each of which comprises a cashflow 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
Y	1/4	-5	-3	2	3	7
	1/4	-4	-3	2	3	1
	1/4	-4	-1	-1	0	1
	1/4	-4	0	1	1	6

⁶Proposition 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.

Instead of integration with respect to time, use 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

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

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 1. □

Proposition 2 characterises First-order Time-Stochastic Dominance for quantile distributions and it can further easily be shown to apply to discrete data.⁷ $F_1^{-1,1}(p, t)$ 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 $H_1^{-1,1}(p, t)$. 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.

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

p	Time period				
	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 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 behaviour *a priori*, especially in the public sector.

⁷Choose 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 1 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.

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 1. □

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 bi-dimensional 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
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:

p	Time period				
	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 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 behaviour, makes it disappear, because relatively greater weight is placed on outcomes with low wealth.

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 non-increasing rate of impatience, 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

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

and there is at least one strict inequality.

Proof. See Appendix 1. □

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
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

p	Time period				
	0	1	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 $H_1^{-1,2} \geq 0$, $\forall p$, $X >_{2TS} Y$.

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

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

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

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

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

From which the statement of the Proposition follows immediately.

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

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

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.⁹ 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.¹⁰ Dominance relations between risky prospects are then characterised for ‘almost’ all decision-makers.

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’, excluding extreme combinations of time and risk preferences so that prospects can still be ranked. In particular, by extending the theory to our bi-dimensional time-risk setup, we define and characterise 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.¹¹

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 in terms of sets of realisations and time-periods, relative to the non-violation of 1TSD, and give this relative measure of violation meaning by linking it with a restriction on time and risk preferences. To see this, begin by defining the set of realisations $z \in [a, b]$ for all t where there is a violation of First-order TSD as S_1^1 :

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

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

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

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*

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

⁹What 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.

¹⁰Tzeng 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.

¹¹We could also characterise Almost Second-order TSD, but when the attention moves in this approach to measuring and interpreting violations of standard classes of function, there is much less to be gained in placing the additional, categorical restriction on time preferences that this entails, going from a first order to a second order with respect to time as outlined in Section 2. Put another way, these can be seen in some respects as alternative methods of restricting preferences.

Proposition 6 (A1TSD). $X \succ_{A1TS} Y$ if and only if, for all $(v, u) \in (V_1 \times U_1)(\gamma_1)$ and $u \in U_1(\varepsilon_{1T})$,
 $NPV_{v,u}(X) \geq NPV_{v,u}(Y)$.

Proof. See Appendix 1. □

The definition of Almost First-order TSD contains two measures of the violation of standard 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. For the sake of an easy exposition, let us explain the latter violation measure first, ε_{1T} . It has in fact the same interpretation in terms of utility as the corresponding violation measure in Leshno and Levy (2002). The difference is that in our bi-dimensional framework it matters over what time-period the violation is measured, and here it is for $t = T$. Adapting their theorem to our context, for every $0 < \varepsilon_{1T} < 0.5$, define the following subset of U_1 :

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

$U_1(\varepsilon_{1T})$ 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 $-\left| \frac{u''(z)z}{u'(z)} \right|$ – such that it cannot be large negative or large positive, and the larger is ε_{1T} the smaller $\left| \frac{u''(z)z}{u'(z)} \right|$ 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 .

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

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

$(V_1 \times U_1)(\gamma_1)$ 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.

All decision makers exhibiting the non-extreme combination of preferences expressed by the discount and utility functions $(v, u) \in (V_1 \times U_1)(\gamma_1)$ and $u \in U_1(\varepsilon_{1T})$ will prefer X to Y if and only if the condition in Proposition 6 is satisfied.

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 standard First-order Time and Second-Order Stochastic Dominance is violated:

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

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

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

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

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

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

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

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 1. □

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 $D_1^2(b, T) \geq 0$, i.e. that the difference between the undiscounted mean values of projects X and Y respectively is at least zero. First, $\gamma_{1,2}$ 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:

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

The set $(V_1 \times U_2)(\gamma_{1,2})$ 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 $\frac{1}{\gamma_{1,2}} - 1$. 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 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 $0 < \varepsilon_{2T} < 0.5$,

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

$U_2(\varepsilon_{2T})$ 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 λ_{1b} and it is equivalent to restricting time preferences as follows:

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

$V_1(\lambda_{1b})$ 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 $\frac{1}{\lambda_{1b}} - 1$. Hence large changes in $v'(t)$ are excluded.

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
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.

5 Practical application

In this section we show how the TSD criteria might be applied in practical contexts. These examples are adapted from an exercise in a well-known textbook on appraisal of public-sector projects (Boardman et al., 2010).

The municipality of New City is considering building a recreation centre. There is an up-front cost of constructing the centre, as well as annual operating costs including staffing and maintenance, which are assumed to be constant in real terms. Analysts have made an estimate of the benefits of the recreation centre in its first year of operation, as well as forecasting real growth of these benefits in future years due to increases in incomes and population. After twenty years of operation, New City plans to sell the land on which the centre has been built. This is the project's scrap value.

Suppose the municipality is presented with a menu of three alternatives (A - C) for the recreation centre, as tabulated in Appendix 2 (these are our extensions to the original exercise). In each of these alternatives, the costs and benefits of building the recreation centre are uncertain, with the various uncertainties characterised by probability distributions. The analyst runs a Monte Carlo simulation, yielding discrete cashflow distributions in discrete time.

Example 1. Consider Options A and B for the recreation centre. These two options are identical in every respect, except that Option A has higher mean construction costs, as well as a higher variance around these costs.

The analyst uses the data in Appendix 2 to perform a Monte Carlo simulation of the cashflows from Options A and B , sampling 1000 times from the probability distributions specified. For each alternative, the resulting set of cashflows is first cumulated across time and then reordered from lowest to highest in each time period in order to obtain $F_1^{-1,1}(p, t)$, which represents Option B , and $G_1^{-1,1}(p, t)$, which represents Option A . Taking the difference between these two quantile distributions gives us $H_1^{-1,1}(p, t)$, which is plotted in Figure 1. Visual inspection confirms that $H_1^{-1,1}(p, t) \geq 0 \forall p \in [0, 1]$ and $t \in [0, 20]$. Therefore, by Propositions 1 and 2, $B >_{1TS} A$ by all impatient decision makers with monotonic non-decreasing preferences, and Option A can be excluded from the efficient set due to its higher costs.

Example 2. Now compare Option B with Option C . Option C has higher mean benefits, but higher operating costs. Analysts also have reason to believe that the mean benefits and operating costs of Option C are correlated (for example, the variable costs of the recreation activities offered by C depend more strongly

Figure 1: $H_1^{-1,1}(p, t)$ for Options A and B compared, where $F_1^{-1,1}(p, t)$ represents Option B and $G_1^{-1,1}(p, t)$ represents Option A.

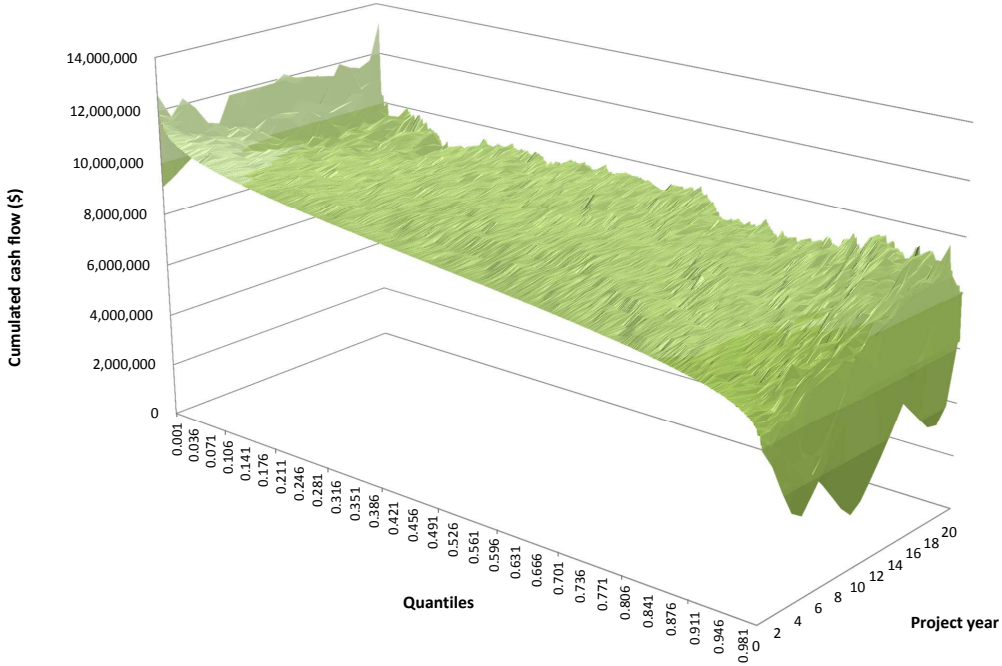
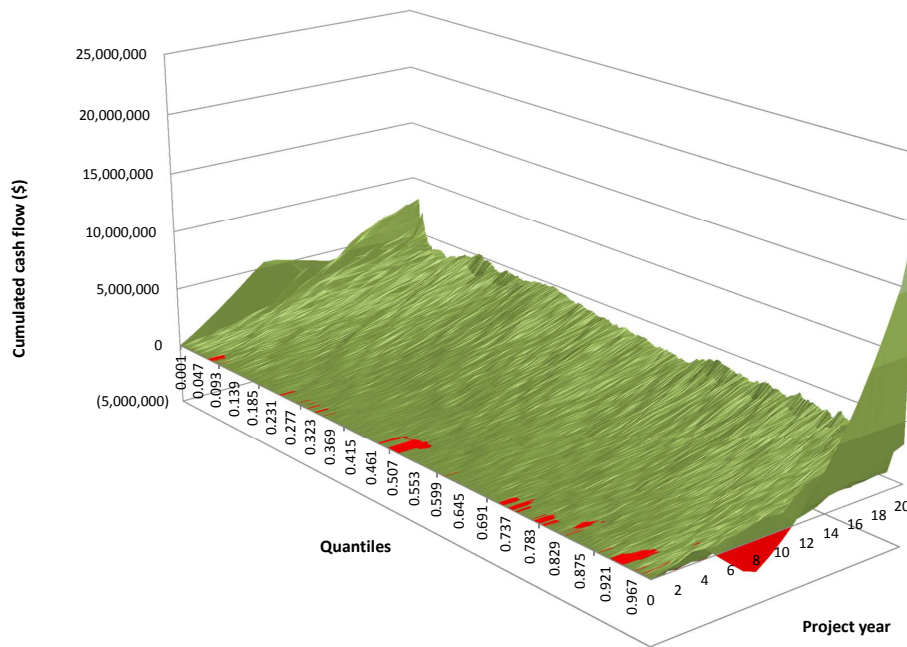


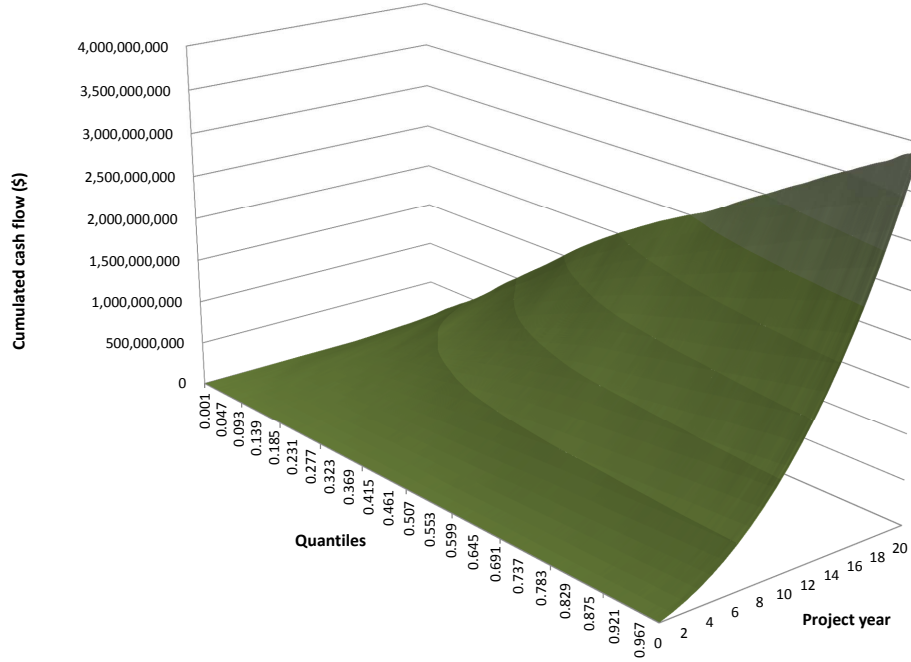
Figure 2: $H_1^{-1,1}(p, t)$ for Options B and C compared, where $F_1^{-1,1}(p, t)$ represents Option C and $G_1^{-1,1}(p, t)$ represents Option B .



on demand/use of the facilities), while in Option B (and A for that matter) it is considered a reasonable approximation to treat them as independently distributed.

Figure 2 plots $H_1^{-1,1}(p, t)$, where $F_1^{-1,1}(p, t)$ represents Option C and $G_1^{-1,1}(p, t)$ represents Option B . While positive in the majority of circumstances, it can be seen that there are violations in this case (shaded red). Therefore First-order TSD cannot be established between these two alternatives. Where can we go from here? One option is to place an additional restriction on the discount function, inspecting whether we can establish *Second-order* Time and First-order Stochastic dominance. $H_2^{-1,1}(p, t)$ is obtained by cumulating the cashflows from each alternative twice with respect to time, reordering these twice-cumulated cashflows from lowest to highest in each time period, and taking the difference between the resulting distributions $F_2^{-1,1}(p, t)$ and $G_2^{-1,1}(p, t)$. However, we will not be able to establish strict TSD in this way. The reason is that a violation of First-order TSD exists in year 1, due to the higher operating costs of Option C , which are relatively more important initially, before benefits growth has its effect. This violation will therefore persist despite cumulation with respect to time (indeed it will persist despite infinitely repeated cumulation across time).

Figure 3: $H_1^{-1,2}(p, t)$ for Options B and C compared.



Instead, let us place an additional restriction on the class of utility functions, limiting the possible space for agreement to risk-averse decision-makers. We inspect whether Option C First-order Time and *Second-order* Stochastic dominates B . This is achieved by cumulating $H_1^{-1,1}(p, t)$ with respect to the consequence space, giving $H_1^{-1,2}(p, t)$. Figure 3 shows that this is positive everywhere, so we can say that $C >_{1T2S} B$.

The property of this example, which makes it possible to establish that $C >_{1T2S} B$, is that the violations of First-order TSD fall predominantly at medium to high quantiles and, critically, not at the very lowest quantiles. That is to say, $H_1^{-1,1}(p, t)$ is more likely to be negative when $F_1^{-1,1}(p, t)$ and $G_1^{-1,1}(p, t)$ are both high. Cumulation with respect to the consequence space can therefore eliminate the violation. The correlation between the higher operating costs and higher mean benefits of Option C relative to Option B is an important factor in producing this effect. With the additional restriction that $u \in U_2$, which excludes risk-seeking behaviour, relatively greater weight is placed on outcomes where the net benefits of the projects are low, where there is at the same time no large violation of First-order TSD.

Let us re-consider the choice between Options B and C . Instead of seeking to establish strict TSD by placing an additional restriction on the utility function, instead quantify *Almost* First-order TSD of Option C over Option B . After all, Figure 2 showed that the violation of First-order TSD appeared to be small

relative to the total volume enclosed between $F_1^{-1,1}(p, t)$ and $G_1^{-1,1}(p, t)$. Doing the necessary calculations as explained in Section 4, $\gamma_1 = 0.0008$ and $\varepsilon_{1T} = 0$, which confirms what is visually apparent. We can identify a space for agreement that $C >_{A1TS} B$. Moreover we can say on the strength of very low γ_1 and zero ε_{1T} that those who do not prefer Option C to B have an extreme combination of time and risk preferences.

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 to our bi-dimensional 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 pay-offs, 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.
- Boardman, A. E., Greenberg, D. H., Vining, A. R., Weimer, D., 2010. *Cost-Benefit Analysis: Concepts and Practice*. Prentice Hall Upper Saddle River, NJ.
- Bøhren, Ø., Hansen, T., 1980. Capital budgeting with unspecified discount rates. *The Scandinavian Journal of Economics*, 45–58.

- Dietz, S., Matei, A., 2013. Is there space for agreement on climate change? a non-parametric approach to policy evaluation.
- 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., 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.
- 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*. Cambridge 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., 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., Stockfish, 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.
- 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 1

A.1 Proof of Proposition 1

Sufficiency:

We want to prove that

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

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

Assume that 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

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

Integrating by parts with respect to z we obtain:

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

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

$$\begin{aligned} \Delta &= \int_0^T v(t) \left[-\int_a^b (-D^1(z, t)) u'(z) dz \right] dt \\ &= \int_0^T \int_a^b v(t) D^1(z, t) u'(z) dz dt. \end{aligned} \tag{1.1}$$

Integrating by parts with respect to t we have:

$$\begin{aligned} \Delta &= \int_a^b \left[D_1^1(z, t) v(t) \Big|_0^T - \int_0^T D_1^1(z, t) v'(t) dt \right] u'(z) dz \\ &= \int_a^b \left[D_1^1(z, T) v(T) - \int_0^T D_1^1(z, t) v'(t) dt \right] u'(z) dz, \end{aligned} \tag{1.2}$$

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

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

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

Starting from Eq. 1.1, 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 a utility function $\tilde{u} \in U_1$ and a discount function $\tilde{v} \in V_1$ for which our supposition implies that $NPV_{\tilde{v},\tilde{u}}(X) < NPV_{\tilde{v},\tilde{u}}(Y)$, thus contradicting the original assumption.

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

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

noting that $\tilde{u}(z) \notin U_1$, 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:

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

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

Substituting these functions into Equation 1.2 we obtain

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

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 $t \in [0, T]$. \square

A.2 Proof of Proposition 2

We need to prove that the following holds:

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

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 $F_1^1(z, t) \leq G_1^1(z, t)$. \square

A.3 Proof of Proposition 3 and Proposition 4

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 :

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

Now integrating by parts with respect to time t ,

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

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

From this last expression 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 $\{v, u\} \in V_2 \times U_2$:

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

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

From here it is easy to note 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

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

This completes the sufficiency part of Proposition 4.

Necessity:

Consider the increasing and concave utility function defined by

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

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

A.4 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_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

$$\begin{aligned} \Gamma &= \int \int_{S_1^1} [D_1^1(z, t)] (-v'(t) u'(z)) dz dt + \\ &\quad \int \int_{\overline{S_1^1}} [D_1^1(z, t)] (-v'(t) u'(z)) dz dt \geq 0. \end{aligned}$$

The first element of Γ is negative and is 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 $\sup_{z \in [a,b] \forall t} \{-v'(t)u'(z)\} = \bar{\theta}$, the minimum value of Γ is

$$\Gamma^* = \bar{\theta} \int \int_{S_1^1} [D_1^1(z, t)] dz dt + \underline{\theta} \int \int_{\bar{S}_1^1} [D_1^1(z, t)] dz dt \geq 0.$$

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

$$\sup[-v'(t)u'(z)] \leq \inf[-v'(t)u'(z)] \frac{\int \int_{\bar{S}_1^1} D_1^1(z, t) dz dt}{\int \int_{S_1^1} D_1^1(z, t) dz dt}$$

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, i.e. 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

$$\begin{aligned} i) \int \int_{S_1^1} [-D_1^1(z, t)] dz dt &> \gamma_1 \int_0^T \int_a^b |D_1^1(z, t)| dz dt \text{ and} \\ ii) \int_{S_1^1, T} [-D_1^1(z, T)] dz &> \varepsilon_{1T} \int_a^b |D_1^1(z, T)| dz. \end{aligned}$$

Let $\bar{\theta}$ and $\underline{\theta}$ be two positive real numbers such that $\gamma_1 = \frac{\underline{\theta}}{\underline{\theta} + \bar{\theta}}$. 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

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

which contradicts the initial assumption and proves that

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

□

A.5 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

$$\begin{aligned} \Delta &= v(T) \left[u'(z) D_1^2(z, T) \Big|_a^b - \int_a^b u''(z) D_1^2(z, T) dz \right] + \\ &+ \int_0^T -v'(t) \left[u'(z) D_1^2(z, t) \Big|_a^b dt - \int_0^T -v'(t) \int_a^b u''(z) D_1^2(z, t) dz dt \right] \geq 0 \\ &v(T) u'(b) D_1^2(b, T) + u'(b) \int_0^T -v'(t) D_1^2(b, t) dt - \\ &- v(T) \int_a^b u''(z) D_1^2(z, T) dz + \int_0^T \int_a^b (-v'(t)) (-u''(z)) D_1^2(z, t) dz dt \geq 0 \\ &v(T) u'(b) D_1^2(b, T) + \Lambda + \bar{E} + \bar{\Gamma} \geq 0. \end{aligned} \tag{6.1}$$

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

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

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 $V_1(\lambda_{1b})$.

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

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

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

$$\bar{\Gamma}^* = \underline{\vartheta} \int_{S_1^2} D_1^2(z, t) dz dt + \bar{\vartheta} \int_{S_1^2} D_1^2(z, t) dz dt.$$

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

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

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

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

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 Eq. 6.1 and using the increasing and concave utility function defined in proving necessity in Proposition 4, the proof proceeds in just the same fashion as for Proposition 6 and is therefore omitted. \square

B Appendix 2

Option A

Assumptions:		
	Mean	Std. deviation
Annual Growth Rate of Benefits	4.0%	0.5%
Construction Cost	\$24,000,000	\$4,000,000
Annual Operations Cost	\$750,000	\$250,000
First Year Benefit	\$1,200,000	\$400,000
Scrap Value	\$4,000,000	\$1,333,333

Year	Construction Cost	Operational Cost	Annual Benefit	Scrap Value	Annual Net Bens
0	\$24,000,000	\$0	\$0	\$0	-\$24,000,000
1	\$0	\$750,000	\$1,200,000	\$0	\$450,000
2	\$0	\$750,000	\$1,248,000	\$0	\$498,000
3	\$0	\$750,000	\$1,297,920	\$0	\$547,920
4	\$0	\$750,000	\$1,349,837	\$0	\$599,837
5	\$0	\$750,000	\$1,403,830	\$0	\$653,830
6	\$0	\$750,000	\$1,459,983	\$0	\$709,983
7	\$0	\$750,000	\$1,518,383	\$0	\$768,383
8	\$0	\$750,000	\$1,579,118	\$0	\$829,118
9	\$0	\$750,000	\$1,642,283	\$0	\$892,283
10	\$0	\$750,000	\$1,707,974	\$0	\$957,974
11	\$0	\$750,000	\$1,776,293	\$0	\$1,026,293
12	\$0	\$750,000	\$1,847,345	\$0	\$1,097,345
13	\$0	\$750,000	\$1,921,239	\$0	\$1,171,239
14	\$0	\$750,000	\$1,998,088	\$0	\$1,248,088
15	\$0	\$750,000	\$2,078,012	\$0	\$1,328,012
16	\$0	\$750,000	\$2,161,132	\$0	\$1,411,132
17	\$0	\$750,000	\$2,247,577	\$0	\$1,497,577
18	\$0	\$750,000	\$2,337,481	\$0	\$1,587,481
19	\$0	\$750,000	\$2,430,980	\$0	\$1,680,980
20	\$0	\$750,000	\$2,528,219	\$4,000,000	\$5,778,219

Option B

Assumptions:		
	Mean	Std. deviation
Annual Growth Rate of Benefits	4.0%	0.5%
Construction Cost	\$15,000,000	\$3,000,000
Annual Operations Cost	\$750,000	\$250,000
First Year Benefit	\$1,200,000	\$400,000
Scrap Value	\$4,000,000	\$1,333,333

Year	Construction Cost	Operational Cost	Annual Benefit	Scrap Value	Annual Net Bens
0	\$15,000,000	\$0	\$0	\$0	-\$15,000,000
1	\$0	\$750,000	\$1,200,000	\$0	\$450,000
2	\$0	\$750,000	\$1,248,000	\$0	\$498,000
3	\$0	\$750,000	\$1,297,920	\$0	\$547,920
4	\$0	\$750,000	\$1,349,837	\$0	\$599,837
5	\$0	\$750,000	\$1,403,830	\$0	\$653,830
6	\$0	\$750,000	\$1,459,983	\$0	\$709,983
7	\$0	\$750,000	\$1,518,383	\$0	\$768,383
8	\$0	\$750,000	\$1,579,118	\$0	\$829,118
9	\$0	\$750,000	\$1,642,283	\$0	\$892,283
10	\$0	\$750,000	\$1,707,974	\$0	\$957,974
11	\$0	\$750,000	\$1,776,293	\$0	\$1,026,293
12	\$0	\$750,000	\$1,847,345	\$0	\$1,097,345
13	\$0	\$750,000	\$1,921,239	\$0	\$1,171,239
14	\$0	\$750,000	\$1,998,088	\$0	\$1,248,088
15	\$0	\$750,000	\$2,078,012	\$0	\$1,328,012
16	\$0	\$750,000	\$2,161,132	\$0	\$1,411,132
17	\$0	\$750,000	\$2,247,577	\$0	\$1,497,577
18	\$0	\$750,000	\$2,337,481	\$0	\$1,587,481
19	\$0	\$750,000	\$2,430,980	\$0	\$1,680,980
20	\$0	\$750,000	\$2,528,219	\$4,000,000	\$5,778,219

Option C

Assumptions:

	Mean	Std. deviation
Annual Growth Rate of Benefits	4.0%	0.5%
Construction Cost	\$15,000,000	\$3,000,000
Annual Operations Cost	\$1,000,000	\$333,333
First Year Benefit	\$1,500,000	\$500,000
Scrap Value	\$4,000,000	\$1,333,333

Year	Construction Cost	Operational Cost	Annual Benefit	Scrap Value	Annual Net Bens
0	\$15,000,000	\$0	\$0	\$0	-\$15,000,000
1	\$0	\$1,000,000	\$1,500,000	\$0	\$500,000
2	\$0	\$1,000,000	\$1,560,000	\$0	\$560,000
3	\$0	\$1,000,000	\$1,622,400	\$0	\$622,400
4	\$0	\$1,000,000	\$1,687,296	\$0	\$687,296
5	\$0	\$1,000,000	\$1,754,788	\$0	\$754,788
6	\$0	\$1,000,000	\$1,824,979	\$0	\$824,979
7	\$0	\$1,000,000	\$1,897,979	\$0	\$897,979
8	\$0	\$1,000,000	\$1,973,898	\$0	\$973,898
9	\$0	\$1,000,000	\$2,052,854	\$0	\$1,052,854
10	\$0	\$1,000,000	\$2,134,968	\$0	\$1,134,968
11	\$0	\$1,000,000	\$2,220,366	\$0	\$1,220,366
12	\$0	\$1,000,000	\$2,309,181	\$0	\$1,309,181
13	\$0	\$1,000,000	\$2,401,548	\$0	\$1,401,548
14	\$0	\$1,000,000	\$2,497,610	\$0	\$1,497,610
15	\$0	\$1,000,000	\$2,597,515	\$0	\$1,597,515
16	\$0	\$1,000,000	\$2,701,415	\$0	\$1,701,415
17	\$0	\$1,000,000	\$2,809,472	\$0	\$1,809,472
18	\$0	\$1,000,000	\$2,921,851	\$0	\$1,921,851
19	\$0	\$1,000,000	\$3,038,725	\$0	\$2,038,725
20	\$0	\$1,000,000	\$3,160,274	\$4,000,000	\$6,160,274

Correlations between project variables

	Construction cost, A	Construction cost, B	Annual Opex, A	Annual Opex, B	1st Year Bens, A	1st Year Bens, C
Construction cost, A	1					
Construction cost, B	0.6	1				
Annual Opex, A	0	0	1			
Annual Opex, B	0	0	0.6	1		
1st Year Bens, A	0	0	0	0	1	
1st Year Bens, C	0	0	0	0.6	0.6	1