Chapter 5

*Probability Discounting and Money Pumps**

ABSTRACT: In response to cases that involve very small probabilities of huge payoffs, some argue that we ought to discount very small probabilities down to zero. However, this chapter shows that doing so violates Independence and Continuity, and as a result of these violations, those who discount small probabilities can be exploited by money pumps. Various possible ways of avoiding exploitation will be discussed. However, echoing the previous chapter, this chapter concludes that the money pump for Independence undermines the plausibility of discounting small probabilities.

On the standard decision theory, a rational agent always maximizes expected utility. However, this seems to lead to counterintuitive choices in cases that involve very small probabilities of huge payoffs. Consider, for example, the following case:[1]

*I wish to thank Andreas Mogensen and Teruji Thomas for valuable feedback and discussions.

[1]Bostrom (2009). This case is based on informal discussions by different people, including Eliezer Yudkowsky (2007). Another case that involves very small probabilities of huge payoffs is the St. Petersburg game. See for example Peterson (2020).

Pascal's Mugging: Someone approaches Pascal and claims to be an Operator from the Seventh Dimension. The stranger promises to perform magic that will give Pascal a thousand quadrillion happy days in the Seventh Dimension if Pascal pays the mugger ten livres—money that the mugger will use for helping orphans in the Seventh Dimension.

Pascal thinks the probability of the mugger telling the truth is very low. However, the potential payoff is so high that the expected utility of paying the mugger is positive. Furthermore, as long as Pascal has a non-zero credence in the proposition that the mugger is able and willing to reward him with any finite amount of utility, the mugger can increase the payoff until the offer has positive expected utility.[2] At some point, maximizing expected utility (with unbounded utilities) requires paying the mugger. And more generally, it leads to

Probability Fanaticism: For any tiny probability $p > 0$, and for 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.[3]

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*. For example, 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 re-

[2]This may not be possible if utility is bounded as standard axiomatizations of expected utility maximization require. See for example Kreps (1988, p. 63) and §1 and §2.1 in Chapter 1 of this thesis.

[3]Wilkinson (2022, p.449).

quired, to do so.[4] There are many ways of making Probability Discounting precise. 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 small probabilities have been discounted down to zero (read as 'the probability-discounted expected utility of X'). Also, let a *negligible probability* be a probability below the discounting threshold, that is, a probability that should be discounted down to zero. Then, one of the simplest versions of Probability Discounting—let's call it *Naive Discounting*—states:[5]

> **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.

Given that Probability Discounting differs from Expected Utility Theory, it has to violate at least one of the following axioms that together entail Expected Utility Theory: Completeness, Transitivity, Independence and Continuity.[6] Furthermore, violating these axioms renders probability discounters vulnerable to exploitation

[4] Smith argues that discounting small probabilities allows one to get a reasonable expected utility for the Pasadena game (see [Nover and Hájek 2004]). On Smith's view, the discounting threshold could be chosen lower than any relevant probability in cases that involve finitely many possible outcomes. So, in effect, discounting small probabilities might not apply to cases involving a finite number of possible outcomes. See Hájek (2014), Isaacs (2016) and Lundgren and Stefánsson (2020) for criticism of discounting small probabilities. Also see Beckstead (2013, ch. 6), Beckstead and Thomas (2020), Goodsell (2021), Russell and Isaacs (2021), Russell (2021) and Wilkinson (2022) for discussions of issues surrounding Probability Fanaticism.

[5] See Chapter 4 for a discussion of some possible versions of Probability Discounting.

[6] von Neumann and Morgenstern (1947), Jensen (1967, pp. 172–182) and Hammond (1998, pp. 152–164). This chapter assumes the von Neumann-Morgenstern framework with its lotteries with given probabilities, rather than the Savage framework, where subjective probabilities must be constructed alongside utilities, requiring the use of a different and more expansive set of axioms.

as there are money-pump arguments for each of these axioms.[7]

This chapter shows that some versions of Probability Discounting, such as Naive Discounting, violate Independence and Continuity. They are therefore vulnerable to exploitation in the money pumps for Independence and Continuity.[8] Here's the structure of the chapter: §1 discusses three ways in which Probability Discounting might violate Continuity. This section also shows that probability discounters are vulnerable to exploitation in a money pump for Continuity. Lastly, it discusses some ways of avoiding exploitation in that case. §2 shows that Probability Discounting violates Independence. As a result, probability discounters are vulnerable to exploitation in a money pump for Independence. §3 discusses possible ways of avoiding exploitation in the Independence Money Pump. It concludes that there is no plausible way to do this. The chapter concludes that the Independence Money Pump greatly undermines the plausibility of Probability Discounting.

1 Continuity

This section discusses three ways in which Probability Discounting might violate Continuity. First, it shows that views that discount probabilities below some discounting threshold violate Continuity. Next, it shows that views that discount

[7]Gustafsson (forthcoming). It has also been argued that even agents who conform to Expected Utility Theory can be exploited in some cases with an infinite series of trade offers. Gustafsson (forthcoming, §8) argues that such agents can avoid exploitation if they use backward induction.

[8]Isaacs (2016) also presents a problem for probability discounters 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 does not help avoid exploitation in the cases discussed in this chapter.

probabilities up to some discounting threshold violate another version of Continuity. Finally, it shows that views that ignore very-small-probability *outcomes* must violate either Continuity or Statewise Dominance.[9] As a result of violating Continuity, Probability Discounting is vulnerable to exploitation in a money pump for Continuity. Some ways of avoiding exploitation in this money pump will be discussed.

1.1 The Continuity Money Pump

As mentioned earlier, Continuity is one of the axioms that together entail Expected Utility Theory. Let $X \succ Y$ mean that X is strictly preferred (or simply 'preferred') to Y.[10] Also, let XpY be a risky prospect with a p chance of prospect X obtaining and a $1 - p$ chance of prospect Y obtaining. Continuity then states the following:

> **Continuity:** If $X \succ Y \succ Z$, then there are probabilities p and $q \in (0, 1)$ such that $XpZ \succ Y \succ XqZ$.

Views that discount probabilities below some threshold violate Continuity. To see how the Continuity violation happens, consider the following prospects:[11]

Continuity Violation:

[9] Instead of ignoring very-small-probability outcomes, one might ignore very-small-probability *states*. See §3 in Chapter 4 of this thesis on State Discounting.

[10] Some prospect X is strictly preferred to another prospect Y when X is weakly preferred to Y, but Y is not weakly preferred to X.

[11] Naive Discounting, Lexical Discounting, State Discounting, Stochastic Discounting and Tail Discounting (discussed in Chapter 4 of this thesis) all violate Continuity in this case (if the discounting threshold is the lowest probability not discounted down to zero).

Prospect A_t Gives probability t of some very good outcome (and otherwise nothing).

Prospect B Certainly gives a good outcome.

Prospect C Certainly gives nothing.

Let t be the discounting threshold. Then, all probabilities less than t will be discounted down to zero, but probabilities at least as great as t will not be discounted. Also, suppose that A_t is better than B, which is better than C; a non-negligible probability of a very good outcome (and otherwise nothing) is better than a certain good outcome, which is better than certainly getting nothing.

Next, consider the following mixed lottery (see table 1):

Prospect A_tpC Gives probability p of A_t and probability $1 - p$ of C (i.e., probability $t \cdot p$ of a very good outcome and otherwise nothing).

Given that t is the discounting threshold, t multiplied by any probability $p < 1$ must be below the discounting threshold. Consequently, $t \cdot p$ is discounted down to zero, and A_tpC only gives a negligible probability of a positive outcome. And, given that B certainly gives a good outcome, B must be better than A_tpC for all probabilities $p \in (0, 1)$. So, now we have that A_t is better than B, which is better than C, but B is better than A_tpC for all probabilities $p \in (0, 1)$—which is a violation of Continuity.

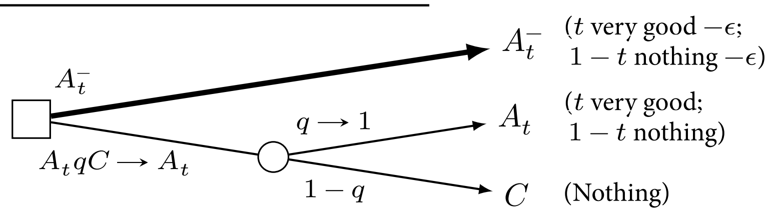
Table 1

A Violation of Continuity

	$p \cdot t$	$1 - p \cdot t$
$A_t pC$	Very good	Nothing
B	Good	Good

There is also a money-pump argument for Continuity. A money-pump argument intends to show that agents who violate some alleged requirement of rationality would make a combination of choices that lead to a sure loss. In so far as vulnerability to this kind of exploitation is a sign of irrationality, Probability Discounting is untenable as a theory of instrumental rationality. The money-pump argument for Continuity goes as follows:[12]

The Continuity Money Pump



$A_t \succ A_t^- \succ A_t pC$ for all probabilities $p \in (0, 1)$.

In this decision tree, the square represents a choice node and the circle represents a chance node. Going up at a choice node means accepting a trade and going down means refusing a trade.[13] The agent starts with $A_t qC$. $A_t qC$ is arbitrarily similar

[12]See Gustafsson (forthcoming, §6). Gustafsson calls this the Lexi-Pessimist Money Pump. Gustafsson (forthcoming, §6) also presents another money pump against preferences that violate Continuity in a different way.

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

to A_t; it results in the same outcome as A_t with a probability arbitrarily close to one. However, no matter how close q is to one, $A_t qC$ will only give a negligible probability of a positive outcome. Next, the agent is offered A_t^- in exchange for $A_t qC$. A_t^- is like A_t except that the agent has some amount ϵ less money. A_t^- gives the threshold probability of a positive outcome, while $A_t qC$ only gives a negligible probability of a positive outcome. Thus, the agent prefers A_t^- over $A_t qC$, no matter how close q is to one. Consequently, the agent accepts the trade. However, this means that the exploiter gets a fixed payment with only an arbitrarily small chance of having to give up anything. The situation is therefore arbitrarily close to pure exploitation.

To summarize, views on which all probabilities below some discounting threshold are ignored violate Continuity, and they are therefore vulnerable to exploitation in the Continuity Money Pump.

1.2 Mixture Continuity

The previous Continuity violation happens because the discounting threshold multiplied by any probability below one results in a probability below the discounting threshold. This happens because the discounting threshold is the lowest probability not discounted down to zero. Hence, the set of non-discounted values is closed (i.e., it is an interval of the form [t, 1]). However, instead of the discounting threshold being the lowest probability not discounted down to zero, it might be the highest probability that is discounted. In that case, there is no lowest non-negligible probability, and the set of non-discounted values is open on one side (i.e., it is an interval of the form (t, 1]). So, A_t will only have positive probability-discounted expected utility if it gives at least a $t + \varepsilon$ probability of a positive outcome, where ε is positive but arbitrarily close to zero. But in that case, one can always find some probability p (that may be very close to one), such that $p(t + \varepsilon) > t$. In other words, for all probabilities above the discounting threshold, there is some probability p such that their product is still above the discounting threshold. Consequently, Probability Discounting can avoid the previous violation of Continuity by letting the discounting threshold be the highest probability discounted down to zero.

However, this view violates another version of Continuity:

> **Mixture Continuity:** For all prospects X, Y and Z, the set of probabilities $\{ p \in [0, 1] \}$ with property $XpZ \succsim Y$ and the set of probabilities $\{ q \in [0, 1] \}$ with property $Y \succsim XqZ$ are closed.[14]

In effect, this principle states that if prospect XpZ is at least as good as prospect Y with some probability p, then there must be some highest and some lowest probability with which XpZ is at least as good as Y. (Similarly, if prospect Y is at least as good as prospect XqZ, then there must be some highest and some lowest probability with which Y is at least as good as XqZ). To see how the view under consideration violates Mixture Continuity, consider the following prospects:[15]

---

[14] This is axiom 2 in Herstein and Milnor (1953, p. 293). Another way to state Mixture Continuity is as follows: If $\lim_{i \to \infty} p_i = p$ and each $Xp_iZ \succsim Y$, then $XpZ \succsim Y$. Similarly, if $\lim_{i \to \infty} p_i = p$ and $Y \succsim Xp_iZ$, then $Y \succsim XpZ$.

[15] This case is also a violation of the following version of Continuity that can be derived from Mixture Continuity (Herstein and Milnor, 1953, pp. 293–294):

> **Continuity (weak-preference):** If $X \succsim Y \succsim Z$, then there is a probability $p \in (0, 1)$ such that $Y \sim XpZ$.

Mixture Continuity Violation:

Prospect A Certainly gives a very good outcome.

Prospect B Certainly gives a good outcome.

Prospect C Certainly gives nothing.

In this case, A is better than B, which is better than C. Moreover, suppose that the very good outcome is sufficiently great so that ApC is at least as good as B for all $p > t$. Given that t is discounted down to zero, it is not the case that AtC is at least as good as B. So, there is no lowest probability p with which ApC is at least as good as B. For all $p > t$, ApC is at least as great as B; when $p = t$, ApC is worse than B. This is a violation of Mixture Continuity.[16]

Furthermore, even though this view avoids the first Continuity violation, it is still vulnerable to the Continuity Money Pump. Let $A_{t+\varepsilon}$ be a prospect that gives a probability $t+\varepsilon$ of a very good outcome (and otherwise it gives nothing). $A_{t+\varepsilon}$ has positive probability-discounted expected utility for all $\varepsilon > 0$, no matter how close ε is to zero. Also, let $A_{t+\varepsilon}pC$ be a prospect that gives a probability $p(t+\varepsilon)$ of a very good outcome (and otherwise it gives nothing). If ε is very close to zero, $A_{t+\varepsilon}pC$ will only have positive probability-discounted expected utility if p is very close to one—otherwise the probability of a positive outcome would be at most t, and thus,

In Mixture Continuity Violation, A is better than B, which is better than C. However, there is no probability $p \in (0, 1)$ such that $B \sim ApC$. When $p > t$, ApC is better than B (we can suppose so); when $p \leq t$, ApC is worse than B because it only gives a negligible probability of a positive outcome.

[16] As before, Naive Discounting, Lexical Discounting, State Discounting, Stochastic Discounting and Tail Discounting all violate Mixture Continuity in this way (if the discounting threshold is the highest probability discounted down to zero).

discounted down to zero. As ε can be arbitrarily close to zero, $A_{t+\varepsilon}pC$ does not have positive probability-discounted expected utility with probabilities arbitrarily close to one; as long as $p(t+\varepsilon)$ is at most t, $A_{t+\varepsilon}pC$ is at most marginally better than nothing. Consequently, even when p is very close to one, probability discounters would be willing to pay some fixed amount in order to trade $A_{t+\varepsilon}pC$ for $A_{t+\varepsilon}$ in the Continuity Money Pump. So, if we fix p, no matter how close to one, we can find a version of the Continuity Money Pump where the exploiter wins with probability p as long as we choose ε sufficiently close to zero. Therefore, an exploiter can get a fixed payment (up to the value of $A_{t+\varepsilon}$) from the agent with only an arbitrarily small chance $(1-p)$ of having to give up anything.

To summarize, views on which probabilities up to some discounting threshold are ignored violate Mixture Continuity. They are also vulnerable to exploitation in the Continuity Money Pump.

1.3 Continuity and Statewise Dominance

As discussed earlier, Continuity violations can be avoided by using an open set of non-discounted probabilities (although this does not help avoid violations of Mixture Continuity). However, I will show that Probability Discounting must violate either Continuity or Statewise Dominance. Consider the following prospects (see table 2):

Continuity or Statewise Dominance Violation:

Prospect D Gives a very good outcome in state 1, a good outcome in state 2 and nothing in state 3.

Prospect D^{--} Gives the same outcome as D minus ϵ in state 1, a good outcome in state 2 and nothing in state 3.

Prospect C Certainly gives nothing.

Let the probability of state 1 be b, which is below the discounting threshold. Also, let the probability of state 2 be a_1, which is above the discounting threshold. Finally, let the probability of state 3 be a_2, which is also above the discounting threshold.

TABLE 2
A VIOLATION OF CONTINUITY
OR STATEWISE DOMINANCE

	State 1 $b < t$	**State 2** $a_1 > t$	**State 3** $a_2 > t$
D	Very good	Good	Nothing
D^{--}	Very good$-\epsilon$	Good	Nothing
C	Nothing	Nothing	Nothing

Given that the only difference between D and D^{--} is what happens in a very-small-probability state (i.e., state 1), D and D^{--} have equal probability-discounted expected utility. However, considering these prospects equally good would be a violation of the following principle:[17]

[17]Note that, strictly speaking, Statewise Dominance is undefined in the framework of decision theory under risk (such as the von Neumann-Morgenstern framework), as this notion belongs to decision theory under uncertainty, where there is an explicit underlying state space (such as in the Savage framework).

> **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 $X \succ Y$.[18]

In state 1, D gives a better outcome than D^{--}, while in the other states, they give the same outcomes. So, by Statewise Dominance, D is better than D^{--}.

To avoid violating Statewise Dominance in this way, probability discounters might use Statewise Dominance to rank prospects that have equal probability-discounted expected utility.[19] D would then be considered better than D^{--}, even though their probability-discounted expected utilities are the same. However, this will lead to a violation of Continuity. Consider the following mixed lottery:

> *Prospect DpC* Gives probability p of D and probability $1 - p$ of C.

Any decrease in the probability of the good outcome in D will make it the case that D's probability-discounted expected utility is less than that of D^{--}. Thus, DpC's probability-discounted expected utility is less than that of D^{--} for all probabilities $p \in (0, 1)$. Consequently, D^{--} is better than DpC for all probabilities $p \in (0, 1)$. Now we have that D is better than D^{--}, which is better than C, but D^{--} is better than DpC for all probabilities $p \in (0, 1)$—which is a violation of Continuity. So, using Statewise Dominance to rank prospects that have equal probability-discounted expected utility leads to a violation of Continuity.

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

[19] See Monton (2019, §7).

However, one might argue that if p is very close to one, the agent should ignore the possibility of obtaining C with DpC. After all, $1 - p$ would then be below the discounting threshold, and C would have a negligible probability. Furthermore, we might rank prospects that have equal probability-discounted expected utility with their expected utilities (without discounting small probabilities).[20] Consequently, DpC would have the same probability-discounted expected utility as D and D^{--}. And, with some values of p, DpC has greater expected utility (without discounting small probabilities). So, if we rank prospects with their expected utilities without discounting in cases where they have equal probability-discounted expected utility, then DpC is better than D^{--} with some probability $p \in (0,1)$. Thus, there is no violation of Continuity; it is not the case that DpC is worse than D^{--} for all $p \in (0,1)$.

However, if one ignores *outcomes* whose associated probabilities are below the discounting threshold, then one should not ignore the possibility of obtaining C; C certainly results in obtaining nothing, and D gives a non-negligible probability of obtaining nothing as well. As the probability of obtaining nothing with DpC is non-negligible, C should not be ignored. Consequently, DpC is worse than D^{--} for all probabilities $p \in (0,1)$—and we have not avoided violating Continuity. As long as one ignores outcomes whose associated probabilities are below the discounting threshold, one must violate Statewise Dominance or Continuity.[21]

[20]See §2 in Chapter 4 of this thesis.

[21]Naive Discounting, Lexical Discounting, Stochastic Discounting and Tail Discounting violate Continuity or Statewise Dominance in this way. However, if one accepts State Discounting, one might be able to ignore the possibility of obtaining C with DpC if its associated state has a negligible probability. See §3 in Chapter 4 of this thesis on State Discounting. State Discounting might

Preferences like this are also vulnerable to a money pump. However, this money pump is not as profitable for the exploiter as the previous one because the exploiter only gets a negligible probability of gaining something. Let D^- be a prospect that is like D except that the agent has less money in state 1 (of table 2), but the outcome in state 1 is still preferred to the outcome of D^{--} in state 1. D, D^- and D^{--} have equal probability-discounted expected utility. But, by Statewise Dominance, we have that D is better than D^-, which is better than D^{--}. The setup of the money pump is similar to the Continuity Money Pump. The agent starts with DqC, which is arbitrarily similar to D. The agent is then offered D^- in exchange for DqC. The agent prefers D^- to DqC, no matter how close q is to one because DqC's probability-discounted expected utility is less than that of D^-; any decrease in the probability of a good outcome in D will result in its probability-discounted expected utility being lower than that of D^-. Thus, the exploiter gets a negligible probability of payment from the agent with only an arbitrarily small chance of having to give up anything. This money pump is not as profitable to the exploiter as the previous ones because there is only a small probability that they will get the payment. However, there is only an arbitrarily small probability that they will lose something, so this scheme is still profitable to the exploiter in expectation.

therefore avoid violating Continuity—at least if the discounting threshold is the highest probability discounted down to zero. However, it would still violate Mixture Continuity in the same way as discussed in §1.2. Furthermore, the different versions of State Discounting violate either Acyclicity or Statewise Dominance. See §3 in Chapter 4 of this thesis.

1.4 Vulnerability to the Continuity Money Pump

Probability discounters are vulnerable to exploitation in the Continuity Money Pump because arbitrarily small increases in probability, from just below the discounting threshold to just above it, can make a large difference to the value of a prospect. One partial solution would be to reduce probabilities just above the discounting threshold, but not all the way down to zero—let's call this *Regressive Discounting*.[22] Probability discounters would still choose A_t^- in the Continuity Money Pump. But they would not be willing to pay as much for it as they would without reducing probabilities above the discounting threshold.

However, even if probabilities above the discounting threshold are reduced, it may be possible to compensate for those reduced probabilities by increasing the utility numbers.[23] So, probability discounters would still pay a significant sum to get A_t^- instead of $A_t qC$. Nevertheless, unlike in the Independence Money Pump (discussed later), at least probability discounters would be paying for something, namely, for a small increase in the probability of a positive outcome (from just below the discounting threshold to just above it). Therefore, this money pump is not as worrisome as the Independence Money Pump.[24] Furthermore, it might be

[22]Reducing probabilities just above the discounting threshold is discussed in Monton (2019, §6.3).

[23]If utility is bounded, the expected utility of a t chance of any positive outcome might be low. However, then Probability Discounting would be redundant, as Expected Utility Theory would no longer have counterintuitive implications in cases that involve very small probabilities of huge payoffs—at least if the upper bound is not very high and the lower bound not very low.

[24]Resolute Choice, Myopic Choice and Self-Regulation (discussed later) do not help in the Continuity Money Pump because this money pump is not dynamic like the Independence Money Pump. Also, Avoid Exploitable Plans and Avoid Dominated Plans (discussed later) do not help avoid exploitation because A_t^- is not dominated by $A_t qC$ as these prospects give slightly different proba-

argued that agents who maximize expected utility with an unbounded utility function are also vulnerable to schemes that are arbitrarily close to exploitation, and, indeed, this is what Pascal's Mugging illustrates. They will accept gambles that are arbitrarily close to a certain loss as long as the payoff in the small-probability state is great enough. However, unlike probability discounters, they will not pay a fixed amount for arbitrarily small changes in probabilities. The Continuity Money Pump illustrates how probability discounters, who wish to ignore very small probabilities, do care a great deal about very small *changes* in probabilities.[25,26]

To summarize, this section discussed three ways in which Probability Discounting might violate Continuity. First, it showed that views that discount probabilities below some threshold violate Continuity. Next, it showed that views that discount probabilities up to some threshold violate Mixture Continuity. Lastly, it showed that views that ignore very-small-probability outcomes must violate either Continuity or Statewise Dominance. Preferences that violate Continuity in these ways are vulnerable to exploitation by a money pump. However, the Continuity Money Pump is not as worrisome as the money pump for Independence because, in the

bilities.

[25] Similarly, Beckstead and Thomas (2020, §3.3) point out that Probability Discounting implies the following principle:

> **Threshold Timidity:** There is some discounting threshold such that, for any finite, positive payoffs x and y, getting x with probability below the threshold is never better than getting y with probability above the threshold—no matter how much better x is than y and no matter how close together the two probabilities may be.

Threshold Timidity states that, close to the discounting threshold, decreasing probability is infinitely more important than increasing expected utility.

[26] One possible response to the objection that probability discounters care about arbitrarily small changes in probabilities is that the discounting threshold is vague.

former, the agent is at least paying for something: a small increase in probability from just below the discounting threshold to just above it. Next, I will discuss the Independence Money Pump, which is a case of pure exploitation.

2 Independence

This section shows that Probability Discounting violates Independence. Then, it shows how violating Independence renders probability discounters vulnerable to exploitation in a money pump for Independence. §3 discusses possible ways of avoiding exploitation in this case.

2.1 A violation of Independence

To see how Probability Discounting violates Independence, consider the following prospects:[27]

Prospect A_q Gives probability q of some very good outcome (and otherwise nothing).

Prospect B Certainly gives a good outcome.

Prospect C Certainly gives nothing.

Let q be a probability that is above the discounting threshold but less than one. Suppose that the very good outcome is sufficiently great so that A_q is better than

[27]Naive Discounting, Lexical Discounting, State Discounting, Stochastic Discounting and Tail Discounting all violate Independence in this case. See §5 and 6 of Chapter 4 on the latter two views and Independence.

B. Next, consider the following mixed lotteries (see table 3):

Independence Violation:

Prospect $A_q pC$ Gives a probability p of A_q and a probability $1-p$ of C (i.e., probability $p \cdot q$ of a very good outcome and otherwise nothing).

Prospect BpC Gives a probability p of B and a probability $1-p$ of C (i.e., probability p of a good outcome and otherwise nothing).

Given that B certainly gives a positive outcome, while A_q gives only a probability q of a positive outcome, we can mix A_q and B with C so that A_q mixed with C (i.e., $A_q pC$) gives only a negligible probability of a positive outcome but B mixed with C (i.e., BpC) gives a non-negligible probability of a positive outcome. This is so because there must be some probability $p \in (0, 1)$ such that the result of q multiplied by p is below the discounting threshold, but p itself is above that threshold. Suppose that the outcomes in question are monetary and that the utility of money equals the monetary amount. Then, there must be some p such that the probability-discounted expected utility of $A_q pC$ is zero, but BpC has positive probability-discounted expected utility. In that case, Probability Discounting judges $A_q pC$ to be worse than BpC.

Table 3

A Violation of Independence

	p		$1-p$
	$p \cdot q$	$p(1-q)$	$1-p$
$A_q pC$	Very good	Nothing	Nothing
BpC	Good	Good	Nothing

Now, we have that A_q is better than B, but $A_q pC$ is worse than BpC for some $p \in (0, 1]$. This is a violation of the following axiom of Expected Utility Theory:

Independence: If $X \succ Y$, then $XpZ \succ YpZ$ for all probabilities $p \in (0, 1]$.[28]

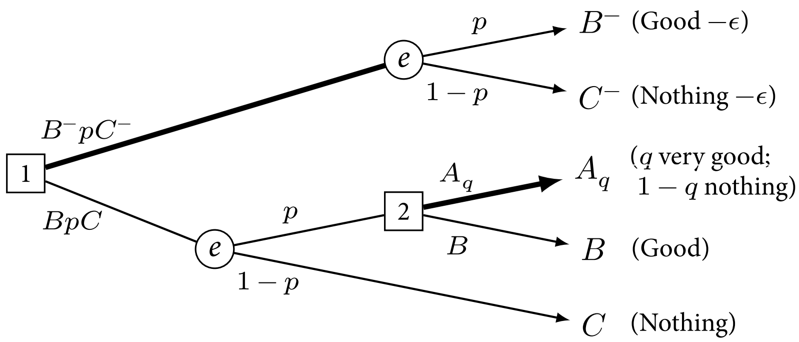
Informally, Independence is the idea that a lottery's contribution to the value of a mixed lottery does not depend on the other lotteries. The previous violation of Independence happens because, by mixing gambles together, one can reduce the probabilities of states or outcomes until their probabilities end up below the discounting threshold. As A_q gives a lower probability of a positive outcome than B does, with some values of p, $A_q pC$ only gives a negligible probability of a positive outcome, while BpC still gives a non-negligible probability.

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

2.2 The Independence Money Pump

Violating Independence renders probability discounters vulnerable to exploitation in the Independence Money Pump. It goes as follows:[29]

THE INDEPENDENCE MONEY PUMP



$A_q \succ B$, and $BpC \succ B^{-}pC^{-} \succ A_q pC$.

The agent starts with prospect BpC: probability p of a good outcome and otherwise nothing. At node 1, the agent is offered a trade from BpC to $B^{-}pC^{-}$. $B^{-}pC^{-}$ is just like BpC except that the agent has less money. If the agent turns down this trade and BpC results in the agent going up at chance node e, then at node 2, the agent will be offered a trade from B (certain good outcome) to A_q (probability q of a very good outcome and otherwise nothing). Both chance nodes depend on the same chance event e.

The agent can use *backward induction* to reason about this decision problem. This means that the agent considers what they would choose at later choice nodes and then takes those predictions into account when making choices at earlier choice

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

nodes.[30] As the agent prefers A_q to B, they would accept the trade at node 2. By using backward induction at node 1, the agent can reason that the prospect of turning down the trade at node 1 is effectively $A_q pC$, and the prospect of accepting the trade is $B^- pC^-$. Given that the agent prefers BpC to $A_q pC$, it seems plausible that there is some price ϵ that they would be willing to pay to get the former instead of the latter. So, the agent pays that price and ends up with $B^- pC^-$. But they have ended up with $B^- pC^-$ even though they could have kept BpC for free had they gone down at both choice nodes. Therefore, they have given up money for the exploiter.[31]

To summarize, this section showed that Probability Discounting violates Independence. This Independence violation happens because, by mixing gambles together, one can reduce the probabilities of states or outcomes until their associated probabilities are below the discounting threshold. As a result of violating Independence, probability discounters are vulnerable to exploitation in the Independence Money Pump. The next section discusses some possible ways of avoiding exploitation in this decision problem.

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

[31] Also, as the chance nodes depend on the same event e, going up at node 1 is statewise dominated by going down at both choice nodes. See Gustafsson (forthcoming, §5).

3 Avoiding exploitation in the Independence Money Pump

This section discusses possible ways of avoiding exploitation in the Independence Money Pump. It argues that none of the standard views, such as Resolute Choice and Self-Regulation, work. It also argues that even if vulnerability to exploitation is not a sign of irrationality, Probability Discounting has untenable implications in a version of the Independence Money Pump that might result in a loss. But before discussing Resolute Choice and Self-Regulation, I will begin by discussing a foolish decision policy that nevertheless gives the right recommendation in the Independence Money Pump.

3.1 Myopic Choice

One decision policy that might help probability discounters is *Myopic Choice*. Myopic Choice advises an agent to choose at each choice node the option that currently seems best with no regard to what one will choose at later choice nodes.[32] But Myopic Choice is unjustifiable. It is irrational not to take one's future choices into account when making decisions. Nevertheless, one might be tempted to accept it as it gives the right recommendation in the Independence Money Pump.[33]

[32] Strotz (1955-1956) and von Auer (1998, p. 111). Myopic Choice is distinct from Naive Choice, on which one should choose the best available plan with no regard to whether one will in fact follow that plan. Similarly as a resolute agent (and unlike a myopic chooser), a naive chooser makes plans. However, unlike a resolute chooser, a naive chooser may not follow such plans.

[33] Myopic Choice is subject to problems that have nothing to do with Probability Discounting. For example, suppose a myopic agent starts with prospect *Bad*, which they can exchange for

If one accepts Myopic Choice, one will turn the offer down at node 1, thinking that one is choosing BpC. But if one ends up in node 2, one will choose A_q over B. So, with Myopic Choice, one can avoid getting money pumped in the Independence Money Pump.

However, probability discounters who use Myopic Choice are vulnerable to monetary exploitation in another decision problem. Recall the earlier prospects A_q, B and C:

> *Prospect* A_q Gives probability q of some very good outcome (and otherwise nothing).
>
> *Prospect* B Certainly gives a good outcome.
>
> *Prospect* C Certainly gives nothing.

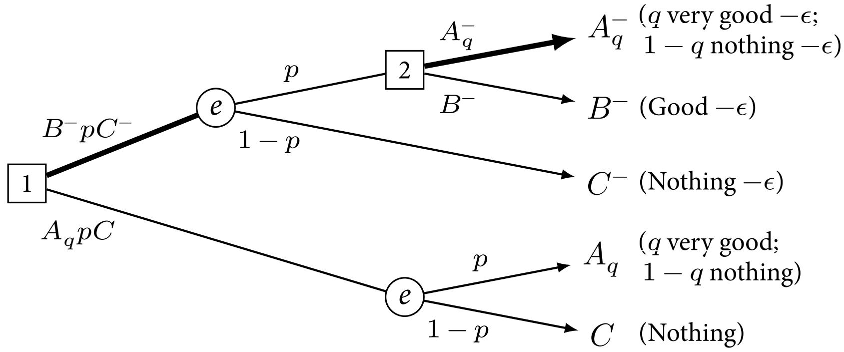
Now consider a reversed version of the Independence Money Pump:[34]

---

prospect *Terrible*. If they decide to do so, they are then offered prospect *Excellent*. A myopic chooser would refuse *Terrible*, and thus, they would end up with *Bad* when they could have had *Excellent*. However, probability discounters might accept some other (more plausible) principle that behaves like Myopic Choice in the Independence Money Pump. For example, they could restrict Myopic Choice to cases that include negligible probabilities.

[34]Note that backward induction gets one out of trouble in this money pump; by backward induction, accepting the trade at node 1 will effectively lead to prospect $A_q^-pC^-$, while rejecting the offer leads to A_qpC. Therefore, one should reject the offer.

The Reverse Independence Money Pump



$A_q \succ A_q^- \succ B^-$, and $B^-pC^- \succ A_qpC \succ A_q^-pC^-$.

Unlike in the Independence Money Pump, this time the agent starts with A_qpC. At node 1, they are offered B^-pC^-, which is like BpC except that the agent has ϵ less money. A myopic agent would choose B^-pC^-, given that they prefer it over A_qpC; there should be some amount ϵ that the agent is willing to pay to get BpC instead of A_qpC, given that they prefer the former. Then, if they end up in node 2, they are offered A_q^- in exchange for B^-. A_q^- is like A_q except that the agent has ϵ less money. Given that the agent prefers A_q over B (and B^-), it is again plausible that they are willing to pay some amount to get A_q instead of B (or B^-). So, the agent accepts the offer. They have now been money pumped; the agent chose $A_q^-pC^-$ even though they could have kept A_qpC for free by refusing the offer at node 1. So, Myopic Choice makes probability discounters vulnerable to exploitation in a reversed version of the Independence Money Pump.[35]

[35] The following case—let's call it *Discounter's Ruin*—also suggests that Probability Discounting should not be combined with Myopic Choice. Let the discounting threshold be (implausibly) just below 0.01. Now consider the following prospects:

3.2 Resolute Choice

Myopic Choice does not help probability discounters avoid monetary exploitation. But perhaps Resolute Choice will? 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.[36] If one accepts Resolute Choice, one can make a plan that one will not trade B for A_q in node 2 of the Independence Money Pump. Even though one would usually prefer A_q over B, one is now committed to keeping B regardless. Consequently, one can safely refuse the trade at node 1, as one is then choosing BpC over $B^{-}pC^{-}$; one will not get money pumped nor end up with the inferior prospect A_qpC.

However, combining Probability Discounting with Resolute Choice gives untenable results in another case. Consider the following prospects:

Prospect C Certainly gives nothing.

Prospect E Gives probability r of some very bad outcome and prob-

---

Discounter's Ruin:

Prospect A Gives a 0.01 chance of \$10 and otherwise −1¢.

Prospect B Gives a 0.01 chance of \$10,000 and otherwise −\$10,

Prospect C Gives a 0.01 chance of \$10,000,000 and otherwise −\$10,000,

and so on for some large but finite number of prospects.
First, an agent is offered A, followed by an offer of B in case the agent wins \$10. As B is better than \$10, the agent would accept the offer. Then, if the agent wins \$10,000 with B, they are offered C. Again, the agent prefers C over \$10,000, so they would accept the offer. And so on for some large but finite number of offers. If one accepts all the offers, one would effectively choose an option that almost certainly results in a negative outcome and gives only a very small probability—a probability way below the discounting threshold—of a positive outcome. Thus, one would, effectively, not discount small probabilities down to zero.

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

ability $1 - r$ of a barely positive outcome.

Prospect F Certainly gives a barely positive outcome.

Let r be a probability above the discounting threshold but less than $1 - r$ (i.e., less than 0.5). Suppose the very bad outcome in E is sufficiently bad so that C is better than E; certainly getting nothing is better than a non-negligible chance of a very bad outcome and otherwise a barely positive outcome.

Next, consider the following mixed lotteries (see table 4):

Independence Violation (Negative):

Prospect CpF Gives a probability p of C and a probability $1 - p$ of F (i.e., probability p of nothing and otherwise a barely positive outcome).

Prospect EpF Gives a probability p of E and a probability $1 - p$ of F (i.e., probability $p \cdot r$ of a very bad outcome and otherwise a barely positive outcome).

Given that r is less than $1 - r$, there must be some (relatively small) probability $p \in (0, 1)$ such that the result of r multiplied by p is below the discounting threshold, but the result of $1-r$ multiplied by p is above the discounting threshold. In that case, the possibility of obtaining a very bad outcome with EpF is ignored. However, given that $p(1 - r)$ is above the discounting threshold, EpF gives a greater probability of a barely positive outcome than CpF.[37] Consequently, EpF is better

[37] This is true whether one ignores very-small-probability outcomes or states. Thus, this argu-

than CpF. But now we have another violation of Independence: C is better than E, but EpF is better than CpF.[38] This violation of Independence happens because the probability of a very bad outcome is above the discounting threshold in E but below the discounting threshold in the mixed lottery EpF. Thus, the possibility of a very bad outcome is not ignored in E, but it is ignored in EpF.

Table 4

Independence Violation (Negative)

	p		$1-p$
	$p \cdot r$	$p(1-r)$	$1-p$
CpF	Nothing	Nothing	Barely positive
EpF	Very bad	Barely positive	Barely positive

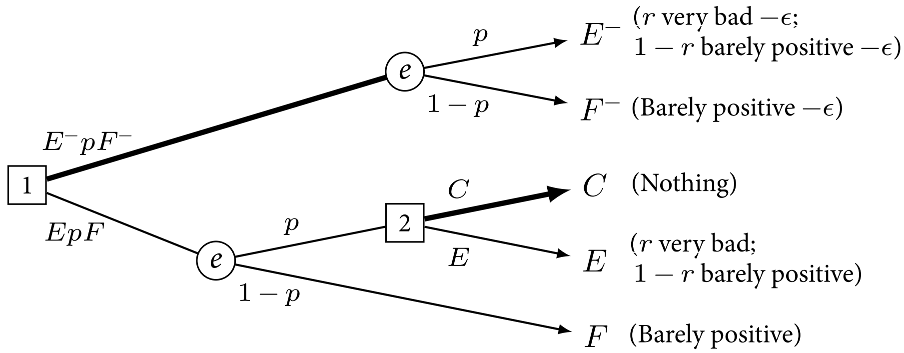
Let's go back to Resolute Choice and the Independence Money Pump. Recall that a probability discounter who uses Resolute Choice would commit to keeping B in node 2 of the Independence Money Pump (and thus avoid getting money pumped). In other words, they would commit to keeping a prospect that certainly gives a good outcome instead of trading it for a non-negligible chance of a very good outcome (and otherwise nothing). This does not seem untenable; one might bite the bullet and accept this implication. However, the same is not true in the

ment applies to all Naive Discounting, Lexical Discounting, State Discounting, Stochastic Discounting and Tail Discounting. If one ignores very-small-probability states and we take columns 2–4 in table 4 to correspond to states, then one ought to ignore column 2 (and not ignore columns 3 and 4). If one ignores very-small-probability outcomes, one ought to ignore the possibility of obtaining a very bad outcome with EpF (and not ignore the possibilities of the other outcomes). Either way, EpF gives a greater probability of a barely positive outcome than CpF.

[38]This violation of Independence is similar to the one discussed in Chapter 4 of this thesis.

following version of the Independence Money Pump:[39]

The Independence Money Pump (Negative)



$C \succ E$, and $EpF \succ E^-pF^- \succ CpF$.

In this case, the agent starts with EpF. At node 1, they are offered E^-pF^-, which is like EpF except that the agent has ϵ less money. If the agent refuses the trade and ends up in node 2, they are offered C in exchange for E. The agent prefers C over E; it is better to certainly get nothing that to choose a prospect that gives a non-negligible probability r of some very bad outcome and otherwise a barely positive outcome. Given that the agent would choose C over E at node 2, by using backward induction at node 1, the agent realizes that the choice is effectively between E^-pF^- and CpF. And, similarly as before, the agent prefers E^-pF^- over CpF, so they accept the offer. But then they have paid for something they could have kept for free.

In this case, a resolute agent can avoid getting money pumped if they commit to keeping E at node 2. However, unlike in the earlier money pump, this time the

[39] As before, the structure of this money pump is from Gustafsson (2021, p. 31n21).

resolute choice seems unreasonable: The agent would choose a prospect that gives a non-negligible probability r of some very bad outcome and otherwise a barely positive outcome over the certainty of getting nothing. Earlier, we assumed that r is above the discounting threshold but less than $1 - r$. So, it could be, for example, 0.49. Then, the agent would choose a prospect that gives a 0.49 probability of a very bad outcome and otherwise a barely positive outcome over certainly getting nothing. Furthermore, note that the very bad outcome can be arbitrarily bad, while the barely positive outcome can be arbitrarily close to getting nothing. No reasonable theory recommends making this choice.

Appeals to Resolute Choice seem to provide a general means of answering dynamic choice arguments against various patterns of preferences. However, Probability Discounting combined with Resolute Choice leads to disastrous results. Thus, Probability Discounting combined with Resolute Choice is untenable as a theory of instrumental rationality.

3.3 Self-Regulation

Another decision policy that has been proposed as a solution to money pumps is *Self-Regulation*.[40] Self-Regulation forbids (if possible) choosing options that may lead via a rationally permissible route to a final outcome that is unchoiceworthy by the agent's own lights.[41] The idea is that one ought not choose options that may (following one's preferences) lead to an outcome that one would not choose

[40] Self-Regulation helps avoid exploitation in money pumps against cyclic preferences. See Ahmed (2017). See Gustafsson (forthcoming, §2) for criticism of Self-Regulation.

[41] Ahmed (2017, p. 1001).

in a direct choice of all final outcomes. Unlike Resolute Choice, Self-Regulation is forward-looking.[42] When an agent's present choices determine the options available to them in the future, they should now choose so that their future choices lead to what they now consider acceptable in light of what is now available.[43] If the agent now wants to avoid some final outcome O, and they know what they are going to do at later choice nodes, then they should (if possible) now choose in such a way that, given those later choices, they will not end up with O.[44]

Self-Regulation in its original formulation does not help in the Independence Money Pumps, as it was intended for money pumps that do not involve chance.[45] However, the Independence Money Pumps involve chance nodes, so the agent does not know what the final holdings will be. One way to adapt Self-Regulation to cases that involve chance is to apply it to plans. A plan specifies 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. Self-Regulation with respect to plans then states the following:

[42]Self-Regulation also differs from backward induction. Suppose that one's preferences are cyclic, so that A is better than B, which is better than C, which is better than A (and also suppose that A^- is better than B). Then, in all decision problems, Self-Regulation forbids (if possible) choosing an option that would lead to A^- via a rationally permissible path because one would not choose A^- in a direct choice of A, B, C and A^- (as A^- is dominated by A). But one might then end up with B, which is worse than A^-. In that case, backward induction would recommend choosing the option that leads to A^- (because A^- is better than B). However, Self-Regulation would recommend choosing the option that leads to B (because A^- would not be chosen in a direct choice of all final outcomes). See Ahmed (2017).

[43]Ahmed (2017, p. 1013).

[44]Ahmed (2017, p. 1003).

[45]Rabinowicz (2021, n. 13) writes: "[H]e [Ahmed, 2017] only shows how self-regulation allows the agents with cyclic preferences to avoid dynamic inconsistency. It is unclear whether and how this approach can be extended to agents who violate Independence."

> **Self-Regulation for Plans (i.e., Avoid Unchoiceworthy Plans):** If possible, one ought not choose options that may (following one's preferences) lead one to follow a plan that one would not choose in a direct choice of all plans (assuming one was able to commit to following some available plan).

Self-Regulation for Plans is a partial characterization of what it means to follow one's preferences: It involves, if possible, not choosing options that may, following one's preferences, lead one to follow an unchoiceworthy plan. A forward-looking choice rule C is self-regulating if and only if it tells you, at each node x, to choose a safe option whenever one is available. An option is 'safe' if and only if subsequently acting in accordance with C will lead you to follow a plan that is permissible at x.

The available plans at node 1 of the Independence Money Pump correspond to prospects $A_q pC$, BpC and $B^- pC^-$. One would not choose $A_q pC$ or $B^- pC^-$ in a direct choice between these plans. Therefore, one should not (if possible) choose any option that may lead via a rationally permissible route to one following $A_q pC$ or $B^- pC^-$. However, both accepting and rejecting the trade at node 1 of the Independence Money Pump lead the agent to follow one of these plans via rationally permissible routes. Rejecting the offer leads one to follow $A_q pC$; accepting the offer leads one to follow $B^- pC^-$. So, Self-Regulation for Plans is silent in this case because it is not possible to make choices that do not lead to unchoiceworthy plans via rationally permissible routes. Thus, Self-Regulation for Plans does not help avoid exploitation in the Independence Money Pump.

3.4 Alternative decision policies

Instead of accepting Self-Regulation for Plans, one might restrict the set of forbidden plans and accept the following decision rule:

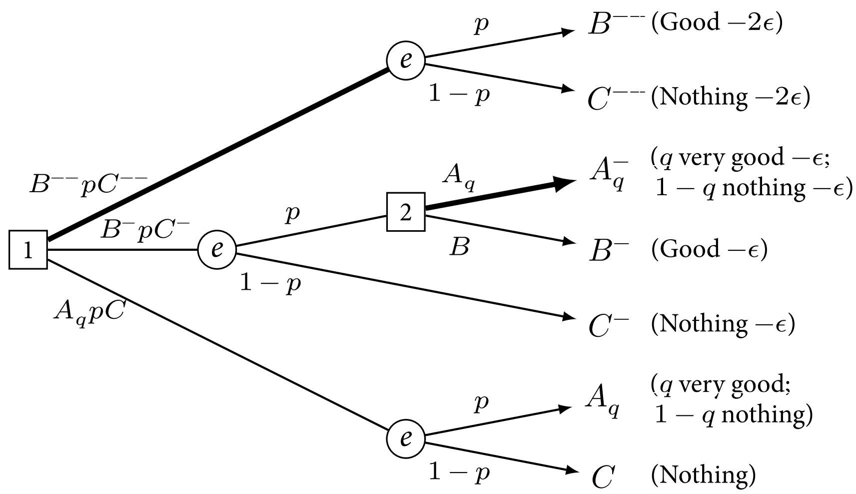
> **Avoid Exploitable Plans:** If possible, one ought not choose options that may (following one's preferences) lead one to pay for a plan that one could keep for free.

Avoid Exploitable Plans forbids accepting the trade at node 1 of the Independence Money Pump because accepting it would be paying for something that one could keep for free. However, Avoid Exploitable Plans does not forbid choosing A_q over B (or C over E) at node 2 because doing so would not be paying for a plan that one could keep for free. Thus, at node 2, an agent using Avoid Exploitable Plans would choose A_q over B (and C over E), given that they prefer the former. So, if one uses Avoid Exploitable Plans, one can avoid getting money pumped and also avoid the conclusion that one should keep E at node 2 of the version of the Independence Money Pump that includes negative payoffs.

But in the following decision problem, someone using Avoid Exploitable Plans would pay a higher price for something they could have obtained cheaper:[46]

[46] It might be objected that expected utility maximizers must also end up worse off than they could have been in some cases with an infinite series of trades. See for example Gustafsson (forthcoming, §8). However, expected utility maximizers might argue that there is a difference between not choosing the best option and paying more than one needs to, as the latter involves freely giving up what one already possesses while the former does not. But this kind of status quo bias may not be rationally justified.

The Three-Way Independence Money Pump



$A_q \succ A_q^- \succ B^- \succ B^{--}$, and $B^-pC^- \succ B^{--}pC^{--} \succ A_qpC \succ A_q^-pC^-$.

In this case, the agent starts with A_qpC. At node 1, they are offered B^-pC^- and $B^{--}pC^{--}$. $B^{--}pC^{--}$ is like B^-pC^- except that the agent has even less money (-2ϵ vs. $-\epsilon$). If the agent chooses B^-pC^- and ends up in node 2, then they are offered A_q^- in exchange for B^-. As the agent prefers A_q^- to B^-, they would accept the offer. So, choosing B^-pC^- at node 1 means effectively choosing $A_q^-pC^-$, given one's later choices. An agent who uses Avoid Exploitable Plans would therefore choose $B^{--}pC^{--}$; they prefer $B^{--}pC^{--}$ over A_qpC and $A_q^-pC^-$, and choosing it does not mean the agent is paying for something they could keep for free (as the agent starts with A_qpC). However, as B^-pC^- is also available, the agent has paid more than they needed to for BpC. They could have paid just ϵ instead of 2ϵ had they chosen B^-pC^- at node 1 (and then kept B^- at node 2).

The focus on avoiding monetary exploitation may be misplaced. Instead, one might prefer adopting a decision rule that forbids all dominated plans whether or not they involve monetary exploitation:[47]

> **Avoid Dominated Plans:** If possible, one ought not choose options that may (following one's preferences) lead one to pay more for a plan that one could obtain for less money.

Avoid Dominated Plans forbids accepting the offer at node 1 of the Independence Money Pump because $B^{-}pC^{-}$ is dominated by BpC. Also, with this decision rule, one can refuse both offers of the Three-Way Independence Money Pump and keep $A_q pC$. One should refuse $B^{--}pC^{--}$ because it is dominated by $B^{-}pC^{-}$. And, one should refuse $B^{-}pC^{-}$ because choosing it means one is effectively choosing $A_q^{-}pC^{-}$, and $A_q^{-}pC^{-}$ is dominated by $A_q pC$. So, one should keep $A_q pC$. Avoid Dominated Plans thus allows an agent to avoid paying too much in this decision problem.

However, Avoid Dominated Plans seems a too narrow decision policy. Self-Regulation for Plans forbids choices that lead to plans that are unchoiceworthy by the agent's own lights. In contrast, Avoid Dominated Plans only forbids choices that lead to dominated plans but allows choices that lead to unchoiceworthy plans (such as $A_q pC$). It seems difficult to motivate such a decision policy. Why would it be irrational to choose an option that leads to a dominated plan (such as $B^{--}pC^{--}$) but not irrational to choose an option that leads to an unchoiceworthy plan (such as

[47] Avoid Dominated Plans is formulated in terms of monetary dominance: One should avoid plans that one can obtain for less money. But one should surely avoid plans that are dominated in other ways as well. More generally, one should avoid plans that are dominated with respect to anything valuable.

$A_q pC$)? Allowing the latter but forbidding the former seems arbitrary. Moreover, it leads one to something that is worse than the dominated plan, namely, $A_q pC$.

Furthermore, if we change the probabilities in the Independence Money Pump slightly, then Avoid Dominated Plans no longer avoids exploitation, at least entirely. Now, instead of $B^- pC^-$, the agent faces $B^- qC^-$, where q is arbitrarily close to p (and $q < p$). Then, given that $B^- qC^-$ and BpC do not give the exact same probabilities of the relevant outcomes, Avoid Dominated Plans no longer forbids accepting the trade at node 1; it is not the case that $B^- qC^-$ is like BpC except that the agent has less money, so Avoid Dominated Plans is silent. Consequently, a probability discounter who uses Avoid Dominated Plans will choose $B^- qC^-$ even though they could have kept BpC for free, and q is arbitrarily close to p. They have therefore given a fixed payment ϵ for an arbitrarily small increase in the probability of a positive outcome. So, Avoid Exploitable Plans is vulnerable to a scheme that is arbitrarily close to exploitation.[48]

3.5 How worrisome are the Independence Money Pumps?

Probability discounters might argue that these money pumps are not worrisome because, for example, the agent only really faces prospects $A_q pC$ and $B^- pC^-$ at node 1 of the Independence Money Pump, given that they would choose A_q at node 2. Thus, given the agent's preferences, in a way BpC is not even available to the agent. So, by choosing $B^- pC^-$, the agent does not end up paying for something they could have kept for free. However, a money-pump argument is supposed

[48]This objection also applies to Avoid Exploitable Plans.

to show that a given set of preferences is irrational because they lead to the agent paying for something they could have kept for free (if they had some other preferences). Therefore, it is not an adequate defense of those preferences that, given those preferences, the agent did not have any other option but to pay for something they could have kept for free. The target of the money pump is the structure of preferences. If one's preferences lead one to pay for something one could have kept for free (if one had some other preferences), then the money pump has succeeded in showing that those preferences are irrational.

Furthermore, even if being exploited is not a sign of irrationality as this argument claims, the violation of Independence in the case that includes negative payoffs (see table 4) is worrisome independently of the exploitation it leads to. This violation of Independence is particularly counterintuitive because EpF is considered better than CpF no matter how bad the very bad outcome is as long as the barely positive outcome is at least slightly positive. Moreover, the agent would choose to lock in a choice of keeping E (at node 2) if that was somehow possible at node 1.[49] This means they would lock in a choice of a prospect that gives a 0.49 probability of a very bad outcome and otherwise a barely positive outcome over certainly getting nothing. This seems irrational. So, even if probability discounters do not accept Resolute Choice, they would still make the same choice of E over C if offered the chance to lock in the choice at node 1.[50] This makes Probability Discounting less

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

[50]See §6 in Chapter 4 of this thesis.

plausible as a theory of instrumental rationality.[51]

To conclude, this section discussed possible ways of avoiding exploitation in the Independence Money Pump. First, it showed that, although Myopic Choice avoids exploitation in the Independence Money Pump, it does not avoid exploitation in the Reverse Independence Money Pump. Resolute Choice, in turn, leads to untenable results in the negative version of the Independence Money Pump, and Self-Regulation for Plans does not avoid exploitation in the Independence Money Pump. An agent who uses Avoid Exploitable Plans would pay too much for a plan in the Three-Way Independence Money Pump. Avoid Dominated Plans solves the Three-Way Independence Money Pump, but it is vulnerable to a scheme that is arbitrarily close to pure exploitation.

It was also argued that locking in the choice of E over C at node 2 of the negative version of the Independence Money Pump is irrational—and that this is something probability discounters would do regardless of whether they accept Resolute Choice or not. So, even if vulnerability to exploitation is not a sign of irrationality, Probability Discounting has untenable implications in the negative version of the Independence Money Pump. All in all, what we learn from these money pumps is that the various possible ways of avoiding exploitation do not ultimately work.[52] In

[51] It is worth pointing out 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).

[52] The money pump arguments against Probability Discounting should be persuasive even for those who reject Independence for other reasons (e.g., due to the Allais paradox), as they might use Resolute Choice to avoid exploitation in the money pumps for Independence. However, as argued above, this solution is not available to probability discounters. In contrast, the Continuity Money Pump is not particularly worrying for probability discounters who already violate Continuity for other reasons.

addition, we learn that Probability Discounting gives untenable implications even if exploitation is not a sign of irrationality.

4 Conclusion

Probability Discounting is one way to avoid fanatical choices in cases that involve tiny probabilities of huge payoffs. However, it faces some serious problems. First, this chapter discussed three ways in which Probability Discounting might violate Continuity. It was shown that views that discount probabilities below some discounting threshold violate Continuity. Secondly, it was shown that views that discount probabilities up to some discounting threshold violate Mixture Continuity. Lastly, it was shown that views that ignore very-small-probability outcomes must violate either Continuity or Statewise Dominance. As a result of these Continuity violations, Probability Discounting is vulnerable to exploitation in the Continuity Money Pump.

In addition to violating Continuity, Probability Discounting also violates Independence, which renders probability discounters vulnerable to exploitation in the Independence Money Pump. Some possible ways of avoiding exploitation in the Independence Money Pump were discussed. However, these either failed to avoid exploitation in some version of the Independence Money Pump or they had otherwise untenable implications. It was also argued that even if vulnerability to exploitation is not a sign of irrationality, Probability Discounting has untenable implications in the negative version of the Independence Money Pump.

To conclude, this chapter has shown that Probability Discounting is vulnerable to exploitation in the money pumps for Independence and Continuity. The former is more worrisome than the latter, and it is difficult to see how Probability Discounting can respond to this challenge.

References

Ahmed, A. (2017), 'Exploiting cyclic preference', *Mind* **126**(504), 975–1022.

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'. Global Priorities Institute Working Paper No.10.
URL: *https://globalprioritiesinstitute.org/nick-beckstead-and-teruji-thomas-a-paradox-for-tiny-probabilities-and-enormous-values/*

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

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.

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

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

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

Hammond, P. J. (1998), Objective expected utility: A consequentialist perspective, *in* S. Barberà, P. J. Hammond and C. Seidl, eds, 'Handbook of Utility Theory Volume 1: Principles', Kluwer, Dordrecht, pp. 143–211.

Herstein, I. N. and Milnor, J. (1953), 'An axiomatic approach to measurable utility', *Econometrica* **21**(2), 291–297.

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.

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.

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.

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

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.

Peterson, M. (2020), The St. Petersburg paradox, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', fall 2020 edn, Metaphysics Research Lab, Stanford University.

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.

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.

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.

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.

von Auer, L. (1998), *Dynamic Preferences, Choice Mechanisms, and Welfare*, Springer-Verlag, Berlin.

von Neumann, J. and Morgenstern, O. (1947), *Theory of Games and Economic Behavior*, 2 edn, Princeton University Press, Princeton.

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