

Peter Vallentyne

Standard Decision Theory Corrected:

Assessing Options When Probability is Infinitely and Uniformly Spread*

Synthese 122 (2000): 261-290.

Abstract: Where there are infinitely many possible [equiprobable] basic states of the world, a standard probability function must assign zero probability to each state—since any finite probability would sum to over one. This generates problems for any decision theory that appeals to expected utility or related notions. For it leads to the view that a situation in which one wins a million dollars if any of a thousand of the equally probable states is realized has an expected value of zero (since each such state has probability zero). But such a situation dominates the situation in which one wins nothing no matter what (which also has an expected value of zero), and so surely is more desirable. I formulate and defend some principles for evaluating options where standard probability functions cannot strictly represent probability—and in particular for where there is an infinitely spread, uniform distribution of probability. The principles appeal to standard probability functions, but overcome at least some of their limitations in such cases.

1. Introduction

Where there are infinitely many possible basic states of the world, a standard probability function can represent a uniform distribution of probability over these states only if it assigns a probability of zero to each state. For if it assigned a positive probability to each state, the total probability would add to more than one. This generates problems for assessing the value of options.

Suppose, for example that options are assessed on the basis of their expected utility. If each state has a probability of zero, then the expected utility of an option that has a payoff of 1000 under

one state, and 0 under all other states, has the same expected utility (namely, zero) as an option that has a payoff of only 1 under that one state, and 0 under all other states. But surely an option that dominates another is more valuable. So there's a problem somewhere.

One response is to deny the possibility of a uniform distribution of probability over an infinite number of states. I argue below that this response is mistaken. Standard (i.e., real-valued) probability functions are not the only ways of representing probability. Both comparative probability rankings and non-standard probability functions (involving infinitesimals) are capable of strictly representing an infinite uniform distribution of probability.

I formulate and defend some principles for evaluating options where standard probability functions cannot strictly represent probability—and in particular for where there is an infinitely spread, uniform distribution of probability. The principles appeal to standard probability functions, but overcome at least some of their limitations in such cases. I defend the principles both in the abstract and because they generate plausible judgements that are endorsed by sound principles based on comparative probability and by sound principles based on non-standard probability functions.

It should be emphasized that the focus of this paper is on "very small" differences. The standard approach yields judgements that are infinitesimally close to the correct judgements. But infinitesimal errors are still errors, and I'm concerned here with exactitude. Readers who are satisfied with mere approximations need proceed no further.

2. The Problem

To help fix our ideas, I shall focus on a lottery in which a positive natural number is randomly

selected from among the infinite number of possibilities, and anyone holding a ticket for the number selected wins a finitely highly valuable prize.¹ For simplicity, throughout I consider only cases with denumerably infinitely many possibilities. The selection process is random in the sense that any two positive integers are equally probable. We shall consider various sets of lottery tickets, and ask whether one set is more valuable than another.

For reasons that will become apparent below, there are two basic cases that we shall consider, depending on whether one set is, loosely speaking, larger than the other. More specifically, I shall appeal to an extended notion of set size defined as follows. One set is larger* than a second just in case the number of elements in the first but not in the second set—its non-common elements—is larger than the number of elements in the second but not in the first. For any two sets, one of them is larger* than the other, except where (1) each set is finite with the same number of elements (and thus the same number of non-common elements), or (2) each set is infinite and each has an infinite number of non-common elements. The following are examples of this notion, and of the cases that we shall consider (where all and only positive integers are possible numbers).

A. Cases where one set of tickets (the second set) is larger* than the other:

A1. Both sets are finite, and one is larger than the other; for example:

- Ticket number 1; vs.
- Ticket number 2 and ticket number 3.

A2. One set is finite, and one is infinite; for example:

- All the tickets with numbers less than or equal to one million; vs.

- All the tickets with numbers that are multiples of 3.

A3. Both sets are infinite and one is a proper subset of the other; for example:

- All the tickets with numbers that are multiples of 6; vs.
- All the tickets with numbers that are multiples of 3.

A4. Both sets are infinite, neither is a proper subset of the other, but one has (finitely or infinitely) more non-common elements than the other; for example:

- Ticket number 1 plus all tickets that are multiples of 4; vs.
- Tickets numbered 2 and 3, plus all tickets that are multiples of 4.

A second example of this case (involving infinitely more non-common elements) is:

- Ticket number 1 plus all tickets that are multiples of 4; vs.
- All tickets that are multiples of 3 or of 4.

B. Cases where neither set is larger* than the other; for example:

B1. Both sets are finite (and therefore the same size):

- Ticket number 1 and ticket number 2; vs.
- Ticket number 3 and ticket number 4.

B2. Both sets are infinite and identical:

- All the tickets with numbers that are multiples of 19; vs.
- All the tickets with numbers that are multiples of 19.

B3. Both sets are infinite, they are not identical, but have the same number of non-common elements [SMALL CORRECTION MADE]:

- All the tickets with numbers that are multiples of 19; vs.

- All the tickets with numbers that are multiples of 3.

Where probability is uniformly distributed, larger* sets of lottery tickets are, I shall argue, more valuable than smaller* sets. Where neither set is larger* than the other, and probability is uniformly distributed, there are three cases. Where two sets are both finite (B1), and neither is larger* than the other, then uncontroversially the two sets are the same size, and thus equally valuable when probability is uniformly distributed. Where two sets are both infinite and identical (B2), then the two sets are, of course, the same size, and equally valuable. Where two sets are both infinite, not identical, and neither is larger* than the other (B3), then, I shall show, the mere fact that probability is uniformly distributed does not determine which is more valuable.

When there is an infinitely spread uniform distribution of probability, the standard ways of evaluating options (sets of tickets) based on standard probability functions fail, I shall now show, to give the correct answer in cases where one set is larger* than another. For concreteness, I shall assume throughout that in finite cases (e.g., with a finite number of possible numbers) options are correctly evaluated on the basis of their expected utility, and show that expected utility evaluations fail to give the right answers in the above cases. Similar points could be made about many other standard ways of evaluating options on the basis of standard probability functions.

So let us see what standard expected utility says about the relative value of two sets of tickets where one set is larger* than the other, and probability is infinitely and uniformly spread. Where both sets are finite (but one is larger* and hence larger, than the other), the standard expected utility approach wrongly says that both sets are equally valuable. For, it says that both

sets have zero probability of winning, and thus equal expected gains (zero). But surely when all numbers are equally probable, a larger finite set is more valuable than a smaller finite set. For example, $\{1,2\}$ is more valuable than $\{1\}$, and also more valuable than $\{4\}$. Furthermore, where one of the sets is infinite and the other is either a finite set or a proper subset (infinite or finite), this approach is inappropriately silent. For, it does not ensure that the probability of an infinite set is greater than that of a finite set, or than that of a proper subset. It allows, for example, that both could have probability zero. But surely, where the elements are equally probable, an infinite set is more valuable than a finite set, and more valuable than a proper subset. For example, $\{3,6,9,12,\dots\}$ [all multiples of 3] is more valuable than $\{1,2,3\}$, and also more valuable than $\{6,12,18,24,\dots\}$ [all multiples of 6]. Finally, the standard approach wrongly claims that two sets are equally valuable when they are both infinite and at least one of them has a finite number of non-common elements. It wrongly claims, for example, that $\{1,2,\dots,1M, 1M+4, 1M+8,\dots\}$ and $\{1M+4, 1M+8, \dots\}$ are equally valuable, since the two sets are, it holds, equally probable (since their non-common elements have probability zero).

So, where there is an infinitely spread uniform distribution of probability, standard expected utility fails to give the correct answers where one set is larger* than another. Some will deny that an infinitely spread uniform distribution is possible. After defending this possibility, I shall formulate some principles that, assuming that in the finite case expected utility correctly evaluates options, judges larger* sets as more valuable in such cases.

3. Standard Probability Functions and Infinitely Spread Uniform Distribution of Probability

I shall first rehearse the familiar argument for why no standard probability function can strictly

represent probability, when probability is uniformly distributed over an infinite number of possible states. Then I shall defend the assumption that such a distribution of probability is possible.

Throughout I shall assume that (for any given decision problem) there is a well-defined set of (basic) states (states of the world, states of nature) such that exactly one of them is realized. Events are sets of states, and are realized whenever one of the members is realized. A uniform distribution of probability is simply a distribution such that the conditional probability has the following two properties: (1) relative to any finite set of states, any two states in that set have the same conditional probability, and (2) if the states have an interval order (e.g., if the states are the selection of a number on a wheel of fortune), then, relative to any finite interval of states, any two interval events [CORRECTION: WHOLLY IN THAT FINITE INTERVAL] having the same length (e.g., the selection of number between 1 and 10 and the selection of a number between 5 and 15) have the same conditional probability. (If a distribution of probability satisfies these two conditions, then, it is easy to show, it also satisfies the two conditions when reformulated in terms of unconditional probability.)

To see the problem with standard probability functions, we need to note that there can be probability judgements without standard probability functions. Probability judgements involve at least a comparative probability ranking of events in terms of one event being at least as probable as another, where this ranking is reflexive (each event is at least as probable as itself) and transitive (if x is at least as probable as y, and y is at least as probable as z, then x is at least as probable as z). Comparative probability need not be complete (either x is at least as probable as y, or vice-versa).² Comparative judgements of probability, relative to a set of background

assumptions, also satisfy, I shall assume, the following conditions:

(1) Non-Trivial Range: (a) The universal event (i.e., the set of all states) is more probable than the empty event (i.e., the empty set), and (b) every event is at least as probable as the empty event, and no more probable than the universal event.

(2) Additivity: B is at least as probable as C iff for any event D having no element in common with B or with C, the union of B and D is at least as probable as the union of C and D.

(3) Regularity: B is equiprobable with the empty event iff B is impossible relative to the background assumptions.

Non-Trivial Range states that the universal event (respectively: the empty event) is non-trivially maximally (respectively: minimally) probable. Additivity states that adding, or removing, the same elements from two events preserves comparative probability, as long as added elements are not already in either event. Regularity says that the only events that are minimally probable are those for which all the elements are impossible relative to the facts. Stated otherwise, any event that is a (genuine) possibility is more probable than the empty event.

Our question is whether a standard probability function can represent comparative probability. A standard (real-valued) probability function is simply a function from events to real numbers inclusively between 0 and 1 such that (1) the probability of the universal event is 1, and (2) if A and B have no elements in common, then the probability of their union is the sum of their respective probabilities. A probability function, p , strictly represents a comparative probability ranking if and only if for any two events, A and B, A is at least as probable as B iff $p(A) \geq p(B)$.

We are finally ready to see that no standard probability function can strictly represent a comparative probability ranking where there is an infinitely spread uniform distribution of probability. For in this case (by assumption) any two states are equally probable. Consequently, a probability function strictly representing the probability ranking would have to assign zero probability to each state (since assigning any finite positive probability would add up to more than one for the entire set of states). But no probability function strictly representing a comparative probability ranking in this case can do this. For (as required by the regularity condition above) only events that are impossible given the facts (i.e., contain only elements that are impossible given the facts) are equiprobable with the empty event (which has zero probability). So no standard probability function strictly represents the comparative probability ranking in this case.

The above argument rests on the usual assumption that probability is defined for every event, where every set of possible states of the world is an event. It is possible for standard probability functions to strictly represent comparative probability if events are understood more sparsely so that not every set of possible states of the world is an event. Very roughly, the idea is to treat as identical sets of possible states that are only "infinitesimally" different.³ Although this is indeed an interesting and important approach, my concern in this paper is precisely with these very small differences, and so throughout we shall assume that every set of world states is an event.

Central to the impossibility result, of course, is the assumption that infinitely many possible states can be equally probable. One response to the above sort of argument is to deny that it is possible for there to be such a uniform probability distribution. But the only motivation

for this move seems to be to ensure that probability judgements are strictly representable by standard probability functions. Standard probability functions are great things, but they aren't capable of representing everything. We can make perfect sense of uniform distributions of probability in terms of comparative probability rankings. A uniform distribution over the positive integers, for example, is a distribution of probability for which any two positive integers are equally probable. If standard probability functions can't capture this feature in certain cases, so much the worse for standard probability functions.

And it's not merely uniform probability distributions that cause the problem. Call a probability distribution over basic states almost uniform just in case there is a positive integer n such that no state is more probable than any set of n distinct states. The intuitive idea is that, although not all states need be equally probable, there is a finite upper limit on how unequal their probabilities can be. No standard probability function can strictly represent an almost uniform distribution of probability when there are an infinite number of possible basic states. For any standard probability function would have to assign zero probability to each possible state (to avoid adding up to more than one), and that violates the regularity condition above (that any possible event is more probable than an impossible one).⁴

Another response to the above argument against the possibility of a standard probability function strictly representing comparative probability rankings in the uniform (or almost uniform) case is to reject the regularity condition in the above characterization of comparative probability.⁵ This condition holds that only events that are impossible given the facts are equiprobable with the empty event. This condition captures the intuition that something that is genuinely possible is more probable than something that is not. Dropping the regularity condition

may be a convenient way of ensuring that standard probability functions can strictly represent comparative probability judgements (by assigning a probability of zero to some possible events), but it is at best a simplifying convenience. Strictly speaking, our comparative probability judgements satisfy this condition. Furthermore, dropping this condition can lead to the judgement that that an option that dominates another (i.e., has a payoff that is always at least as good, and sometimes better) is equally valuable with the latter. For example, for a lottery over the natural numbers where any two natural numbers are equally probable, assigning each number zero probability has the result that holding a ticket for the number 3 is equally valuable (since it has the same expected utility) as holding that ticket plus a million other tickets. Given the compelling plausibility of the view that an option that dominates another is more valuable, this is a compelling objection to assigning zero probability to possible events.

So in cases of infinitely spread uniform probability, no standard utility function can strictly represent comparative probability. And the example given above of a dominating option having the same expected utility as the dominated option shows that in such cases standard expected utility does not strictly represent value. As acknowledged at the beginning of the paper, this problem has little practical import. It arises only where there are an infinite number of possible states of the world, and the error involved is infinitesimally small. It is, I claim, nonetheless important to be clear that in such cases standard probability functions do not strictly represent probability, and expected utility function do not strictly represent value.

Note that I have not challenged the correctness of standard probability function judgements when the judgement is that one set is more probable than another. Nor have I challenged the correctness of expected utility value judgements when the judgement is that one

set is more valuable than another. I have only challenged claims of two sets being equally probable, or equally valuable. Indeed, I agree that a standard probability function, p , can partially represent comparative probability, where this means that if $p(A) > p(B)$, then A is more probable than B (but not necessarily vice versa). And I agree that expected utility can partially represent comparative value, where this means that if the expected utility of A is greater than that of B, then A is more valuable than B (but not necessarily vice versa).

How, then, are options to be evaluated in the infinite uniform probability cases? I shall now formulate some principles for the evaluation of options in such cases.

4. Evaluating Lotteries When Probability is Infinitely and Uniformly Spread

The problem with standard probability functions, we have seen, is that sometimes they assign equal probability to unequally probable events. And the problem with expected utility based on these probability functions is that sometimes it assigns equal value to unequally valuable options. The problem comes from the attempt to represent probability by standard probability functions. I shall now identify some principles that correct for at least some of the errors generated.

The principles I shall introduce appeal to the notion of the restriction of a possibility set (restriction of a lottery) to a given set of states (numbers). Any probabilities involved are conditionalized on the assumption that the genuine possibilities are restricted to the specified set.⁶ Relative to a given restriction, states that are not in the specified set are not possible, and thus have zero probability. For example, the restriction of a lottery with infinitely spread uniform probability to the set $\{1,2,3\}$ is a lottery in which one of the three numbers is selected, and each number has the same probability (1/3).

The rough idea of the principles is that, if for some finite restriction of the lottery, one set of tickets is more valuable than a second set, and that set of tickets is more valuable for all finite superset restrictions (i.e., that contain all the possible states of the initial restriction and more), then the first set is more valuable than the second (tout court). More precisely:

Radical Superset Betterness: If there is a finite set of states (numbers) such that, relative to any restriction of the possibility set (lottery) to any finite superset, one option (set of lottery tickets) is more valuable than a second, then the first option is more valuable than the second.⁷

Note that this principle says nothing about how options are to be evaluated in finite cases. It simply tells how to extend judgements about finite cases to infinite cases. We are assuming throughout that, in cases with a finite number of possibilities, options are evaluated on the basis of expected utility, but Radical Superset Betterness applies to whatever principles of evaluation are plausible for finite cases.

In the remainder of this section, I first show that Radical Superset Betterness generates the intuitively correct answers with respect to our problem cases. I then note an incorrect answer that it gives in a different sort of case, and identify a modification to the principle. In later sections I show that the judgements about our problem cases are endorsed by plausible principles based on comparative probability and by plausible principles based on non-standard probability. I then show how the corrected version of Radical Superset Betterness can be strengthened and discuss the general implications of this approach.

Radical Superset Betterness entails that, for lotteries where probability is (denumerably) infinitely and uniformly spread, a larger* set is more valuable than a smaller* one. To show this we must show that in such cases there is a finite set of numbers such that relative to a lottery restricted to any finite superset, the larger* set is more valuable than the smaller* one. Let n be the number of non-common elements in the smaller* set. The definition of larger* (i.e., the set of non-common elements in one set is larger than the set of non-common elements in the other set) ensures that n is finite, and that the larger* set has at least $n+1$ non-common elements. Take any set that includes at least $n+1$ non-common elements of the larger* set. Any finite superset of this set will have more elements from the larger* set than from the smaller* set (since the smaller set has only n non-common elements). Thus, for our uniform lottery, relative to all restrictions of the lottery to such supersets, the larger* set is has a higher probability of containing the winning number, and thus is more valuable (because it has a higher expected utility).

To make this concrete, let's consider the four cases involving larger* sets that I introduced above. One case is where both sets are finite, and one is larger than the other; for example, $\{1\}$ vs $\{2,3\}$. Here, relative to our lottery restricted to any finite superset of $\{2,3\}$, the larger* set is more probable, and thus more valuable, than the smaller* set. A second case is where one set is finite, and one is infinite; for example $\{1,2, \dots, 1 \text{ million}\}$ vs. $\{3,6,9, \dots\}$ [multiples of 3]. Here, relative to our lottery restricted to any finite superset of one million and one multiples of 3, the larger* set is more probable, and thus more valuable, than the smaller* set. A third case is where both sets are infinite and one is a proper subset of the other; for example, $\{6,12,18,\dots\}$ [multiples of 6] vs. $\{3,6,9, \dots\}$ [multiples of 3]. Here, relative to our lottery restricted to any finite set of numbers containing at least one odd multiple of 3, the larger*

set is more probable, and thus more valuable, than the smaller* set. Finally, the fourth case is where both sets are infinite, neither is a proper subset of the other, but one has more non-common elements than the other; for example, $\{1,4,8,12,\dots\}$ [ticket 1 plus all tickets that are multiples of 4] vs. $\{2,3,4,8,12,\dots\}$ [tickets 2 and 3, and all tickets that are multiples of 4]. Here, relative to our lottery restricted to any finite set of numbers containing 2 and 3, the larger* set is more probable, and thus more valuable, than the smaller* set.

So, assuming that expected utility correctly evaluates options in the finite cases, Radical Superset Betterness yields intuitively plausible judgements in these cases. In later sections I will show that these judgements are correct by showing that they are entailed by plausible principles of evaluation based on comparative probability and based on non-standard probability.

There is, however, a problem with Radical Superset Betterness. For simplicity we have been focusing on a lottery with a single prize, but of course the principle should be able to handle cases where the prize varies with the number selected. Consider a lottery, with uniform probability over the natural numbers, which gives a certain prize if 1 is selected, a prize of $1/2$ that value if 2 is selected, and in general a prize of $(1/2)^{n-1}$ the original value if n is selected. (Thus, the values of the prizes of the natural numbers in their natural order are: 1, $1/2$, $1/4$, $1/8$, etc.) Compare the value of the following two sets of tickets: $\{1\}$ vs. $\{\text{all natural numbers other than } 1\}$. According to Radical Superset Betterness, the former is more valuable. For, relative to the lottery restricted to $\{1\}$, $\{1\}$ has a greater expected utility, and this remains so for all finite superset restrictions. For example, relative to $\{1,2,3\}$, $\{1\}$ has an expected value of $1/3$, and $\{2,3\}$ has an expected value of $1/4$ [$(1/3) \times (1/2) + (1/3) \times (1/4)$]. But, although the expected value of $\{1\}$ is always greater than that of $\{\text{all natural numbers other than } 1\}$, at the limit the two

expected values are the same. For a finite number of possibilities, n , the expected value of $\{1\}$ is $1/n$, and the expected value of the $\{\text{all natural numbers other than } 1\}$ is $[1-(1/2)^{n-1}]/n$. The difference of these two values is $(1/2)^{n-1}/n$, and as n goes to infinity, this diminishes very quickly. Radical Superset Betterness gets this case wrong.

Radical Superset Betterness needs to be weakened so as to be silent in cases where the two expected values are strictly mathematically equal. Consider, then, the following weaker principle:

Superset Betterness: If there is a strictly positive real number k (for a given value measurement scale) and a finite set of states (numbers) such that, relative to a restriction of the possibility set (lottery) to any finite superset of size n , one option (set of lottery tickets) is more valuable than a second, and the value of the first option is at least k/n units better than that of the second option, then the first option is more valuable than the second.⁸

Unlike Radical Superset Betterness, this principle does not say that $\{1\}$ is more valuable than $\{\text{all natural numbers other than } 1\}$ in the uniform lottery over natural numbers with prizes of 1 for 1, $1/2$ for 2, $1/4$ for 3, ... etc. For any set of size n that includes 1, the expected value of $\{1\}$ is $1/n$, and the expected value of $\{\text{all natural numbers other than } 1\}$ is at most $[1-(1/2)^{n-1}]/n$. The difference between the two is thus at most $(1/2)^{n-1}/n$, and no matter how small a number k one picks, $(1/2)^{n-1}/n$ will be less than k/n for sufficiently large n . Hence Superset Betterness is rightly silent in this case.

Superset Betterness also delivers the desired judgement that a larger* set is more valuable than a smaller* one in our uniform probability case. It judges a larger finite set as more valuable. For example, {1,2} is deemed more valuable than {3}, since the difference in expected values, for a set of size n that includes 1-3 will be $1/n$, and by letting k be .9 (for example), this is greater than k/n . It judges an infinite set as more valuable than a finite set. For example, {all natural numbers other than 1} is deemed more valuable than {1}, since the difference in expected values, for a set of size n that includes at least 2 will be at least $1/n$, and by letting k be .9 this is greater than k/n . It judges an infinite set as more valuable than an infinite proper subset. For example, {all multiples of 3} is deemed more valuable than {all multiples of 6}, since the difference in expected values, for a set of size n that includes at least 3 will be at least $1/n$, and by letting k be .9 (for example) this is greater than k/n . Finally, for two distinct infinite sets at least one of which has only a finite number of non-common elements, it judges the set with more non-common elements to be more valuable. For example, {1,2, all multiples of 4} is deemed more valuable than {3, all multiples of 4}, since the difference in expected values, for a set of size n that includes at least 1 and 2 will be at least $1/n$, and by letting k be .9 (for example) this is greater than k/n .

I shall show below that these judgements are endorsed by sound principles of assessment based on comparative probability. First, however, we should note that Superset Betterness is silent about the ranking of distinct infinite sets neither of which is larger* than the other.

5. Uniform Probability and Distinct Infinite Sets Neither of Which is Larger*

When probability is infinitely and uniformly spread, Superset Betterness entails that a larger* set

of tickets is more valuable, but it is silent when two distinct sets are infinite and neither is larger* than the other. Here I show that this is so and that such silence is appropriate.

Superset Betterness is silent in cases where two distinct sets are infinite, neither is larger* than the other, and no particular probability distribution is specified other than that any two numbers are equally probable. Suppose, for example, that we are comparing the set of all the multiples of 3 to the set of all multiples of 19. (Both have an infinite number of non-common elements, and so neither is larger* than the other.) No matter what finite set you consider, (1) you can expand it by adding enough multiples of 3 so that, relative to the lottery restricted to the expanded set, the set of multiples of 3 has a higher expected utility than the set of multiples of 19, and (2) you can expand it by adding enough multiples of 19 so that, relative to the lottery restricted to the expanded set, the set of multiples of 19 has a higher expected utility than the set of multiples of 3. Consequently, Superset Betterness is silent. And that's a good thing. For, as I shall argue below, there is no determinate answer to what is more valuable in such a case.

To see this it will be useful to think of the lottery as involving an idealized line of fortune (like a wheel of fortune except linear and infinitely long). The line is divided into equal intervals with each positive integer assigned to exactly one interval. A ball is rolled and it randomly stops somewhere along the line. (This is all idealized, of course. We might suppose that the amount of friction is randomly selected, and this determines where the ball stops.) When the ball stops, the point directly below the center of the ball determines which interval is selected. If it lands on the dividing line, then it is deemed to have stopped on the lower number. By assumption, the ball and line are fair in the sense that one finite interval (e.g., 1-4 inclusive) is at least as probable (as a stopping point for the ball) as a second interval (e.g., 6-9 inclusive) just in case the length of the

former is at least as great as the latter.

Now, to our question: Are the multiples of 3 more probable than the multiples of 19? If the numbers are arranged along the line in their natural order, then the answer is yes. The answer can, however, be no, if the numbers are suitably arranged in a different way. Consider, for example, the following line:

L1: 1, 2, 19, 4, 5, 2x19, 7, 8, 3x19, 10, 11, 4x19, 13, 14, 5x19, 16, 17, 6x19, 3, 20, 7x19,

This is simply the positive integers in their natural order, but with each multiple of 3 permuted with the corresponding multiple of 19 (2x3 is permuted with 2x19, etc.). For this line of fortune, the multiples of 19 are more probable than the multiples of 3.

This shows very clearly that the mere fact that there is a uniform distribution of probability over the positive integers does not determine the answer to the question of whether the multiples of 3 are more probable than the multiples of 19 in our original lottery problem. For the original problem told us nothing about how the numbers were selected except that each positive integer was equally probable with every other positive integer. It all depends on which line of fortune is used to select a number. The assumption of equiprobability is compatible with an infinite number of lines of fortune.

Of course, if the original case is modified by the addition of the specification that the positive integers are selected by a particular line of fortune, then we may have determinate answers to questions about the relative probability of two distinct infinite sets neither of which is larger* than the other. We shall consider below how Superset Betterness can be strengthened to

deal with such cases. Such information, however, is not provided in our original case (which specified only that any two numbers are equally probable).

Thus, where probability is infinitely and uniformly spread, the silence of Superset Betterness in cases where two distinct sets are infinite and neither is larger* than the other is appropriate.

6. Comparative Probability

In this section, I shall identify two weak principles for evaluating options that appeal only to comparative probability, and show that they jointly agree with Superset Betterness judgements that a larger* set of tickets is more valuable than a smaller* set in our lottery case.

Comparative probability rankings have no problem with infinite uniform distributions of probability. They simply hold that any two basic states are equally probable, and that if states have an interval order (e.g., if states correspond to selections of numbers on a line of fortune), then any interval event is equally probably with any other interval event of equal length. For generality, I shall identify some principles below that apply to lotteries with several distinct possible prizes (e.g., different prizes for different numbers) even though our simple lottery has only one prize.

The following is a highly plausible, and widely accepted, principle for assessing lotteries. (Recall that throughout I use “possible” for “genuinely possible”, i.e., more probable, given the context, than the empty event.)

Basic State Prize Dominance: If, for two options (sets of tickets), (1) for each possible state (selection of positive integer) the prize won by the first option is at least as valuable as the prize won by the second option, and (2) for at least one possible state, the prize won by the first option is more valuable than that of the second, then the first option is more valuable than the second.

This principle ensures that a set of tickets is more valuable than a proper subset, but it does not cover the other cases of one set being larger* than another. Fortunately, there is a second, widely accepted principle that enables us to cover those cases. Understand a probability-preserving payoff-shift of a given option to be an option obtained by shifting the payoffs under some of the states to some other equally probable states. Call such a shift finite just in case it involves only a finite number of such states, and infinite otherwise. In our uniform lottery, for example, the option {1,3,4} is a finite probability-preserving payoff-shift of {1,5,6}. For they have the same payoff structure except that the payoff under 5 has been shifted to 3 and the payoff under 6 has been shifted to 4. A plausible and generally accepted principle is:

Basic State Finite Shifting: If one option is a finite probability-preserving payoff-shift of another, then the two options are equally valuable.

In our lottery over the positive integers, the conjunction of Basic State Prize Dominance and Basic State Finite Shifting entails that a larger* set of tickets is more valuable than a smaller* set. For a smaller* set has a finite number, n , of non-common elements. By Basic State Finite Shifting, it is equally valuable with the set of its common elements plus any n of the non-

common elements in the larger* set (by finite shifting). And by Basic State Prize Dominance, the larger* set is more valuable than this latter set (since it is a proper superset and thus wins the same prizes in all states in which the smaller* set wins, and wins in some other states as well).

So these two widely accepted principles agree with the judgements of Superset Betterness in our lottery case where one set is larger* than another. Of course, those who reject expected utility as the basis for evaluation in finite cases may reject one of both of these principles. Basic State Finite Shifting is a separability principle, and thus those inclined to reject the separability (or independence) that underlies expected utility theory will be inclined to reject it. In this paper, however, I'm granting for the sake of argumentative concreteness that separability, and the expected utility criterion of value, are plausible in the finite case. Superset Betterness is also applicable in conjunction with other principles of evaluation for finite cases, and I believe that it is plausible principle in conjunction with any sensible such principle. But in this paper I limit myself to defending something like Superset Betterness in conjunction with expected utility evaluations in finite cases.

Although Basic State Prize Dominance is highly plausible, it is violated in our lottery case by standard expected utility theory combined with standard probability (assigning zero probability to each state). For this combination says that the probability of selecting any one number is zero, and thus that the expected utility of both $\{1,2\}$ and $\{1\}$ is zero, even though the former wins whenever the latter does and also wins sometimes when the latter doesn't. This violation of Basic State Prize Dominance is one of the main problems generated by standard probability theory.

The above two principles are restricted versions of stronger principles that are often

accepted. They are restricted to (basic) state by state comparisons, and not applicable to event by event comparisons for arbitrary partitions of the states into events. And the shifting principle is restricted to a finite number of shifts. In the Appendix I note some problems with stronger versions of these principles.⁹ Here we can simply note that these two principles do not generate any judgements for our lottery when two distinct sets are infinite and neither set is larger* than the other. Basic State Prize Dominance is silent, since neither is a proper subset of the other. Basic State Finite Shifting is silent because there are an infinite number states under which they have different payoffs (infinite number of non-common elements). This silence is good, since, as we saw, the mere fact that the lottery involves uniform probability does not determine which of two such infinite sets of is more probable.

7. Non-Standard Probability Functions

Since the work of Abraham Robinson in the 1960s, it has been recognized that there are (or at least that there can be posited) infinite, and infinitesimal, numbers that—unlike Cantor's infinite numbers—obey the usual laws of addition and multiplication.¹⁰ If H is a non-standard infinite number, then $H+1$ is greater than H , and less than $2H$. Furthermore, $1/H$ is an infinitesimal that is greater than $1/(H+1)$ and less than $2/H$. Such infinitesimals, it is well known, can be used as probabilities to deal with the problems at hand. Here, as above, we continued to assume that expected utility theory correctly evaluates options in the finite case.

A non-standard probability function is a function from events to non-standard numbers inclusively between 0 and 1 such that (1) the probability of the universal set is 1, and (2) if A and B have no elements in common, the probability of their union is the sum of their respective

probabilities. A non-standard probability function is exactly like a standard one, except that it can take as values numbers that are infinitesimally greater, or less, than any standard number strictly between 0 and 1.

Non-standard probability functions can strictly represent infinitely uniform probability. For they can assign an arbitrarily selected infinitesimal probability (e.g., $1/H$ for some non-standard infinite H) to each possibility. For any such assignment, the probability of a possibility will be greater than that of an impossibility (namely zero). And unlike the case with standard probability functions, the sum over all possibilities will not be greater than 1. Such probability functions, like standard ones that assign zero probability to possibilities, are not countably additive (the probability of a union of infinitely many disjoint sets is not necessarily the sum of their respective probabilities), and this imposes certain limitations. But non-standard probability functions are an improvement over non-standard ones, and solve, as we shall now see, the problems at hand.

According to non-standard probability, when probability is infinitely and uniformly spread, a larger* set is more probable than a smaller* one. For let the probability of the set of common elements be c , and let the number of non-common elements in the smaller* set be s . From the definition of larger*, we know that s must be finite, and so non-standard probability of the non-common elements of the smaller* set can be set at s/H , for some arbitrary non-standard infinite H . And the non-standard probability of the non-common elements in the larger* set is greater than s/H (since the larger* set has more than s elements). Thus, the probability of the larger* set (p plus something greater than s/H) is greater than the probability of the smaller* set (p plus s/H).

Thus, where probability is infinitely and uniformly spread, a larger* set has a greater non-standard expected utility (i.e., involving infinitesimals), and is thus more valuable, than the smaller* set. Non-standard probability thus yields the same judgements as Superset Betterness in our lottery cases where one set is larger* than another.

This result is not surprising, given that non-standard probability functions (as is well known) can strictly represent comparative probability, and that non-standard expected utility evaluations satisfy Basic State Prize Dominance and Basic State Finite Shifting. As we saw in the previous section, comparative probability judgements that satisfy these two conditions judge larger* sets as more valuable than smaller* sets, when probability is uniformly distributed. The appeal to non-standard probability functions and expected utility evaluations merely reinforces this point in a concrete way.

Where no specification is given of how the lottery numbers are generated (other than that any two positive are equally probable), non-standard probability is rightly silent about which of two infinite sets, neither of which is larger* than the other, is more probable. For when defined only over the integers (but not over all hyperintegers) non-standard probability is not countably additive. Consequently, nothing follows automatically about which of two such infinite sets is more probable, or more valuable.

So non-standard probability theory combined with expected utility theory agrees with Superset Betterness that, when probability is uniformly spread over a denumerably infinite number of states, an option with a larger* set is more valuable than one with a smaller* set. And it agrees that, where both sets are infinite and neither is larger*, there the issue is underdetermined when no further specification of probability is given.

8. Strengthening the Superset Principles

Superset Betterness does not presuppose that basic states have any order or distance metric. It treats basic states of affairs like balls drawn from an urn. Sometimes, however, basic states do have a natural distance metric (as when they correspond to numbers on a line or wheel of fortune), and in such cases the principle can be strengthened. Suppose, then, that a line of fortune is divided into intervals of finite (but not necessarily equal) length, and that there is a 1-1 correspondence between basic states and these intervals. One finite interval (and hence set of basic states) is at least as probable as a second just in case it is at least as long. In this case we can reformulate our principles so that they appeal—not to arbitrary finite sets of basic states, but only—to finite intervals of basic states. And since we are focusing on cases where the states have the structure of the natural numbers (which have a first element), we shall focus on initial intervals, which are simply intervals that include the interval of the first basic state.¹¹ ($\{1,2,3\}$ is an initial interval, $\{2,3\}$ is a non-initial interval, and $\{1,3\}$ is not an interval.) Consider, then:

Interval Superset Betterness: If there is a strictly positive real number k (for a given value measurement scale) and a finite initial interval of states (numbers) such that, relative to the restriction of the possibility set (lottery) to any longer finite initial interval of length n , one option (set of lottery tickets) is more valuable than a second, and the value of the first option is at least k/n units better than that of the second option, then the first option is more valuable than the second.

As we saw above, for infinitely spread uniform probability, Superset Betterness judged all larger* sets as more valuable, but was silent when both sets have an infinite number of non-common elements. It made no judgement, for example, as to whether the set of all multiples of 3 is more valuable than the set of all multiples of 19. The strengthened principle is also silent if no information is provided about how the lottery is generated. If, however, full information is provided about the arrangement of basic states (numbers) on a line of fortune, then the new principle may make a judgement. If, for example, numbers are arranged in their normal order (and, given the assumption of uniform probability, correspond to intervals of a equal length), then the strengthened principle says that the set of multiples of 3 is more valuable than the set of multiples of 19. For, relative to any finite initial interval containing at least 1-3, the expected value of the multiples of 3 will be at least $\frac{3}{5}$ (as in $\{1,2,3,4,5\}$), and the expected value of the multiples of 19 will be at most $\frac{1}{19}$. Hence the expected value of the former will always be at least $\frac{52}{95}$ units greater than the expected value of the latter.

Of course, if the line of fortune used to generate the lottery has a different structure, a different answer may emerge. For example, the strengthened principle will judge the set of multiples of 19 more valuable if the line of fortune has the same structure as above except with the location of each multiple of 3 permuted with the location of the corresponding multiple of 19. Furthermore, for some lines of fortune, the new principle will still be silent. Consider the following line of fortune (in which semicolons are used to highlight groupings and do not indicate anything about the nature of the sequence):

L2: 1, 3; 2, 19, 2x19; 4, 2x3, 3x3, 4x3; 5, 3x19, ...6x19; ...

On this line the numbers appear in the following order. For brevity, call a number that is neither a multiple of 3 nor a multiple of 19 a non-targeted number. The line starts with the lowest non-targeted number followed by 3. Next comes the lowest remaining non-targeted number followed by 19 and the second multiple of 19. Next comes the lowest remaining non-targeted number followed by the three lowest remaining multiples of 3. Next comes the lowest remaining non-targeted number followed by the four lowest remaining multiples of 19. And so on.

For this line of fortune Interval Superset Betterness is silent about whether the set of all multiples of 3 is more valuable than the set of all multiples of 19. For no matter what finite initial interval one considers, relative to the lottery restricted to some longer finite initial intervals (e.g., ones that end with a non-targeted number preceded by a multiple of 3) the set of multiples of 3 has a higher expected utility, and relative to the lottery restricted to some other longer finite initial intervals (e.g., ones that end with a non-targeted number preceded by a multiple of 19) the set of multiples of 3 has a lower expected utility. Hence the principle is silent. This silence is appropriate in such cases. For there is no reason to judge one set as more valuable than the other.

Where states have an interval structure like that of the natural numbers, Interval Superset Betterness is, I claim, a plausible principle for assessing options.

So far, we have only considered principles for judging one set as more valuable than another. We can supplement these principles with principles for when one set is equally valuable with another. At a minimum, we can endorse:

Superset Equivalence: If there is a finite set of states (numbers) such that, relative to the restriction of the possibility set (lottery) to any finite superset, one option (set of lottery tickets) is equally valuable as a second, then the first option is equally valuable as the second.

If the states have an interval structure like that of the natural numbers, then this can be further strengthened to:

Interval Superset Equivalence: If for every strictly positive real number k (for a given value measurement scale) there is a finite initial interval of states (numbers) such that, relative to the restriction of the possibility set (lottery) to any longer finite initial interval of length n , neither option (set of lottery tickets) is more than k/n units more valuable than the other, then the two options are equally valuable.

These two principles are uncontroversially correct, and they never disagree with standard theory. (They each follow from Basic State Finite Shifting.) In our uniform probability case where neither set is larger* than the other, they judge equally sized finite sets as equally valuable, and judge any set as equally valuable as itself, but are silent where the sets are distinct and infinite.

These two principles are very plausible as far as they go, but they both fail to judge as equally valuable options the values of which converge in a certain manner (e.g., $\{1\}$ vs $\{\text{all natural numbers other than } 1\}$ where there prizes for the numbers in their natural order are: 1, $1/2$, $1/4$, $1/8$, etc.). They can, I believe, be plausibly strengthened to:

Strong Superset Equivalence: If for every strictly positive real number k (for a given value measurement scale) there is a finite set of states (numbers) such that, relative to the restriction of the possibility set (lottery) to any finite superset, neither option (set of lottery tickets) is more than k/n units more valuable than the other, then the two options are equally valuable.

Interval Strong Superset Equivalence: If for every strictly positive real number k (for a given value measurement scale) there is a finite initial interval of states (numbers) such that, relative to the restriction of the possibility set (lottery) to any long finite initial interval, neither option (set of lottery tickets) is more than k/n units more valuable than the other, then the two options are equally valuable.

These are like the original two versions, except that instead of requiring that the two options be exactly equally valuable for all sufficiently inclusive finite supersets (initial intervals), they require only that the values of options converge in the appropriate way (differ by less than k/n). They ensure that $\{1\}$ is equally valuable with $\{\text{all the natural numbers other than } 1\}$ is the case where all numbers are equally probable and the prizes for the numbers in their natural order are $1/2, 1/4, 1/8, \text{ etc.}$

I have been unable to find any plausible further strengthening of these principles. It would be implausible, for example, to hold that two sets are equally valuable just in case neither is more valuable according to Superset Betterness (or the interval version thereof). This principle would

generate a contradiction in conjunction with Superset Betterness. To see this, consider L2 (discussed above)

L2 (repeated): 1, 3; 2, 19, 2x19; 4, 2x3, 3x3, 4x3; 5, 3x19, ...6x19; ...

According to the proposal, relative to this line of fortune, {all multiples 3} is equally valuable as {all multiples of 19}, and the latter is equally valuable to {all multiples of 3 greater than 3}. (In both cases the cumulative relative frequencies of the members of the sets alternate in size on L2.) But {all multiples of 3 greater than 3} is a proper subset of {all multiples of 3}, and hence is less valuable according to Superset Betterness.

Perhaps, however, there is some other sort of strengthening that would be plausible. I leave that as an open question.

9. Superset Principles Applied to the General Case

So far we have focused on the superset principles applied to the case where probability is uniformly spread over denumerably many basic states. And we have assumed for concreteness that options are evaluated on the basis of expected utility in finite cases. We will now discuss some of the implications in the general case.

It is easy to show that Interval Strong Superset Equivalence (as well as the regular and non-interval versions) preserves the reflexivity, symmetricity and transitivity of the relation of being equally valuable, and that Interval Superset Betterness (and its non-interval version) preserve asymmetricity and transitivity of the relation of being more valuable. That is, if these

relations have those properties in the finite case, then they have them in the infinite case as well. For brevity, I shall show that this is so for Superset Betterness only.

Superset Betterness preserves the asymmetry of the relation of being more valuable, since it can't be both that (1) there is a finite set of numbers such that the first set of tickets is more valuable than the second relative to all finite superset restricted lotteries, and also that (2) there is a finite set of numbers such that the second set of tickets is more valuable than the first relative to all finite superset restricted lotteries. Likewise, it preserves the transitivity of being more valuable, since if Superset Betterness judges one set as more valuable than a second, and a second as more valuable than a third, then the union of the two initial finite set of numbers giving rise to these judgements is a set that ensures that the first is more valuable than the third (assuming transitivity in the finite case).

As indicated above Interval Strong Superset Equivalence is the strongest plausible superset principle that I have been able to identify for establishing when two sets are equally valuable. Its combination with Interval Superset Betterness is incomplete in that some pairs of options are incomparable relative to these principles. For example, relative to L2 above (in which multiples 3 and 19 alternate indefinitely in terms of the their cumulative relative frequencies on the line), both these principle are silent. So these principles do not generate complete ranking of options. Although it may be possible to strengthen these principles somewhat, I doubt very much that any plausible complete ranking is possible. For once one starts taking small differences very seriously (as we have been), completeness seems unachievable.

Some might claim that the lack of completeness is a shortcoming of the superset approach. Let us agree that completeness would nice, but if (as with standard theory) the cost of

achieving completeness includes judging a set of genuine possibilities as equally valuable as a proper subset, then the theoretical cost is too high. It is better to have an incomplete set of sound principles than a complete set of unsound principles. Of course, for most practical purposes the unsoundness involved in judgements based on standard probability (which are complete) is of little importance, but theoretically this unsoundness is of the utmost importance.

Another objection to Interval Superset Betterness is that it is incompatible with standard continuity conditions. For it holds, in our infinitely spread uniform lottery, that $\{1,2,3, \dots\}$ —which is sure to win—is more valuable than $\{1,2\}$, and that $\{1,2\}$ is more valuable than $\{1\}$, but it denies that there is a real number p such that $\{1,2\}$ is equally valuable as a p chance of holding $\{1,2,3, \dots\}$ and a $1-p$ chance of holding $\{1\}$. Zero is too small, and any positive real number is too large (since any non-infinitesimal probability of winning is better than the infinitesimal chance of winning that $\{1,2\}$ offers). Again, it would certainly be nice to have continuity, but there is little reason to think that it is a requirement of rational valuation (it rules out perfectly coherent lexical orderings, for example). Furthermore the view that in our uniform probability case a set is more valuable than a proper subset is far more compelling than the view that value is continuous.

So, Strong Superset Equivalence and Superset Betterness (and their interval versions) ensure the satisfaction of the core formal properties of transitivity and symmetry for the relation of being equally valuable, and of transitivity and asymmetry for the relation of being more valuable. This is true no matter how probability is distributed (and not merely where probability is uniform). More generally, the abstract plausibility of these principles does not depend on probability being uniformly distributed. The general idea is (roughly) simply that whatever

judgements hold in an appropriate manner for all sufficiently inclusive supersets (of the appropriate type), hold tout court. That has nothing to do with how probability is distributed.

It is worth noting the implications of the superset principles in cases where regular probability functions can strictly represent probability. One case is where there are only a finite number of (genuine) possibilities (e.g., probability is fully distributed among 1 to 100). In this case, the combination of Superset Betterness with Strong Superset Equivalence (and a fortiori for the combination of their interval versions) yields a complete ranking that fully agrees with standard theory. For one just takes any finite set including all the finite number of possibilities, and the judgements made relative to that set are complete and do not change with expansion (all other elements will have zero probability).

A second case where regular probability functions can strictly represent probability is where there are an infinite number of (genuine) possibilities, but the cumulative probabilities decrease asymptotically with convergent cumulative total probability (e.g., $1/2$, $1/4$, $1/8$, etc.). Here things are more complex. First, it is possible in such cases for the expected value to be infinite. For example, if the probability of number natural number n being selected is $(1/2)^{n-1}$ and the value of the prize when n is selected is 2^n , then any infinite set will have an infinite value ($1+1+1+1\dots$). Here Interval Superset Betterness will judge some infinite sets as more valuable than others (e.g., any set is more valuable than a proper set).¹² So in this case Interval Superset Betterness has some implications that standard theory does not have. Suppose, however, that we restrict our attention to cases where the expected values are finite. This assumption combined with the assumption that probability can be strictly represented by a regular probability function ensures that relative to any line of fortune, any two options, and any positive real number k , there

is some sufficiently long initial interval such that either (1) relative to a restriction of the possibility set to any longer initial interval of length n , one option is at least k/n units more valuable than the other, or (2) relative to a restriction of the possibility set to any longer initial interval of length n , neither option is more than k/n units more valuable than the other. Thus in this case, the combination of Interval Superset Betterness and Interval Strong Superset Equivalence is complete. Furthermore, it agrees exactly with the judgements of standard theory.

Before concluding, we should recall that, although I have assumed for illustration that in the finite case options are evaluated on the basis of their expected values, nothing in the superset principles commits one to this. For the superset principles simply tells one how to extend judgements about finite cases to cases with an infinite number of possibilities. I conjecture that these principles can be plausibly combined with a wide range (perhaps all) theories of value in the finite case.

10. Conclusion

When there is an infinitely spread uniform distribution of probability over states, standard probability functions must assign zero probability to each state. This generates false probability judgements (e.g., that a possible state is equally probable with an impossible one), leads to false judgements about what options are more valuable (e.g., that $\{1,2\}$ is equally valuable as $\{1\}$), and, when no further specification is given about probability, produces inappropriate silence about what options are more valuable (e.g., about whether $\{\text{all multiples of } 3\}$ is more valuable than $\{3\}$).

I formulated and defended some principles as a way of correcting the shortcomings of

appealing to standard probability functions for the evaluation of options. These principles are completely silent about how options should be evaluated when there are only a finite number of possible states. They merely provide ways of extending the evaluations to cases where there are infinitely many possible states. The superset betterness principles are all based on the core idea that, if there is some finite set of states such that, relative to any restriction of the possibilities to finitely many states including at least those states, one option is more valuable than a second, and the difference in the two values does not decrease in a certain specified manner, then the first option is more valuable than the second (tout court). This principle does not hold as a matter of logic. There is no contradiction in denying this principle. But for the purposes of evaluating options it is, I claim, highly plausible.

Interval Superset Betterness and Interval Superset Equivalence are, I have argued, highly plausible. In the case where probability is uniformly spread over denumerably many states, very weak and generally accepted principles based on comparative probability confirm their judgements, as do evaluations based on non-standard expected utility (with infinitesimals). Interval Superset Equivalence can, I believe, be plausibly strengthened to Interval Strong Superset Equivalence (which judges sets as equally valuable if their values converge in an appropriate manner). The combination of these betterness and equivalence principles, however, does not ensure that the ranking of options is complete. Some of this incompleteness is, I have claimed, inevitable once one takes small differences seriously (e.g., no judgement should be made about the multiples of 3 vs. the multiples of 19 relative to L2, the line of fortune in which multiples of 3s and of 19s alternate in their cumulative relative frequency).

Although I have focused on the evaluation of options, the problems discussed concern—

not just standard probability theory, but—standard measure theory in general. A suitable generalization of the principles in measure-theoretic terms is, I would argue, plausible. But that's a topic for another occasion. In any case, we should recall that standard measure theory—and standard probability theory in particular—involves only infinitesimal errors, and for most practical purposes those errors don't really matter. I hope, however, that I have made it plausible that we can, and should, have a theoretically more correct treatment of these matters.

APPENDIX

Here I note that we should reject some (generally accepted) principles of evaluation based on comparative probability that are stronger than those considered in the text. In the text I appealed to Basic State Prize Dominance and to Basic State Finite Shifting to justify the claim that larger* sets are more valuable than smaller* ones when probability is uniformly spread over denumerably many states. These principles involve two restrictions. First, each appeals to state by state comparisons, and not to events (sets of states) under any arbitrary partition of states into events (into mutually exclusive and exhaustive ways the world could be). Second, the shifting principle only holds when there are a finite number of states involved in the shifting (and not when there are any infinite number). I explain here why the principles would not be plausible without most of these restrictions.

Consider first the shifting principle. We must reject this principle without the restriction to the finite number of shiftings:

Basic State Unrestricted Shifting: If one option is a (finite or infinite) probability-preserving payoff-shift of the other, then the two options are equally valuable.

This principle must be rejected because it entails that some sets are equally valuable with some of their proper subsets. It says, for example, that the {all multiples of 3} is equally valuable with {all multiples of 6}. For the state by state payoff vector of the latter can be transformed to that of the former by an infinite number of probability-preserving payoff-shiftings. Let the prize

in our lottery be a payoff of P. Take the payoff vector of the multiples of 6, and permute the payoffs for this set under (the selection of) number 3 and the number 6 (permuting is a kind of double shift). This produces a payoff of P under 3 and 0 under 6. Then permute the payoffs under number 6 and 12. This produces a payoff of P under 3 and under 6. One continues in this fashion, ensuring that successive multiples of 3 have the payoff P, and with an infinite number of such shifts, the payoff structure of the multiples of 6 can be transformed into that of the multiples of 3. The unrestricted version of the shifting principle would then judge these two sets of tickets as equally valuable, but, as we have seen, that would be a mistake. So the unrestricted shifting principle must be rejected.

Consider now the dominance principle. Basic State Prize Dominance requires that payoffs be compared state by state. A strengthened version of this principle replaces this requirement with a requirement that payoffs have the right sort of dominance event by event for some partition of states into events (i.e., into mutually exclusive and exhaustive sets of states):

Event Prize Dominance: If, for two options (sets of tickets), there is a partition of states into events such that (1) relative to each event the first option is at least as valuable as the second option, and (2) relative to at least one event, the first option is more valuable than the second, then the first option is more valuable than the second.

Until we know how to assess options relative to events (as opposed to relative to states) this principle tells us little. The next principle addresses this issue:

Event Relative Probability Dominance: If, for a given event (e.g., selection of one of a given set of numbers) and two options (sets of tickets), (1) for every prize it is at least as probable given the event that the first option will win that prize or better than that the second option will so win, and (2) for some prize it is more probable given that event that the first option will win that prize or better than that the second option will so win, then relative to that event the first option is more valuable than the second.

This principle says, for example, that in our lottery case, relative to the event {3,4,5} (i.e., relative to the selection of one of those numbers), the set of tickets {3,4} is more valuable than the set {5,6,7,8,9}. For the former has a 2/3 chance of winning the prize and the latter only a 1/3 chance.

As far as I can tell, there is no problem with this principle when relativized to the universal event (i.e., not really relativized at all). But the conjunction of the general principle with Event Prize Dominance must definitely be rejected. For consider the following two different partitions of the basic states in our lottery:

Partition P: $\langle \{1,3,6\}, \{2,9,12\}, \{4,15,18\}, \{5,21,24\}, \{7,27,30\}, \dots \rangle$

Partition P': $\langle \{1,19,38\}, \{2,57,76\}, \{3,95,114\}, \{4,133,152\}, \dots \rangle$

Each of these is a partition of the basic states of our original lottery into events. Each set in P includes two multiples of 3 and one other number. Each set in P' includes two multiples of

19 and one other number. Assume that any two numbers are equally probable. Relative to any event in P, the set of tickets numbered with all multiples of 3 is more probable to win the prize than the set of tickets numbered with all multiples of 19. Thus, Event Relative Probability Dominance says that the set of multiples of 3 tickets is more valuable relative to each such event in P. And Event Prize Dominance then says that the set of multiples of 3 tickets is more valuable (tout court) than the set of multiples of 19. But these two principles also entail that the set of multiples 3 is less valuable. For relative to any event in P', the set of tickets numbered with all multiples of 19 is more probable to win the prize than the multiple of 3 set of tickets. Thus, Event Relative Probability Dominance says that the set of multiples of 19 is more valuable relative to each such event in P'. And Event Prize Dominance then says that the set of tickets that are numbered multiples of 19 is more valuable (tout court) than the set of multiples of 3, which contradicts the result at which we just arrived.

The problem is that when there are an infinite number of possibilities, the set of basic states can be partitioned in radically different ways (e.g., so that, relative to every set in one partition, multiples of 19 are more probable than multiples of 3, and also so that relative to every set in another partition they are less probable). This does not generate problems in our uniform probability case when one set is larger* than the other, but it does when both sets are infinite and neither is larger*.

So the conjunction of Event Prize Dominance and Event Relative Probability Dominance must be rejected. One possible approach is to keep the latter, and revise the former by adding a clause requiring the non-existence of a partition of states satisfying the two conditions but with the options interchanged. This revised condition would judge one option as more valuable than a

second if relative to some partition the first prize dominates the second, and relative to no partition does the second prize dominate the first.

For the purposes of this paper, however, the important question of how to strengthen Basic State Prize Dominance and Basic State Finite Shifting need not be resolved. For those principles are weak and plausible, and they are enough to support the case for the Superset Betterness Principle.

References

- Carnap, R.: 1962, "The Aim of Inductive Logic," in E. Nagel, P. Suppes, and A. Tarski (eds.) Logic, Methodology, and Philosophy of Science, Stanford University Press, Stanford.
- de Finetti, B.: 1972, Probability, Induction, and Statistics Wiley, New York.
- Fishburn, P.: 1986, "The Axioms of Subjective Probability," Statistical Science **1**, 335-358.
- Kadane, J., Schervish, M. and Seidenfeld, T.: 1986, "Statistical Implications of Finitely Additive Probability," in P. Goel and A. Zellner (eds.) Bayesian Inference and Decision Techniques, North Holland, Amsterdam.
- Keisler, H.J.: 1976, Elementary Calculus, Prindle, Weber & Schmidt, Boston.
- Koopman, B.O.: 1940, "The Axioms and Algebra of Intuitive Probability," Annals of Mathematics, Ser. 2, **41**, 269-292.
- Levi, I.: 1980, The Enterprise of Knowledge: An Essay on Knowledge, Credal Possibility, and Chance, MIT Press, Cambridge, MA.
- McCall, S. and Armstrong, D.M.: 1989, "God's Lottery", Analysis **49**, 223-224.
- Robinson, A.: 1966, Non-Standard Analysis, North Holland, Amsterdam.
- Savage, L.: 1972, The Foundations of Statistics, Dover Publications, New York.
- Seidenfeld, T., and Schervish, M.: 1983, "A Conflict Between Finite Additivity and Avoiding Dutch Book," Philosophy of Science **50**, 398-412.
- Shimony, A.: 1955, "Coherence and the Axioms of Confirmation," Journal of Symbolic Logic **20**, 8-20.

Skyrms, B.: 1995, "Strict Coherence, Sigma Coherence, and the Metaphysics of Quantity,"

Philosophical Studies **77**, 39-55.

Vallentyne, P. and Kagan, S.: 1997, "Infinite Utility and Finitely Additive Value Theory",

Journal of Philosophy **94**, 5-26.

Notes

* Work on this paper was supported in part by a fellowship grant from the American Council of Learned Societies. I've benefited from comments from John Broome, Luc Lauwers, Storrs McCall, Trenton Merricks, Gene Mills, Teddy Seidenfeld, Howard Sobel, Roy Sorensen, and Paul Weirich.

¹ This is the lottery about which McCall and Armstrong (1989) raise puzzles.

² Clearly, if comparative probability is not complete, it will not be representable by a probability function. For a statement, and interesting discussion, of an additional condition (strong additivity) that ensures such representability, see section 2 of Fishburn (1986).

³ This approach is developed by Skyrms (1995).

⁴ More generally still, regular probability functions are incapable of representing cases where there are an infinite number of genuine possibilities one of which is equally or less probable than all the others. In addition, regular probability functions can't represent certain cases where the probability of states in some specified order decreases indefinitely. They can't represent, for example, a probability distribution of the natural numbers, where 2 is $2/3$ as probable as 1, 3 is $3/4$ as probable as 2, 4 is $4/5$ as probable as 3, and so on.

⁵ The regularity is rejected by Savage (1972, p. 79), de Finetti (1972, p. 91), and Levi (1980, p. 110). The condition is endorsed by Koopman (1940), Shimony (1955), and Carnap (1962).

⁶ I assume here that the conditional probability of some number (e.g., 3) being selected from among a finite set of numbers that includes it (e.g., {3,6}) is well defined and positive (e.g., 1/2). For our equi-probable lottery case, this requires the existence of a conditional probability function that is not derivable from the unconditional one (since the derivation would require division by the unconditional probability that one of a given finite set of numbers is selected, and that is 0).

⁷ For discussion of some closely related ideas, see Vallentyne and Kagan (1997). That paper focuses on aggregations of value (e.g., for utilitarianism) that are infinitely valuable, where the locations are people or times. It does not explicitly address probability theory.

⁸ As background, here are two other weakenings that one might consider, but which are needlessly weak. We could simply require that, for some strictly positive k , one option always is always at least k units better than the other. This would, however, block the judgement that a finite set is more valuable than a proper subset. For the difference in expected value between the two can be made arbitrarily small by increasing the possibility set enough. Alternatively, we could require that, for some strictly positive k , the ratios of the two expected values is never in the range $[1-k, 1+k]$. This would not block the judgement in the above case, but it would block the judgement that an infinite set is more valuable than the subset that results by removing finitely many elements. For the ratio of the two values can be made arbitrarily close to one by increasing the possibility set enough.

⁹ For discussion of the related issues of non-conglomerability and finite additivity without countable additivity, see Seidenfeld and Schervish (1983) and Kadane, Schervish, and Seidenfeld (1986).

¹⁰ See Robinson (1966). For an extremely intuitive and accessible introduction to non-standard math, see Ch.1 of Keisler (1976).

¹¹ If we were dealing with all integers (positive and negative), then we would need to proceed slightly differently. We would need to appeal to symmetric superintervals of a given interval, where these are intervals covering the original interval and expanding an equal distance in both directions. See Vallentyne and Kagan (1997) for discussion.

¹² See Vallentyne and Kagan (1997) for an in-depth discussion of principles for assessing options that have infinite value.