Contents lists available at SciVerse ScienceDirect

Journal of Economic Behavior & Organization

Fairness, risk preferences and independence: Impossibility theorems☆

Drew Fudenberg [a],[*], David K. Levine [b]

[a] *Department of Economics, Harvard University, 1805 Cambridge St, Cambridge, MA 02138, United States*
[b] *Department of Economics, Washington University in St. Louis, St Louis Federal Reserve Bank and European University Institute, United States*

ARTICLE INFO

ABSTRACT

The most widely used economic models of social preferences are specified only for certain outcomes. There are two obvious methods of extending them to lotteries. If we do so by expected utility theory, so that the independence axiom is satisfied, our results imply that the resulting preferences do not exhibit *ex ante* fairness. If we do so by replacing certain outcomes with their expected utilities for each individual, so that individual risk preferences are preserved, then *ex ante* fairness may be preserved, but neither the independence axiom nor *ex post* fairness is satisfied. Both ex ante and ex post fairness can be satisfied but then the individual does not have well defined preferences over own lotteries.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

If people have both a preference for fairness and continuous preferences, they are willing to make a small sacrifice for a more egalitarian outcome. One way to create "fairness" in an *ex ante* sense is to flip a coin and reverse the roles. The use of lotteries to allocate indivisible rewards and costs (such as the draft lottery) is evidence that *ex ante* fairness is often a concern. There is also considerable experimental evidence that agents care about *ex ante* fairness, see for example, Bolton et al. (2005), Krawcyck and Le Lec (2006), Bolton and Ockenfels (2009) and Kircher et al. (2009). We show that there is a conflict between *ex ante* fairness and the independence axiom. In particular, we point out that five leading theories of outcome-based social preferences for fairness, those of Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Cox and Sadirij (2004), and Andreoni and Miller (2002) all fail to reflect concerns for *ex ante* fairness when they are extended to lotteries using expected utility theory as is implicit in some of this work.

One contribution of this paper is the formulation of very weak notions of *ex ante* fairness that are clearly inconsistent with the independence axiom. Our results are a formalization and generalization of examples and results from the social choice literature, starting from the Diamond (1967), Machina (1989) example of a parent or social planner who strictly prefers to use a coin flip to allocate an indivisible good to two (other) agents and thus violates the axioms of Harsanyi (1955). The social choice literature, which adopts Harsanyi's "impartial-observer" viewpoint, has responded to this example by relaxing independence in various ways.[1] The conflict between independence and *ex ante* fairness has also been noted in the

---

  [1] For example Karni and Safra (2002) make use of only a partial version of independence in their study of the use of lotteries to solve indivisibility problems. Grant et al. (2010) also recognize this when they reconcile the Diamond paradox with Harsanyi's social choice theory: they do so by means of

behavioral literature, for example Kircher et al. (2009) give a verbal argument indicating that *ex ante* fairness is inconsistent with utilitarianism, and implicitly with the independence axiom.

We examine formal and fairly weak implications of fairness, using only the domain of "coin flip" lotteries, namely binary lotteries where each outcome has probability 1/2.[2] We define "fairness for you" as a willingness to sacrifice my payoff for your benefit and "fairness for me" as a willingness to reduce your utility to achieve higher utility for me. Our first point is that if preferences satisfy either of these conditions then the independence axiom must be violated. As a consequence, if social preferences over deterministic allocations are extended to lotteries by treating the associated utility functions as expected utility, *ex ante* fairness is violated both for you and for me.

We also show that if preferences over lotteries are transitive and state-independent, and the induced preferences over own lotteries are "individualistic" in the sense of not depending on other people's consumption, then a weak form of *ex post* fairness is necessarily violated. This is true, for example, for Fehr–Schmidt preferences extended to lotteries by replacing money income with expected money income. More generally, preferences that depend only on the expected value of individual utilities, such as those in Grant et al. (2010), may be *ex ante* fair, but cannot be *ex post* fair. Moreover, the "obvious" extensions of standard models of fairness and social preference used in experimental research also fail to incorporate both *ex ante* and *ex post* fairness, though there are however relatively straightforward variants of these models that do reflect both concerns; of course these variants necessarily violate the independence property as well as the individualistic risk preference property. The same is true of Borah's (2011) representation of preferences reflecting a trade-off between ex-ante and ex-post fairness. We conclude that experimental research on social preferences should pay more attention to preferences over lotteries, and that decision theorists should then consider what classes of tractable preferences are broadly consistent with the resulting data.

2. Ex ante fairness for you and for me

There are two players who we refer to as "me" and "you." We consider *certain outcomes* $(m, y) \in \Re_+^2$ that can be interpreted as money for me and money for you. We also consider simple lotteries generated by tossing a fair coin with equal 50% probability of H(eads) or T(ails). We call this a *coin flip* and it can be written as $((m^H, y^H), (m^T, y^T)) \in \Re_+^2 \times \Re_+^2$.

We are interested in "my" preferences over certain outcomes and coin flips given as a complete order $\succsim$ over $\Re_+^2 \cup \Re_+^4$, with the derived strict ordering $\succ$ and indifference relation $\sim$.[3] We develop several properties that such preferences might have, and investigate their relationship with each other and with various forms of social preference in the literature.

Property 1 (Independence). *If* $(m_1, y_1) \succsim (m_2, y_2)$ *then for any* (m^H, y^H) *we have* $((m^H, y^H), (m_1, y_1)) \succsim ((m^H, y^H), (m_2, y_2))$.

Notice that since the role of the two certain prospects may be reversed, the same holds for the indifference relationship. Any expected utility theory must satisfy Property 1, and some non-expected utility theories satisfy the axiom as well. It is weaker than the usual independence axiom, in the sense that it needs to hold only for fair coin flips; in particular since prospect theory models do not distort the probabilities {0, 1/2, 1} they satisfy this axiom.[4] On the other hand at an intuitive level it is clear that the property conflicts with the idea that people might prefer lotteries that are ex ante fair.

The most immediate conflict between the independence axiom and fairness comes from an *ex ante* version of fairness, so we start from that. Below we show there is also a contradiction between independence and *ex post* notions of fairness under the standard continuity assumption.

To motivate our definition of *ex ante* fairness, suppose the agent weakly prefers (8, 5) to the more egalitarian (7, 7), and strictly prefer the fairer coin-flip lottery ((1000, 0), (7, 7)) to the less fair ((1000, 0), (8, 5)). Such preferences satisfy the following condition, regardless of preferences over other lotteries.

Property 2 (Ex Ante Fairness for You). *There is a* $y_2 > y_1, m_2 > m_1$[5] *so that* $(m_1, y_2) \precsim (m_2, y_1)$ *and an* $m^H > m_2, y^H < y_1$ *such that* $((m^H, y^H), (m_1, y_2)) \succ ((m^H, y^H), (m_2, y_1))$.

This says that there is a sacrifice to make you better off that I would not make when comparing two deterministic outcomes, but I would make the sacrifice in the context of a coin flip if the other outcome is sufficiently less fair for you.[6]

---

[2] Note that the domain we study does not allow preferences to depend on past actions, and that "fairness" here refers to fair allocations and not, for example, reciprocity.

[3] It seems natural to assume that the order over certain outcomes is consistent with the order over coin flips, but we do not assume this as it is not required for our results.

[4] Herstein and Milnor (1953) derive the usual independence axiom from this axiom and a continuity assumption. If we replace weak preference with strict preference, this is what Machina (1989) refers to as "mixture separability."

[5] In this and the next property we use larger subscripts as a reminder of larger values.

[6] Suppose $m_2 = 1, y_1 = 4, m^H = 2, y^H = 3$. Then the outcome (2, 3) may seem more fair than (1, 4). However it might be "less fair" for you, even though it is farther from an equal division. Equal division *is* identified with fairness in many economic models, but in general depends on the units in which things are measured, and one can imagine situations where I think it is fair that you get four times as much as I do. Here we take the approach of being agnostic about what division is "fair." However, we could add to Property 1 the restriction that $m_2 > y_1$ (to reflecting the notion that fairness means equal division) without any substantive change in our results.

The paragraph at the top of the footnote area reads: a social welfare function that is a non-linear function of the expected utility of different individuals, and so violates the independence axiom. Grant et al. (2011) show provide additional conditions that restrict these preferences to utilitarianism.

It is easy to see that there are no preferences satisfying Properties 1 and 2, as Property 1 requires that there are no lotteries for which the conditions in Property 2 hold.

In particular our example of *ex ante* fairness violates the independence property.

Ex ante fairness for me is just the flip side of *ex ante* fairness for you.

Property 3 (Ex Ante Fairness for Me). *There is a $y_2 > y_1$, $m_2 > m_1$ so that $(m_1, y_2) \succsim (m_2, y_1)$ and an $m^H < m_1$, $y^H > y_2$ such that $((m^H, y^T), (m_2, y_1)) \succ ((m^H, y^T), (m_1, y_2))$.*

This says that there is an opportunity to enrich myself at your expense I would not take, but I am concerned about *ex ante* fairness for me in the sense that I would exploit it in the context of a coin flip if the other outcome is "sufficiently less fair" for me.

As with Property 2, it is obvious that Property 3 conflicts with Property 1.

3. Economic models of social preference

We now discuss four different social preferences from the literature that reflect a concern for fairness.[7] Each of them describes choices under certainty. One method of extending them to uncertainty is to treat them as expected utility functions, and evaluate lotteries by their expected value; in this case Properties 2 and 3 must be violated. As we will see, alternative extensions to uncertainty, such as replacing lotteries over income with the expected value of income, run in to different problems.

3.1. Fehr and Schmidt

The Fehr and Schmidt (1999) social preferences, in our notation, are given by $U(m, y) = m - \alpha \max\{y - m, 0\} - \beta \max\{m - y, 0\}$ with $0 \leq \beta < \alpha$. That is, if you are getting more than me I dislike it, and if I am getting more than you, I also dislike that, although not as much as I dislike you getting more than me. Although Fehr and Schmidt do not explicitly say this, it is implicit from their discussion and analysis that this is an expected utility function, so the independence property (Property 1) is satisfied. Thus this version of their preferences fails both Property 2 and Property 3. Neilson (2006) provides an axiomatization of Fehr Schmidt preferences under certainty.

Now consider the alternative of extending the preferences by replacing income with expected income. We will show that if $\alpha > \beta \geq 0$ that Properties 2 and 3 are satisfied, while of course the independence property is not. Let $m_1 = \alpha$, $y_2 = \alpha + 1$, and for any γ let $m_2 = \gamma\beta$, $y_1 = \gamma(\beta - 1)$, so that $(m_1, y_2) \sim (m_2, y_1)$ since both yield zero utility. Define $c(x) = \alpha \max(x, 0) + \beta \min(x, 0)$. Then

$$\Delta(m^H, y^H) \equiv 2U((m^H, y^H), (m_1, y_2)) - 2U((m^H, y^H), (m_2, y_1)) = m^H + m_1 - c(m^H + m_1 - y^H - y_2)$$

$$- (m^H + m_1 - c(m^H + m_2 - (y^H + y_2))) = \alpha - \gamma\beta + c(y^H - m^H - \gamma) - c(y^H - m^H + 1)$$

If $y^H - m^H < -1$, with $m^H > m_2$, $y^H < y_1$

then

$$\Delta = \beta(y^H - m^H - \gamma) - \beta(y^H - m^H + 1) + \alpha - \gamma\beta = \alpha - \beta - 2\gamma\beta$$

For $\alpha > \beta$ we can choose $\gamma < (\alpha - \beta)/2\beta$ so that $\Delta > 0$, so Property 2 is satisfied. Notice if $\beta = 0$ we have $\Delta > 0$ for any γ. Intuitively, since I care about fairness for me a coin flip that is unfair to you lets me ignore the fact that in the alternative outcome you get more than me. Notice that $m_2 > y_1$ and $m_1 < y_2$ so that even if we added these constraints to the property, it would still be satisfied by these Fehr–Schmidt preferences.

To show that Property 3 is satisfied let $y^H - m^H$ be greater than β, with $m^H < m_1$, $y^H > y_2$. Then

$$\Delta = \alpha(y^H - m^H - \gamma) - \alpha(y^H - m^H + 1) + \alpha - \gamma\beta = -\gamma(\alpha + \beta)$$

so $\Delta < 0$.

Notice that the results concerning Properties 2 and 3 are not symmetric, and in particular Property 3 is satisfied even if $\beta > \alpha$. The reason for this is that the Fehr and Schmidt utility function does not treat you and me symmetrically so that "fairness to me" may be satisfied even if the agent is unconcerned about "fairness to you".

3.2. Bolton and Ockenfels

The Bolton and Ockenfels (2000) preferences are given by $U(m, y) = v(m, m/(m + y))$ where v is twice differentiable, increasing and concave in the first argument, and concave with a maximum at 1/2 in the second argument.

---

[7] Cox et al. (2008) is somewhat related but considers axioms for "being more altruistic than" as opposed to preferences for fairness and does not consider the role of lotteries.

Here my utility depends on my share of the total as well as the amount of money I receive. Notice that is closely related to a variation of Fehr and Schmidt preferences in which differences are measured relative to the total[8]

$$U(m, y) = m - \alpha \max \left\{ \frac{(y - m)}{(y + m)}, 0 \right\} - \beta \max \left\{ \frac{(m - y)}{(y + m)}, 0 \right\}$$

This differs from Bolton and Ockenfels only in that U is not differentiable.[9] As is the case with Fehr and Schmidt if we use the independence axiom to extend these preferences to uncertainty then Properties 2 and 3 are violated.

Alternatively we can either use expected money income for both parties, or we can use expected money income for me and my expected share. In either case the independence property will be violated. Indeed, under the assumptions of Bolton and Ockenfels, we may find $y_1 < 1 < y_2$ so that $v(1, 1/(1 + y_1)) = v(1, 1/(1 + y_2))$. Moreover, if $s \notin \{1/(1 + y_1), 1/(1 + y_2)\}$ then $v(1, s) \neq v(1, 1/(1 + y_1))$. Independence then implies the coin flip between $(1, y_1)$ and $(1, y_2)$ must be indifferent to both of the outcomes of the coin flip. The utility of the coin flip is $v(1, s)$, and since neither the share of expected income nor the expected share takes on the values $\{1/(1 + y_1), 1/(1 + y_2)\}$, the coin flip is not indifferent two the certain outcomes.

Both the Fehr and Schmidt and the Bolton and Ockenfels preferences exhibit spite or egalitarianism in the sense that a Pareto inferior allocation may be preferred if it is fairer. In particular, for both preferences, my utility decreases in your income when $y > m$, while when $y < m$ my utility increases in your income. The former case implies a willingness for me to pay to reduce your income. This is one possible notion of fairness, but not one required in our properties.

The remaining preferences we discuss are monotone and not egalitarian.

3.3. Charness and Rabin

The Charness and Rabin (2002) preferences are given in our notation by $U(m, y) = (1 - \gamma)m + \gamma(\delta \min(m, y)) + (1 - \delta)(m + y))$, where $0 \leq \gamma$, $\delta \leq 1$. That is to say, they are a weighted average of my money income, the least income either of us have, and the social total. It is the dependence on the least income of either of us that gives rise to a concern for fairness. Naturally if we extend these preferences to uncertainty using the independence axiom Properties 2 and 3 must fail. On the other hand the extension using expected money payoffs violate independence since $\min(E\tilde{m}, E\tilde{y})$ obviously does so.

3.4. Cox and Sadirij

Cox and Sadirij (2004) specify preferences that depend on who is getting more in a non-linear way. In our notation, their preferences are given by

$$U(m, y) = (1 - \theta_- 1(m < y) - \theta_+ 1(m \geq y))m^\alpha + (\theta_- 1(m < y) + \theta_+ 1(m \geq y))y^\alpha$$

where $0 < \alpha < 1$, $0 \leq \theta_+ < 1$, $0 \leq \theta_- \leq \theta_+$, $\theta_- \leq 1 - \theta_+$. The interpretation is that the weights on m^α, y^α depend on how fair the allocation is. Because of decreasing marginal utility in m, y, it is natural to interpret this as an expected utility function, in which case Properties 2 and 3 are violated.

3.5. Andreoni and Miller

Andreoni and Miller (2002) consider preferences over m, y of the form $U(m, y) = \text{sgn}(\alpha)[m^\alpha + \delta y^\alpha]$, $\alpha \leq 1$. As in the case of Cox and Sadirij, it is difficult to interpret this other than as a Von Neumann–Morgenstern utility function so Properties 2 and 3 are violated. Notice in the limiting Leontief case when $\delta = 1$ this model is a special case of Charness and Rabin.

Andreoni and Miller preferences are a special case of preferences $U(m, y)$ that are strictly concave and strictly increasing in both arguments. Since they reflect a concern for the welfare of the other player they are clearly altruistic. Assuming also symmetric utility for m and y implies an equal division will always be preferred when 1–1 transfers are available. One interpretation is that this reflects a concern for fairness. Alternatively we could argue that concern for fairness must involve spite as is the case with Fehr and Schmidt or Bolton and Ockenfels – but this rules out many preferences such as those of Charness and Rabin that some might feel exhibit fairness.[10] In the case of lotteries a much clearer distinction can be drawn between altruism and fairness. For example, consider Andreoni and Miller preferences $U(m, y) = m^{1/2} + y^{1/2}$. One interpretation is that they exhibit concern for altruism and not fairness. Of course the expected utility extension does not

---

[8] It is not clear that measuring money relative to the total is desirable since it is not clear what the base amounts are supposed to be. For example if the amounts are receipts in a laboratory experiment, they may sum to zero or even negative.

[9] There is a greater difference between these two preferences when applied to settings where the number of other people the agent cares about varies. Engelmann (in press) points out the pitfalls in extending these preferences to add a concern for total surplus.

[10] Yet a third interpretation of fairness might be that I care about the minimum of our two incomes as is the case with Fehr and Schmidt and Charness and Rabin. That rules out Andreoni and Miller preferences – yet Leontief preferences are a limiting case of their preferences.

exhibit *ex ante* fairness in the sense that it does not satisfy Properties 2 and 3. However the extension to lotteries defined by $U(\tilde{m}, \tilde{y}) = (E\tilde{m})^{1/2} + (E\tilde{y})^{1/2}$ does.[11]

4. *Ex post* fairness and risk preference

The models discussed in the previous section specify social preferences only over certain outcomes and are representable by means of a utility function. If we extend preferences to lotteries by taking the expected utility the independence axiom is necessarily satisfied, and so very weak notions of *ex ante* fairness for both you and me must fail. As we noted, an alternative procedure for extending preferences to lotteries is to replace certain income with its expected value. This allows for *ex ante* fairness, while the expected-utility extension does not. However, this method of relaxing independence does not allow for *ex post* fairness, and more broadly any preferences that have well defined risk preferences for me do not allow *ex post* fairness.

In particular, consider

Property 4 (Ex Post Fairness). *There exist $m_1 < m_2$ and $y_1 < y_2$ such that* $((m_1, y_1), (m_2, y_2)) \succ ((m_2, y_1), (m_1, y_2))$.

This says that for some lottery the more egalitarian coin flip where we are both better off at the same time is preferred to the less egalitarian coin flip where our fates are opposite. For example, given a choice between $(1, 1, 0, 0)$ and $(1, 0, 0, 1)$ many people might prefer $(1, 1, 0, 0)$. This captures the common observation that "misery likes company" as well as ideas of status competition and relative consumption.[12]

Taking, say, the Fehr–Schmidt preferences, and extending them to lotteries using expected income, allows for ex-post fairness but has the perhaps undesirable implication that agents are risk neutral; we next examine the extent to which a concern for *ex post* fairness is consistent with (a) well defined notions of preferences over lotteries that allow for risk aversion and (b) a concern for *ex ante* fairness.

Property 5 (Correlation Invariance). *If* $((m_1, y_1), (m_2, y_2)) \succ ((m_3, y_1), (m_4, y_2))$ *then* $((m_1, y_2), (m_2, y_1)) \succ ((m_3, y_2), (m_4, y_1))$.

This property says that if we hold fixed the two possible outcomes for you, then my preferences over lotteries on my own consumption are independent of the states in which you receive each of the outcomes, so that my preferences ignore the correlation between my consumption and yours. However, Property 5 does not require that preferences over my consumption are the same regardless of the level of your consumption, i.e. it allows for preferences such as $((3, 1), (1, 1)) \succ ((1, 1), (1, 1))$ and $((3, 0), (1, 0)) \prec ((1, 0), (1, 0))$. The following stronger condition rules this out and also implies Property 5:

Property 5' (Weakly Individualist Risk Preference). *If* $((m_1, y_1), (m_2, y_2)) \succ ((m_3, y_1), (m_4, y_2))$ *then for all y_3, y_4,* $((m_1, y_3), (m_2, y_4)) \succ ((m_3, y_3), (m_4, y_4))$.

This property is satisfied by the utility representation obtained by Grant et al. (2010) for an impartial observer:

$$U(\tilde{m}, \tilde{y}) = v_m(Eu_m(\tilde{m})) + v_y(Eu_y(\tilde{y}))$$

where v_m and v_y are both continuous and increasing. This utility function allows for a form of *ex ante* fairness, as when the v's are concave, I am more willing to give you utility when I am better off. However, it is not consistent with a descriptive model of social preferences, as it rules out any concern for *ex post* fairness.

More generally, any preferences with a utility representation of the form $U(\tilde{m}, \tilde{y}) = V(Eu_m(\tilde{m}), Eu_y(\tilde{y}))$, with V increasing in its first argument, satisfy Property 5': additive separability is not necessary. Even more generally any utility function of the form $U(\tilde{m}, \tilde{y}) = W(V(\tilde{m}), \tilde{y})$ where V is scalar-valued satisfies this property: we conjecture that if Property 5' is strengthened to hold on a wider domain along with the standard assumptions of continuity that give rise to a utility representation, then such a representation is always possible. Conversely, when Property 5' fails, an agent whose choices only alter the distribution of own outcome still needs to consider the allocations of others in making his decision.

Notice, moreover, preferences can satisfy both Property 5' (and *a fortiori* Property 5) and also *ex ante* fairness for both you and me, and in particular the Grant et al. preferences can have this property.[13] In addition, if the Fehr and Schmidt utility function is extended to uncertainty by taking the expected value of income, then it satisfies Property 5'. However, under a very mild state-independence condition, any preferences that satisfy Property 5 violate ex-post fairness. Suppose the agent prefers $((1,1),(0,0))$ to $((1,0),(0,1))$; then Property 5 implies the agent also prefers $((1,0),(0,1))$ to $((0,0),(1,1))$ so that the agent cares about whether the high consumption occurs on H or T.

---

[11] This is shown in footnote 12.

[12] Harel et al. (2005), Fleurbaey (2010), and Grant et al. (2010) consider notions of ex-post fairness for an outside observer in settings where preferences depend only on the expected individual utilities; their conditions are assumed to hold at "many" lotteries as opposed to our condition which must hold at least one.

[13] This is true for example if individual utility functions are $u_m(m) = m$, $u_y(y) = y$, with v_m strictly increasing, $v'_m(M) \to 0$ as $M \to \infty$ and v_y strictly increasing. To see why Property 2 is satisfied, the assumptions guarantee the existence of an $(X, Y) \sim (1, 0)$ with $X < 1$ and $Y > 0$. Consider the coin flips $A = ((1, 0), (M, -1)$ and $B = ((X, Y), (M, -1)$. Compute $U(A) = v_m(.5 + .5M) + v_y(-.5)$, $U(B) = v_m(.5(M + X)) + v_y(.5Y - .5)$. Since $v'_m(M) \to 0$ as $M \to \infty$ also $v_m(.5 + .5M) - v_m(.5(M + X)) \to 0$. On the other hand while $v_y(.5Y - .5) - v_y(-.5) > 0$. Thus when M is large enough B is strictly preferred. For Property 3, consider $(X, Y) \sim (1, 0)$ with $X > 1$ and $Y < 0$. Then as $M \to \infty$ we again have $v_m(.5 + .5M) - v_m(.5(M + X)) \to 0$ and now $v_y(.5Y - .5) - v_y(-.5) < 0$ so A is strictly preferred.

Property 6 (State Independence). $((m_1, y_1), (m_2, y_2)) \sim ((m_2, y_2), (m_1, y_1))$.

Proposition.   *There are no transitive preferences satisfying* Properties 4, 5 and 6.

Proof.   By Property 4, let $((m_1, y_1), (m_2, y_2)) \succ ((m_2, y_1), (m_1, y_2))$. Property 5 implies $((m_1, y_2), (m_2, y_1)) \succ ((m_2, y_2), (m_1, y_1))$. But $((m_2, y_2), (m_1, y_1)) \sim ((m_1, y_1), (m_2, y_2))$ and $((m_1, y_2), (m_2, y_1)) \sim ((m_2, y_1), (m_1, y_2))$ by Property 6. We conclude that $((m_2, y_1), (m_1, y_2)) \succ ((m_1, y_1), (m_2, y_2))$, which contradicts the irreflexive property of strict preference.

Roughly speaking the situation is this. Given preferences for fairness under certainty $U(m, y)$, their extension to lotteries by taking expected utility $Eu(\tilde{m}, \tilde{y})$ must violate *ex ante* fairness for both me and you. If instead we extend the preferences to lotteries by taking the expected value of individual income $u(E\tilde{m}, E\tilde{y})$ or taking the expected value of any function of individual income, we violate *ex post* fairness. We can view the former extension as exhibiting a preference for *ex post* fairness only and the second as exhibiting a preference for *ex ante* fairness only. By combining the two, we may easily get preferences for both *ex post* and *ex ante* fairness, though if Property 6 is satisfied Property 5 must be violated.

Consider, for example, defining $U(m, y)$ to be the Fehr and Schmidt functional form, and extend these to uncertainty using $U(\tilde{m}, \tilde{y}) \equiv \gamma EU(\tilde{m}, \tilde{y}) + (1 - \gamma)U(E\tilde{m}, E\tilde{y})$. (Notice that this reduces to the original utility function for deterministic outcomes.) Recall that Properties 2 and 3 involve strict preference and we already showed $U(E\tilde{m}, E\tilde{y})$ satisfies Properties 2 and 3 for $\alpha > 0$. It follows that for γ sufficiently small, $U(\tilde{m}, \tilde{y})$ must also satisfy Properties 2 and 3. Property 6 is obviously satisfied, so we shall show that Property 4 is satisfied even when γ is small. Consider in particular $m_1 < m_2$ and $y_1 < y_2$ where $m_1 = y_1$, $m_2 = y_2$. Then

$$U((m_1, y_1), (m_2, y_2)) - U((m_2, y_1), (m_1, y_2)) = \left(\frac{\gamma}{2}\right)(U(m_1, m_1) + U(m_2, m_2) - U(m_2, m_1) - U(m_1, m_2))$$

$$= \left(\frac{\gamma}{2}\right)(\alpha + \beta)(m_2 - m_1) > 0$$

so that Property 4 is indeed satisfied provided that $\gamma > 0$ and $\alpha > 0$.

Borah (2011) provides an axiomatic characterization of a more complicated form of preferences that allow ex-ante and ex-post fairness: He proposes that the utility of a lottery p on (m, y) is the expected value of $(1 - \sigma)w(m, y) + \sigma w(m, p_y)$ where σ is a constant, p_y is the marginal distribution of the lottery on y, and w is a function on $M \times P_y$ that need not be linear in p_y.

5. Conclusion

The independence axiom is inconsistent with even very weak notions of *ex ante* fairness. Existing models of fairness do not focus on the role of lotteries, and the preferences analyzed in Fehr and Schmidt (1999), Bolton and Ockenfels (2009), Charness and Rabin (2002), Cox and Sadirij (2004) and Andreoni and Miller (2002) are defined for certain outcomes, without specifying how they are to be extended to lotteries. There are two issues this raises. First, it is not easy to distinguish fairness from altruism in the case of purely deterministic outcomes. Second, there are two obvious ways of extending preferences to lotteries.

One way to extend preferences from certain outcomes to lotteries is the standard one of treating the certainty utility as a Von Neumann–Morgenstern utility function. In this case the independence axiom is satisfied, but then *ex ante* fairness for both me and you are violated. Alternatively preferences can be extended by replacing the certain individual utility with its expected value, in which case the independence axiom is violated. In the later case, for example, in the social welfare theory of Grant et al. (2010), not only is independence violated, but *ex post* fairness is ignored. Hence the standard models of fairness and social preference used in experimental research and their obvious extensions to lotteries do not incorporate both *ex ante* and *ex post* fairness. As we have seen, it is relatively straightforward to construct preferences that do satisfy both conditions, in part because we deliberately formulated very weak notions of fairness to make the impossibility results more sharp.

References

Andreoni, J., Miller, J., 2002. Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.
Bolton, G., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. American Economic Review 90, 166–193.
Bolton, G., Brandts, J., Ockenfels, A., 2005. Fair procedures: evidence from games involving lotteries. Economic Journal 115, 1054–1076.
Bolton, G., Ockenfels, A., 2009. A Comment on 'Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. Mimeo.
Borah, A., 2011. Other-Regarding Preferences and Concerns for Procedure. Mimeo.
Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics 117, 817–869.
Cox, J., Sadirij, V., 2004. Direct Tests of Social Preferences and a New Model. Mimeo University of Arizona.
Cox, J.C., Friedman, D., Sadiraj, V., 2008. Revealed altruism. Econometrica 76 (1), 31–69.
Diamond, P., 1967. Cardinal welfare, individualistic ethics, and intertemporal comparisons of utility: comment. Journal of Political Economy 75, 765–766.
Engelmann, D., 2012. How not to extend models of inequality aversion. Journal of Economic Behavior and Organization 81, 599–605.
Fehr, E., Schmidt, K., 1999. A theory of fairness, competition and cooperation. Quarterly Journal of Economics 114, 817–868.
Fleurbaey, M., 2010. Assessing risky social situations. Journal of Political Economy 118, 649–680.
Grant, S., Kajii, A., Polak, B., Safra, Z., 2010. Generalized utilitarianism and Harsanyi's impartial observer. Econometrica 78, 1939–1971.

Grant, S., Kajii, A., Polak, B., Safra, Z., 2011. Equally-distributed equivalent utility: ex post egalitarianism and utilitarianism. Journal of Economic Theory, Forthcoming.

Harel, A., Safra, Z., Segal, U., 2005. Ex-post egalitarianism and legal justice. The Journal of Law Economics and Organization 21, 57–75.

Harsanyi, J., 1955. Cardinal welfare, individualistic ethics, and interpersonal comparison of utility. Journal of Political Economy 63, 309–321.

Herstein, I.N., Milnor, J., 1953. An axiomatic approach to utility. Econometrica 2, 291–297.

Karni, E., Safra, Z., 2002. Individual sense of justice: a utility representation. Econometrica 70, 263–284.

Kircher, P., Ludwig, S., Sandroni, A., 2009. Fairness: A Critique to the Utilitarian Approach. Mimeo.

Krawcyck, M., Le Lec, F., 2006. Social Decisions Under Risk: Evidence from the Probabilistic Dictator Game. Mimeo.

Machina, N., 1989. Dynamic consistency and non-expected utility models of choice under uncertainty. Journal of Economic Literature 27, 1622–1668.

Neilson, W., 2006. Axiomatic reference-dependent behavior toward others and toward risk. Economic Theory 28, 681–692.