Chapter 4

How to Discount Small Probabilities[*]

ABSTRACT: Maximizing expected value leads to counterintuitive choices in cases that involve tiny probabilities of huge payoffs. In response to such cases, some have argued that we ought to discount very small probabilities down to zero. In this chapter, I discuss how exactly this view can be formulated. I begin by showing that less plausible versions of discounting small probabilities violate dominance. Then, I show that more plausible formulations of this view avoid these dominance violations, but instead, they violate the axiom of Independence—and in a particularly counterintuitive way. As a result of this violation, those who discount small probabilities can be exploited by a money pump. Lastly, I discuss one possible way of avoiding exploitation by this money pump.

Orthodox decision theory claims that a rational agent always maximizes expected utility. However, this seems to imply counterintuitive choices in cases that involve very small probabilities of huge payoffs. In these cases, an option can be great in

[*]I wish to thank Tomi Francis, Andreas Mogensen, Teruji Thomas and the audience of the Global Priorities Institute seminar for valuable feedback and discussions.

expectation, even if the probability of obtaining a valuable outcome is tiny, as long as this valuable outcome is great enough. One example of such a case is *Pascal's Mugging*:[1]

> **Pascal's Mugging:** A stranger approaches Pascal and claims to be an Operator from the Seventh Dimension. He promises to perform magic that will give Pascal an extra thousand quadrillion happy days in the Seventh Dimension if he pays the mugger ten livres—money that the mugger will use for helping very many orphans in the Seventh Dimension.

Pascal thinks that the probability of the mugger telling the truth is very low. However, the potential payoff is so high that Pascal is forced to conclude that the expected utility of paying the mugger is positive. Furthermore, if Pascal gives a non-zero probability to the proposition that the mugger can reward him with any finite amount of utility, then the mugger can always increase the payoff until the offer has positive expected utility.[2] Consequently, maximizing expected utility (with unbounded utilities) requires paying the mugger—which seems counterintuitive.[3]

Another case that involves tiny probabilities of huge payoffs is the St. Petersburg game, a version of which was originally proposed by Nicolaus Bernoulli.[4] This

[1]Bostrom (2009). This case is based on informal discussions by various people, including Eliezer Yudkowsky (2007*b*).

[2]Contrary to this, see Hanson (2007), Yudkowsky (2007*a*) and Baumann (2009).

[3]This may not hold if utilities are bounded as standard axiomatizations of expected utility maximization (such as the von Neumann-Morgenstern utility theorem) require. See Kreps (1988, p. 63).

[4]The game was simplified by Gabriel Cramer in 1728 and published by Daniel Bernoulli in 1738. See Pulskamp (n.d.) and Bernoulli (1954).

game is played by flipping a fair coin until it lands on heads. The prize of this game is \$$2^n$, where n is the number of coin flips. This game has infinite expected monetary value, so agents who maximize expected monetary value would pay any finite amount to play the game. However, this seems counterintuitive.[5] Furthermore, if this game's (monetary) value is infinite, one would value it higher than any of its possible finite payoffs, which seems irrational.[6]

In response to cases like this, some have argued that we ought to discount very small probabilities down to zero—let's call this *Probability Discounting*. Nicolaus Bernoulli first proposed this idea in response to the St. Petersburg game. He writes: "[T]he cases which have a very small probability must be neglected and counted for nulls, although they can give a very great expectation. [...] This is a remark which merits to be well examined."[7] Recently, Smith (2014) and Monton (2019) have also defended the idea of Probability Discounting. Monton (2019) argues that one ought to discount very small probabilities down to zero, while Smith (2014) argues that it is rationally permissible—but not required—to do so.[8] However, we

[5] Pulskamp (n.d., p. 6). Daniel Bernoulli (cousin of Nicolaus Bernoulli) argues that, due to the diminishing marginal utility of money, one should not pay any finite sum to play the St. Petersburg game. See Bernoulli (1954). However, one can change the game slightly to bypass this objection by changing the prize from money to something with no diminishing marginal utility, such as perhaps days of life. See Monton (2019, p. 2). Relatedly, Menger (1967, pp. 217–218) shows that if utilities are unbounded, one can always create a *Super St-Petersburg game*, in which the payoffs grow sufficiently fast so that the expected utility of the game is infinite. See also Samuelson (1977, §2).

[6] Huemer (2016, pp. 34–35) and Russell and Isaacs (2021).

[7] Pulskamp (n.d., p. 2). Other proponents of Probability Discounting include, for example, Buffon and Condorcet. See Hey et al. (2010) and Monton (2019, pp. 16–17).

[8] Smith argues that discounting small probabilities down to zero is a way of getting a unique expected value for the Pasadena game. See Nover and Hájek (2004). See Hájek (2014) and Isaacs (2016) for criticism of discounting small probabilities. Also, see Beckstead (2013, ch. 6), Beckstead and Thomas (2020), Goodsell (2021), Russell and Isaacs (2021) and Russell (2021) for discussions

do not yet have a well-specified and plausible theory that tells us how to discount small probabilities. As Monton writes: "I don't have a perfectly rational, reasonable decision theory to hand you just yet (sorry)."[9]

This chapter discusses how Probability Discounting can be formulated and what the most plausible version of it might look like. §1 discusses a simple version of Probability Discounting on which one should conditionalize on outcomes associated with tiny probabilities not occurring. I show that this view faces a problem with individuating outcomes, and it also violates Statewise Dominance. §2 discusses a version of Probability Discounting that considers very-small-probability outcomes as tiebreakers when prospects would otherwise be equally good. I show that this view also violates Statewise Dominance. §3 discusses a version of Probability Discounting on which one should conditionalize on very-small-probability *states* not occurring. I discuss three ways of specifying this view. I show that one violates Stochastic Dominance and Acyclicity within choice sets, another violates Pairwise Acyclicity, Contraction and Expansion Consistency and Stochastic Dominance, and the last violates Statewise Dominance. §4 discusses more plausible versions of Probability Discounting that avoid the earlier violations of dominance and Acyclicity. §5 shows that these views violate the axiom of Independence—and

---

of related issues, and see Wilkinson (2022) for a defense of *Probability Fanaticism*:

> **Probability Fanaticism:** For any probability $p > 0$ and any finite utility u, there is some large enough utility U such that probability p of U (and otherwise nothing) is better than certainty of u.

In this context, 'otherwise nothing' means retaining the status quo or baseline outcome.

[9] Monton (2019, p. 15).

in a particularly counterintuitive way. As a result of this violation, those who discount small probabilities are vulnerable to exploitation by a money pump for Independence.[10] Lastly, §6 discusses one possible way of avoiding exploitation by this money pump. I conclude that Probability Discounting faces significant problems that undermine its plausibility as a theory of instrumental rationality.

1 Naive Discounting

This section discusses a version of Probability Discounting on which one should conditionalize on outcomes associated with tiny probabilities not occurring. However, I show that this view faces the *Outcome Individuation Problem*, and it also violates Statewise Dominance. Therefore, it is implausible as a theory of instrumental rationality.

According to Probability Discounting, an agent is rationally required or permitted to discount very small probabilities down to zero. On this view, there is some discounting threshold t such that probabilities below this threshold are discounted down to zero, but probabilities at least as great as this threshold are not discounted.[11,12] But when are probabilities small enough to be discounted? Or, as

[10] Isaacs (2016) also presents a problem for Probability Discounting in a dynamic context, to which Smith (2016) and Monton (2019) respond by arguing that relevantly similar choices ought to be evaluated collectively. This response is not available in the Independence Money Pump I will discuss later.

[11] Alternatively, this threshold probability t and probabilities below it are discounted, while the probabilities above t are not discounted. Note that this threshold might also be vague.

[12] Smith (2014) holds that the threshold might not apply to simple prospects, that is, prospects that assign a non-zero probability to only finitely many outcomes. Also, Smith does not argue for one universal threshold applicable in all situations. Instead, he maintains that this threshold may

Buffon writes: "[O]ne can feel that it is a certain number of probabilities that equals the moral certainty, but what number is it?"[13] Some possible discounting thresholds have been suggested. For Buffon and Condorcet, the discounting thresholds were 1/10,000 and 1/144,768 (respectively), while for Monton, this threshold is approximately 1 in 2 quadrillion.[14] As Monton argues, the discounting threshold is plausibly subjective. There is no objective answer to Buffon's question. Instead, it is up to each individual where the discounting threshold is.[15]

So, on this view, one should discount small probabilities—but small probabilities of *what*? This chapter discusses versions of Probability Discounting that ignore very-small-probability outcomes or states.[16] I will begin with the former views. There are many ways of ignoring outcomes associated with small probabilities. One way to ignore the very-small-probability outcomes of some prospect $\mathcal{P}_1$ would be to treat $\mathcal{P}_1$ as interchangeable with a prospect $\mathcal{P}_2$, which really does

be different in different situations.

[13] Hey et al. (2010, p. 256).

[14] Buffon's discounting threshold was the probability of a 56-year-old man dying in 24 hours—an outcome reasonable people typically ignore. See Monton (2019, pp. 8–9). Condorcet's discounting threshold was the difference between the probability that a 47-year-old man would die in one day and the probability that a 37-year-old man would. See Monton (2019, pp. 16–17). Monton's discounting threshold is between $1/2^{50}$ and $1/2^{51}$, as he treats the probability of getting tails at least 50 times in a row (with a fair coin) as rationally negligible. Monton (2019, p. 17).

[15] The subjectivity of the discounting threshold may be reasonable for individuals' rational preferences. However, it seems less so in the context of ethics when we are asking which prospects are better or worse.

[16] Whether one ignores very-small-probability outcomes or states makes a difference in some cases; a very-small-probability state might result in some outcome that overall has a non-negligible probability (when one also considers the other states). If one ignores very-small-probability states, one would discount down to zero (or at least decrease) the probability of this outcome. In contrast, if one ignores very-small-probability outcomes, one would not discount down to zero nor decrease the probability of this outcome.

assign probability zero to these outcomes.[17] However, $\mathcal{P}_2$ cannot assign the same probabilities as $\mathcal{P}_1$ to the remaining outcomes. Otherwise, the sum of all the probabilities assigned to outcomes of $\mathcal{P}_2$ would be less than one.[18] Instead, the probabilities assigned by $\mathcal{P}_2$ can be obtained from those assigned by $\mathcal{P}_1$ by conditionalizing on the supposition that some outcome of non-negligible probability occurs, where 'non-negligible' means a probability that is at least as great as the discounting threshold.[19]

Let $X \succsim Y$ mean that X is at least as preferred as Y. Also, let $EU(X)_{pd}$ denote the expected utility of prospect X when tiny probabilities have been discounted down to zero (read as 'the probability-discounted expected utility of X', where 'pd' stands for 'probability-discounted'). A *prospect* is taken to be a situation that may result in different outcomes with different probabilities. Then, one of the simplest versions of Probability Discounting—let's call it *Naive Discounting*—states:

> **Naive Discounting:** For all prospects X and Y, $X \succsim Y$ if and only if $EU(X)_{pd} \geq EU(Y)_{pd}$, where $EU(X)_{pd}$ and $EU(Y)_{pd}$ are obtained by conditionalizing on the supposition that some outcome of non-negligible probability occurs.[20]

On Naive Discounting, one should conditionalize on very-small-probability outcomes not occurring—but what counts as an 'outcome'? In particular, Naive

[17] Smith (2014, p. 478).

[18] Smith (2014, p. 478).

[19] Smith (2014, p. 478).

[20] Note that $EU(X)_{pd}$ and $EU(Y)_{pd}$ are obtained by conditionalizing, potentially, on different events not occurring.

Discounting faces the following problem:[21]

> **Outcome Individuation Problem:** If we individuate outcomes with too much detail, all outcomes have negligible probabilities. Is there a privileged way of individuating outcomes that avoids this?

The most obvious non-arbitrary way of individuating outcomes is by their utilities:[22]

> **Individuation by Preference:** Outcomes should be distinguished as different if and only if one has a preference between them.

Following this principle, each final utility level that a prospect might result in is considered a distinct outcome, and the possibilities of these outcomes are ignored if their associated probabilities are below the discounting threshold. For example, this view recommends against paying the mugger in Pascal's Mugging because the probability of obtaining an outcome as good as a thousand quadrillion happy days is very unlikely.

However, individuating outcomes by their utilities might result in ignoring all possible outcomes of some prospect if all its final utility levels are very unlikely.

[21] See also Beckstead and Thomas (2020, p. 13).

[22] If an agent is indifferent between winning $1 and eating an apple, on this view these would be considered the same outcome. Suppose the total probability of winning $1 and eating an apple is above the discounting threshold. In that case, these possibilities are not ignored, even if both winning $1 and eating an apple have negligible probabilities. Contrast Individuation by Preference with a similar principle presented by Broome (1991, p. 103):

> **Principle of Individuation by Justifiers:** Outcomes should be distinguished as different if and only if they differ in a way that makes it rational to have a preference between them.

In response to such cases, agents might lower their discounting thresholds until at least some outcomes have non-negligible probabilities. However, in cases where all outcomes have a zero probability, it is not possible to do so (except, of course, by not discounting at all).[23] Imagine, for example, an ideally shaped dart thrown on a dartboard, where each point results in a different utility. The probability that the dart hits a particular point may be zero. But one should not ignore every possible outcome of throwing the dart. Nevertheless, one might argue that we need not worry about cases where all outcomes have a zero probability because they are rare in practice. In all (or near all) cases we care about, some outcomes have non-zero probabilities.

Some might be satisfied with the above solution to the *Outcome Individuation Problem*. However, besides this problem, Naive Discounting also violates dominance. Let $X \succ Y$ mean that X is strictly preferred (or simply 'preferred') to Y. Then, Naive Discounting violates the following dominance principle:[24]

> **Statewise Dominance:** If the outcome of prospect X is at least as preferred as the outcome of prospect Y in all states, then $X \succsim Y$. Furthermore, if in addition the outcome of X is strictly preferred to the outcome of Y in some possible state, then $X \succ Y$.

Statewise Dominance is very plausible.[25] If some prospect is sure to turn out at

[23] Beckstead and Thomas (2020, pp. 12–13).

[24] Savage (1951, p. 58) and Luce and Raiffa (1957, p. 287).

[25] Russell (2021, p. 13) writes on (strict) Statewise Dominance: "What if Statewise Dominance fails? In that case, I'm not sure what we're doing when we compare how good prospects are. [...] [W]hat we ultimately care about is how well things turn out; choosing better prospects is supposed

least as well as another prospect, but it might turn out better, then that prospect should be better.

To see why Naive Discounting violates Statewise Dominance, consider the following prospects (see table 1):[26]

Naive Statewise Dominance Violation:

Prospect A Gives \$1,000,000 in state 1 and nothing in state 2.

Prospect B Gives nothing in both states.

Suppose the probability of state 1 is below the discounting threshold. After conditionalizing on the supposition that some outcome of non-negligible probability occurs, A is substituted by B. One would then be indifferent between A and B, even though the outcomes of A and B are equally good in state 2, but the outcome of A is better than the outcome of B in state 1. Therefore, Naive Discounting violates Statewise Dominance.

TABLE 1
A VIOLATION OF STATEWISE DOMINANCE

	State 1 $p <$ threshold	**State 2** $1 - p$
A	\$1,000,000	\$0
B	\$0	\$0

to guide us toward achieving better outcomes. In light of this, if dominance reasoning is wrong, then I don't want to be right. If A is sure to turn out better than B, then this tells us precisely the thing that betterness-of-prospects is supposed to be a guide to."

[26]Monton (2019, pp. 20–21) discusses a similar dominance violation. He proposes that Probability Discounting be supplemented with dominance. On discounting small probabilities and dominance violations, also see Isaacs (2016), Smith (2016), Lundgren and Stefánsson (2020, pp. 912–914) and Beckstead and Thomas (2020, §2.3).

To summarize, Naive Discounting states that one should conditionalize on not obtaining very-small-probability outcomes. This view faces the *Outcome Individuation Problem*, which can be solved by individuating outcomes by their utilities (except in cases where all outcomes have a zero probability). However, Naive Discounting also faces another problem: It violates Statewise Dominance. This undermines its plausibility as a theory of instrumental rationality.[27]

2 Lexical Discounting

This section discusses a version of Probability Discounting that treats very-small-probability outcomes as tiebreakers when prospects would otherwise be equally good. This view avoids the previous violation of Statewise Dominance. However, I show that it violates Statewise Dominance in another case.

There is a straightforward solution to the previous case. Probability Discounting can avoid the earlier violation of Statewise Dominance if outcomes whose probabilities are below the discounting threshold are treated as tiebreakers. Then, A is better than B because A and B have equal probability-discounted expected utility but, in addition, A gives a negligible probability of a positive outcome (while B does not). More generally, in tied cases, prospects can be compared by their ex-

[27] As shown in Chapter 2 of this thesis, Expected Utility Theory also violates Statewise Dominance, on pain of violating Continuity. Monton (2019, §7) argues that violations of Statewise Dominance should not count against Probability Discounting, given that Expected Utility Theory violates Statewise Dominance too. Later in §4, I discuss versions of Probability Discounting that do not violate Statewise Dominance.

pected utilities without any discounting (like Expected Utility Theory would do).

On this proposal, prospects are first ranked by their probability-discounted expected utilities. Then, in cases of ties, these prospects are ranked by their expected utilities without discounting small probabilities. Formally this view—let's call it *Lexical Discounting*—states the following:

> **Lexical Discounting:** For all prospects X and Y, $X \succsim Y$ if and only if
>
> - $EU(X)_{pd} > EU(Y)_{pd}$ or
> - $EU(X)_{pd} = EU(Y)_{pd}$ and $EU(X) \geq EU(Y)$,
>
> where $EU(X)_{pd}$ and $EU(Y)_{pd}$ are obtained by conditionalizing on the supposition that some outcome of non-negligible probability occurs.[28]

It is slightly misleading to say that Lexical Discounting is a form of discounting small probabilities down to zero because small probabilities and their associated utilities are considered in cases of ties. The outcomes whose probabilities are (at and) above the discounting threshold just take lexical priority over the very-small-probability outcomes.[29]

[28] As before, note that $EU(X)_{pd}$ and $EU(Y)_{pd}$ are obtained by conditionalizing, potentially, on different events not occurring.

[29] It might be argued that because some small probabilities are much smaller than others, one should have multiple discounting thresholds that form probability ranges, where higher probability ranges take lexical priority over the lower ones. Beckstead and Thomas (2020, p. 24 n. 19) point out that Probability Discounting faces some of the same problems as Probability Fanaticism if it uses very-small-probability outcomes as tiebreakers. Having multiple discounting thresholds may help probability discounters avoid these problems. For brevity, I will only discuss views on which there is just one discounting threshold.

However, Lexical Discounting also violates Statewise Dominance. To see how this violation happens, consider the following case (table 2):

Lexical Statewise Dominance Violation:

Prospect A Gives \$10 in states 1 and 2, \$100 in state 3, and nothing in state 4.

Prospect B Gives \$10 in state 1, \$100 in states 2 and 3, and nothing in state 4.

The only difference between these prospects is that A gives \$10 in state 2, while B gives \$100 in that same state. The probability of states 1 and 4 is 0.49, and the probability of states 2 and 3 is 0.01. For simplicity, let the discounting threshold be (implausibly) 0.03. Then, all probabilities below 0.03 should be discounted down to zero, while probabilities at least as great as 0.03 should not be discounted down to zero. Let's also assume that the utility of money equals the monetary amount.

TABLE 2
A VIOLATION OF STATEWISE DOMINANCE

	State 1	**State 2**	**State 3**	**State 4**
p	0.49	0.01	0.01	0.49
A	\$10	\$10	\$100	\$0
B	\$10	\$100	\$100	\$0

In this case, A gives a 0.5 probability of \$10 and a 0.01 probability of \$100 (and otherwise nothing). So, A's probability-discounted expected utility is $EU(A)_{pd} \approx 5.05$ after conditionalizing on not obtaining \$100 (as its associated probability is below the discounting threshold).[30] B in turn gives a 0.49 probability of \$10 and a 0.02 probability of \$100 (and otherwise nothing). B's probability-discounted expected utility is $EU(B)_{pd} = 5$ after conditionalizing on not obtaining \$100 with it.[31] Given that the former is greater than the latter, A is better than B according to Lexical Discounting. However, as mentioned above, the only difference between A and B is that A gives \$10 in state 2, while B gives \$100 in that same state. Therefore, Lexical Discounting—too—violates Statewise Dominance.

This violation of Statewise Dominance happens because when one conditionalizes on not obtaining \$100 with A (state 3), the probability of state 3 is divided between states 1, 2 and 4. However, when one conditionalizes on not obtaining \$100 with B (states 2 and 3), the probability of states 2 and 3 is divided between states 1 and 4. Therefore, the probability of obtaining nothing is greater with B than with A after ignoring the possibility of obtaining \$100.

To summarize, Lexical Discounting states that outcomes whose probabilities are (at or) above the discounting threshold take lexical priority over very-small-probability outcomes in determining prospects' betterness ranking—very-small-probability outcomes are only treated as tiebreakers. However, like Naive Discounting, Lexical Discounting also violates Statewise Dominance. This makes it a less plausible candidate for a theory of instrumental rationality.

[30] $0.5/(1 - 0.01) \cdot 10 \approx 5.05$.

[31] $0.49/(1 - 0.02) \cdot 10 = 5$.

3 State Discounting

This section discusses a version of Probability Discounting on which one should conditionalize on very-small-probability states not occurring. Three versions of this view are presented. I show that one violates Stochastic Dominance and Acyclicity within choice sets, another violates Pairwise Acyclicity, Contraction and Expansion Consistency and Stochastic Dominance, and the last one violates Statewise Dominance.

3.1 Pairwise and Set-Dependent State Discounting

Again, there is a straightforward solution to the previous violation of Statewise Dominance. Earlier it was assumed that one should ignore (except in cases of ties) the possibility of obtaining outcomes associated with tiny probabilities. However, one might instead ignore very-small-probability *states*—call this view *State Discounting*.[32] One can also make a lexical version of this view:

> **State Discounting** For all prospects X and Y, $X \succsim Y$ if and only if
>
> - $EU(X)_{pd} > EU(Y)_{pd}$ or
> - $EU(X)_{pd} = EU(Y)_{pd}$ and $EU(X) \geq EU(Y)$,
>
> where $EU(X)_{pd}$ and $EU(Y)_{pd}$ are obtained by conditionalizing on the supposition that no state of negligible probability occurs.

[32] This view naturally captures the idea that one should ignore very small *changes* in probabilities instead of very small (absolute) probabilities. Thus, it allows one to ignore the possibility of making a difference to some outcome if the probability of doing so is negligible. See §4 in Chapter 6 of this thesis.

State Discounting recommends against paying the mugger in Pascal's Mugging because the state in which the mugger delivers a thousand quadrillion happy days is very unlikely to occur. In the previous violation of Statewise Dominance, State Discounting tells one to ignore states 2 and 3 as their associated probabilities are below the discounting threshold. Consequently, A and B have equal probability-discounted expected utility (as they give the same outcomes in states 1 and 4). However, B has greater expected utility without discounting, so it is better than A (assuming a lexical version of State Discounting). Thus, State Discounting avoids the previous violation of Statewise Dominance.[33]

However, notice that State Discounting faces an analogous problem to the *Outcome Individuation Problem*, namely, the

> **State Individuation Problem:** If one individuates states with too much detail, all states have negligible probabilities. Is there a privileged way of individuating states that avoids this?

As before, a possible solution is to individuate states by the utilities of their outcomes.[34]

There are different views about how states should be partitioned. On another version of State Discounting, prospects are always compared two at a time, and the possible states of the world are partitioned for every pairwise comparison sepa-

[33] However, as I will show later, one version of State Discounting violates Statewise Dominance in this case.

[34] As before, one problem with this is that, in some cases, all states might have probabilities below the discounting threshold. One could lower the threshold in such cases. However, this will not solve the problem in cases where all states have a zero probability.

rately. Alternatively, one could compare all available options at once and partition the states for every choice set separately. Let's call these views *Pairwise State Discounting* and *Set-Dependent State Discounting*, respectively (the difference between these views is illustrated with an example later).

> **Pairwise State Discounting:** States are partitioned by comparing two prospects at a time.

> **Set-Dependent State Discounting:** States are partitioned by comparing all available prospects at once.

Although these views avoid the earlier violations of Statewise Dominance, they violate the following principle instead:

> **Acyclicity:** If $X_1 \succ X_2 \succ \cdots \succ X_n$, then it is not the case that $X_n \succ X_1$.

To see why these views violate Acyclicity, consider the following case:

> **Acyclicity Violation:** A random number generator returns a number between 1 and 100.
>
> *Prospect A* Gives \$1000 with numbers 1 and 2 (probability 0.02); otherwise, it gives nothing.
>
> *Prospect B* Certainly gives \$10 no matter what number comes up.
>
> *Prospect C* Gives \$1000 with number 1 (probability 0.01) and otherwise it gives \$1.

Let the discounting threshold be 0.02. First, compare A and B. Individuating states by the utilities of their outcomes results in two states as shown in table 3. A is better than B because neither state has a non-negligible probability, and A's expected utility is greater than that of B.[35] Next, compare B and C. In this case, individuating states by the utilities of their outcomes results in states shown in table 4. As the probability of state 1* is below the discounting threshold, one should ignore the possibility of state 1* occurring. Once one does that, B is better than C, as it gives a better outcome in state 2* ($10 vs. $1).

TABLE 3

A IS BETTER THAN B

Output	State 1 1 or 2 (p=0.02)	State 2 3 to 100 (p=0.98)
A	$1000	$0
B	$10	$10

TABLE 4

B IS BETTER THAN C

Output	State 1* 1 (p=0.01)	State 2* 2 to 100 (p=0.99)
B	$10	$10
C	$1000	$1

Now we have that A is better than B, which is better than C. It follows by Acyclicity that C is not better than A. However, when we compare A and C pairwise, we notice that C is better than A. In this case, individuating states by the utilities of their outcomes results in states shown in table 5. As states 1** and 2** have probabilities below the discounting threshold, the agent should ignore the possibilities of these states. Moreover, when the agent does that, C is better than A because it gives a better outcome in state 3**. So, we have a violation of Acyclicity: A is better than B, which is better than C, which is better than A.

[35] $EU(A)_{pd} = 0.02 \cdot 1000 = 20$ and $EU(B)_{pd} = 10$.

Table 5

C is better than A

Output	**State 1**** 1 (p=0.01)	**State 2**** 2 (p=0.01)	**State 3**** 3 to 100 (p=0.98)
A	\$1000	\$1000	\$0
C	\$1000	\$1	\$1

Let's now go back to Pairwise and Baseline State Discounting. If we partition states for each pair of options in a way that depends on the particular two options being compared (in line with Pairwise State Discounting), then State Discounting violates Acyclicity within choice sets. Consequently, it is not clear what one ought to choose when all A, B and C are available, as there is no most-preferred alternative.[36] However, if we partition states in a way that depends on the overall choice set (in line with Set-Dependent State Discounting), then there is no violation of Acyclicity within choice sets (see table 6). In this case, states 1*** and 2*** have probabilities below the discounting threshold, so one should ignore the possibilities of these states. And when one does that, B is the best prospect as it gives the best outcome in state 3***, and C is the second-best prospect as it gives a better outcome than A in that state.

However, Set-Dependent State Discounting violates Acyclicity across choice sets (as shown in tables 3, 4 and 5). In particular, it was shown that Set-Dependent State Discounting violates Pairwise Acyclicity, that is, it violates Acyclicity when we compare two options at a time (when each choice set only includes two options).

[36]Fishburn (1991, p. 116).

Table 6

No violation of Acyclicity

Output	**State 1******* 1 (p=0.01)	**State 2******* 2 (p=0.01)	**State 3******* 3 to 100 (p=0.98)
A	\$1000	\$1000	\$0
B	\$10	\$10	\$10
C	\$1000	\$1	\$1

It is odd that adding or removing options can influence which events one ignores. For example, when comparing A and B, Set-Dependent State Discounting does not ignore the possibility of the random number generator outputting number 1 or 2. However, when C is also available, Set-Dependent State Discounting ignores these possibilities. Consequently, the value of A decreases significantly when C is also available, as one then ignores the possibility of obtaining \$1000 with A.

This case shows that Set-Dependent State Discounting violates the following principles that many find plausible:[37]

> **Contraction Consistency:** For all prospects X and Y, if it is permissible to choose X from the set $\{X, \dots, Y\}$, then it is permissible to choose X from any subset of the set $\{X, \dots, Y\}$.

[37] Sen (1977, pp. 63–66). More generally, Contraction Consistency implies Acyclicity. Suppose that one violates Acyclicity. Then, there is a sequence of prospects such that $X_1 \succ X_2 \succ \dots \succ X_n \succ X_1$. Suppose that some prospect X_i is chosen from the choice set that includes all these prospects. Next, consider the choice set that includes only X_i and X_{i-1} (if $X_i = X_1$, then this choice set includes X_1 and X_n). Given that $X_{i-1} \succ X_i$, one would now choose X_{i-1} (or X_n if $X_i = X_1$). This is a violation of Contraction Consistency. Thus, if a view does not violate Contraction Consistency, then it does not violate Acyclicity. See Sen (1977, p. 67).

Strong Expansion Consistency: For all prospects X, Y and Z, if it is permissible to choose X from the set $\{X, \dots, Y\}$, then if it is permissible to choose Y from the set $\{X, \dots, Y, \dots, Z\}$, it is permissible to choose X from the set $\{X, \dots, Y, \dots, Z\}$.

Set-Dependent State Discounting violates Contraction Consistency as it is permissible (indeed rationally required) to choose B when all A, B and C are options. However, when only A and B are options, it is no longer permissible to choose B (because then one is rationally required to choose A as one no longer ignores the possibility of obtaining $1000 with A). On the other hand, Set-Dependent State Discounting violates Strong Expansion Consistency because it is permissible (indeed rationally required) to choose A when only A and B are options. But when A, B and C are options, it is permissible to choose B but not permissible to choose A.

Next, let $X = \{x_1,\ p_1; \dots;\ x_n,\ p_n\}$ stand for prospect X that gives non-zero probabilities $p_1, p_2, \dots, p_n$ of outcomes $x_1, x_2, \dots, x_n$. Then, both versions of State Discounting violate the following principle:[38]

Stochastic Dominance: Prospect $X = \{x_1,\ p_1; \dots;\ x_n,\ p_n\}$ is pre-

[38] Buchak (2013, p. 42). More precisely, the definition given here is for *first-order stochastic dominance*, an idea that was introduced to statistics by Mann and Whitney (1947) and Lehmann (1955), and to economics by Quirk and Saposnik (1962). The name 'first-degree stochastic dominance' is due to Hadar and Russell (1969, p. 27).

ferred to prospect $Y = \{y_1,\ q_1; \ldots;\ y_n,\ q_n\}$ if, for all outcomes o,

$$\sum_{\{i \mid x_i \succsim o\}} p_i \geq \sum_{\{j \mid y_j \succsim o\}} q_j,$$

and for some outcome u,

$$\sum_{\{i \mid x_i \succsim u\}} p_i > \sum_{\{j \mid y_j \succsim u\}} q_j.$$

A violation of Stochastic Dominance happens if, for all outcomes, some prospect X gives an at least as high probability of an at least as great outcome as some other prospect Y does, and for some outcome, X gives a greater probability of an at least as great outcome as Y does—yet Y is judged better than or equally as good as X.

To see why both versions of State Discounting violate Stochastic Dominance, consider the following case:

Two Coins:

Prospect A Gives $10 if a coin lands on heads (probability 0.5), nothing if it lands on tails (probability 0.49), and $100 if it lands on the edge (probability 0.01).

Prospect B Gives $10 if another coin lands on heads (probability 0.49), nothing if it lands on tails (probability 0.49), and $100 if it lands on the edge (probability 0.02).

Let the discounting threshold be 0.03. These prospects give the same probabilities of the same outcomes as the prospects in *Lexical Statewise Dominance Violation* (table 2), but instead of four states, we now have nine different states due to having two coins. Let 'H' stand for 'heads', 'T' for 'tails' and 'E' for 'edge'. Also, let '(X, Y)' stand for the first coin landing on 'X' and the second one on 'Y'. All states in which either coin lands on the edge have probabilities below the discounting threshold (given that the probabilities of either coin landing on the edge are alone already below the discounting threshold). Only four states have probabilities above the discounting threshold: (H, H), (H, T), (T, H) and (T, T) (see table 7).

TABLE 7
A VIOLATION OF STOCHASTIC DOMINANCE

	State 1	**State 2**	**State 3**	**State 4**
	H, H	H, T	T, H	T, T
p^*	*0.253*	*0.253*	*0.247*	*0.247*
A	\$10	\$10	\$0	\$0
B	\$10	\$0	\$10	\$0

p^*=probability conditional on one of states 1, 2, 3 or 4 occurring.

After conditionalizing on one of these four states occurring, the probability-discounted expected utility of A is greater than that of B: Now the only difference between these prospects is that A gives \$10 in state 2 (and nothing in state 3), while B gives \$10 in state 3 (and nothing in state 2), and state 2 has a greater probability than state 3.[39] Thus, A is better than B according to both versions of State Discounting. However, this is a violation of Stochastic Dominance. Before dis-

[39] A's probability-discounted expected utility is $EU(A)_{pd} \approx 5.05$. B's probability-discounted expected utility, in turn, is $EU(B)_{pd} = 5$.

counting, both A and B give a 0.51 probability of at least \$10, but B gives a greater probability of at least \$100 (0.02 vs. 0.01). So, for all outcomes, B gives an at least as high probability of an at least as great outcome as A does, and for some outcome, B gives a greater probability of an at least as great outcome as A does. Thus, B stochastically dominates A, and Pairwise and Set-Dependent State Discounting violate Stochastic Dominance as they claim that A is better than B.[40]

3.2 Baseline State Discounting

According to the previous versions of State Discounting, states might be partitioned differently depending on what other options are available. This leads to a violation of Acyclicity. However, states might also be partitioned in a way that does not depend on the other available options. This can be done by comparing each prospect to some baseline or status quo prospect—let's call this *Baseline State Discounting*.

> **Baseline State Discounting:** States are partitioned by comparing every prospect to a status quo prospect (each separately).[41]

[40]Note that the prospects in *Two Coins* are stochastically equivalent with the earlier prospects in *Lexical Statewise Dominance Violation*; both prospects give the same probabilities of the same outcomes. However, both Pairwise and Set-Dependent State Discounting judged B as better than A in the earlier case, but A as better than B in this case. Thus, on these views, the probabilities and the utilities associated with them are not the only determinants of the value of prospects. It also matters which states result in the different outcomes and what the probabilities of those states are. In general, Stochastic Equivalence and Statewise Dominance imply Stochastic Dominance. See Russell (2021, §2).

[41]On Baseline State Discounting, one might sometimes ignore some events e_1 and e_2 when comparing some prospect X to the status quo prospect, but not ignore them when comparing another prospect Y to the status quo prospect. This can happen if both the status quo prospect and prospect Y result in the same outcome with e_1 as with e_2, but prospect X results in a different

According to this view, one should ignore the very-small-probability states of some prospect X when states are partitioned by comparing X to a baseline or status quo prospect, which corresponds to doing nothing.

However, Baseline State Discounting violates Statewise Dominance in the same way as Lexical Discounting does. Consider again *Lexical Statewise Dominance Violation* (table 2). This time, let's specify the events that result in each outcome:

Random Number: A random number generator returns a number between 1 and 100.

Prospect A Gives \$10 if a number between 1 and 50 is returned, \$100 if number 51 is returned, and nothing if a number between 52 and 100 is returned.

Prospect B Gives \$10 if a number between 1 and 49 is returned, \$100 if 50 or 51 is returned, and nothing if a number between 52 and 100 is returned.

In this case, the baseline prospect is (presumably) certainly getting nothing. When A is compared to this baseline prospect, state individuation by utilities results in three states as shown in table 8. As the probability of state 2 is below the discounting threshold of 0.03, the possibility of this state is ignored. Consequently, the

outcome with e_1 than with e_2. Consequently, e_1 and e_2 result in two different states when prospect X is compared to the status quo, but only one state when Y is compared to the status quo. If the combined probability of these states is above the discounting threshold, but the probabilities of these states taken individually are below the discounting threshold, then e_1 and e_2 will get ignored with prospect X but not with prospect Y.

probability-discounted expected utility of A is $EU(A)_{pd} \approx 5.05$.[42]

Table 8
A vs. the Baseline

Output	**State 1** 1–50 (p=0.5)	**State 2** 51 (p=0.01)	**State 3** 52–100 (p=0.49)
A	$10	$100	$0
Baseline	$0	$0	$0

Next, compare B to the baseline prospect. This time state individuation by utilities results in the three states shown in table 9. Again, the probability of state 2* is below the discounting threshold, so the possibility of this state is ignored. Then, the probability-discounted expected utility of B is $EU(B)_{pd} = 5$.[43]

Table 9
B vs. the Baseline

Output	**State 1*** 1–49 (p=0.49)	**State 2*** 50 or 51 (p=0.02)	**State 3*** 52–100 (p=0.49)
B	$10	$100	$0
Baseline	$0	$0	$0

A's probability-discounted expected utility is greater than that of B (5.05 vs. 5), so A is better than B. However, B statewise dominates A because the only difference between these prospects is that A gives $10 if the random number generator returns the number 50, while B gives $100 in that case. Consequently, Base-

[42] $EU(A)_{pd} = 0.5/0.99 \cdot 10 \approx 5.05$.

[43] $EU(B)_{pd} = 0.49/0.98 \cdot 10 = 5$.

line State Discounting violates Statewise Dominance when states are partitioned in the usual way corresponding to possible states of the world (such as 'number 50 is returned'). This violation of Statewise Dominance happens because the possible states of the world that Baseline State Discounting ignores are not the same for every prospect. For example, when comparing A to the baseline prospect, the possibility of the random number generator returning the number 50 is not ignored, but when B is compared to the baseline prospect, this possibility is ignored (tables 8 and 9).

To summarize, instead of ignoring very-small-probability outcomes, Probability Discounting might ignore very-small-probability states. State Discounting faces the *State Individuation Problem*, which can be solved by individuating states by the utilities of their outcomes. I have discussed three ways of formulating State Discounting. Pairwise State Discounting always compares two options at a time, even if the choice set includes other options as well. It ignores very-small-probability states in every pairwise comparison. However, Pairwise State Discounting violates Stochastic Dominance and Acyclicity within choice sets. Set-Dependent State Discounting compares all available options simultaneously and ignores very-small-probability states in every choice set. This view violates Pairwise Acyclicity, Contraction and Expansion Consistency and Stochastic Dominance. Finally, Baseline State Discounting ignores very-small-probability states of some prospect X, when states are partitioned by comparing X to a baseline prospect. This view violates Statewise (and hence also Stochastic) Dominance. To conclude, all three versions of State Discounting violate plausible principles of rationality, which undermines their plausibility as theories of instrumental rationality.[44]

4 Stochastic and Tail Discounting

This section discusses more plausible versions of Probability Discounting that avoid the earlier violations of dominance (and Acyclicity). However, §5 shows that these views violate the axiom of Independence and are therefore vulnerable to exploitation by a money pump.

4.1 Stochastic Discounting

One version of Probability Discounting—let's call it *Absolutist Stochastic Discounting*—works like this: To obtain the probability-discounted expected utility of a prospect, first add the lowest possible positive utility, weighted by the probability of getting at least that much utility.[45] Next, add the difference between the lowest utility and the next lowest utility, weighted by the probability of getting at least the higher amount of utility. Then, add the difference between this utility and the next lowest utility, weighted by the probability of getting at least that much utility, and so on until the next probability is below the discounting threshold.[46] Then, ignore the

[44]Someone might adopt a view on which one should first filter one's options by Statewise and Stochastic Dominance and then choose following some version of Probability Discounting from amongst the remaining options. This view avoids the dominance violations, but it also seems *ad hoc*. However, some may find the benefit of a greater fit with our intuitions worth the cost in terms of simplicity.

[45]Note that this view requires an objective zero point on the utility-scale.

[46]This is similar to an alternative way of calculating the expected utility of a prospect discussed by Buchak (2014, p. 1100).

rest of the utility levels (whose probabilities are below the discounting threshold). Negative utilities are then treated similarly, and their expectation is summed with the expectation of positive utilities to obtain the value of a prospect. (This is equivalent to calculating the probability-discounted expected utility of a prospect in the same way as Expected Utility Theory would calculate expected utilities with the following exception: The greatest positive and negative utilities [whose utility levels have negligible cumulative probability] have been replaced by the greatest positive or negative utility whose utility level has a non-negligible cumulative probability [respectively for positive and negative utilities]).[47]

According to Absolutist Stochastic Discounting, there is an objective neutral level. On this view, one should ignore the possibility of very high or very low utility levels when the cumulative probability of ending up with at least or at most that much utility (respectively for positive and negative utilities) is negligible. This view recommends against paying the mugger in Pascal's Mugging if there is only a tiny probability that one will get an outcome at least as good as a thousand quadrillion happy days. However, it does not recommend against paying the mugger if there is a non-negligible probability of obtaining an outcome that is at least as great as a

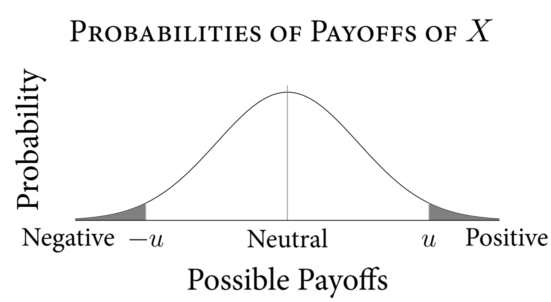
[47]That is, the following formula for calculating the probability-discounted expected utility of positive outcomes of prospect X is equivalent to the formulation given later (see *Positive outcomes*):

$$EU(X)_{pd,\ pos} = \sum_{i=1}^{m} p(E_i)u(x_i) + \Big(\sum_{i=m+1}^{n} p(E_i)\Big)u(x_m),$$

where x_m is the greatest positive utility that has a non-negligible cumulative probability, and x_n is the greatest positive utility possible with prospect X. (Negative utilities are treated similarly.)

thousand quadrillion happy days for some reason unrelated to the mugger's offer.[48]

Suppose that some prospect X has possible outcomes whose values are normally distributed with a mean of zero (when the outcomes are ordered in terms of betterness). Absolutist Stochastic Discounting tells one to ignore the highest positive and the lowest negative utility levels of X. This is equivalent to substituting the values of the outcomes in the grey areas (see the graph below) with the values of u and $-u$ (respectively for positive and negative outcomes), where u and $-u$ are the best positive and the worst negative utility levels that have non-negligible cumulative probabilities.[49] For example, suppose X gives a negligible probability p of $u + \epsilon$ utility. Then, when calculating the probability-discounted expected utility of X, $u + \epsilon$ is substituted with u (that is, the contribution of a p chance of $u + \epsilon$ utility to the probability-discounted expected utility of X is $p \cdot u$.).

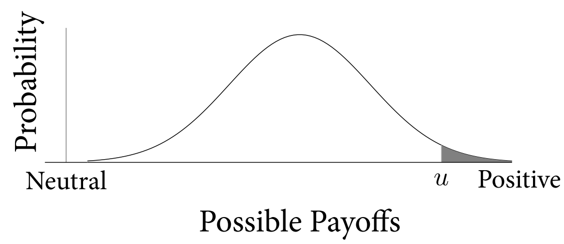


[48]For example, an agent who thinks there is a non-negligible probability of going to Heaven would not ignore the possibility of a great payoff in Pascal's Mugging. More generally, such an agent would not discount small probabilities very often (if ever); the non-negligible probability of going to Heaven makes it the case that there is a non-negligible probability of ending up with at least u amount of utility for all positive values of u.

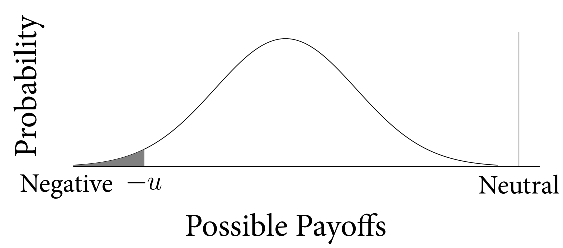
[49]One is not simply ignoring the possibilities of the outcomes in the grey areas because their probabilities contribute to the cumulative probabilities of the utility levels with lower magnitudes.

The following prospects Y and Z can only result in positive or negative outcomes, respectively. Consequently, Absolutist Stochastic Discounting tells one to ignore Y's highest positive utility levels and Z's lowest negative utility levels. This is equivalent to substituting the values of the best positive outcomes of Y (the grey area in the left image below) with u and the values of the worst negative outcomes of Z (the grey area in the right image below) with $-u$ where u and $-u$ are the best positive and the worst negative utility levels of Y and Z (respectively) that have non-negligible cumulative probabilities.

Probabilities of Payoffs of Y
(Certain positive outcome)



Probabilities of Payoffs of Z
(Certain negative outcome)



Call the versions of Probability Discounting that have the same structure as Absolutist Stochastic Discounting *Stochastic Discounting*. Besides Absolutist Stochastic Discounting, there is another way of understanding Stochastic Discounting.

This view is similar to Baseline State Discounting because it compares each prospect to a baseline prospect—so it might be called *Baseline Stochastic Discounting*. On this view, one calculates the amount by which the baseline/status quo utility level is increased or decreased by the different possible outcomes of a prospect. Then, to obtain the probability-discounted expected utility of a prospect, one first adds the lowest possible gain (i.e., positive change to the baseline), weighted by the probability of gaining at least that much. Next, one adds the difference between the lowest gain and the next lowest gain, weighted by the probability of gaining at least the higher amount. Then, one adds the difference between this gain and the next lowest gain, weighted by the probability of gaining at least that much, and so on, until the next probability is below the discounting threshold. Then, one ignores the rest of the possible gains (whose probabilities are below the discounting threshold). Losses (i.e., negative changes to the baseline) are then treated similarly, and their expectation is summed with the expectation of gains to obtain the value of a prospect.[50]

Unlike the previous version of Stochastic Discounting, this view recommends against paying the mugger in Pascal's Mugging, even if there is a non-negligible probability of gaining an equally great payoff for some reason unrelated to the mugger's offer. This is so because once one has 'subtracted' the baseline prospect from

[50]One can also make a version of Stochastic Discounting that is analogous to Pairwise State Discounting in that it compares prospects to other available prospects pairwise—call this *Pairwise Stochastic Discounting*. On this view, one considers the utility difference in each state between two prospects and ignores the largest differences when the cumulative probability of states with differences at least that large is negligible. (Again, one does not entirely ignore these differences because the probabilities of these differences contribute to the cumulative probability of lesser differences.)

the mugger's offer, gains at least as great as a thousand quadrillion happy days have a negligible cumulative probability.

On both versions of Stochastic Discounting, the probability-discounted expected utility of positive outcomes is calculated as follows (here 'positive outcomes' are either gains if one accepts Baseline Stochastic Discounting or final utilities if one accepts Absolutist Stochastic Discounting):[51]

Positive outcomes: For all prospects X, such that X gives non-zero probabilities of positive outcomes
$X_{pos} = \{E_1,\ x_1;\ E_2,\ x_2; \dots;\ E_m,\ x_m; \dots;\ E_n,\ x_n\}$, and
$0 < u(x_1) \leq \dots \leq u(x_m) \leq \dots \leq u(x_n)$, the probability-discounted expected utility of positive outcomes of X is

$$\begin{aligned} EU(X)_{pd,\ pos} = &\Big(\sum_{i=1}^{n} p(E_i)\Big)u(x_1) + \Big(\sum_{i=2}^{n} p(E_i)\Big)(u(x_2) - u(x_1)) \\ &+ \Big(\sum_{i=3}^{n} p(E_i)\Big)(u(x_3) - u(x_2)) \\ &+ \dots + \Big(\sum_{i=m}^{n} p(E_i)\Big)(u(x_m) - u(x_{m-1})), \end{aligned}$$

[51] Technically, this formulation requires the following qualifications: If all probabilities of positive utility levels are non-negligible, then in order to obtain $EU(X)_{pd,\ pos}$, one simply sums up the positive utilities weighted by their probabilities (without discounting). And if all probabilities of positive utility levels are negligible, then $EU(X)_{pd,\ pos} = 0$ (on Baseline Stochastic Discounting) or the value of the baseline (on Absolutist Stochastic Discounting). Furthermore, this formula assumes that, amongst the possible positive utility levels, there is one that is the lowest. But this may not always be true. Consider for example a St. Petersburg-style prospect in which the possible payoff halves with each additional coin flip (i.e., it gives probability 1/2 of 2 utilities, probability 1/4 of one utility, probability 1/8 of 0.5 utilities and so on.) One can calculate the probability-discounted expected utility of such prospects as discussed on p. 159, that is, the same way as Expected Utility Theory would do with the following exception: The greatest positive utilities (whose utility levels have negligible cumulative probability) are replaced by the greatest positive utility whose utility level has a non-negligible cumulative probability.

where

$$\sum_{i=m}^{n} p(E_i) \geq t > \sum_{i=m+1}^{n} p(E_i),$$

where t is the discounting threshold.

The probability-discounted expected utility of prospect X is then obtained by summing the probability-discounted expected utilities of its positive and negative outcomes:[52]

$$EU(X)_{pd} = EU(X)_{pd,\ pos} + EU(X)_{pd,\ neg}.$$

Stochastic Discounting can use very-small-probability utility levels as tiebreak-

[52]The probability-discounted expected utility of negative outcomes is calculated as follows:

Negative outcomes: For all prospects X, such that X gives non-zero probabilities of negative outcomes $X_{neg} = \{E_{-1},\ x_{-1};\ E_{-2},\ x_{-2}; ...;\ E_{-m},\ x_{-m}; ...;\ E_{-n},\ x_{-n}\}$, and $0 > u(x_{-1}) \geq ... \geq u(x_{-m}) \geq ... \geq u(x_{-n})$, the probability-discounted expected utility of negative outcomes of X is

$$\begin{aligned} EU(X)_{pd,\ neg} =& \Big(\sum_{i=-1}^{-n} p(E_i)\Big)u(x_{-1}) + \Big(\sum_{i=-2}^{-n} p(E_i)\Big)(u(x_{-2}) - u(x_{-1})) \\ &+ \Big(\sum_{i=-3}^{-n} p(E_i)\Big)(u(x_{-3}) - u(x_{-2})) \\ &+ \cdots + \Big(\sum_{i=-m}^{-n} p(E_i)\Big)(u(x_{-m}) - u(x_{-m+1})), \end{aligned}$$

where

$$\sum_{i=-m}^{-n} p(E_i) \geq t > \sum_{i=-m-1}^{-n} p(E_i),$$

where t is the discounting threshold.

ers to rank prospects with equal probability-discounted expected utility. It can then be stated as follows:

Stochastic Discounting: For all prospects X and Y, $X \succsim Y$ if and only if

- $EU(X)_{pd} > EU(Y)_{pd}$ or
- $EU(X)_{pd} = EU(Y)_{pd}$ and $EU(X) \geq EU(Y)$,

where, for all prospects X, it holds that

$$EU(X)_{pd} = EU(X)_{pd,\ pos} + EU(X)_{pd,\ neg}.$$

The following case illustrates the difference between these versions of Stochastic Discounting:[53]

Absolutist vs. Baseline Stochastic Discounting:

Prospect A Gives a 0.01 probability of $-\$1000$ (the agent loses \$1000) and otherwise \$10.

Prospect B Certainly gives \$1.

Let the discounting threshold be 0.02. Absolutist and Baseline Stochastic Discounting treat this case similarly if the agent does not have any money when facing this choice. Both versions imply that the agent should ignore the possibility of losing

[53]Note that in this example prospects are defined in terms of monetary gains and losses rather than final consequences, such as wealth levels.

\$1000 with A. Consequently, A is better than B because its probability-discounted expected utility is greater.[54]

Next, suppose the agent already possesses \$2000. Then, A gives a 0.01 probability of ending up with \$1000 and a 0.99 probability of ending up with \$2010. With B, the agent certainly ends up with \$2001. Baseline Stochastic Discounting treats this case similarly as the case where the agent does not have any money on their bank account. In contrast, Absolutist Stochastic Discounting calculates the values of these prospects using the amounts of money the agent could end up with. So, on this view, the agent ought *not* ignore the possibility of losing \$1000 with A; if they lose \$1000, then they will end up with \$1000 overall, and the probability of ending up with \$1000 or more is 1. So, according to Absolutist Stochastic Discounting, B is better than A because because its probability-discounted expected utility is $EU(B)_{pd} = 2001$, while A's probability-discounted expected utility is $EU(A)_{pd} \approx 2000$.[55]

Absolutist Stochastic Discounting has the (possible) disadvantage of requiring an objective neutral utility level. Baseline Stochastic Discounting does not require one because it ignores very large changes to the baseline—the baseline serves the same purpose as the objective neutral level on the absolutist view. Also, Absolutist Stochastic Discounting entails that sometimes one might not ignore the possibility of a huge loss even if there is only a tiny probability of it occurring. This can happen when the agent ends up with a positive outcome regardless of the loss and the

[54] $EU(A)_{pd} = 0.99 \cdot 10 = 9.9$ and $EU(B)_{pd} = 1$.

[55] $EU(A)_{pd} = 1000 + 0.99(2010 - 1000) \approx 2000$.

probability of obtaining an outcome that is at least as good as that is non-negligible. Similarly, it can also happen if the probability of ending up with a worse utility level is non-negligible for some reason not related to the prospect in question. So, Absolutist Stochastic Discounting sometimes lets tiny probabilities of huge losses dictate one's course of action (and similarly for payoffs). Therefore, it does not capture the motivation behind Probability Discounting as well as Baseline Stochastic Discounting does.[56]

Now, recall the earlier violations of Statewise and Stochastic Dominance (*Lexical Statewise Dominance Violation*, *Random Number* and *Two Coins*):

> *Prospect A* Gives a 0.5 probability of \$10 and a 0.01 probability of \$100 (and otherwise nothing).
>
> *Prospect B* Gives a 0.49 probability of \$10 and a 0.02 probability of \$100 (and otherwise nothing).

[56]It is also worth pointing out the following features of Absolutist Stochastic Discounting: First, egoistic agents who are offered the "same" prospects (when one ignores the baseline utility levels) and have the same discounting threshold can reach different conclusions about which option is best. This can happen when they have different baseline utility levels, because one of these agents might end up with an overall positive utility level in a state where the other agent ends up with a negative one. Secondly, if Absolutist Stochastic Discounting takes into account past value, then what happened in the past can influence what altruistic agents now ought to do. For example, if one learns that the past was much better than one thought, then the overall moral value of the world would be much higher. Consequently, one might no longer ignore some tiny probability of a large loss because even if the loss occurred, the value of the world would still be positive (and there is a non-negligible probability of obtaining an outcome that is at least as good as that). This is similar to the Egyptology objection to the Average View in population ethics. See McMahan (1981, p. 115) and Parfit (1984, p. 420). Also, even if one only takes into account future value (or value in one's future light cone), what happens in distant places can affect what altruistic agents here ought to do. Wilkinson (2022, §6) shows that views that reject Probability Fanaticism must violate separability or Stochastic Dominance. Absolutist Stochastic Discounting violates the former. Given that Baseline Stochastic Discounting ignores background uncertainty (and thus satisfies separability), it must sometimes violate Stochastic Dominance.

Again, let the discounting threshold be 0.03. Unlike the earlier versions of Probability Discounting, both versions of Stochastic Discounting imply that B is better than A. B gives a 0.51 probability of at least \$10 and a 0.02 probability of at least \$100, so its probability-discounted expected utility is $EU(B)_{pd} = 5.1$.[57] A in turn gives a 0.51 probability of at least \$10 and a 0.01 probability of at least \$100, so its probability-discounted expected utility is also $EU(A)_{pd} = 5.1$.[58] As A and B have equal probability-discounted expected utility, these prospects are then compared by their expected utilities without discounting. Consequently, B is better than A, and both versions of Stochastic Discounting avoid the earlier violations of Statewise and Stochastic Dominance.

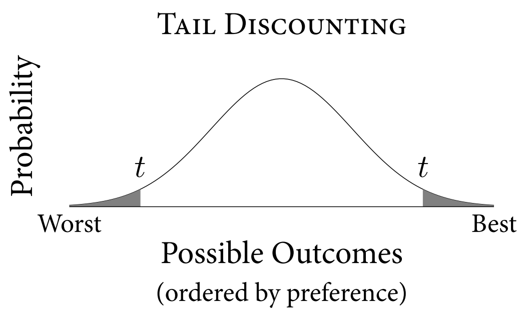
4.2 Tail Discounting

There is a similar view to Absolutist Stochastic Discounting called *Tail Discounting*.[59] According to Tail Discounting, one should ignore both the left and the right 'tails' of the distribution of possible outcomes of some prospect X when these outcomes are ordered by one's preference. Suppose the possible outcomes of some prospect are normally distributed when they are ordered from the least to the most preferred. Then, Tail Discounting advises one to ignore the grey areas under the curve:

[57] $EU(B)_{pd} = 0.51 \cdot 10 = 5.1$.

[58] $EU(A)_{pd} = 0.51 \cdot 10 = 5.1$.

[59] Beckstead and Thomas (2020, §2.3).



Call the outcomes that fall in the middle of the distribution of possible outcomes 'normal outcomes'. An outcome is normal if and only if there is a non-negligible probability of getting at least and at most as good an outcome.[60] Tail Discounting then states the following:

> **Tail Discounting:** For all prospects X and Y, $X \succsim Y$ if and only if
>
> - $EU(X)_{pd} > EU(Y)_{pd}$ or
> - $EU(X)_{pd} = EU(Y)_{pd}$ and $EU(X) \geq EU(Y)$,
>
> where $EU(X)_{pd}$ and $EU(Y)_{pd}$ are obtained by conditionalizing on the supposition that some normal outcome occurs.[61]

[60] For example, in the St. Petersburg game, all payoffs up to some large payoff o are normal, where o depends on one's discounting threshold.

[61] Formally this view states the following:

> **Tail Discounting (formal):** In order to determine $EU(X)_{pd}$, first order the possible outcomes of some prospect X from the least to the most preferred. Then, conditionalize on obtaining some outcome in the middle part of the distribution such that the following necessary conditions hold for all outcomes o that are not ignored:
>
> i The probability of obtaining an outcome that is at least as good as o is above the discounting threshold and
>
> ii the probability of obtaining an outcome that is at most as good as o is above the discounting threshold.
>
> If some outcome o fulfills the above necessary conditions, and

Tail Discounting has the advantage over Absolutist Stochastic Discounting that it does not require an objective neutral level. However, similarly to Absolutist Stochastic Discounting, Tail Discounting recommends paying the mugger in Pascal's Mugging if there is a non-negligible probability of obtaining an outcome at least as great as a thousand quadrillion happy days. This is because then a thousand quadrillion happy days falls in the middle part of the distribution of possible outcomes, which is not ignored.[62]

Again, recall the earlier violations of Statewise and Stochastic Dominance (*Lexical Statewise Dominance Violation*, *Random Number* and *Two Coins*). Tail Discounting also implies that B is better than A. B gives nothing with a 0.49 probability, \$10 with a 0.49 probability and \$100 with a 0.02 probability. Consequently, its probability-discounted expected utility is $EU(B)_{pd} \approx 5.1$.[63] A in turn gives

---

- the probability of obtaining an outcome that is better than o is below the discounting threshold, then decrease the probability of obtaining o until the total discounted probability of outcomes that are at least as good as o equals the discounting threshold (and conditionalize to make sure the remaining probabilities add up to 1), and
- if the probability of obtaining an outcome that is worse than o is below the discounting threshold, then decrease the probability of obtaining o until the total discounted probability of outcomes that are at most as good as o equals the discounting threshold (and conditionalize to make sure the remaining probabilities add up to 1).

[62] One might also make a version of Tail Discounting similar to Baseline Stochastic Discounting. On this view—let's call it *Baseline Tail Discounting*—one compares every prospect to a baseline prospect as follows: First, calculate the difference in utilities a prospect makes in each state of the world (compared to the baseline prospect). Then, order these differences from the greatest loss to the greatest gain. Then, ignore the right and left tails of this distribution by conditionalization. Also, one can make a version of Tail Discounting similar to Pairwise Stochastic Discounting (i.e., *Pairwise Tail Discounting*). On this view, one compares prospects pairwise instead of comparing every prospect to a baseline prospect.

[63] $EU(B)_{pd} = (0.49 - 0.01)/0.94 \cdot 10 \approx 5.1$. The divisor '0.94' comes from subtracting the

nothing with probability 0.49, \$10 with probability 0.5 and \$100 with probability 0.01. So, its probability-discounted expected utility is also $EU(A)_{pd} \approx 5.1$.[64] As A and B have equal probability-discounted expected utility, these prospects are then compared by their expected utilities without discounting. B has greater expected utility than A without discounting, so B is better than A—and Tail Discounting avoids the earlier violations of Statewise and Stochastic Dominance.

To summarize, I have discussed three versions of Probability Discounting in this section. Absolutist Stochastic Discounting states that one should ignore the possibility of a very high (or a very low) utility level in cases where the cumulative probability of such utility levels is below the discounting threshold. Baseline Stochastic Discounting works similarly, but it operates on gains and losses instead of final utilities. Finally, Tail Discounting states that one should ignore the 'tails' of the distribution of possible outcomes of some prospect when these outcomes are ordered from the least to the most preferred. All these views avoid the earlier violations of Statewise and Stochastic Dominance (and Acyclicity). However, next, I will raise a diachronic problem for these views.

---

discounting threshold of 0.03 from both tails of the distribution. '0.01' is subtracted from 0.49 to make sure that the discounting threshold of 0.03 is ignored in the right tail as well; the probability of obtaining \$100 is 0.02, so merely ignoring the possibility of \$100 would mean one ignores the '0.02 part' of the right tail. More generally, on Tail Discounting, one discounts a little bit of each 'tail' with every prospect (until the discounting threshold is ignored from both tails). See footnote 61.

[64] $EU(A)_{pd} = (0.5 - 0.02)/0.94 \cdot 10 \approx 5.1$.

5 Independence

This section shows that Stochastic and Tail Discounting violate the axiom of Independence.[65] As a result of this violation, these views are vulnerable to exploitation by a money pump. In the next section, I discuss one possible way of avoiding exploitation by this money pump.

5.1 A violation of Independence

Both Stochastic and Tail Discounting violate the axiom of Independence. Let XpY be a risky prospect with a p chance of prospect X obtaining and a $1-p$ chance of prospect Y obtaining. Then, Independence states that

> **Independence:** If $X \succ Y$, then $XpZ \succ YpZ$ for all probabilities $p \in (0, 1]$.[66]

Informally, Independence is the idea that every outcome contributes to the value of a prospect in a way that does not depend on the alternative outcomes.

The basic problem for Probability Discounting is that by mixing gambles, one can arbitrarily reduce the probabilities of different states or outcomes within the compound lottery until these probabilities end up below the discounting threshold. Therefore, mixtures of gambles can end up being valued differently than the gambles that are mixed together. For example, consider the following case:

[65] As these views differ from Expected Utility Theory, they must violate at least one of its axioms. In addition to violating Independence, they also violate Continuity. See §1 in Chapter 5 of this thesis.

[66] Jensen (1967, p. 173).

A Violation of Independence:

Prospect A Certainly gives nothing.

Prospect B Gives a 0.5 probability of $1 and otherwise −$1,000,000.

Prospect C Certainly gives $1.

Next, let $p = 0.02$. Then, we have the following mixed prospects (see table 10):

Prospect ApC Gives a 0.98 probability of $1 and otherwise nothing.

Prospect BpC Gives a 0.99 probability of $1 and a 0.01 probability of −$1,000,000.

TABLE 10
A VIOLATION OF INDEPENDENCE

p	0.01	0.01	0.98
ApC	$0	$0	$1
BpC	−$1,000,000	$1	$1

First, consider what Stochastic Discounting says about these prospects (Baseline and Absolutist Stochastic Discounting treat this case similarly if the agent possesses nothing when making this choice). Let the discounting threshold be 0.02. ApC gives a 0.98 probability of gaining at least $1 (and otherwise nothing), so its probability-discounted expected utility is $EU(ApC)_{pd} = 0.98$. BpC, in turn, gives a 0.99 probability of gaining at least $1 and a 0.01 probability of losing at least $1,000,000. The probability of losing at least $1,000,000 is below the discounting threshold, so this possibility is ignored. Thus, BpC's probability-discounted expected utility is $EU(BpC)_{pd} = 0.99$. So, according to Stochastic Discounting, BpC is better than ApC, given that its probability-discounted expected utility is greater than ApC's.

However, the difference between them is that ApC gives a 0.02 probability of nothing, while BpC gives a 0.01 probability of gaining $1 and a 0.01 probability of losing $1,000,000 instead (columns 1 and 2 in table 10). Note that BpC is better than ApC no matter how bad the negative outcome is (in column 1) as long as the good outcome (in column 2) is at least slightly positive.

Next, consider what Tail Discounting says about these prospects. Now let the discounting threshold be 0.01. Then, Tail Discounting also implies that BpC is better than ApC. After ignoring both tails of the distribution of possible outcomes of ApC, it probability-discounted expected utility is $EU(ApC)_{pd} \approx 0.99$.[67] And after ignoring the tails of the distribution of possible outcomes of BpC, its probability-discounted expected utility is $EU(BpC)_{pd} = 1$.[68] Thus, we have that BpC is better than ApC.

Some might consider this implication already worrisome on its own, but it is also a violation of Independence—and there is a money pump against theories that violate Independence.[69] Both Stochastic and Tail Discounting consider A better than B. It is better to get nothing certainly than to take a 50–50 gamble between gaining $1 and losing $1,000,000. Thus, we have the following violation of Inde-

[67] $(0.98 - 0.01)/0.98 \cdot 1 \approx 0.99$.

[68] $(0.99 - 0.01)/0.98 \cdot 1 = 1$.

[69] See Gustafsson (forthcoming, §5).

pendence:

$A \succ B$, and $BpC \succ ApC$ for some probability $p \in (0, 1]$.

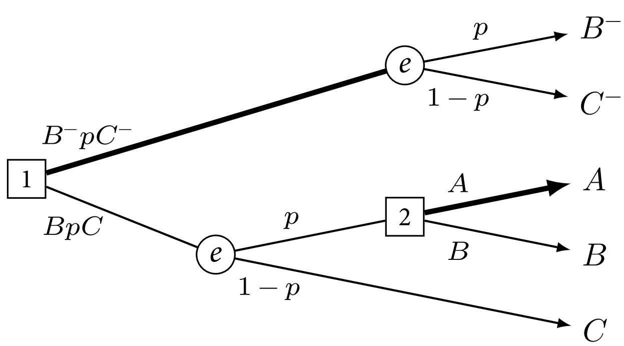
This renders Stochastic and Tail Discounting vulnerable to exploitation by a money pump for Independence.

5.2 The Independence Money Pump

A money-pump argument intends to show that agents who violate some alleged requirement of rationality are vulnerable to making a combination of choices that leads to a sure loss. If vulnerability to this kind of exploitation is a sign of irrationality, then Stochastic and Tail Discounting are untenable as theories of instrumental rationality.

Consider the following decision problem:[70]

THE INDEPENDENCE MONEY PUMP



$A \succ B$, and $BpC \succ B^-pC^- \succ ApC$.

[70]This money pump is from Gustafsson (2021, p. 31 n21). Also see Gustafsson (forthcoming, §5), Hammond (1988*b*, pp. 292–293) and Hammond (1988*a*, pp. 43–45).

In this decision tree, the squares represent choice nodes, and the circles represent chance nodes. Going up at a choice node means accepting a trade, and going down means refusing a trade.[71] The agent starts with prospect BpC: a 0.99 probability of \$1 and a 0.01 probability of $-$\$1,000,000. At node 1, they are offered a trade from BpC to $B^{-}pC^{-}$, which is like BpC except that the agent has ϵ less money. If the agent turns down this trade and BpC results in the agent going up at the chance node e, the agent will be offered a trade from B to A at node 2. Both chance nodes depend on the same chance event e.

An agent can use *backward induction* to reason about this case. This means that they consider what they would choose at later choice nodes and take those predictions into account when making choices at earlier choice nodes.[72] As the agent prefers A to B, they would accept the trade at node 2. They would rather certainly get nothing than take a 50–50 gamble between gaining \$1 and losing \$1,000,000. Then, by using backward induction at node 1, the prospect of turning down the trade is effectively ApC, and the prospect of accepting the trade is $B^{-}pC^{-}$. Given that the agent prefers BpC over ApC, there must be some price they would be willing to pay to get the former rather than the latter. So, there is some ϵ amount of money such that the agent prefers $B^{-}pC^{-}$ over ApC. Then, for some ϵ, they go up at node 1. However, they then end up with $B^{-}pC^{-}$, even though they could have kept BpC had they gone down at both choice nodes.[73] They have given up money

[71] Rabinowicz (2008, p. 152).

[72] Selten (1975) and Rosenthal (1981, p. 95).

[73] Note that if one accepts a baseline or pairwise version of Stochastic or Tail Discounting, then the fact that the agent starts with BpC might matter. For example, if BpC is considered the baseline prospect, then according to the baseline versions, the value of ApC is calculated by comparing

for the exploiter.

Furthermore, as the chance nodes depend on the same event, the prospect of going up at node 1 is statewise dominated by the prospect of going down at both choice nodes. Let a *plan* be a specification of a sequence of choices to be taken by an agent at each choice node that can be reached from that node while following this specification. Stochastic and Tail Discounting advise that an agent make a plan that results in a worse outcome in every state than another available plan. This is a violation of a sequential version of Statewise Dominance:

> **Sequential Statewise Dominance:** If the outcome of plan X is at least as preferred as the outcome of plan Y in all states, and the outcome of X is strictly preferred to the outcome of Y in some possible state, then $X \succ Y$.

Moreover, by choosing $B^{-}pC^{-}$, the agent has paid to give up their power to choose A rather than B if event e occurs. The agent has therefore paid for having their freedom of choice taken away from them.[74]

To summarize, Stochastic and Tail Discounting violate Independence—and in a particularly counterintuitive way.[75] The violation of Independence is particu-

it to BpC in every state. In the state in which event e does not happen, both prospects result in C. Thus, the value of ApC in that state is zero. Compared to BpC, ApC gives a 0.01 probability of gaining a million, a 0.01 probability of losing \$1, and otherwise, it gives nothing. Consequently, its probability-discounted expected utility is $EU(ApC)_{pd} = 0$. These prospects are then compared by their expected utilities without discounting. Consequently, ApC is better than BpC—and we have avoided the Independence violation. However, if the agent does not start with BpC but instead is offered BpC, then they would choose $B^{-}pC^{-}$ for the reasons explained in the main text. They have therefore chosen a dominated prospect.

[74] Rabinowicz (2021, pp. 530–531).

[75] Stochastic and Tail Discounting also violate the following version of Independence:

larly counterintuitive because BpC is considered better than ApC no matter how bad the negative outcome (losing $1,000,000) is. In addition, this implication renders those who accept these views vulnerable to exploitation in the Independence Money Pump. Next, I will discuss one possible way of avoiding exploitation in this case.

6 Avoiding exploitation in the Independence Money Pump

Those who accept Stochastic or Tail Discounting might avoid exploitation in the Independence Money Pump if they use policies of decision-making that prevent dynamic inconsistency. One such decision policy is Resolute Choice.[76] However, I will show that using Resolute Choice in the Independence Money Pump leads to

---

Independence for Constant Outcome: For all probabilities $p \in (0, 1]$, $XpU \succ YpU$ if and only if $YpV \succ XpV$ (McClennen, 1990, p. 45).

In addition to the earlier prospects ApC and BpC, consider the following prospects:

A Violation of Independence for Constant Outcome:

Prospect ApD Gives a 0.98 probability of −$1,000,000 and otherwise nothing.

Prospect BpD Gives a 0.99 probability of −$1,000,000 and a 0.01 probability of $1.

Both Stochastic and Tail Discounting imply that BpC is better than ApC, but ApD is better than BpD. For example, according to Stochastic Discounting, $EU(ApD)_{pd} = 0.98 \cdot (-1,000,000) = -980,000$ and $EU(BpD)_{pd} = 0.99 \cdot (-1,000,000) = -990,000$ (with a discounting threshold of 0.02). There is also a money-pump argument against preferences like this. See Gustafsson (forthcoming, §5) and Raiffa (1968, pp. 83–85).

[76]Backward induction was initially proposed as a decision policy to avoid exploitation. However, as we saw earlier, backward induction got one in trouble in the Independence Money Pump.

untenable results. Furthermore, I will argue that even if there is a way of avoiding exploitation in the Independence Money Pump, Stochastic and Tail Discounting cannot escape the untenable implication they face if combined with Resolute Choice. This makes them (and Probability Discounting more generally) less plausible as theories of instrumental rationality.

A resolute agent chooses in accordance with any plan they have adopted earlier as long as nothing unexpected has happened since the adoption of the plan.[77] Resolute Choice violates Decision-Tree Separability, which states that what is rational at a choice node does not depend on what has happened in the past—only the future matters. With Resolute Choice, one can commit to choosing BpC at node 1 of the Independence Money Pump and then stick to that plan. One then makes a plan that one will not trade B for A at node 2, even though one would usually prefer the latter prospect over the former one. But this seems wrong. Choosing B over A would mean choosing a 0.5 probability of losing $1,000,000 and otherwise gaining $1 over certainty of nothing. No reasonable view recommends this.

However, someone might object that the whole point of Resolute Choice is that, by adhering to a plan, the agent makes choices that they would view as unreasonable if they occurred outside the scope of the plan as stand-alone decisions. Therefore, the agent agrees that no reasonable view sanctions the choice if the choice happens outside a plan. Their view is that such a choice can be reasonable when licensed by adhering to the best available plan. However, choosing a 0.5 probabil-

[77] Strotz (1955-1956) and McClennen (1990, pp. 12–13). See Steele (2007), Steele (2018) and Gustafsson (forthcoming, §7) for criticism of Resolute Choice.

ity of losing \$1,000,000 and otherwise gaining \$1 over certainty of nothing is beyond the scope of what is reasonable even for someone who is resolute. We might also change the example so that one loses arbitrarily much instead of losing just \$1,000,000. Furthermore, the probability of this loss can be arbitrarily close to 1.[78] It would not be rational to commit to choosing that prospect over certainty of nothing. So, Resolute Choice is untenable in combination with Stochastic and Tail Discounting as a solution to the Independence Money Pump.

Stochastic and Tail Discounting violate the axiom of Independence in a particularly counterintuitive way. This case is, therefore, worrisome independently of the exploitation. The violation of Independence is particularly counterintuitive because BpC is considered better than ApC no matter how bad the negative outcome ($-$\$1,000,000) is as long as the good outcome (\$1) is at least slightly positive.

Suppose that at node 1 of the Independence Money Pump, the agent is offered another option: to lock in their choice at node 2 without knowing whether e has happened. Stochastic and Tail Discounting would recommend locking in the choice of B because then, at node 1, the agent faces BpC, which is better than ApC and $B^{-}pC^{-}$. However, this seems wrong. First, this would mean that the agent willingly avoids costless information by locking in their choice at node 2 without knowing whether e has happened.[79] More importantly, they would lock in the

[78]For example, let BpC be a prospect that gives a $0.02 - \epsilon$ probability of losing arbitrarily much; otherwise, it gives \$1 (probability $0.98 + \epsilon$). Then, BpC would still be better than ApC because it gives a higher probability of \$1. However, prospect B would almost certainly give an arbitrarily large loss and only a small probability of \$1.

[79]More generally, agents who violate Independence avoid costless information. See for example Wakker (1988), Hilton (1990) and Machina (1989, p. 1638–1639).

choice of a lottery that gives a 0.5 probability of $-\$1{,}000{,}000$ and otherwise \$1 (B) over certainty of nothing (A). Limiting one's future choices in this way seems irrational. Even if the agent does not accept Resolute Choice, they would still lock in the same choice of B over A if offered the option at node 1. This makes Probability Discounting less plausible even if some decision policy helps probability discounters avoid exploitation in the Independence Money Pump. Even absent exploitation, choosing B over A, or locking in the choice of B over A, seems irrational.[80]

To summarize, those who accept Stochastic or Tail Discounting might be able to avoid exploitation in the Independence Money Pump if they use policies of decision-making that prevent dynamic inconsistency. I have argued that these views give untenable recommendations if combined with Resolute Choice. I also argued that even if there is a way of avoiding exploitation in the Independence Money Pump), Stochastic and Tail Discounting cannot avoid the untenable result they face if combined with Resolute Choice. This makes them—and Probability Discounting more generally—less plausible as theories of instrumental rationality.[81]

[80]However, some might argue that saying that it is irrational to lock in B over A in this context simply amounts to saying that it is irrational to prefer BpC to ApC—which is begging the question.

[81]What should one do now? One could, for example, bite the bullet and accept one version of Probability Discounting discussed in this chapter, find a more plausible version of Probability Discounting, bound utilities, conditionalize on one's knowledge before maximizing expected utility (see for example Francis and Kosonen [n.d.]) or accept Probability Fanaticism (see for example Beckstead and Thomas [2020] and Wilkinson [2022]). However, note that, independently of Probability Discounting, agents with unbounded utilities are also vulnerable to money pumps because they violate countable generalizations of the Independence axiom. See Russell and Isaacs (2021).

7 Conclusion

Maximizing expected utility implies counterintuitive choices in cases that involve tiny probabilities of huge payoffs. In response to such cases, some have argued that we should deviate from Expected Utility Theory by discounting small probabilities to zero. I have discussed how exactly this view can be formulated. First, I argued that less plausible versions of Probability Discounting violate dominance. More specifically, I showed that Naive Discounting, Lexical Discounting and Baseline State Discounting violate Statewise Dominance. I also showed that Pairwise State Discounting violates Stochastic Dominance and Acyclicity within choice sets and that Set-Dependent State Discounting violates Pairwise Acyclicity, Contraction and Expansion Consistency and Stochastic Dominance.

Then, I showed that more plausible versions of Probability Discounting, namely Stochastic Discounting and Tail Discounting, avoid these dominance violations. However, they violate the axiom of Independence and do so in a particularly counterintuitive way. As a result of this violation, those who accept these views can be exploited in the Independence Money Pump. I then argued that these views cannot use Resolute Choice to avoid exploitation because this would have untenable implications. Lastly, I argued that even if there is a way of avoiding exploitation in the Independence Money Pump, Stochastic and Tail Discounting cannot avoid the untenable result they face if combined with Resolute Choice. This makes them—and Probability Discounting more generally—less plausible as theories of instrumental rationality. All in all, I have discussed possible ways of formulating Probability

Discounting. All of these theories have significant problems, and it is yet to be seen whether there is a perfectly rational, reasonable decision theory that deviates from Expected Utility Theory by discounting small probabilities down to zero.

References

Baumann, P. (2009), 'Counting on numbers', *Analysis* **69**(3), 446–448.

Beckstead, N. (2013), On the overwhelming importance of shaping the far future, PhD thesis, Rutgers, the State University of New Jersey.

Beckstead, N. and Thomas, T. (2020), A paradox for tiny probabilities and enormous values. Unpublished manuscript. Global Priorities Institute Working Paper No.10.
URL: *https://globalprioritiesinstitute.org/nick-beckstead-and-teruji-thomas-a-paradox-for-tiny-probabilities-and-enormous-values/*

Bernoulli, D. (1954), 'Exposition of a new theory on the measurement of risk', *Econometrica* **22**(1), 23–36.

Bostrom, N. (2009), 'Pascal's Mugging', *Analysis* **69**(3), 443–445.

Broome, J. (1991), *Weighing Goods: Equality, Uncertainty and Time*, Blackwell, Oxford.

Buchak, L. (2013), *Risk and Rationality*, Oxford University Press, Oxford.

Buchak, L. (2014), 'Risk and tradeoffs', *Erkenntnis* **79**(6), 1091–1117.

Fishburn, P. C. (1991), 'Nontransitive preferences in decision theory', *Journal of Risk and Uncertainty* **4**(2), 113–134.

Francis, T. and Kosonen, P. (n.d.), Ignore outlandish possibilities. Unpublished manuscript.

Goodsell, Z. (2021), 'A St Petersburg paradox for risky welfare aggregation', *Analysis* **81**(3), 420–426.

Gustafsson, J. E. (2021), 'The sequential dominance argument for the Independence axiom of Expected Utility Theory', *Philosophy & Phenomenological Research* **103**(1), 21–39.

Gustafsson, J. E. (forthcoming), *Money-Pump Arguments*, Cambridge University Press, Cambridge.

Hadar, J. and Russell, W. R. (1969), 'Rules for ordering uncertain prospects', *The American Economic Review* **59**(1), 25–34.

Hájek, A. (2014), 'Unexpected expectations', *Mind* **123**(490), 533–567.

Hammond, P. J. (1988*a*), 'Consequentialist foundations for expected utility', *Theory and Decision* **25**(1), 25–78.

Hammond, P. J. (1988*b*), 'Orderly decision theory: A comment on professor Seidenfeld', *Economics and Philosophy* **4**(2), 292–297.

Hanson, R. (2007), ‘Pascal’s Mugging: Tiny probabilities of vast utilities’.

URL: *https://www.lesswrong.com/posts/a5JAiTdytou3Jg749/pascal-s-mugging-tiny-probabilities-of-vast-utilities?commentId=Q4ACkdYFEThA6EE9P*

Hey, J. D., Neugebauer, T. M. and Pasca, C. M. (2010), Georges-Louis Leclerc de Buffon’s ‘Essays on moral arithmetic’, *in* A. Sadrieh and A. Ockenfels, eds, ‘The Selten School of Behavioral Economics: A Collection of Essays in Honor of Reinhard Selten’, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 245–282.

Hilton, R. W. (1990), ‘Failure of Blackwell’s Theorem under Machina’s generalization of expected-utility analysis without the independence axiom’, *Journal of Economic Behavior & Organization* **13**(2), 233–244.

Huemer, M. (2016), *Approaching Infinity*, Palgrave Macmillan, New York.

Isaacs, Y. (2016), ‘Probabilities cannot be rationally neglected’, *Mind* **125**(499), 759–762.

Jensen, N. E. (1967), ‘An introduction to Bernoullian utility theory: I. Utility functions’, *The Swedish Journal of Economics* **69**(3), 163–183.

Kreps, D. M. (1988), *Notes on the Theory of Choice*, Westview Press, Boulder.

Lehmann, E. L. (1955), ‘Ordered families of distributions’, *The Annals of Mathematical Statistics* **26**(3), 399–419.

Luce, R. D. and Raiffa, H. (1957), *Games and Decisions: Introduction and Critical Survey*, Wiley, New York.

Lundgren, B. and Stefánsson, H. O. (2020), 'Against the *De Minimis* principle', *Risk Analysis* **40**(5), 908–914.

Machina, M. J. (1989), 'Dynamic consistency and non-expected utility models of choice under uncertainty', *Journal of Economic Literature* **27**(4), 1622–1668.

Mann, H. B. and Whitney, D. R. (1947), 'On a test of whether one of two random variables is stochastically larger than the other', *The Annals of Mathematical Statistics* **18**(1), 50–60.

McClennen, E. F. (1990), *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge University Press, Cambridge.

McMahan, J. (1981), 'Problems of population theory', *Ethics* **92**(1), 96–127.

Menger, K. (1967), The role of uncertainty in economics, *in* M. Shubik, ed., 'Essays in Mathematical Economics: In Honor of Oskar Morgenstern', Princeton University Press, Princeton, pp. 211–231.

Monton, B. (2019), 'How to avoid maximizing expected utility', *Philosophers' Imprint* **19**(18), 1–24.

Nover, H. and Hájek, A. (2004), 'Vexing expectations', *Mind* **113**(450), 237–249.

Parfit, D. (1984), *Reasons and Persons*, Clarendon Press, Oxford.

Pulskamp, R. J. (n.d.), Correspondence of Nicolas Bernoulli concerning the St. Petersburg game. Unpublished manuscript. Accessed through:

https://web.archive.org/.

URL: *http://cerebro.xu.edu/math/Sources/NBernoulli/correspondence_petersburg_game.pdf*

Quirk, J. P. and Saposnik, R. (1962), 'Admissibility and measurable utility functions', *The Review of Economic Studies* **29**(2), 140–146.

Rabinowicz, W. (2008), Pragmatic arguments for rationality constraints, *in* M. C. Galavotti, R. Scazzieri and P. Suppes, eds, 'Reasoning, Rationality, and Probability', CSLI, Stanford, CA, pp. 139–163.

Rabinowicz, W. (2021), Between sophistication and resolution—wise choice, *in* R. Chang and K. Sylvan, eds, 'The Routledge Handbook of Practical Reason', Routledge, Abingdon, Oxon ; New York, NY, pp. 526–540.

Raiffa, H. (1968), *Decision Analysis: Introductory Lectures on Choices under Uncertainty*, Addison–Wesley, Reading, MA.

Rosenthal, R. W. (1981), 'Games of perfect information, predatory pricing and the chain-store paradox', *Journal of Economic Theory* **25**(1), 92–100.

Russell, J. S. (2021), 'On two arguments for fanaticism'. Global Priorities Institute Working Paper 17–2021.

URL: *https://globalprioritiesinstitute.org/on-two-arguments-for-fanaticism-jeff-sanford-russell-university-of-southern-california/*

Russell, J. S. and Isaacs, Y. (2021), 'Infinite prospects', *Philosophy and Phenomenological Research* **103**(1), 178–198.

Samuelson, P. A. (1977), 'St. petersburg paradoxes: Defanged, dissected, and historically described', *Journal of Economic Literature* **15**(1), 24–55.

Savage, L. J. (1951), 'The theory of statistical decision', *Journal of the American Statistical Association* **46**(253), 55–67.

Selten, R. (1975), 'Reexamination of the perfectness concept for equilibrium points in extensive games', *International Journal of Game Theory* **4**(1), 25–55.

Sen, A. (1977), 'Social choice theory: A re-examination', *Econometrica* **45**(1), 53–88.

Smith, N. J. J. (2014), 'Is evaluative compositionality a requirement of rationality?', *Mind* **123**(490), 457–502.

Smith, N. J. J. (2016), 'Infinite decisions and rationally negligible probabilities', *Mind* **125**(500), 1199–1212.

Steele, K. (2007), Precautionary Decision-Making: An Examination of Bayesian Decision Norms in the Dynamic Choice Context, PhD thesis, University of Sydney.

Steele, K. (2018), Dynamic decision theory, *in* S. O. Hansson and V. F. Hendricks, eds, 'Introduction to Formal Philosophy', Springer, Cham, pp. 657–667.

Strotz, R. H. (1955-1956), 'Myopia and inconsistency in dynamic utility maximization', *The Review of Economic Studies* **23**(3), 165–180.

Wakker, P. (1988), 'Nonexpected utility as aversion of information', *Journal of Behavioral Decision Making* **1**(3), 169–175.

Wilkinson, H. (2022), 'In defence of fanaticism', *Ethics* **132**(2), 445–477.

Yudkowsky, E. (2007*a*), 'A comment on Pascal's Mugging: Tiny probabilities of vast utilities'.
URL: *https://www.lesswrong.com/posts/a5JAiTdytou3Jg749/pascal-s-mugging-tiny-probabilities-of-vast-utilities?commentId=kqAKXskjohx4SSyp4*

Yudkowsky, E. (2007*b*), 'Pascal's Mugging: Tiny probabilities of vast utilities'.
URL: *http://www.overcomingbias.com/2007/10/pascals-mugging.html*