Countable Additivity, Dutch Books, and the Sleeping Beauty Problem[1]

Currently, it appears that the most widely accepted solution to the Sleeping Beauty problem is the one-third solution.[2]Another widely held view is that an agent's credences should be countably additive.[3] In what follows, I will argue that these two views are incompatible, since the principles that underlie the one-third solution are inconsistent with the principle of countable additivity (hereafter, CA). I will then argue that this incompatibility is a serious problems for thirders, since it undermines one of the central arguments for their position.

I will begin, in the first section, by considering the view that violations of CA are not a special problem for thirders, since halfers and thirders alike are committed to an indifference principle that is incompatible with CA. And I will argue, to the contrary, that no indifference principle to which halfers *or* thirders are committed is inconsistent with CA. In the second section, I will argue that thirders are nonetheless committed to the denial of CA, because, when the thirders' position is generalized in the most natural way, the resulting principle is inconsistent with CA. In the third section, will show that this resulting principle, which I call the Generalized Thirder Principle, follows from the premises that underlie each of the main arguments for the one-third solution. Hence, thirders who endorse any of these arguments must accept the Generalized Thirder Principle, and must therefore reject CA. And in the two concluding sections, I will argue that the thirders' position is seriously weakened by its incompatibility with CA. For while thirders have argued that Dutch book

considerations support their position, the incompatibility between their position and CA implies that precisely the opposite is the case. This incompatibility, I argue, undermines the diachronic Dutch book argument given in favor of the one-third solution, and it also gives rise to a synchronic Dutch book argument against the one-third solution. And while many prominent thirders would reject the synchronic Dutch book argument against their view, I argue that their reasons for doing so are unsound.

1 Countable Additivity and Indifference

In the original Sleeping Beauty problem, a fully rational agent, named Beauty, participates in an experiment in which she falls asleep on Sunday night and is then awoken a number of times that is determined by the toss of a fair coin (*heads*: once, *tails*: twice). If the coin lands tails, then on Monday night, all the memories she has formed since the beginning of the experiment are erased. Hence, all the experimental awakenings are mutually indistinguishable. A *thirder* is someone who holds that on Monday morning, Beauty's credence in the proposition that the coin landed heads should be one-third. A *halfer*, by contrast, is someone who holds that on Monday morning, Beauty's credence in this proposition should be one-half.

We can think of the original Sleeping Beauty problem as one instance of a general class of Sleeping Beauty problems. Let us define a *Sleeping Beauty problem* as a problem in which a fully rational agent, Beauty, will undergo one or more mutually indistinguishable awakenings, and in which the number of awakenings she will undergo is determined by the outcome of a random process. Let S be a partition of alternative hypotheses concerning the outcome of this random process. Beauty knows the objective chances of each hypothesis in S, and she also knows how many time she will awaken conditional on each of these hypotheses, but she has no other relevant information. The problem is to determine how her credence

should be divided among the hypotheses in S when she first awakens. The original Sleeping Beauty problem is the instance of this class of problems in which the random process that determines how many times Beauty awakens is the tossing if a fair coin, in which S consists in *Heads* (the hypothesis that the coin came up heads) and *Tails* (the hypothesis that the coin came up tails), and in which Beauty awakens once given *Heads* and twice given *Tails*.

One place where thirders and halfers generally agree is that when Beauty awakens on Monday morning in the original Sleeping Beauty problem, she should have the same credence in the centered proposition that can be expressed "the coin landed tails and I have awoken exactly once" as she has in the centered proposition that can be expressed "the coin landed tails and I have awoken exactly twice." This claim is generally justified by appealing to a principle of indifference. Just how this principle should be formulated is controversial, but one might suppose that any indifference principle that would support this particular claim would also support the following, more general claim:

***Sleeping Beauty Indifference* (SBI).** In any Sleeping Beauty problem, for any hypothesis $h \in S$,upon first awakening, Beauty should have equal credence in each of the awakening possibilities associated with h.

where an *awakening possibility associated with* h is a centered proposition of the form ⌜hypothesis h is true, and I have awoken exactly x times⌝ for any positive integer x that is no greater than the total number of times Beauty will awaken if h is true.

If thirders and halfers are indeed committed to SBI, then it seems they are committed to denying the following principle:

***Countable Additivity* (CA).** For any set of countably many centered or uncentered propositions, any two of which are incompatible, rationality requires that one's credences in the propositions in this set sum to one's credence in their disjunction.

For consider a Sleeping Beauty problem in which Beauty will undergo an infinite series of mutually indistinguishable awakenings no matter what. It follows from SBI that in this case, upon first awakening, Beauty should have the same credence in each of the infinitely many possibilities of the form "I have awoken exactly x times." Let k be this uniform credence. If k equals zero, Beauty's credences in these possibilities will sum to zero. And if k exceeds zero, her credences in these possibilities will have no finite sum. Either way, they will not sum to her credence in their disjunction, which is one. Hence, she will violate Countable Additivity.[4]

This argument is open to a number of objections. While it is true that many halfers and thirders do in fact accept some sort of principle of indifference that applies in Sleeping Beauty problems, being a halfer does not commit one to any such principle, as their position can be defended without appealing to any indifference principle, nor to any principles that entail such a principle. The halfer can defend the claim that Beauty's credence in *Heads* should be one-half upon awakening on Monday morning, on the grounds that before falling asleep on Sunday night her credence should be one-half, and awakening on Monday morning provides Beauty with no information that should lead her to change her credence in *Heads*.[5] And in giving this argument, the halfer can remain neutral on how her credence should be divided between the two awakening possibilities in which *Tails* is true.

Nor does being a thirder commit one to SBI. For while some of the arguments for the one-third solution do indeed appeal a principle of indifference, they needn't appeal to anything as strong as SBI. For they can instead appeal to a weaker principle:

Finitistic Sleeping Beauty Indifference **(FSBI).** For any Sleeping Beauty problem, and any hypothesis $h \in S$, if the number of times Beauty awakens conditional on h is finite, then upon first awakening, she must have equal credence in each of the

awakening possibilities associated with h.

Since, in the original Sleeping Beauty problem, Beauty awakens only twice if the coin lands *Tails*, FSBI suffices to ground the claim that her credence should be divided equally between the awakening possibilities associated with *Tails*. And this principle is not, by itself, inconsistent with CA.[6] Moreover, as we shall see in the next section, there are several arguments for the one-third solution that do not appeal to any principle of indifference whatsoever. And these other arguments can likewise be constructed on the basis of finitistic principles which, while entailing FSBI, do not entail SBI.

A further objection that can be raised to the above argument is that it rests on a misformulation of the principle of countable additivity. Even if halfers and thirders accept the unrestricted indifference principle, SBI, they can claim that their position is perfectly compatible with countable additivity, so long as the latter principle is understood, in the standard way, as pertaining to our credences over ordinary (i.e., uncentered) propositions. The conflict between SBI and countable additivity arises only when the latter principle is taken to range not only over ordinary propositions, but also over *centered* propositions of the form "I have awoken exactly n times." And so, arguably, the proper lesson to be drawn from the above argument is that scope of countable additivity should be understood as ranging only over uncentered propositions.

I conclude, therefore, that neither halfers nor thirders are committed to a principle of indifference that is incompatible with countable additivity. Nonetheless, as I will argue in the next two sections, there is a further principle, to which thirders, but not halfers, are committed, which is incompatible with CA. Hence, while halfers can coherently accept CA, thirders cannot.

2 Countable Additivity and the Generalized Thirder Principle

In order to show that thirders are committed to rejecting of countable additivity, we must determine how the one-third solution to the original Sleeping Beauty problem generalizes to other Sleeping Beauty problems. Let us say that a Sleeping Beauty problem is *standard* just in case Beauty is guaranteed to awaken finitely many times no matter what, or in other words, in which, for every outcome of the random process that determines how many times Beauty awakens, the number of times she awakens conditional on this outcome is a positive integer. The most natural way to generalize the one-third solution to the class of standard Sleeping Beauty Problems is as follows:

Generalized Thirder Principle. In any standard Sleeping Beauty problem, upon first awakening, Beauty's credence in any given hypothesis in S must be proportional to the product of the hypothesis' objective chance and the number of times Beauty will awaken conditional on this hypothesis.

fwhere S is the partition of hypotheses that determine how many times she awakens. Thus, in the original Sleeping Beauty problem, while *Heads* and *Tails* have the same objective chance, the number of times Beauty awakens conditional on *Heads* is only half the number of times she awakens conditional on *Tails*. Hence, by the Generalized Thirder Principle, upon first awakening, Beauty's credence in *Heads* should be one-half her credence in *Tails*. Since these credences must sum to one, her credence in *Heads* must be $\frac{1}{3}$.

It will be useful to express this principle formally. For any hypothesis $i \in S$, let $Ch(i)$ be the objective chance that hypothesis i is true, and let $N(i)$ be the number of times Beauty awakens if i is true. Let P be the Beauty's credence function upon first awakening. The

Generalized Thirder Principle states that in any standard Sleeping Beauty problem,

$$\text{For all } i,\ j \in S,\ \frac{P(i)}{P(j)} = \frac{N(i)Ch(i)}{N(j)Ch(j)} \text{ whenever } Ch(j) > 0$$

In the next section, I will argue that thirders are indeed committed to this principle. But first, I will demonstrate that this principle gives rise to violations of CA in one particular standard Sleeping Beauty problem, namely

Sleeping Beauty in St. Petersburg. This is a Sleeping Beauty problem in which the random process that determines how many times Beauty awakens is a series of coin tosses in which a fair coin is tossed repeatedly until it first lands heads. The set S consists in all the alternative hypotheses concerning the length of this series of coin tosses. Thus, for every positive integer, x, S will include a *finite sequence hypothesis* according to which the length of the sequence of coin tosses is x; if this hypothesis is true, Beauty will awaken 2^x times. S will also include the *infinite sequence hypothesis*, or the hypothesis that the coin never comes up heads and hence the sequence is infinite; if this hypothesis is true, Beauty will awaken exactly once.

Since, conditional on every hypothesis in S, the number of times Beauty awakens is a positive integer, this problem is a standard Sleeping Beauty problem, and so the Generalized Thirder Principle applies to it.

In this problem, for any finite sequence hypothesis, h, where x is the length of the sequence of coin tosses according to h, $N(h) = 2^x$, and $Ch(h) = \frac{1}{2^x}$, since this is the objective chance that a fair coin first lands heads on the x^{th} toss. Therefore, $N(h)Ch(h) = \frac{2^x}{2^x} = 1$. Hence, it follows from the Generalized Thirder Principle that for any two finite sequence hypotheses,

i and j, belonging to S,

$$\frac{P(i)}{P(j)} = \frac{N(i)Ch(i)}{N(j)Ch(j)} = \frac{1}{1} = 1$$

Thus, according to the Generalized Thirder Principle, upon first awakening, Beauty must have the same credence in all the finite sequence hypotheses in S. Let k be this uniform credence.

Now let z be the infinite sequence hypothesis. Since the objective chance of an infinite sequence of tosses landing tails is zero, $Ch(z) = 0$. Therefore, by the Generalized Thirder Principle, for any finite sequence hypothesis, h,

$$\frac{P(z)}{P(h)} = \frac{N(z)Ch(z)}{N(h)Ch(h)} = \frac{1\times\ 0}{1} = 0$$

Thus, according to the Generalized Thirder Principle, upon first awakening, Beauty must have a credence of 0 in the infinite sequence hypothesis. And so her credence in the disjunction of the finite sequence hypotheses must be 1. Since, therefore, there is a countably infinite number of finite sequence hypotheses, and Beauty must have the same credence in each, her credences in these hypotheses cannot sum to her credence in their disjunction. And so, according to the Generalized Thirder Principle, Beauty must violate Countable Additivity. And so the two principles are incompatible.

Note, further, that partition with respect to which the violation of CA arises is a partition among ordinary or uncentered propositions concerning the outcome of a sequence of coin tosses. And so even if CA is interpreted as ranging over only uncentered propositions, it will still be incompatible with the Generalized Thirder Principle. Therefore, if we can show that thirders are committed to the Generalized Thirder Principle, we will thereby have shown that they must reject CA.

3 Why Thirders are Committed to the Generalized Thirder Principle

I will now argue that the Generalized Thirder Principle follows from the assumptions that underlie the standard arguments for the one-third solution to the original Sleeping Beauty problem. Let us begin by considering the simplest such argument.

3.1 The Long-Run Frequency Argument

Elga presents his first and simplest argument for the one-third solution thus:

> Imagine the experiment repeated many times. Then in the long run, about $\frac{1}{3}$ of the wakings would be *Heads-wakings*—wakings that happen on trials in which the coin lands *Heads*. So on any particular waking, [Beauty should] have credence $\frac{1}{3}$ that that waking is a *Heads-waking*, and hence have credence $\frac{1}{3}$ in the coin's landing *Heads* on that trial. This consideration remains in force in the present circumstance, in which the experiment is performed just once.[7]

This argument seems to assume the following:

Frequency. In any standard Sleeping Beauty problem, where V is the random variable whose value determines how many times Beauty awakens, and S is the set of alternative hypotheses concerning the value of this variable, if Beauty is fully rational, then upon awakening her credence in any hypothesis $h \in S$ will be equal to the expected long-run frequency of awakenings that occur on trials in which V has the value predicted by h, assuming the experiment were repeated many times (holding constant the objective chance of each value of V and the number of times Beauty awakens conditional on each such value).

And this assumption entails the Generalized Thirder principle. For in any standard Sleeping Beauty problem defined by a partition, S, if the experiment were repeated many times in the manner just described, then for any hypothesis $h \in S$ occur on trials in which the random process has the outcome predicted by h would be proportional to the product of the objective chance of this outcome and the number of times Beauty awakens conditional on this outcome. Hence, this expected frequency will be proportional to the product of $Ch(h)$ and $N(h)$. And so anyone who accepts Frequency is committed to the Generalized Thirder Principle.[8]

3.2 The Argument from the Principal Principle

After presenting the frequency argument, Elga gives his main argument for the one-third solution. In order to show that this argument commits him to the Generalized Thirder Principle, I will reconstruct his argument on the basis of the weakest general assumptions possible. On this reconstruction, the argument is based on Finitistic Sleeping Beauty Indifference (FSBI) in addition to the following four assumptions:

Finite Additivity. If an agent is fully rational, then for any set of finitely many propositions, any two of which are incompatible, her credence in the propositions in this set will sum to her credence in their disjunction.

Principal Principle. If an agent is fully rational, and she knows the objective chance of a hypothesis, h, and possesses no information that is inadmissible in relation to h, then her credence in h will be equal to its known objective chance.

First Awakening Admissability. In any standard Sleeping Beauty problem, upon first awakening, if Beauty learns only that she has awoken exactly once, then she possesses

no information that is inadmissible with respect to any hypothesis in S.

Limited Conditionalization. If an agent is fully rational, and the only change in her epistemic situation is that she learns some centered or uncentered proposition, q, in which she antecedently had positive credence, then she will update her credence in any uncentered proposition, h, by conditionalizing on q. Hence, if P_1 and P_2 represent her credence functions before and after learning q,

$$P_2(h) = P_1(h \mid q) = \frac{P_1(h \;\&\; q)}{P_1(q)} \text{ whenever } P_1(q) > 0$$

Now let P represent Beauty's credence function in the original Sleeping Beauty problem upon first awakening. And let P^+ represent what her credence function would be one minute after first awakening if she were then to learn that she had awoken only once. When Beauty first awakens, her credence will be divided among the following three possibilities:

T1: The coin landed tails and I have awoken exactly once;

T2: The coin landed tails and I have awoken exactly twice;

H1: The coin landed heads and I have awoken exactly once.

One minute after first awakening, Beauty knows that H1 is true just in case *Heads* is true. And from the Principal Principle, in conjunction with First Awakening Admissibility, it follows that her credence in the latter must be equal to its known objective chance, which is $\frac{1}{2}$. Hence, $P^+(\text{H1}) = .5$. Further, since all that happens to Beauty during the first minute after she awakens is that she learns ($\text{H1} \vee \text{T1}$), and since her prior credence in this disjunction is greater than 0, we can apply Limited Conditionalization. Thus, $P^+(\text{H1}) = P(\text{H1} \mid \text{H1} \vee \text{T1})$.

Therefore, $P(\mathrm{H1} \mid \mathrm{H1} \vee \mathrm{T1}) = .5$, which implies that $P(\mathrm{H1}) = P(\mathrm{T1})$. And it follows from FSBI that $P(\mathrm{T1}) = P(\mathrm{T2})$. Hence, $P(\mathrm{T1}) = P(\mathrm{T2}) = P(\mathrm{H1})$. And by Finite Additivity, these three credences must sum to 1. Therefore, $P(\mathrm{H1}) = \frac{1}{3}$.

I will now show that the five premises that underlie this argument entail the Generalized Thirder Principle. In any standard Sleeping Beauty problem defined by a partition, S, let P represent Beauty's credence function upon first awakening, and let P^+ represent the credence function she would have one minute after awakening if she were then to learn that she had awoken only once. For any positive integer, x, let w_x be the centered proposition that can be expressed thus: "I have awoken exactly x times." And let us define an *awakening possibility* as any centered proposition of the form $\ulcorner h \,\&\, w_x \urcorner$, where h belongs to S, and x is between 1 and $N(h)$ (the number of times Beauty awakens conditional on h).

Now from Principal Principle, in conjunction with First Awakening Admissibility, it follows that when Beauty learns that she has awoken exactly once, her credence in each hypothesis in S must be equal to its objective chance. That is,

$$\text{For all } h \in S,\ P^+(h) = Ch(h)$$

And by Limited Conditionalization:[9]

$$\text{For all } h \in S,\ P^+(h) = P(h \mid w_1) = \frac{P(h \,\&\, w_1)}{P(w_1)}$$

Therefore,

$$\text{For all } h \in S,\ Ch(h) = \frac{P(h \,\&\, w_1)}{P(w_1)}$$

Now in a standard Sleeping Beauty problem, for any hypothesis $h \in S$, the number of awakening possibilities associated with h, namely $N(h)$, is finite. So it follows from FSBI

that Beauty will have equal credence in each of these possibilities. And by Finite Additivity, these must sum to her credence in h. Hence,

$$\text{For all } h \in S,\ \ P(h) = N(h)P(h \,\&\, w_1)$$

From from the last two equations,

$$\text{For all } h \in S,\ \ P(h) = N(h)Ch(h)P(w_1)$$

And this entails the Generalized Thirder Principle, namely:

$$\text{For all } i,\ j \in S,\ \frac{P(i)}{P(j)} = \frac{N(i)Ch(i)}{N(j)Ch(j)} \text{ whenever } Ch(j) > 0$$

3.3 The Commutativity Argument

Another standard argument for the one–third solution is one that Cian Dorr and Frank Arntzenius arrived at independently.[10] On its most natural reconstruction, this argument is based on the same assumptions as Elga's, with the following exceptions: it does not involve the controversial assumptions of First Awakening Admissibility, but instead involves these two assumptions:

Equal Number Admissability. In any standard Sleeping Beauty problem in which Beauty awakens the same number of times regardless of which hypothesis in S is true, awakening provides her with no information that is relevant with respect to any hypothesis in S. And so upon first awakening in such a problem, she possesses no information that is inadmissible with respect to any hypothesis in S.

Commutativity. For any two Sleeping Beauty problems that are defined by the same partition, S, if Beauty is fully rational, and if the set of awakening possibilities in the first problem is a subset of the set of awakening possibilities in the second problem, then for any hypothesis $h \in S$, Beauty's credence in h will be the same in the first problem, immediately after first awakening, as it would be in the second problem, one minute after first awakening, if she were then to learn that some possibility in this subset obtains.[11]

The commutativity argument for the one–third solution involves a variant of the original Sleeping Beauty problem in which Beauty awakens twice regardless of whether the coin lands heads or tails. Thus, it involves a fourth awakening possibility:

H2: The coin landed heads and I have awoken exactly twice.

In this variant of the problem, let P_D represent Beauty's credence function when she first awakens. And let P_D^+ represent what her credence function would be one minute later if she were then to learn that H2 does not obtain. From the Principal Principle, together with Equal Number Admissibility, it follows that $P_D(\textit{Heads}) = \frac{1}{2}$ And since *Heads* is equivalent to the disjunction of H1 and H2, it follows from Finite Additivity that $P_D(\mathrm{H1}) + P_D(\mathrm{H2}) = \frac{1}{2}$. And from FSBI, it follows that $P_D(\mathrm{H1})$ and $P_D(\mathrm{H2})$ must be equal to one another, and so they must each equal $\frac{1}{4}$. We can therefore infer the value of $P_D^+(\mathrm{H1})$ on the basis of Limited Conditionalization, as follows:

$$P_D^+(\mathrm{H1}) = P_D(\mathrm{H1} \mid \neg\mathrm{H2}) = \frac{P_D(\mathrm{H1} \mid \neg\mathrm{H2})}{P_D(\neg\mathrm{H2})} = \frac{P_D(\mathrm{H1})}{P_D(\neg\mathrm{H2})} = \frac{1/4}{3/4} = \frac{1}{3}$$

And so it follows from Commutativity that Beauty's credence in Heads must likewise be $\frac{1}{3}$ in the original Sleeping Beauty problem immediately after first awakening.

I will now show how the Generalized Thirder Principle can be derived from the premises of the commutativity argument. For any arbitrary standard Sleeping Beauty problem defined by a partition, S, let P represent Beauty's credence function upon first awakening, and let P^+ represent the credence function she would have one minute after first awakening if she were then to learn that she had awoken only once. Let us call this arbitrary Sleeping Beauty problem *Problem 1*, and let *Problem 2* be a variant of this problem in which Beauty will awaken exactly once regardless of which hypothesis in S is true. And let P_2 represent Beauty's credence function in Problem 2 immediately after first awakening.

Since, in Problem 2, Beauty will awaken exactly once no matter what, it follows from the Principal Principle, together with Equal Number Admissibility, that

$$\text{For all } h \in S, \;\; P_2(h) = Ch(h)$$

Furthermore, the set of awakening possibilities in Problem 2 is a subset of the set of awakening possibilities in Problem 1. And if, in Problem 1, one minute after awakening, Beauty were to learn that she has awoken exactly once, this would be equivalent to learning that some awakening possibility in this subset obtains. And so it follows, by Commutativity, that

$$\text{For all } h \in S, \;\; P^+(h) = P_2(h)$$

Hence, from the last two equations,

$$\text{For all } h \in S, \;\; P^+(h) = Ch(h)$$

As we saw in the previous section, from the above equation, along with Limited Conditionalization, Finite Additivity, and FSBI, we can derive the Generalized Thirder Principle.

3.4 Dutch Book Argument

A final argument for the one-third solution is based on the premise that a fully rational agent is never vulnerable to a *Dutch book strategy*, or in other words, that such an agent will never be willing to accept each bet in a sequence of bets that is guaranteed to result in a loss. Several authors have shown than if Beauty's credence in *Heads*, upon awakening, is $\frac{1}{2}$, then she will be vulnerable to such a strategy.[12] But we can show, more generally, that if her credence in *Heads* is anything other than $\frac{1}{3}$, she will be vulnerable to such a strategy. The argument requires the following two premises:

Diachronic Dutch Book Invulnerability. If an agent is fully rational, then there is no legitimate diachronic Dutch book strategy by which she is guaranteed to lose money.

Every Awakening Legitimacy. In standard Sleeping Beauty problems, a diachronic Dutch book strategy is legitimate so long as the same bets are offered to Beauty every time she awakens and so long as, in other respects, the protocol followed by the bookie is not affected by any information that Beauty is lacks.

Let x be Beauty's credence in *Heads* upon awakening. Assume x is greater than $\frac{1}{3}$. In this case, the Dutch book strategy to which Beauty is vulnerable will consist in offering Beauty a bet on *Tails* before she falls asleep at the outset of the experiment, and then a bet on *Heads* each time she awakens during the experiment. The first bet, on *Tails*, will cost \$2, and will pay \$4 if the coin comes up tails. If we define the *payoff* of a bet as what it pays minus what it costs, then the payoff of this bet will be \$2 if *Tails* and −\$2 if *Heads*. Since, at the outset

of the experiment when Beauty is offered the bet on *Tails*, her credence in *Tails* is $\frac{1}{2}$, she will be willing to accept this bet. Then, every time she awakens, she is offered a follow-up bet on *Heads* with a payoff of $\$(3-3x)$ if *Heads* and $-\$3x$ if *Tails*. Since, upon awakening, her credence in *Heads* is x, she will be willing to accept this bet. Now if the coin comes up heads, then she will lose the bet on *Tails*, for a payoff of $-\$2$. And she will be awoken exactly once, and so she will be offered only one bet on *Heads*. She will win this bet, for a payoff of $\$(3-3x)$. But since we are assuming that x is greater than $\frac{1}{3}$, this payoff must be less than \$2. And so the net payoff of the two bets will be negative. If, on the other hand, the coin comes up tails, then she will win the bet on *Tails*, for a payoff of \$2. And she will be awoken twice, and so on two occasions she will be offered, and accept, a bet on *Heads*. Since she will lose both bets, each bet have a payoff of $-\$3x$. And since x is greater than $\frac{1}{3}$, the two bets together will result in a loss of more than \$2. Thus, she will lose more on the follow up bets than she made on the initial bet, and so, once again, she will incur a net loss.

Now assume that x is less than $\frac{1}{3}$. In this case, the Dutch book strategy will consist in first offering Beauty a bet on *Heads* at the outset of the experiment, with a payoff of \$2 if *Heads* and $-\$2$ if *Tails*. Then, every time she awakens, she is offered a bet on *Tails* with a payoff of $\$3x$ if *Tails* and $\$(3x-3)$ if *Heads*. If the coin comes up *Heads*, then she will win the bet on *Heads*, for a payoff of \$2. She will then be offered a bet on *Tails* only once, which she will accept and lose, for a payoff of $\$(3x-3)$. Since we are assuming that x is less than $\frac{1}{3}$, the loss from the bet on *Tails* will be more than \$2, and so there will be a net loss. If, on the other hand, the coin comes up tails, then she will lose the bet on *Heads*, for a payoff of $-\$2$. And she will win both the bets on *Tails*, for a combined payoff $\$6x$. But since x is less than $\frac{1}{3}$, the payoff of these two bets will be less than \$2, and so the net payoff from all the bets will again be negative.

It follows from Every Awakening Legitimacy that the Dutch book strategy that figures in the above argument is legitimate, and hence that if Beauty's credence in *Heads*, upon awakening, is anything other than $\frac{1}{3}$, then there is legitimate Dutch book strategy by which she is guaranteed to lose money. And so if we assume Every Wakening Admissibility, along with Dutch Book Strategy Invulnerability, we can conclude that unless Beauty's credence in *Heads* is $\frac{1}{3}$, she is irrational.

On the basis of the same two premises, we can derive the Generalized Thirder Principle. In any standard Sleeping Beauty problem defined by a partition, S, suppose that when Beauty first awakens she violates the Generalized Thirder Principle. Since Beauty's evidence is the same every time she awakens, we may assume that she has the same credences every time she awakens; let P represent these credences. Thus, there will be two hypotheses, i and j, belonging to S such that either

$$\frac{P(i)}{P(j)} < \frac{N(i)Ch(i)}{N(j)Ch(j)}$$

or

$$\frac{P(i)}{P(j)} > \frac{N(i)Ch(i)}{N(j)Ch(j)}$$

If the former inequality obtains, let A be hypothesis i and let B be hypothesis j. But if the latter inequality obtains, let A be hypothesis j and let B be hypothesis i. Thus,

$$\frac{P(A)}{P(B)} < \frac{N(A)Ch(A)}{N(B)Ch(B)}$$

it follows that

$$\frac{P(A)}{P(B)} < \frac{Ch(A)/N(B)}{Ch(B)/N(A)}$$

And hence it follows that for some sufficiently small value of Δ,

$$\frac{P(A)}{P(B)} = \frac{\frac{Ch(A)}{N(B)} - \Delta}{\frac{Ch(B)}{N(A)} + \Delta} \qquad (*)$$

We can now show that Beauty will be vulnerable to the following Dutch book strategy. Before she falls asleep at the outset of the experiment she is offered two bets. The first is a bet on the disjunction of A and B. The second is a bet on A that is conditional on the disjunction of A and B. Furthermore, every time she awakens, she will be offered a follow-up bet on B that is conditional on the disjunction of A and B. The following table indicates the payoffs for each of these three bets, depending on whether A is true, or B is true, or neither A nor B is true.

	If A	If B	If neither A nor B
Bet 1	$\$(1 - Ch(A \text{ or } B))\Delta$	$\$(1 - Ch(A \text{ or } B)\Delta$	$-\$Ch(A \text{ or } B)\Delta$
Bet 2	$\$Ch(B)$	$-\$Ch(A)$	$0
Follow-Up	$-\$\left(\frac{Ch(B)}{N(A)} + \Delta\right)$	$\$\left(\frac{Ch(A)}{N(B)} - \Delta\right)$	$0

Since, at the outset of the experiment, Beauty's credences in A and B are $Ch(A)$ and $Ch(B)$, respectively, she will regard the first two bets as fair, and hence should be willing to accept both. And, given equation $(*)$, every time she awakens and is offered a follow up bet, she should regard this bet as fair, and should therefore be willing to accept it. And so she should be willing to accept all the bets. But is she does so, then she is guaranteed to lose money.

For suppose A is true. In this case, Beauty will win the first and second bet, and she will lose the follow up bet every time it is offered, namely $N(A)$ times. Hence, the total payoff will be

$$\$(1 - Ch(A \ \vee B))\Delta \ + \ Ch(B) \ + \ N(A)\left(\frac{Ch(B)}{N(A)} + \Delta\right) = \ \$(1 - N(A) - Ch(A \ \vee B)\Delta)$$

And since $N(A)$ must be at least one, $1 - N(A) - Ch(A \text{ or } B)$ must be negative, and so the net payoff from all the bets must be negative. Now suppose B is true. In this case, Beauty will win the first bet, lose the second bet, and win the follow-up bet every time it is offered, namely $N(B)$ times. Hence, the total payoff will be

$$\$(1 - Ch(A \ \vee B))\Delta \ - \ Ch(A) \ + \ N(B)\left(\tfrac{Ch(A)}{N(B)} - \Delta\right) = \ \$(1 - N(B) - Ch(A \ \vee B)\Delta)$$

And since $N(B)$ must be at least one, $1 - N(B) - Ch(A \vee B)\Delta$ must be negative, and so the net payoff from all the bets must be negative. Suppose, finally, that neither A nor B is true. In this case, Beauty will lose the first bet, and the all the other bets will be called off, and so she will suffer a net loss. Hence, Beauty is guaranteed a net loss.

Thus, given the two premises that underlie the Dutch book argument for the one-third solution, it follows that if Beauty violates the Generalized Thirder Principle, then there is a legitimate Dutch book strategy by which she is guaranteed to lose money, and hence that she is irrational.

We have seen that the Generalized Thirder Principle follows from the premises that underlie the main arguments for the one-third solution. And so the Generalized Thirder Principle is not only the most natural generalization of the one-third solution. It is a generalization that the thirder cannot easily avoid.

4 Why the Denial of CA Undermines the Dutch Book Argument for Thirdism

In this section, I will argue that it is a serious problem for thirders that their position is incompatible with CA, since this incompatibility undermines one of the central arguments for

their position, namely the Dutch book argument. I will begin by explaining why the Dutch book argument for the one-third position initially appears to be a very strong argument. And I will then show that, given the incompatibility between the thirders' position and CA, one cannot coherently defend this position on the basis of a Dutch book argument. For a Dutch book argument can be made for CA that is at least as strong as any Dutch book argument for the one-third solution.

Recall, from the previous section, that the one-third solution can be derived from the following two premises:

Diachronic Dutch Book Invulnerability. If an agent is fully rational, then there is no legitimate diachronic Dutch book strategy by which she is guaranteed to lose money.

Every Awakening Legitimacy. In standard Sleeping Beauty problems, a diachronic Dutch book strategy is legitimate so long as the same bets are offered to Beauty every time she awakens and so long as, in other respects, the protocol followed by the bookie is not affected by any information that Beauty lacks.

One way to reject this argument would be to reject the first premise. If one responds in this way, then one must reject any Dutch book argument for any solution to the Sleeping Beauty problem, since any such argument would need to involve a diachronic Dutch book strategy. An alternative response would be to reject the second premise. And one could reject this premise with some plausibility. For one might plausibly hold that a legitimate Dutch book strategy cannot exploit any information that the agent lacks. But if, in the original Sleeping Beauty problem, the bookie offers Beauty a bet on *Heads* (or on *Tails*) every time she awakens, then the number of bets that are offered to Beauty will depend on whether *Heads* or *Tails* is true. Hence, a Dutch book strategy in which Beauty is offered a bet on *Heads* (or on *Tails*) a different number of times, depending on whether *Heads* or *Tails*

is true, would appear to be a strategy that exploits relevant information that Beauty lacks, namely, whether *Heads* or *Tails* is true. Hence, arguably, such a strategy is illegitimate. Thus, it could be plausibly argued that Every Awakening Legitimacy is false.

One might instead hold that in Sleeping Beauty contexts, a legitimate Dutch book strategy is one in which the number of times a given bet is offered to the agent does not depend on any relevant information the agent lacks. This would be true, for example, if Beauty were offered a fixed set of bets on one and only one occasion when she awakens. Hence, one might propose that the assumption of Every Awakening Legitimacy be replaced by the following.

Single Awakening Legitimacy. In standard Sleeping Beauty problems, a diachronic Dutch book strategy is legitimate if, regardless of how many awakenings Beauty undergoes, she is offered a fixed set of bets on exactly one of these awakenings, and so long as, in other respects, the protocol followed by the bookie is not affected by any information that Beauty lacks.

And this assumption, one might hold, supports not the one-third solution, but rather the one-half solution. For suppose that when Beauty awakens her credence in *Heads* is less than one-half. In this case, she will be vulnerable to the following strategy. At the outset of the experiment, she is offered a bet that returns \$1 if *Heads*, for which she will be willing to pay \$.5. And on exactly one occasion when she awakens, she is offered a follow-up bet that returns \$1 if *Tails*, for which she will be willing to pay more than \$.5. And so she will pay a total of more than \$1 for a pair of bets which together are guaranteed to return \$1, thereby incurring a sure loss. Suppose, on the other hand, that when Beauty awakens her credence in *Heads* is more than one-half. In this case, she will be vulnerable to a parallel strategy in which she pays \$.5 for a bet on *Tails* at the outset of the experiment, and then pays more than \$.5 for a bet on *Heads* on exactly one occasion when she awakens. And this strategy will

likewise result in a sure loss for Beauty. Therefore, assuming Single Awakening Legitimacy, the only way Beauty can avoid being vulnerable to a legitimate Dutch book strategy is for her credence in *Heads* upon awakening to be one-half.

One might think, therefore, that the Dutch book argument for the one-third solution is not very strong. For one might think that the assumption of Every Awakening Legitimacy is no more plausible than the assumption Single Awakening Legitimacy, and one might think that the latter supports not the one-third solution but rather the one-half solution. But this would be a mistake. Properly understood, Single Awakening Legitimacy, like Every Awakening Legitimacy, supports the one-third solution. For what the above argument shows is that if Beauty is to avoid being vulnerable to a Dutch book strategy in which she is offered a follow-up bet on one occasion when she awakens, then *when she is offered this follow-up bet* her credence in *Heads* must be one-half. But this does not entail that *when she awakens* her credence in *Heads* must be one-half. For, in the scenario in which Beauty is offered a follow-up bet on exactly one occasion when she awakens, Beauty's being offered such a bet on a given awakening will provide evidence for *Heads*. And so if her credence in *Heads* is to be one-half upon being offered the bet, it must be less than one-half upon awakening.[13]

Indeed, in the scenario in which Beauty is guarenteed to receive a follow-up bet only once, when we take into account the evidence provided by Beauty's being offered such a bet, we can show that her credence in *Heads*, immediately after awakening, must be one-third. Let P be Beauty's credence function immediately upon awakening. And let B be the centered proposition that can be expressed thus: "I am offered a bet on this awakening." If *Heads* is true, then Beauty will be offered a bet on the only occasion when she awakens. And so, when Beauty awakens, her credence that she will be offered a bet on this awakening, conditional on *Heads*, must be one. That is, $P(B \mid \textit{Heads}) = 1$. But if *Tails* is true, then Beauty will be

offered a bet on only one of the two occasions when she awakens. And so it follows from the Finitistic Sleeping Beauty Indifference Principle that when she awakens, her credence that she will be offered a bet on this awakening, conditional on *Tails*, must be one-half. That is, $P(B \mid \mathit{Tails}) = .5$. We have seen that if Beauty is to avoid being vulnerable to a Dutch book, then upon being offered the follow-up bet, her credence in *Heads* must be .5. And so $P(\mathit{Heads} \mid B) = .5$. But by Bayes' Theorem,

$$P(\mathit{Heads} \mid B) = \frac{P(B \mid \mathit{Heads})P(\mathit{Heads})}{P(B \mid \mathit{Heads})P(\mathit{Heads}) \ + \ P(B \mid \mathit{Tails})P(\mathit{Tails})}$$

$$= \frac{1 \cdot P(\mathit{Heads})}{1 \ \cdot \ P(\mathit{Heads}) \ + \ .5 \ \cdot \ (1 - P(\mathit{Heads}))}$$

It follows that if $P(\mathit{Heads} \mid B) = .5$, then $P(\mathit{Heads}) = \frac{1}{3}$. Therefore, assuming Single Awakening Legitimacy, the only way for Beauty to avoid being vulnerable to a legitimate Dutch book strategy is for her credence in *Heads*, upon awakening, to be one-third.

It would seem, therefore, that diachronic Dutch book arguments unambiguously support the one-third solution, since they seem to support this solution regardless of whether we accept Every Awakening Legitimacy or Single Awakening Legitimacy. However, this support is undermined by the incompatibility between the thirders' position and Countable Additivity. For it follows from this incompatibility that there are contexts in which satisfying the Generalized Thirder Principle will make Beauty vulnerable to a legitimate Dutch book strategy, on either of the two conceptions of such a strategy that we have considered.

The basic problem is that if Beauty satisfies the Generalized Thirder Principle, then, as we have seen, she will violate CA. Consequently, her credences will be *non-conglomerable*, in the sense that her unconditional credence in some proposition will exceed the upper bound for

her credence in this proposition conditional on every element in some countable partition.[14] And in virtue of having non-conglomerable credences, she will be vulnerable to a Dutch book strategy involving finitely many bets.

For an illustration of this problem, consider the following case:

Three Coins. This is just like the problem of Sleeping Beauty in St.Petersburg, except that on Sunday evening before the experiment begins, three coins are tossed: a nickel, a dime and a quarter. The nickel is tossed repeatedly until it lands heads, and the same is done with the dime, but the quarter is tossed only once. Let n be the length of the sequence of nickel tosses; let d be the length of the sequence of dime tosses; and let $Heads_Q$ and $Tails_Q$ be the propositions that the quarter came up heads, and that the quarter came up tails, respectively. After all the coins are tossed, a number, X, is secretly written onto a piece of paper. If $Heads_Q$, then the number that is written on the piece of paper is the length of the sequence of nickel tosses; that is, $X = n$. But if $Tails_Q$, then $X = d$. Beauty is then awoken 2^n times. And one minute after each awakening, she is shown the value of X.[15]

Since the number of times Beauty awakens depends on n, but not on d, thirders must hold that when Beauty awakens on the first morning of the experiment, assuming she is fully rational, there will be a change in her credence in the possible values of n, but no change in her credences in the possible values of d. In particular, for every positive integer value of x, her new credence in the proposition that $n = x$ will be 0, while her credence in the proposition that $d = x$ will remain $\frac{1}{2^x}$. But assuming $Heads_Q$, for any positive integer, x, $X = x$ just in case $n = x$. And so, for every positive integer, x, her credence in the proposition that $X = x$, conditional on $Heads_Q$, must be 0. Similarly, assuming $Tails_Q$, for any positive integer, x, $X = x$ just in case $d = x$. And so her credence in the proposition

that $X = x$, conditional on $Tails_Q$, must be $\frac{1}{2^x}$. Since she has positive credence in each of the countably many possible values of x, it follows that for each of these values of x, her credence in $Heads_Q$, conditional on the proposition that $X = x$, must be 0. And yet, since she has no evidence that is relevant to the outcome of the quarter toss, it follows from the Principal Principle that her unconditional credence in $Heads_Q$ must be equal to its objective chance, namely be one-half. Thus, her unconditional credence in $Heads_Q$ will significantly exceed her credence in $Heads_Q$ conditional on every possible value of X. Hence, her credences will be non-conglomerable.

As a result, she will be vulnerable to a Dutch book strategy. For since, upon first awakening, Beauty's credence in $Heads_Q$ is $\frac{1}{2}$, she will be more than happy to pay, say, \$.49 for a bet that pays \$1 if $Heads_Q$. But after she learns the value of X, her credence in $Heads_Q$ will be 0, and so she will then be more than happy to pay, say, \$.99 for a bet that pays \$1 if $Tails_Q$. As a result, she will be sure to lose \$.48.

The Dutch book strategy to which Beauty will be vulnerable can be spelled out in either of two ways. In the first version of the strategy, Beauty is offered the sequence of bets just described (the bet on *Heads* before she learns the value of X, and the bet on *Tails* after she learns this) on exactly one of the mornings when she awakens. This version of the strategy will count as legitimate according to Single Awakening Legitimacy. And this version of the strategy will result in Beauty's suffering a sure loss of \$.48. In the second version, Beauty is offered this sequence of bets on every morning when she awakens. This version will count as legitimate according to the assumption of Every Awakening Legitimacy. And this version will result in her suffering a sure loss of \$.48 every time she awakens. Thus, on either of the views we have considered concerning what constitutes a legitimate Dutch book strategy, if Beauty satisfies the Generalized Thirder Principle, then there will be contexts in which she

is vulnerable to such a strategy.

Note, further, that these strategies involve offering only finitely many bets. For in the first version, in which Beauty is offered the sequence of bets on exactly one morning when she awakens, the total number of bets she is offered is two. And in the second version, in which Beauty is offered this sequence of bets on every morning when she awakens, the total number of bets she is offered will be twice the total number of awakenings. And since the total number of awakenings is guaranteed to be finite, the total number of bets is likewise guaranteed to be finite.

Thus, while Dutch book arguments initially appear to give strong support to the thirders' position, this appearance is illusory. One cannot coherently accept the premises that underlie the Dutch book arguments for the one-third solution. That is, we cannot coherently accept the conjunction of Dutch Book Strategy Invulnerability and Every Awakening Admissibility. Nor can we coherently accept the conjunction of Dutch Book Strategy Invulnerability and Every Awakening Admissibility. For, as we saw in the previous section, the former pair of premises implies that a rational agent must always satisfy the Generalized Thirder Principle. And we could derive the same conclusion from the second pair of premises, so long as we take into account the evidential significance of being offered the follow-up bets. However, as we have just seen, each of these pairs of premises implies that in the case of Three Coins, a rational agent cannot satisfy the Generalized Thirder Principle, for in doing so she would violate CA, and thereby be vulnerable to a legitimate Dutch book strategy. And so we cannot coherently accept either of these pairs of premises.

Thus, the incompatibility between CA and the Generalized Thirder Principle undermines the Dutch book argument for the latter. But the thirder might object that this incompatibility cuts both ways: since Dutch book arguments appear to support both CA and the

Generalized Thirder Principle, the incompatibility of these two principles calls into questions the general applicability of Dutch book arguments in Sleeping Beauty contexts. Thus, the thirder may argue that since Dutch book arguments provide the primary source of support for CA, the incompatibility between CA and the Generalized Thirder principle calls into question the applicability of CA in Sleeping Beauty contexts.

This response, however, overlooks an important asymmetry between the Dutch book arguments for CA and the Dutch book arguments for the one-third solution. For while the only Dutch book arguments for the one-third solution are diachronic, CA can be supported by either a diachronic or a synchronic Dutch book argument. In other words, while the one-third solution can be supported only by a Dutch book argument involving a sequence of bets that are offered at different times, CA can also be supported by a Dutch book argument involving only simultaneous bets. Thus, while it may be true that *diachronic* Dutch book arguments, or at least the applicability of such arguments in Sleeping Beauty contexts, are called into question by the fact that such arguments seem to support both the Generalized Thirder Principle and CA, this does nothing to impugn *synchronic* Dutch book arguments, which support CA but which do not support the Generalized Thirder Principle. And, independently of this dialectic, synchronic Dutch book arguments appear to be on a stronger footing than diachronic Dutch book arguments.[16] And so if synchronic Dutch book arguments are the only such arguments left standing, it seems we should conclude that Dutch book arguments unambiguously favor CA, and hence the rejection of the Generalized Thirder Principle.

However, some very prominent thirders deny that CA receives any support from synchronic Dutch book arguments. Frank Arntzenius, Adam Elga and John Hawthorne maintain that that the standard synchronic Dutch book arguments for CA have no force, since

these arguments involve infinitely many bets and "Dutch book arguments have no force in infinite cases."[17] Thus, before concluding that Dutch book arguments support CA, and hence the rejection of the Generalized Thirder Principle, we should consider this opposing view.

5 Have Thirders Undermined the Dutch Book Argument for Countable Additivity?

In this section, I will begin by presenting a synchronic Dutch book argument for CA. The argument I will sketch will be stronger than the more standard Dutch book arguments for CA, since it will involve a set of bets, each one of which the agent regards not merely as fair, but as favorable.[18] I will then examine an objection to Dutch book arguments for CA that has been leveled by Arntzenius, Elga and Hawthorne.

In order to provide a Dutch book argument for CA, it will suffice to consider the case in which an agent violates Countable Additivity without violating Finite Additivity, since Finite Additivity is generally accepted as a requirement of rationality, and since it can be defended on the basis of a synchronic Dutch book argument involving only finitely many bets. For any agent who violates Countable, but not Finite, Additivity, there will be a partition, Q, of countably many propositions, q_1, q_2, etc., such that the agent's credences in these propositions sum to less than her credence in their disjunction, h. Let Δ be the margin by which her credence in this disjunction exceeds the sum of her credences in the disjuncts. That is,

$$P(h) \ = \ \sum_n P(q_n) \ + \ \Delta$$

We can construct a Dutch book against such an agent, consisting of a set, S, of countably

many bets, b_0, b_1, b_2, etc, with the following payoff schedules. Let b_0 be a bet on h whose payoff is $\$(1 - P(h) + \frac{\Delta}{3})$ if h is true, and $\$(\frac{\Delta}{3} - P(h))$ if h is false. This bet therefore has a positive expected payoff of $\frac{\Delta}{3}$. For every value of n greater than 0, let b_n be a bet on the negation of q_n whose payoff is $\$(P(q_n) + \varepsilon_n)$ if q_n is false, and $\$(P(q_n) + \varepsilon_n - 1)$ if q_n is true, where the values of the ε_n's are all positive and sum to $\frac{\Delta}{3}$. Call these the *negative bets*. Thus, each negative bet, b_n, will have a positive expected payoff of ε_n.

Since every bet in S has a positive expected payoff, the agent is committed to accepting every bet in S. But if she accepts every bet in S, then she is guaranteed to lose money. For if h is true, then she will win the bet on h, for a payoff of $(1 - P(h) + \frac{\Delta}{3})$, and she will lose one of the negative bets and win all the rest. Together, the negative bets will pay

$$\$\Big(\Big(\sum_n P(q_n)\Big) + \Big(\sum_n \varepsilon_n\Big) - 1\Big) = \$\big(P(h) - \Delta + \tfrac{\Delta}{3} \ - \ 1\big) = \$\big(P(h) \ - \ \tfrac{2}{3}\Delta \ - \ 1\big)$$

And so the net payoff from all these bets together will be $-\$\frac{\Delta}{3}$.

Suppose, on the other hand, that h is false. In this case, the agent will lose the bet on h, for a payoff of $\$(\frac{\Delta}{3} - P(h))$, and win all the negative bets, which together will pay

$$\$\Big(\sum_n P(q_n) \ + \ \sum_n \varepsilon_n\Big) = \$\big(P(h) - \Delta + \tfrac{\Delta}{3}\big) = \$\big(P(h) \ - \ \tfrac{2}{3}\Delta)\big)$$

And so, once again, the net payoff from all the bets together will be $-\$\frac{\Delta}{3}$. Since the agent is guaranteed to lose money regardless of whether h is true or false, she is vulnerable to synchronic Dutch book. Thus, assuming that vulnerability to such a Dutch book indicates irrationality, anyone who violates CA is irrational.

Arntzenius, Elga and Hawthorne reject this argument, since they reject the inference from the claim that every bet in S has a positive expected value to the conclusion that the

agent is committed to accepting all the bets in S. In order to undermine this inference, they present the following case.

Satan's Apple. Satan has cut a delicious apple into infinitely many pieces, labeled by the natural numbers. Eve may take whichever pieces she chooses. If she takes merely finitely many of the pieces, then she suffers no penalty. But if she takes infinitely many of the pieces, then she is expelled from the Garden for her greed. Either way, she gets to eat whatever pieces she has taken.[19]

For any given slice of the apple, α, if, apart from α, Eve takes finitely many pieces of the apple, then her taking piece α will result in her having one more piece of apple to enjoy in the Garden. And if, apart from α, she takes infinitely many pieces of the apple, then her taking piece α will result in her having one more piece to enjoy as she is expelled from the Garden. Hence, Eve regards the choice of each piece of the apple as favorable when considered on its own, in the sense that, for any given piece, α, *no matter what combination of other slices of the apple she takes*, the expected gain from taking these other pieces along with α is greater than the expected gain from taking these other pieces alone. Nonetheless, it seems clear that she is not thereby committed to taking all the slices of apple. For it would seem that she could rationally take a large but finite number of slices of apple, and leave infinitely many slices behind—doing so would be far more rational than taking every slice, and thereby being expelled from the Garden.

Arntzenius, Hawthorne and Elga infer that we should reject the *Deal Agglomeration Principle*, according to which, if an agent regards each deal in a set of infinitely many deterministic deals as favorable when considered on its own, she is thereby committed to regarding the entire package of deals as favorable, and hence to accepting all the deals in this package. But the Deal Agglomeration Principle appears to stand on an equal footing with

the *Bet Agglomeration Principle*, which states that if each bet, b, in a set of bets is favorable when considered on its own (in the sense that, *no matter what combination of other bets the agent takes*, the expected utility of taking these other bets along with b is greater than the expected utility from taking these other bets alone) then the agent is committed to regarding the entire package of bets as favorable, and hence to accepting all the bets in this package. Since the case of Satan's Apple shows that we must reject the Deal Agglomeration Principle, they infer that we should also reject the Bet Agglomeration Principle. They conclude that since the Dutch book argument for CA rests on the Bet Agglomeration Principle, the latter argument is unsound.

The reason that the Deal and Bet Agglomeration Principles have considerable prima facie plausibility is that they follow from a principle of dominance:

Unrestricted Dominance. For any fully rational agent, s, any pair of options, ϕ and ψ, and any partition, E, if s strongly prefers ϕ to ψ conditional on every element of E, then s strongly prefers ϕ to ψ unconditionally.

(where an agent *strongly prefers* ϕ to ψ just in case she does not prefer ψ to ϕ and she is not indifferent between the two options). For if an agent regards a deal, d, as favorable when considered on its own, in the sense defined above, then she will strongly prefer accepting this deal to declining this deal conditional on every possible combination of the other deals she might accept. Hence, by Unrestricted Dominance, she is rationally required to have a strong unconditional preference for accepting d over declining d, and so we may infer that she is rationally required to accept d. It follows that if an agent regards each deal in a set of deals as favorable when considered on its own, then she cannot rationally fail to accept any of the deals in this set, and she is thus rationally required to accept them all. And so we can derive the Deal Agglomeration Principle from Unrestricted Dominance. And, by similar

reasoning, we can derive the Bet Agglomeration Principle from Unrestricted Dominance.

What the case of Satan's Apple shows, according to Artnzenius, et., al, is that the principle of dominance, on which both the agglomeration principles derive their intuitive support, must be rejected. In the case of Satan's Apple, for any piece, α, of the apple, conditional on every possible combination of other slices of the apple Eve might take, she should strongly prefer taking α to declining α. And so it follows from Unrestricted Dominance that if Eve is fully rational, then for every slice of the apple, she will strongly prefer taking it to declining it unconditionally. Hence, assuming that a fully rational agent does what she unconditionally prefers to do, it follows that if Eve is fully rational, she will take every slice of the apple. Thus, Unrestricted Dominance implies that if Eve is fully rational, she will make a set of simultaneous choices that will result in her being expelled from the Garden. And this implication is unacceptable. Thus, "the first lesson of Satan's Apple" according to Arntzenius, et. al., is that "in infinite cases, rationality does not require one to choose one's dominant options."[20] On their view, the case of Satan's Apple shows that we must reject Unresticted Dominance, thereby undermining the Bet Agglomeration Principle, and with it any Dutch book arguments involving infinitely many bets, including the Dutch book argument for CA. We should conclude, on their view, that dominance principles, agglomeration principles, and Dutch book arguments are applicable at most in finitistic cases.

I believe this is the wrong lesson to be drawn from the case of Satan's Apple. While this case does show that the properly formulated principle of dominance can't apply to all partitions, it doesn't follow that we should restrict this principle by limiting it to finite partitions. For one thing, restricting the principle in this way will not result in an adequate principle, since there are plenty of counterexamples to Unrestricted Dominance involving only finite partitions. Suppose, for example, that Russ would enjoy playing Russian Roulette, but

all things considered he would rather not play this game, since he values his life. Nonetheless, he would rather play the game, and be shot in the head, than not play the game and be shot in the head. And he would likewise rather play the game and not be shot in the head than not play the game and not be shot in the head. Hence, conditional on every element of the finite partition {*Russ is shot in the head*, *Russ is not shot in the head*}, Russ strongly prefers playing Russian Roulette to not playing it. And yet he does not prefer playing Russian Roulette to not playing it unconditionally. And so Unrestricted Dominance implies, falsely, that his preferences are irrational.

A properly formulated dominance principle would need to avoid implications of this kind. The standard view among causal decision theorists (among whom Arntzenius, et. al., number themselves) is that the dominance principle should be restricted to partitions that the agent regards as causally independent of her choice among the options under consideration. Hence, Causal decisions theorists will accept something like the following principle

Causal Dominance. If a fully rational agent is faced with a pair of options, ϕ and ψ, and if she regards the elements of some countable partition, E, as causally independent of her choice, and if she strongly prefers ϕ to ψ conditional on every element of E, then she strongly prefers ϕ to ψ unconditionally.[21]

This principle, I will now argue, suffices to ground standard synchronic Dutch book arguments, including the argument for CA we considered above. And this principle, unlike Unrestricted Dominance, does not have counterintuitive implications in the case of Satan's Apple.

The reason that Causal Dominance suffices to ground standard Dutch book arguments is that the choice situations that figure in such arguments are ones in which an agent is presented with a set of simultaneous choices that she can make, and should make, independently

of one another. Consequently, she should regard her choices as having the kind of causal independence that is required for the applicability of the principle of Causal Dominance.

Let us be more precise. Let S be a set of bets that constitutes a Dutch book, and let b_1, b_2, etc., be the constituent bets. Let π be a plan that specifies, for each bet in S, whether to accept or decline this bet. For any bet, b_n, in S, let $I^{\pi}_{b_n}$ represent the payoff the agent will receive from bet b_n if she adopts plan π—thus, if π prescribes accepting b_n, then $I^{\pi}_{b_n}$ will be equal to the payoff of this bet, whereas if π prescribes declining b_n, then $I^{\pi}_{b_n}$ will be zero. Now the utility of carrying out plan π is simply the sum of the payoffs of the individual bets it prescribes accepting. That is,

$$U(\pi) = I^{\pi}_{b_1} + I^{\pi}_{b_2} + I^{\pi}_{b_3} + \ldots$$

Thus, assuming the bets in S are independent (in the sense that whether one wins a given bet does not depend on which other bets in S one accepts), it follows that for any bet, $b_n \in S$, the agent's choice as to whether to accept b_n will affect only one term in this sum, namely the nth term. Thus, in deciding whether to accept b_n, she need only be concerned with maximizing the expected value of this term. And, given this independence among the bets, the expected value of this nth term is not affected by her choice concerning any other bet in S. Consequently, for any bet, $b_n \in S$, her choice as to whether to accept this bet should be made independently of the choices she makes regarding the other bets in S.

Let A_{b_n} be the set of alternatives the agent has concerning which subset of the bets in S to accept other than b_n. Since her choices concerning the bets in S should be made independently, it follows that which alternative she chooses in A_{b_n} should not be affected by her choice concerning b_n. And she should recognize this. Hence she should regard the elements of A_{b_n} as causally independent of her choice concerning b_n. And so it follows from

Causal Dominance that if she prefers accepting b_n to declining it conditional on every element of A_{b_n}, then she must prefer accepting b_n to declining it unconditionally.

Therefore, it follows from Causal Dominance that for any set, S, of independent bets, if an agent regards each bet in this set as favorable when considered on its own (in the sense that she prefers accepting each bet to declining it conditional on every possible combination of choices she might makes concerning the other bets in this set), then she must prefer accepting each bet in this set to declining it unconditionally. Hence, she will be rationally required to accept every bet in this set. Thus, from the principle of Causal Dominance, we can derive the Bet Agglomoration Principle, so long as the latter is restricted to sets of independent bets. And the sets of bets that figure in standard Dutch book arguments, including the Dutch book argument for CA, are sets of independent bets. Therefore, the Dutch book argument for CA can be grounded in the Causal Dominance principle.

Moreover, the Causal Dominance priniple, unlike the Unrestricted Dominance principle, does not yield the unacceptable result that in the case of Satan's Apple, Eve is committed to taking every piece of the apple. For in this case, the total utility of Eve's set of choices concerning which pieces of the apple to take cannot be represented as the sum of the utilities of her choices concerning the individual pieces of the apple. Rather, the utility of a given plan, or set of choices, is determined by two things: its effect on how much of the apple Eve gets to eat, and its effect on whether Eve is expelled from the Garden. Plausibly, the value of the former effect can be represented as a sum of values corresponding to the individual choices. For any piece, α_n, of the apple, let $I^{\pi}_{\alpha_n}$ represent the value of the satisfaction Eve would get from α_n if she were to adopt plan π—thus, if π prescribes taking α_n, then $I^{\pi}_{\alpha_n}$ will be positive, whereas if π prescribes not taking α_n, then $I^{\pi}_{\alpha_n}$ will be zero. And plausibly, the value of the effect of plan π on how much of the apple Eve gets to consume will be the

sum of the $I^{\pi}_{\alpha_n}$'s. But the value of the effect of plan π on whether Eve is expelled from the Garden, which we can denote G^{π}, cannot be represented as a sum of values corresponding to the individual choices. And so the total utility of plan π will be as follows.

$$U(\pi) = G^{\pi} + I^{\pi}_{\alpha_1} + I^{\pi}_{\alpha_2} + I^{\pi}_{\alpha_3} + \dots$$

In deciding which subset of pieces of the apple to take, Eve must be concerned not only with the effect of alternative total plans on the terms $I^{\pi}_{\alpha_1}, I^{\pi}_{\alpha_2}, I^{\pi}_{\alpha_3}$, etc., but also with the effect of alternative total plans on the remaining term, G^{π}. And the value of this term depends on her entire set of choices taken as a whole. Hence, she cannot rationally make her choices concerning the individual pieces of the apple independently. Rather, she must make each of these choices in the context of forming a total plan. Eve should therefore recognize that, for any piece, α_n, of the apple, had she made a different choice concerning α_n, she would have done so in the context of choosing an alternative total plan that may have involved making different choices concerning the other pieces of the apple. And so she should not regard her choice concerning which subset of pieces of the apple to take other than α_n as causally independent of her choice concerning α_n. Therefore, we cannot infer from the principle of Causal Dominance that since Eve prefers taking α_n to declining α_n conditional on every set of choices she might make concerning the other pieces of the apple, she must prefer taking α_n to declining it unconditionally. And so this principle does not imply that Eve is committed to taking every piece of the apple.

Since the principle of Causal Dominance suffices to ground the Dutch book argument for CA, and since this principle is invulnerable to the arguments given by Arntzenius, Elga, and Hawthorne, it appears that the Dutch book argument for CA is on solid ground. Hence, given the incompatibility between CA and the thirders' position, there appears to be a strong

Dutch book argument *against* the thirders' position. And, as we saw in the previous section, this incompatibility also implies that there cannot be a successful Dutch book argument *in favor* of the thirders' position. Thus, if the Dutch book approach to grounding the principles of epistemic rationality is well motivated, then there is strong reason not to be a thirder. And conversely, if the thirders' position is well motivated—if, for example, any of the arguments considered in sections 3.1 through 3.3, above, is sound—then there is reason to be skeptical of the Dutch book approach to understanding epistemic rationality.

References

Arntzenius, F. (2002). Reflections on Sleeping Beauty. *Analysis*, 62:53.

Arntzenius, F. (2003). Some Problems for Conditionalization and Reflection. *The Journal of Philosophy*, 100:356–370.

Arntzenius, F., Elga, A., and Hawthorne, J. (2004). Bayesianism, Infinite Decisions, and Binding. *Mind*, 113:251–283.

Bradley, D. and Leitgeb, H. (2006). When Betting Odds and Credences Come Apart: More Worries for Dutch Book Arguments. *Analysis*, 66:119.

Christensen, D. (1991). Clever Bookies and Coherent Beliefs. *The Philosophical Review*, 100:229–247.

De Finetti, B. (1972). *Probability, Induction and Statistics.* New York: J. Wiley.

Dorr, C. (2002). Sleeping Beauty: In Defence of Elga. *Analysis*, 62:292.

Draper, K. and Pust, J. (2008). Diachronic Dutch Books and Sleeping Beauty. *Synthese*, 164:281–287.

Elga, A. (2000). Self-Locating Belief and the Sleeping Beauty Problem. *Analysis*, 60:143.

Hitchcock, C. (2004). Beauty and the Bets. *Synthese*, 139:405–420.

Horgan, T. (2004). Sleeping Beauty Awakened: New Odds at the Dawn of the New Day. *Analysis*, 64:10.

Howard, J. (2006). Countable Additivity and the Foundations of Bayesian Statistics. *Theory and Decision*, 60:127–135.

Howson, C. (2008). De Finetti, Countable Additivity, Consistency and Coherence. *The British Journal for the Philosophy of Science*, 59:1.

Joyce, J. (1999). *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Lewis, D. (2001). Sleeping Beauty: Reply to Elga. *Analysis*, 61:171.

Monton, B. (2002). Sleeping Beauty and the Forgetful Bayesian. *Analysis*, 62:47.

Nozick, R. (1969). Newcomb's Problem and Two Principles of Choice. In Essler, W., Putnam, H., and Stegmuller, W., editors, *Essays in Honor of Carl G. Hempel*. Dordrecht: D. Reidel.

Schervish, M., Seidenfeld, T., and Kadane, J. (1984). The Extent of Nonconglomerability of Finitely Additive Measures. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 66:205–226.

Seminar, O. (2008). An Objectivist Argument for Thirdism. *Analysis*, 68:149–155.

Titelbaum, M. (2008). The Relevance of Self-Locating Beliefs. *The Philosophical Review*, 117:555.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

Weatherson, B. (2005). Should We Respond to Evil with Indifference? *Philosophy and Phenomenological Research*, 70:613–635.

Weintraub, R. (2004). Sleeping Beauty: A Simple Solution. *Analysis*, 64:8.

Williamson, J. (1999). Countable Additivity and Subjective Probability. *The British Journal for the Philosophy of Science*, 50:401–416.

Notes

[1]Acknowledgments omitted for blind review.

[2]This solution is defended in Arntzenius (2003), Dorr (2002), Draper and Pust (2008), Elga (2000), Hitchcock (2004), Horgan (2004), Monton (2002), Seminar (2008), Titelbaum (2008), and Weintraub (2004).

[3]See, for example, Howard (2006), Walley (1991), and Williamson (1999).

[4]For a very similar argument against a more general indifference principle, see Weatherson (2005).

[5]See Lewis (2001).

[6]In Weatherson (2005); pp. 630–633, there is an argument that suggests that even FSBI may lead to violations of CA. But no halfer would be persuaded by this argument, since it assumes a highly controversial view of how the Principal Principle is to be applied, a view that any halfer would reject. This view is analogous to what I call "First Awakening Admissibility" in section 3.2, below.

[7]See Elga (2000); pp. 143–144.

[8]Indeed, quite apart from the intermediary step of the Generalized Thirder Principle, it is clear that the Frequency assumption gives rise to a violation of CA in the problem of Sleeping Beauty in St. Petersburg.

For in this problem, for every possible length of the sequence of coin tosses, the expected long-run frequency of awakenings that occur on trials in which the sequence of coin tosses has this length will be 0. And so, according to Frequency, for every hypothesis $h \in S$, upon first awakening, Beauty's credence in h must be 0.

[9]This application of Conditionalization assumes that $P(w_1)$ is greater than 0. But if this were not the case, then there would automatically be a violation of CA. For it follows from the thirder's assumptions that upon first awakening, Beauty's credence in w_1 should be at least as great as her credence in w_n, for any other value of n. Hence, if she has zero credence in the possibility that she has awoken exactly once, she must also have zero credence in every one of the countably infinite number of alternative times she may have awoken, thereby violating CA.

[10]See Arntzenius (2003), pp. 293–294; see also Dorr (2002), pp. 363–364.

[11]Arntzenius (2003) explicitly appeals to a principle of Commutativity in his argument. Dorr informs me on his current view, Commutativity will hold only on the assumption that Beauty knows that the universe contains infinitely many experiences, and that her current evidence is instantiated only in the present experiment. On Dorr's current view, it is only on these conditions that the one-third solution to the original Sleeping Beauty problem is precisely correct. Dorr grants, however, that in his original argument he assumed Commutativity.

[12]See, for example, Arntzenius (2002), Draper and Pust (2008), and Hitchcock (2004).

[13]In Arntzenius (2002), and in Bradley and Leitgeb (2006), it is argued that in Sleeping Beauty problems, Beauty's betting odds may come apart from her credences. Here I am making the different point that even if her betting odds for a bet on a given proposition should match her credence in this proposition upon being offered the bet, her betting odds may differ from the credence she had in this proposition prior to being offered the bet, since her being offered the bet may rationally alter her credence in this proposition.

[14]The connection between countable additivity and conglomerability was first discussed in De Finetti (1972); p. 99. See also Schervish et al. (1984).

[15]This example is based on one discussed in De Finetti (1972) and in Howson (2008).

[16]See, for example, Christensen (1991).

[17]Arntzenius et al. (2004); p. 251.

[18]Suppose we define a Dutch book, for a given agent, as a set of bets such that the agent regards each bet in this set as *fair* and taking every bet in this set would result in a sure loss. In this case, we can give a

synchronic Dutch book argument for *unrestricted* additivity, that is, for the principle that a rational agent's credences in the elements of any partition, countable or uncountable, must sum to her credence in their disjunction. We can avoid this problem, however, if we insist that the bets in a Dutch book be considered *favorable* by the agent to whom they are offered.

[19]Arntzenius et al. (2004); p. 262.

[20]Arntzenius et al. (2004); p. 264.

[21]See Joyce (1999), p. 151. For an early statement of a view of this kind, see Nozick (1969), p. 228.