Preference Heterogeneity, Extended Preferences and Social Welfare

Matthew Adler, Duke University

December 2014

Background

- A SWF ranks outcomes (and policies) as a function of interpersonally comparable well-being numbers ("utilities"). x at least as good as y iff $(w_1(x), \ldots, w_N(x)) \succcurlyeq^{k\text{-}M} (w_1(y), \ldots, w_N(y))$.
- The SWF is a powerful tool for the ethical/moral/social assessment of policies. It can, to some extent, be mimicked by BCA with distributional weights.
- With heterogeneous preferences, we should allow w(.) to take account of attributes *and* preferences. $w_i(x) = w(a_i(x), R_i(x))$.
- However, much of the SWF literature ignores preference heterogeneity
  - Canonical optimal tax model: $w_i(x) = w(c_i(x), l_i(x))$
  - Climate change scholarship: $w_i(x) = w(c_i(x))$
  - Distributional weights, as recommended by UK Treasury "Green book," based upon log consumption utility function
- I will discuss one approach ("extended preferences") for constructing interpersonally comparable well-being measure that allows heterogeneous prefs.

Extended Preferences

- The account I present here is developed in Adler (2012; 2013; 2015).
- The account builds upon and generalizes John Harsanyi's theory of well-being measurement. Harsanyi used the name "extended preferences." So as to honor his pioneering work, and acknowledge my deep debt to him, I use that name as well.
- That said, there are important differences between my account and Harsanyi's. Mine:
  - Allows a diversity of well-being measures. $w(.) = w^k(.)$, with k a particular ethical deliberator ("social planner")
  - Accommodates non-shareable attributes
  - Explicitly incorporate k's judgments of well-being differences. "Bernoulli," which says that w(.) is a vNM function w/r/t the well-being ranking of lotteries, is an optional rather than mandatory axiom
  - Is not committed to utilitarianism. While Harsanyi saw w(.) as the input into a utilitarian SWF, my account is agnostic about the functional form of the SWF, and indeed I have argued elsewhere for prioritarianism
- Time permitting, I will compare "extended preferences" to "equivalent income"

Well-Being First

- In this approach, the ethical deliberator k first develops an account of well-being, and matching measures of well-being $w^k(.)$and $v^k(.)$, and then identifies her SWF $\succcurlyeq^{k\text{-}M}$ by virtue of axiomatic constraints on the ranking of well-being vectors (e.g., monotonicity, anonymity, separability, continuity, Pigou-Dalton in well-being, invariance to ratio rescalings)
- This is a defensible approach (if not the only one). It corresponds to much of the philosophical literature, where one subliterature focuses on the nature of well-being independent of moral assessment (preference-based versus hedonic versus objective), and another on the structure of moral assessment independent of the specific nature of well-being (the debate between utilitarians, prioritarians, sufficientists, and egalitarians).
- A different approach: the recent social choice literature on "fair social orderings," whereby axioms regarding moral rankings of vectors of attribute bundles characterize an "emergent" well-being measure

Axioms for an SWF (in terms of well-being vectors)

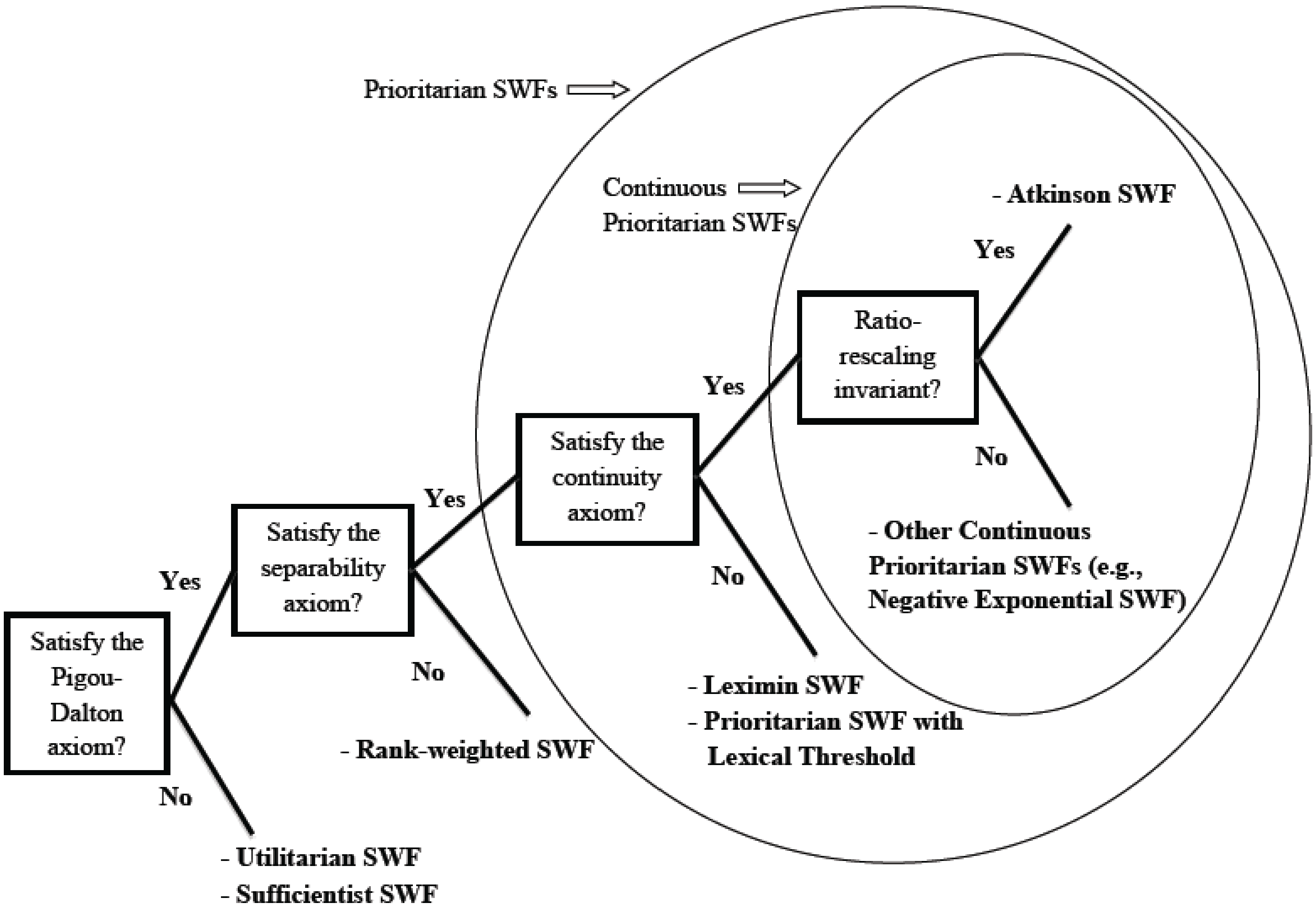
- Monotonicity/Pareto: (3, 4, 10, 13) $\succ^M$ (3, 4, 10, 12)
- "Anonymity": (7, 12, 4, 60) $\sim^M$ (12, 60, 4, 7)
- Pigou-Dalton: (3, 6, 8, 12) $\succ^M$ (3, 4, 10, 12)
- Separability:

  (7, 100, 100, 7) $\succcurlyeq^M$ (4, 100, 100, 12) iff

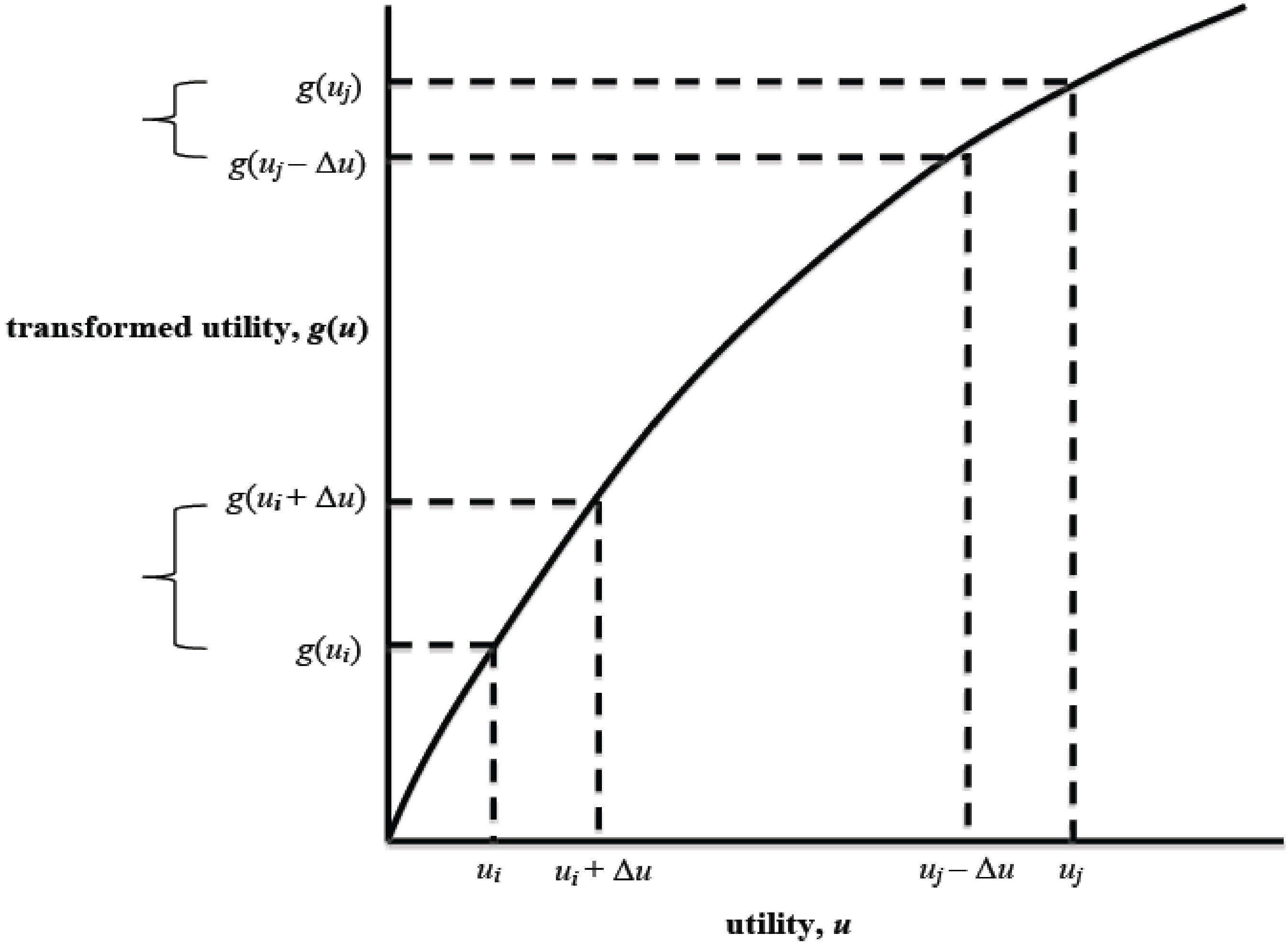
  (7, 7, 7, 7) $\succcurlyeq^M$ (4, 7, 7, 12)
- Continuity: If (1, 3, 50000, 50000) $\succ^M$ (1, 3, 6, 8), then (1, 3±ε, 50000, 50000) $\succ^M$ (1, 3, 6, 8) for ε sufficiently small
- Ratio rescaling invariance:

  (10, 12, 17, 20) $\succ^M$ (10, 10, 20, 20) iff (50, 60, 85, 100) $\succ^M$ (50, 50, 100, 100)

The Universe of Paretian, Anonymous SWFs



The Prioritarian SWF



The ex ante Pareto principles and stochastic dominance

	Status Quo			Policy		
	A	*B*	*EU*	*A*	*B*	*EU*
Jim	10	90	50	50-ε	50-ε	50-ε
June	90	10	50	50-ε	50-ε	50-ε

Any prioritarian SWF will say that the policy is sure to yield a better outcome; and yet the policy is ex ante Pareto inferior. Set ε = zero to produce conflict with ex ante Pareto indifference

Key Ideas

- Let **A** be a set of attribute bundles {a, a*…} and **R** a set of (ordinal and risk preferences, "tastes") over **A**. **R** = {R, R*, …}. Then each "history" h is a pairing (a, R), with **H** the set of all histories.
- The deliberator k makes structured judgments regarding the well-being levels of histories (a ranking of **H**), and regarding the well-being differences between histories (a ranking of **H** x **H**). These are represented by $w^k(.)$. $w^k(h) \geq w^k(h^*)$ iff $h \succcurlyeq^k h^*$. $w^k(h) - w^k(h^*) \geq w^k(h^+) - w^k(h^{++})$ iff $(h, h^*) \succcurlyeq^{k\text{-Diff}} (h^+, h^{++})$.
- The deliberator also makes judgments regarding lotteries over **H**, represented by a vNM function $v^k(.)$. $\sum_{\mathbf{H}} \pi_L(h)v^k(h) \geq \sum_{\mathbf{H}} \pi_{L*}(h)v^k(h)$ iff $L \succcurlyeq^{k\text{-Lott}} L^*$. Note that, because histories are degenerate lotteries, it must be that $v^k(h) \geq v^k(h^*)$ iff $h \succcurlyeq^k h^*$
- These judgments are <u>non-paternalistic</u>, deferring to the ordinal and risk tastes embedded in histories
  - <u>Ordinal deference</u>: $(a, R) \succcurlyeq^k (a^*, R)$ iff a R a*
  - <u>Risk deference</u>: If l and l* are lotteries over **A**, and L and L* corresponding lotteries over **H** by combining every attribute bundle with the very same R, then $L \succcurlyeq^{k\text{-Lott}} L^*$ iff l R l*

Ordinal Deference: The Core of a Preference-Respecting Account of Well-Being

- If Jim favors chocolate over vanilla, and Sue favors vanilla over chocolate, then (if preferences are respected) the well-being of Jim-having-chocolate is greater than that of Jim-having-vanilla, but the well-being of Sue-having-vanilla is greater than that of Sue-having-chocolate. That is, if k's well-being judgments are non-paternalistic, (chocolate, R_{Jim}) $\succ^k$ (vanilla, R_{Jim}) but (vanilla, R_{Sue}) $\succ^k$ (chocolate, R_{sue}).
- Thus $w^k(.)$ must take the form $w^k(a, R)$, not $w^k(a)$, as would an "objective index." Obviously, a measure of the form $w^k(a)$ would have a unitary ranking of chocolate and vanilla, failing to differentiate having ice cream with Jim's tastes and having ice cream with Sue's.

Well-Being Differences

- Economists are often suspicious of well-being differences as a primitive
- However, a substantial literature in decision theory characterizes the features of a coherent and numerically representable difference ordering. Krantz et al (2007); Kobberling (2006)
- An established psychological literature elicits difference judgments
- Folk moral psychology sees nothing problematic in comparing well-being differences. The well-being difference between my being nourished versus starving is larger than the well-being difference between your owning a Mercedes versus a Ford
- Nor do academic philosophers. E.g., Parfit's characterization of "prioritarianism"

Well-Being Differences

- A coherent difference ordering $\succcurlyeq^{Diff}$ on a set **S** x **S** and corresponding ordering $\succcurlyeq$ on **S** satisfies substantive axioms:

  Reversal: (s, r) $\succcurlyeq^{Diff}$ (t, u) iff (u, t) $\succcurlyeq^{Diff}$ (r, s). Separability: If (s, r) $\succcurlyeq^{Diff}$ (t, r) then (s, r*) $\succcurlyeq^{Diff}$ (t, r*) for all r*. Neutrality: (s, s) $\sim^{Diff}$ (t, t) for all s, t. Concatenation: If (s, r) $\succcurlyeq^{Diff}$ (s', r') and (r, t) $\succcurlyeq^{Diff}$ (r', t'), then (s, t) $\succcurlyeq^{Diff}$ (s', t') . Linkage: s $\succcurlyeq$ t iff (s, t) $\succcurlyeq^{Diff}$ (t, t)
- $\succcurlyeq^{Diff}$ and $\succcurlyeq$ can be numerically represented by a single function on **S** if we add two technical axioms.
- If the deliberator's ordering of histories and differences (on **H** and **H** x **H**) satisfy these axioms, we end up with $w^k(.)$
- Non-paternalism axioms and (optional or mandatory) axioms regarding the nexus between $w^k(.)$ and $v^k(.)$ allow the deliberator to make further progress in identifying $w^k(.)$

Identifying $v^k(.)$

- Recall that R embodies a risk preference w/r/t lotteries over **A**. Let $u^R(.)$ be a vNM function representing R. By non-paternalism, $v^k(.)$ defers to R in ranking lotteries over **H** with R embedded. Thus, by vNM theory, there must be scaling factors $s^k(R)$ and $t^k(R)$ such that:

$$v^k(a, R) = s^k(R)\, u^R(a) + t^k(R)$$

- The scaling factors for all tastes under consideration can, in turn, be determined by picking a baseline taste R and then, for every other taste R*, identifying two "points of contact": k judges (a, R) equally good as (a^+, R^*) and (a', R) equally good as (a^{++}, R^*).
- These "across-taste" judgments are normative judgments which, together with the normative commitment to non-paternalism, allow the deliberator to identify $v^k(a, R)$ for all attribute bundles and tastes.
- "High-low" rule: the simplest version of this. (Decanq and Neumann 2015)

The Role of the Scaling Factors

Attribute bundle (a)	$u^{R^*}(a)$	$u^{R^{**}}(a)$
a^+	1	1
a^{++}	3	5
a^{+++}	5	4

Attribute bundle (a)	$s^k(R^*)=50$ $t^k(R^*)=0$ $v^k(a, R^*)$	$s^k(R^{**})=1$ $t^k(R^{**})=0$ $v^k(a, R^{**})$
a^+	50	1
a^{++}	150	5
a^{+++}	250	4

Attribute bundle (a)	$s^k(R^*)=1$ $t^k(R^*)=0$ $v^k(a, R^*)$	$s^k(R^{**})=1$ $t^k(R^{**})=10$ $v^k(a, R^{**})$
a^+	1	11
a^{++}	3	15
a^{+++}	5	14

Attribute bundle (a)	$s^k(R^*)=1$ $t^k(R^*)=3$ $v^k(a, R^*)$	$s^k(R^{**})=100$ $t^k(R^{**})= -99$ $v^k(a, R^{**})$
a^+	4	1
a^{++}	6	401
a^{+++}	8	301

From $v^k(.)$ to $w^k(.)$

- Because $v^k(.)$ and $w^k(.)$ both represent well-being levels, there is a fundamental linkage between them. $w^k(h) = w^k(h^*)$ iff $v^k(h) = v^k(h^*)$. $w^k(h) > w^k(h^*)$ iff $v^k(h) > v^k(h^*)$. Thus in general:

$$w^k(a, R) = F[v^k(a,R)] = F[s^k(R)u^R(a) + t^k(R)]$$

  with F some increasing function.

- If the deliberator adopts the “Bernoulli” axiom, then F(.) is just linear or (w.l.o.g.) the identity function

$$w^k(a, R) = s^k(R)u^R(a) + t^k(R)$$

- <u>“Bernoulli”</u>. The deliberator is “risk neutral” in well-being. She ranks lotteries in terms of their expected $w^k(.)$ values. For any L that gives well-being level W for certain, and L* giving a 50/50 chance of $W+\Delta W$ or $W- \Delta W$, the deliberator judges L and L* equal for well-being

A Summary

- Given a set **A** of attribute bundles and **R** of ordinal and risk tastes over the bundles, the deliberator can arrive at a measure $w^k(.)$ of well-being levels and differences over (a, R) combinations, i.e., "histories"—a measure that allows for heterogeneous tastes—via the following normative judgments about well-being: Non-paternalism; Bernoulli; and "Points of Contact" (two across-taste well-being level judgments for each taste compared to a baseline taste, most simply the high-low rule).
- If so, $w^k(a, R) = s^k(R)\, u^R(a) + t^k(R)$. This yields a $w^k(.)$ unique up to an affine transformation, which can be rendered unique up to a ratio transformation by picking a "zero history", e.g., one no better than nonexistence and setting $w^k(h^{zero}) = 0$
- Topics for additional research: (1) A theory of the "points of contact"; (2) What is the F(.) function if "Bernoulli" not adopted?

Objections

- If the deliberator applies her SWF under risk in "ex post" manner, then $v^k(.)$ plays no direct role in her ethical choices. Why, then, is this part of the theory?

  Ex post utilitarian SWF: $\mathbf{E} \sum w_i$ Ex post prioritarian SWF: $\mathbf{E} \sum g(w_i)$

  Ex ante utilitarian SWF: $\sum F(\mathbf{E}(v_i))$ Ex ante prioritarian SWF: $\sum g(F(\mathbf{E}(v_i)))$

Answer: A reflective deliberator will want to "test" her judgments of well-being levels and differences by being sure they cohere with the judgments she would make regarding other aspects of well-being (lotteries), even if such judgments do not figure directly in ethical evaluation.

- Why shouldn't R be "richer," reflecting difference-rankings over attribute bundles as well as ordinal and risk tastes? $d^R(.)$ represents R's ranking of differences between attribute bundles, just as $u^R(.)$ R's ranking of lotteries

Answer: A topic for research. Requiring that $v^k(.)$ defer to risk tastes and $w^k(.)$ defer to difference tastes may yield an impossibility, where $v^k(a,R) = v^k(a,R^*)$ but not $w^k(a, R) = w^k(a, R^*)$.

Equivalent Incomes

- $e^k(a, R) = e((c, b), R) = c^{equiv}$ s.t. $(c^{equiv}, b^{ref\text{-}k})$ I (c, b). (Fleurbaey & Blanchet 2013; Fleurbaey 2015)
- A potential measure of well-being? Like ext. preferences (w^k), e^k is sensitive to heterogeneous preferences. (Decanq & Neumann 2015, using GSOEP to compare these measures to each other and to three measures that ignore heterogeneous preferences: income, SWB, and an index of goods)
- The choice of $e^k(.)$ involves various normative judgments (e.g., choice of $b^{ref\text{-}k}$), but so does $w^k(.)$
- $e^k(.)$ like $w^k(.)$ satisfies ordinal deference: $e^k(a, R) \geq e^k(a^*, R)$ iff a R a*
- Two differences from $w^k(.)$: $e^k(.)$ is linear in income at the reference bundle of non-income attributes; and $e^k(.)$ doesn't differentiate between risk tastes. It gives the same value to (a, R) and (a, R*) if R and R* have the same ordinal ranking of bundles.
- If $e^k(.)$ really measures well-being levels and differences, then $G(e^k(.))$, for some G increasing, should be a vNM function representing the well-being ranking of lotteries. But $G(e^k(.))$ will violate risk deference

Equivalent Incomes and Risk Deference

- Assume R and R* are the same ordinal tastes, but R prefers bundle a to a 50/50 lottery between a' and a'', while R* has the opposite preference. A well-being measure satisfying risk deference will prefer (a, R) to 50/50 lottery over (a', R), (a'', R), but prefer a 50/50 lottery over (a', R*), (a'', R*) to (a, R*). Since $e^k(.)$ assigns the same values to (a, R), (a', R) and (a'', R) as (a, R*), (a',R*) and (a'', R*), respectively, it is obviously impossible for the expected value of $G(e^k(.))$ to rank the lotteries consistent with risk deference.
- By contrast, note that since $v^k(.)$ by construction will be such that $v^k(a,R) > .5v^k(a',R) + .5v^k(a'', R)$, but $v^k(a,R^*) < .5v^k(a',R) + .5v^k(a'',R)$. $w^k(.) = F(v^k(.))$, and thus there is some $G(w^k(.))$, i.e., $G = F^{-1}$, s.t. that the expected value of $G(w^k(.))$ ranks the lotteries consistent with risk deference.
- Note that $w^k(.)$ does not assign the same values to (a, R), (a', R) and (a'', R) as (a, R*), (a',R*) and (a'', R*), respectively.
- The relation between a candidate measure of well-being and the lottery rankings it allows is one test the deliberator can use to evaluate measures. The anti-paternalist deliberator might think: I want a ranking of the well-being levels and differences of histories that coheres with an anti-paternalist ranking of history lotteries.