Rationality applied: resolving the two envelopes problem

Christian Hugo Hoffmann[1]

Abstract

The *Two Envelopes Problem* is a beautiful and quite confusing problem in decision theory which is ca. 35 years old and has provoked at least 150 papers directly addressing the problem and displaying a surprising variety of different responses. This paper finds decisive progress in an approach of Priest and Restall in 2003, contends that the recent papers having appeared since did not really go beyond that paper, argues further that Priest's and Restall's solution is still not complete, and proposes a completion of their solution. If the analysis is correct, this work has the potential of laying the Two Envelopes Problem at rest.

Keywords Two envelopes problem · Decision theory · Paradox · Probability · Methodology

1 Introduction

Variants of the following problem have acquired a certain renown (e.g. Nalebuff, 1989): You are presented with a choice between two sealed envelopes. Each envelope contains some quantity of money, which can be of any positive real magnitude. You know that one envelope contains twice as much money as the other, but you do not know which contains the larger sum and which is the smaller. You choose one of them—call it "*A*" and the other "*B*". You can keep the money in A (a_1), whose numerical value you do not know at this stage, or you can switch A for B (a_2). You wish to maximise your money. What should you do?

This is the standard formulation of the *Two Envelopes Problem*, a much studied 'paradox' in decision and probability theory, mathematical economics, logic and philosophy. Time and again a new analysis is published in which the author(s) claim (s) finally to explain what actually goes wrong in this puzzle. Each author is eager to

---

✉ Christian Hugo Hoffmann
christian@hoffmann-economics.com

1   Institute for Technology Assessment and Systems Analysis (ITAS), Karlsruhe Institute of Technology (KIT), Karlstraße 11, 76133 Karlsruhe, Germany

emphasize what is new and exceptional in her or his approach and is inclined to conclude that earlier approaches did not get to the root of the matter.

This paper addresses the Two Envelopes Problem from a different perspective which will not only enrich the debate but might also resolve the issue. By all means, it will become clear that the presented Two Envelopes Problem does not deserve to be called a paradox (and certainly not an unresolved paradox, as some writers still insist on claiming). Rather, there has only been an easily solvable decision problem —where anything except indifference has no rational ground—serving as one possible starting point for calling basic assumptions of decision theory into question.

The structure of this paper, consisting of three parts, is as follows. First, I begin with a short summary of various replies to the Two Envelopes Problem that have been given in the literature by consulting a variety of experts. Their arguments lose cogency, however, when we critically analyse them and depart from the special frameworks they adopted. In Sect. 3, I turn to a more promising approach to tackle the Two Envelopes Problem. This approach is found in a proposal by Priest and Restall (2003). In the final Sect. 4, I will then correct some small flaws and add some new insights, including a clear structuring of the problem, an explanation for where the reasoning of Priest and Restall may go astray, and including a sophisticated conclusion regarding my own proposed solution to the issue.

2 Tales of two envelopes and the cacophony among tale-tellers

It seems there are already different satisfactory solutions or at least approaches to solutions for *different* Two Envelopes Problems on the market. Beginning with the original description of the situation (given above), which is clearly *under-determined*, I will classify solutions which are either too simple or too specific. Insofar, they will turn out to be unsatisfactory as far as the original problem is concerned. More precisely, I would like to briefly discuss the following replies to the Two Envelopes Problem:

The psychologist's reply: If the psychologist is consulted, he will probably point to the empirical fact that some (or most) people are risk-averse—they tend to prefer a_1 −, others are risk-prone—they prefer a_2. Now, if the decision-maker was risk-averse, he would sometimes prefer a safe option to a risky one even if the safe option offers a lower expectation of money. It turns out that if the expected value of a_2 exceeds the expected value of a_1 and if the agent is risk averse *enough*, he will refuse to trade A for B—without there being a problem of irrational behaviour or a paradox (Broome, 1995, p. 9).

Be that as it may, on the one hand, I would like to proceed *in the most general way*, i.e. to avoid (specific) assumptions about the decision maker's risk attitude. On the other hand, in presenting the *standard formulation* of the decision problem, I already said that, of two options, the agent would always prefer the one that offers him a greater expectation of money. This implies that the decision-maker is *risk neutral* about money. Therefore, I do not need to go into the details of the psychologist's argument (cf. also Broome, 1995).

An economist's reply: He might appeal to the fact that money is discrete, i.e. with a minimum or maximum. It seems to be natural to assume that any rational, minimally informed person recognises that the amount of money in the world is finite, with a smallest unit in any currency (cf. Jackson et al., 1994). Taking this into account, one might make the case for a game theoretic analysis (e.g. Nalebuff, 1989):

Suppose that the Two Envelopes decision problem is given as a strategic game where envelope A is handed to player *Ali* and envelope B is handed to player *Baba*. Further, Ali and Baba are allowed to look privately at the amount of money in their own envelope. Then they are given an opportunity to trade envelopes if both agree. Besides, it is *common knowledge* that a hidden coin was flipped and that if it came up heads, twice the amount of money of A was placed in envelope B and that if it came up tails, only half the original sum was put in B. Finally, the arbitrator or game master can never put more than MAX in envelope A (which follows from the fact that the world's money supply has an upper bound).

> Then the amount of money in Baba's envelope, y, must lie in the range [0, 2 MAX]. Baba knows that when he finds y between MAX/2 and 2 MAX, the coin must have landed heads. Ali cannot have more than MAX, so that Ali's envelope must contain $x = y/2$. Therefore, Baba would never trade if he finds y > MAX/2, since his expected gain is negative. Ali then reasons that if she has between MAX/4 and MAX, she should not trade. Why? If Baba's envelope is larger, his y must be between MAX/2 and 2 MAX and thus he will refuse to trade. The only time Baba would be willing to trade is when his envelope has between MAX/8 and MAX/2, in which case Ali loses money by trading. Similarly, once Baba recognizes that Ali won't trade when her envelope contains anything between MAX/4 and MAX, Baba should not want to trade when his envelope contains anything between MAX/8 and MAX/2. The reasoning continues inductively so that neither Ali nor Baba would ever want to trade. (Nalebuff, 1989, p. 176)

This argument shows by means of *backward induction*, a method used to compute subgame perfect Nash equilibria in sequential games. Cf. Neumann and Morgenstern (1944), the impossibility of trade when Ali and Baba agree that the maximum possible sum in either envelope is bounded by some high number.

Is this a (the) correct solution for the Two Envelopes Problem? Should both players stick with the money in their envelope? Be that as it may for the moment, the economist's reply is unsatisfactory for several, already visible reasons. First, the question would have to be settled on whether or not it is possible to recover the Two Envelopes Problem by arguing that the upper bound should be considered infinite and thus non-existent. Second, the crucial difference between the Two Envelopes *decision problem*, where there is only one agent involved, and the Two Envelopes *game*, where two players with conflicting interests interact and whose decisions affect each other, cannot be neglected. And, moreover, the original Two Envelopes Problem was modified by the economist. He introduced further assumptions (e.g. a specific mechanism by which the money was selected and deposited into the two envelopes) which go beyond the original description of the situation. So if he solved

the (better: his) Two Envelopes Problem at all, his argument is not relevant for my purposes.

The Bayesian replies: Most attempts to solve the Two Envelopes Problem are in a Bayesian spirit (e.g. Broome, 1995; Christensen & Utts, 1992; Clark & Shackel, 2000; Linzer, 1994; McGrew et al., 1997; Nalebuff, 1989; Zabell, 1988a, 1988b). As I see it, the advocates of these approaches maintain that to treat the decision problem precisely, one needs to know more about how the various amounts that may be in the envelopes are chosen. Since the original statement of the Two Envelopes Problem is vague here—in the description of the original decision situation, there is no information given on a) which precise sums of money could have been put into A and B and b) the probabilities with which these sums are chosen −, they go into the questions of what the possible and relevant states of the world are and of what the probabilities of each of the possible and relevant states are. To proceed they fill in the details in some way. A very popular way is this:

Opening envelope A: Suppose that the agent opens envelope A before deciding whether to exchange. Now that he knows exactly how much money is in it, which should he prefer, this money (i.e. a_1), or the money in envelope B (i.e. a_2)?

Note, for a start, that, relative to certain items of background information, new information provided by opening envelope A can make it rational to pick a_1 (a_2). In general, the basic idea is that, if one knows a prior probability distribution over the size of the cheques in the envelopes, or at least enough about it, then, by employing *Bayes' Law*

$$p(H|E) = \frac{p(H) \times p(E|\mathrm{H})}{p(H) \times p(E|H) + \mathrm{p}(H) \times \mathrm{p}(E|H)},$$

one can compute a posterior probability distribution, given the evidence provided by opening envelope A. The posterior probability distribution thus generated provides the basis for the maximum-expectation computation which leads the agent to the best decision (cf. Clark & Shackel, 2000; Dietrich & List, 2005; Sobel, 1994).

However, if one has no such information at hand, if one does not know a prior probability distribution over the size of the cheques in the envelopes, then "[…] there is no way one can compute posterior probabilities, and so use these in a computation of expectation" (Priest & Restall, 2003, p. 8):

Now if one opens the envelope and discovers, say, \$10, the only two possibilities left with non-zero probability are ‹10, 5› and ‹10, 20›. But since one has no information about the prior probabilities of these two possibilities, one cannot compute their posterior probabilities. In particular, one cannot argue that, since there are two possibilities left, each has probability ½. Thus, for example, if the prior probability distribution was such that ‹10, 5› and ‹5, 10› each had probability ½, whilst everything else had probability 0 (which is consistent with our information), then the posterior probabilities of these two options [i.e. of ‹10, 5› and ‹10, 20›, C.H.] are 1 and 0, respectively. On the other hand, if it was such that ‹10, 20› and ‹20, 10› each had probability ½, whilst everything else had probability 0, then the posterior probabilities of these two

options [i.e. of ‹10, 5› and ‹10, 20›, C.H.] are 0 and 1, respectively. To claim that the relevant posterior probabilities are a half each is, therefore, fallacious. (Priest & Restall, 2003, p. 8f.)

Probability distributions over (an infinite vs. finite number of) states: In the situation just considered, there is insufficient information to compute the relevant expectations. The same situation arises, even before opening envelope *A*, if one supposes that Bayesian considerations are still relevant to rational choice. Bayesians hold that, regardless whether the decision-maker is allowed to open envelope *A*, he must have some prior probability distribution concerning the size of the cheques in the two envelopes. "The probabilities in question cannot be objective, of course; not enough about such probabilities is known" (Priest & Restall, 2003, p. 9). Indeed, they are subjective and the operative notion of probability here (and, presumably, in classical decision theory in general) is *epistemic* in the following sense: "Calculations of expected [values, C.H.] require epistemic probability, and epistemic probability depends on the information that one (i.e. the subject) has" (Katz & Olin, 2007, p. 906).

Now, in accordance with the original description of the Two Envelopes situation, assume that the decision-maker does not know how much is in his envelope *A*. In Bayesians' view, he believes only that the envelopes contain amounts in accord with his probability distribution over the possible states of the world. Ask whether the number of states is finite or infinite.

- Jackson et al. (1994) and Sobel (1994) discuss examples for bounded cases. That is to say, there is only a finite number of possible and relevant states and a finite number of possible sums of money in the two envelopes, respectively, because the agent knows that there is only a finite amount of money in the world. According to Jackson et al. (1994), such an individual will have a prior probability distribution concerning the total sum of money in the two envelopes, with the consequence that for some amount in *A*, it is not equally likely that, when *A* has *that* value, *B* has two or half times *that* value. From this, they infer in the end that, for some values of the content of *A*, it is reasonable to switch, for others it is not. However, suppose, for example, that the agent's choice will be based on the toss of a fair coin. In that case, it seems clear that the likelihood that he picked the envelope containing the larger, or the smaller, sum is 0.5. It is hard to see how the fact that the world's money supply has an upper and a lower bound – or that he has a prior probability distribution concerning the total sum of money in the two envelopes – could have any bearing on this. Notice, furthermore, that, if there is an upper bound on the amounts that may be in the envelopes – as seen before, the stipulation of an upper bound goes beyond the conditions of the original Two Envelopes Problem anyway –, it can be shown that the Two Envelopes Problem is trivially escaped (Norton, 1998, p. 37). To face the actual Two Envelopes Problem, Norton (1998) concludes, therefore, that one should move on with the unbounded case.

- Technically, things are a little more complicated when it comes to the theoretical question of distributions over infinite sets of values (cf. Broome, 1995; Castell &

Batens, 1994; Chalmers, 2002). The basic idea, however, is simple: The prior probability distributions in question are over an infinite space.[1]

But let us skip a detailed characterisation of unbounded cases and pay more attention to a solid objection to the Bayesian approach to the Two Envelopes problem *in general*. To put it in a nutshell, there is no one distribution over all the possibilities that recommends itself since there is simply no uniform distribution:

> There are infinitely many equally good distributions consistent with our knowledge. We may nullify any argument to the effect that one should switch or keep based on a probability distribution by pointing out that there are equally good distributions that recommend the opposite. We have already seen this in the case in which we open the envelope.
> (Priest & Restall, 2003, p. 9)

All the agent knows about the prior probability distribution is that all the possibilities with non-zero probability are of the form ‹q_1, q_2›, where $q_1 = 2q_2$, or $2q_1 = q_2$. On this basis, it is obsolete to postulate (as Jackson et al., 1994 do) that he must have a *specific* prior probability distribution concerning the size of the cheques in A and B.[2]

One final remark on this angle: Of course, the decision-maker might be a *complete* subjectivist about the matter: Whatever probability distribution does, in fact, reflect your degrees of belief, go with that and pick a_1 or a_2 correspondingly. However, if one is *such* a subjectivist, there is no point in dealing with *rational* choice anymore—suppose, for example, that you had to bet on one of the two possibilities in terms of the relative sizes of the two envelopes: *Given what you actually know*, would you really have a preference for betting on one over the other?

The logician's reply: Another possibility to deal with the Two Envelopes Problem, one might suppose, is to abandon appeal to Bayesian considerations and probability (distributions) altogether, in favour of some other principle of decision-making. In fact, as Smullyan (1992) points out, probability is really quite inessential to the heart of the Two Envelopes Problem; he presents it as a *logical paradox*, i.e. "two contrary, or even contradictory, propositions to which we are led by apparently sound arguments. The arguments are considered sound because, when used in other contexts, they do not seem to create any difficulty" (Heijenoort, 1967). Here are the propositions derived by Smullyan (1992, p. 174):

Proposition 1: *The amount you will gain by trading, if you do gain, is greater than the amount Proposition 1: you will lose, if you do lose.*

Proposition 2: *The two amounts are really the same.*

---

[1] Consider, for example, the following states which correspond to possible amounts of money in A and B, respectively: (1, 2), (2, 1), (2, 4), (4, 2), (4, 8), (8, 4), …, (2^n, 2^{n+1}), (2^{n+1}, 2^n), … Norton (1998) offers some probability distribution over these states: $p_n = P(2^n, 2^{n+1}) = P(2^{n+1}, 2^n) = (1/2)(1-k) k^n$. Cf. p. 38.

[2] Albers et al. (2005) confirm this conclusion (p. 106): "The Bayesian solutions obtained, however, are not applicable because they depend on the unknown function f. The really interesting element of this two-envelope problem is, however, that *nothing is known* about f (even not if such an f actually exists).".

He proves both of them in the following way: To establish Proposition 1, "let A be the amount you are now holding, and then the other envelope either contains $2A$ or $A/2$. If you gain by swapping, you gain $2A-A=A$, and if you lose by trading, you lose $A-A/2=A/2$. Since A is greater than $A/2$, the proposition is proved." (Smullyan, 1992, p. 174).

To prove Proposition 2, "let D be the difference between the two amounts in the envelopes. If you gain by swapping, you gain D and if you lose by swapping you lose D. Since D equals D, the proposition is proved." (Smullyan, 1992, p. 174). Both proofs seem to be sound, but it cannot be the case that both are correct.

This is a version of the Two Envelopes Problem without probabilities (Cf. also Chase, 2002; Chen, 2007). Only the (unknown) absolute values in the envelopes and for gains and losses play a role in Smullyan's (1992) formulation of the problem. To identify an act (i.e. a_1 or a_2) as rational (or as the one which maximises the agent's money), the amount the decision-maker will gain by trading, if he does gain, is compared to the amount he will lose, if he does lose. No expected values and no probabilities are needed to come to a rational solution for the decision problem.

So, from Smullyan's point of view, the difficult topic of probabilities for what the amount in the envelopes is, is irrelevant. Instead, the problem in the Two Envelopes Problem turns out to be that some of the terms that are used in the initial description of the decision situation, in general, and some of the terms that are used in both proofs (by Smullyan, 1992), in particular, are *ambiguous*; i.e. they can be interpreted in different ways. Speaking for the initial Two Envelopes Problem, the crucial ambiguous phrase is that "one envelope contains twice as much money as the other"; speaking for Smullyan's version of the Two Envelopes Problem, the ambiguous terms are "the amount you will gain by trading, if you do gain" and "the amount you will lose by trading if you do lose". What do these phrases refer to? This all depends on what you mean by them and under what circumstances the agent gains and loses, respectively.

During the course of this paper, the importance of the difference between interpretations of the Two Envelopes decision situation or of particular terms and the importance of different solutions for different formulations of the Two Envelopes Problem, respectively, will become clear.

However, up to this point and after this rough sketch, it can be summarised that the issue discussed in the Two Envelopes Problem apparently cannot be settled unless something additional (i.e. further assumptions) is incorporated to define the problem more precisely: "The literature shows that there are many ways to 'solve' this problem but, in absence of additional information, these solutions cannot be regarded as satisfactory" (Albers et al., 2005, p. 90). But this conclusion is a little

premature because there does exist a promising way to handle the *original* Two Envelopes Problem.

3 The approach by Priest and Restall (2003)[3]

"We have just solved this paradox."—Priest and Restall.

In opposition to the articles listed above, Priest and Restall (2003) focus on the original description of the Two Envelopes situation; according to them, there is no need for introducing further assumptions, which modify the initial problem. While others try to settle the question of why it seems to be rational to choose a_2 (or why this conclusion is mistaken, respectively), they show that, if one poses this question, one tacitly presupposes a specific formulation of the problem. In contrast to this, Priest and Restall (2003) identify *different* formulations which they regard as, *prima facie*, equally good and which lead to different results. By distinguishing three different forms of reasoning about the Two Envelopes situation (and corresponding mechanisms), Priest and Restall (2003) argue that the basic set-up of this problem basically allows three different interpretations/specifications (related to three different explanations) which overall should make the decision-maker indifferent between a_1, a_2 and the alternative of being indifferent between a_1 and a_2. Here are their three forms of reasoning (Priest & Restall, 2003).

Form 1: Let n be the minimum of the quantities in the two envelopes. Then there are two possibilities, which can be depicted in this manner (Matrix 1):

	Possibility 1	Possibility 2
Your Envelope A	n	$2n$
Other Envelope B	$2n$	n

By the principle of indifference, the probability of each possibility is ½. The expected value of a_1 is

$$1/2 \times n + 1/2 \times 2n = 3n/2.$$

The expected value of a_2 is

$$1/2 \times 2n + 1/2 \times n = 3n/2.$$

---

[3] An almost identical version of their text, the authors published in 2008.

Conclusion: Be indifferent between a_1 and a_2, choose a_3: $p \times a_1 + (1-p) \times a_2$ (for $p \in [0, 1]$).[4]

Form 2: Let x be the amount of money in envelope A. Then there are two possibilities, which can be depicted in this manner (Matrix 2):

	Possibility 1	Possibility 2
Your Envelope A	x	x
Other Envelope B	$2x$	$x/2$

By the principle of indifference, the probability of each possibility is ½. The expected value of a_1 is

$$1/2 \times x + 1/2 \times x = x$$

The expected value of a_2 is

$$1/2 \times 2x + 1/2 \times x/2 = 5x/4$$

Conclusion: Choose a_2.

Form 3: Let y be the amount of money in envelope B. Then there are two possibilities, which can be depicted in this manner (Matrix 3):

	Possibility 1	Possibility 2
Your Envelope A	$2y$	$y/2$
Other Envelope B	y	y

By the principle of indifference, the probability of each possibility is ½. The expected value of a_1 is

$$1/2 \times 2y + 1/2 \times y/2 = 5y/4$$

The expected value of a_2 is

$$1/2 \times y + 1/2 \times y = y.$$

Conclusion: Choose a_1.

---

[4] a_3 consists of infinitely many mixed strategies available to the decision-maker. In the theory of games a player is said to use a *mixed strategy* whenever he or she chooses to randomise over the set of available actions (here: a_1 and a_2).

According to Priest and Restall ([2003]), all three answers give the right solution, in three different circumstances. The point is, however, that the relevant reasoning determining what the decision-maker ought to do to maximise his payoff is *under-determined* by the original description of the situation:

> "The correct way to reason, in the sense of maximising your return given the possibilities – which, after all, is the aim of each kind of reasoning – depends on the process [unknown to the decision-maker, C.H.] by which the money ends up in the envelopes. For each form of reasoning there are mechanisms such that, if *that* mechanism was employed, the reasoning delivers the correct answer." (Priest & Restall, [2003], p. 4)

Here are their three examples reflecting three different types of processes by which the conditions of the Two Envelopes Problem can be implemented (Priest & Restall, [2003], p. 4):

> **Mechanism 1:** A number, n, is chosen in any way one likes. One of the two envelopes is chosen by the toss of a fair coin, and n is put in that envelope; $2n$ is put in the other. (Let us say that A contains n if, and only if, the coin lands tails.)
> **Mechanism 2:** A number, x, is chosen in any way one likes. That is put in your envelope A. Either $2x$ (tails) or $x/2$ (heads) is then put in the other envelope B, depending on the toss of a fair coin.
> **Mechanism 3:** A number, y, is chosen in any way one likes. This is put in the other envelope B. Either $2y$ (tails) or $y/2$ (heads) is then put in your envelope A depending on the toss of a fair coin.

It is easy to see that the three different forms of reasoning are correct for each of the corresponding mechanism once one has seen the three possibilities. In other words, the decision-maker should apply Form 1 or 2 or 3, respectively, for computing expectations if (s)he knew that the money was selected and deposited into the two envelopes as described in the corresponding mechanism. Less explicitly, Priest and Restall ([2003]) say something stronger than that. Apart from the conditional "If Mechanism i was employed, then Form i of reasoning is correct" (for $i=1, 2, 3$), the inverse conditional also seems to hold: "If Mechanism i was *not* employed, then Form i of reasoning is *not* correct" (cf. Priest & Restall, [2003], p. 5). In short, they maintain that each of the following three biconditionals is true:

> (Mechanism 1 was employed) $\leftrightarrow$ (Argument 1/Form 1 of reasoning is correct)
> (Mechanism 2 was employed) $\leftrightarrow$ (Argument 2/Form 2 of reasoning is correct)
> (Mechanism 3 was employed) $\leftrightarrow$ (Argument 3/Form 3 of reasoning is correct)

If the decision-maker knew that, for instance, Mechanism 2 was employed, (s)he would clearly know what (s)he ought to do, namely to choose a_2 (in accordance with Form 2). Because, speaking for a circumstance well-suited to Form 2, "[…] the two possibilities countenanced in that form of reasoning [i.e. $\langle x, 2x \rangle$ and $\langle x, x/2 \rangle$; C.H.] match up precisely with the different outcomes of Mechanism 2 […]. This cannot be said of Mechanism 1 or Mechanism 3" (Priest & Restall, [2003], p. 7).

If someone has any doubt about this, intuitions about the scenarios can be checked by a series of trials: "[…] a sequence of trials is generated employing, for example, Mechanism 2: Adopting the policy of switching comes out 5/4 ahead of the policy of keeping in the long term (and changing at random comes out 9/8 ahead)" (Priest & Restall, 2003, p. 5). Or, put another way, a situation where the conditions of the Two Envelopes Problem are implemented by employing Mechanism 2 seems essentially the same as one in which the agent has the opportunity to accept a bet on a fair coin at pay-off double or half. However, since the original description of the situation does not include the relevant information about which mechanism wasm in fact employed, the decision-maker, according to Priest and Restall (2003), should be indifferent between a_1, a_2 and a_3 (by the principle of indifference). In other words, he should pick a *mixed strategy*.[5]

This is the solution offered by Priest and Restall (2003), who claims to have solved the (paradigm version of the) Two Envelopes Paradox. So far so good; but why, it may fairly be asked, have many authors called the analysed decision problem a paradox?

Many authors refer to the Two Envelopes Problem as a paradox.

(i)   because they primarily consider the *switching argument* (i.e. Form 2 of reasoning) when they analyse the Two Envelopes decision problem. This argument is supposed to be paradoxical, "because the same reasoning would lead you to switch back again if you did switch" (Broome, 1995, p. 6).
Others identify a paradox within the decision problem under consideration.

(ii)   because they confront the apparently sound Form 1 of reasoning (or the initial intuition according to which there is nothing to choose between the two envelopes to start with) with the apparently likewise plausible Form 2 of reasoning (Arntzenius & McCarthy, 1997, p. 42). One has two apparently valid arguments with conflicting conclusions, then.

In any case, Priest and Restall (2003) defeat these attempts to create a paradox. They point out that it is tried to produce a Two Envelopes *Paradox* by "giving reasoning of Form 2 in a context where Mechanism 1 is deployed. This, of course, gives the wrong results" (p. 280).[6]

In particular, ad (i), Form 2 of reasoning is perfectly fine if the conditions of the Two Envelopes Problem are implemented by using the corresponding Mechanism 2. *Then* it is rational to choose a_2 (see below) and it would *not* be rational to 'switch back again if you did switch': Under these circumstances, regardless of whether possibility 1 or 2 materialises, envelope A contains x and the decision-maker would

---

[5] In this case, the mixed strategy is the assignment of a probability of 1/3 to each main strategy (i.e. a_1, a_2 and a_3). This allows for the decision-maker to randomly select a pure strategy (i.e. a_1, a_2).

[6] Norton (1998) serves as an alarming example: "In the exchange paradox, a game is played in which a randomly chosen amount of money is placed in one envelope and twice that amount in a second. The envelopes are then shuffled and randomly assigned to two players such that each player has an equal chance of receiving the first envelope. The players are then given the option of swapping. Player 1 reasons: My envelope will contain some amount of money—x dollars, say. There is a probability 1/2 that player 2's envelope has 2× and probability.
1/2 that it has x/2. Therefore my expectation in swapping is […]" (p. 34f.).

be justified in expecting a higher amount of money in B. So he should trade A for B and keep B afterwards since, before and after doing a_2, A has the value of x and B has the expected value of $1{,}25x$. Under other circumstances, i.e. if Mechanism 2 is not employed, the agent, following Priest and Restall (2003), should not even switch A for B in the first place. Besides, it is not legitimate to only consider Form 2 of reasoning because this form of reasoning is not implied by the original description of the situation – in particular, one envelope contains twice the amount of the other, but the decision-maker does not know which one is which.

On the other hand, ad (ii), although the conclusion from Form 2 (the agent should switch) is not compatible with the conclusion from Form 1 (the agent has no reason to switch), both answers give the agent the right solution, but, as seen, under two *different* circumstances. The authors clarify that Form 2 is not weaker than Form 1 of reasoning. Besides, it is not legitimate to ignore Form 3 of reasoning, which would be correct if Mechanism 3 was employed: If all three formulations of the Two Envelopes Problem and corresponding mechanisms, respectively, are taken into consideration, there is only one thing for the rational decision-maker to do: Priest and Restall (2003) conclude that he should be indifferent between a_1, a_2 and a_3 (or Form 1, Form 2, and Form 3, respectively). Insofar, they are convinced to have not only solved the Two Envelopes Paradox, but also the decision-makers Two Envelopes decision problem.

4 A case against Priest and Restall (2003)

4.1 Decision problem buried

To judge what one must do to obtain a good or avoid an evil, it is necessary to consider not only the good and the evil in itself, but also the probability that it happens or does not happen; and to view geometrically the proportion that all these things have together.

The Port-Royal Logic, 1662.

It is one thing to point out that the switching argument, which has attracted much attention in the literature on Two Envelopes Problem, is not paradoxical at all and to outline how to solve Two Envelopes Paradoxes. Priest and Restall (2003) succeeded with these tasks. However, it is quite another thing to solve the original decision problem. To treat the Two Envelopes Problem more precisely (from now on: TEP), I may put the conditions for TEP as follows:

(C1)   There are two sealed envelopes that contain some quantity of money, which can (C1) be of any positive real magnitude.
(C2)   The decision-maker, who wishes to maximise his money, knows that one envelope (C2) contains twice as much money as the other.
(C3)   The decision-maker randomly selected an envelope – this envelope is called "A", (C2) the other envelope is called "B" –, without knowing which envelope contains the (C2) larger sum and which the smaller.
(C4)   Now, the decision-maker has two alternatives for action:

(C4) a_1: Keep envelope A.
(C4) a_2: Switch envelope A for envelope B.
(C5)   He gets the sum of money in the envelope that he finally selects.

That is to say, if, and only if, these five necessary conditions are fulfilled, which together form a sufficient condition, then there is a (original) TEP. A decision situation where (C1)–(C5) hold is called a Two Envelopes situation. By contrast, I only use the phrase "Two Envelopes Paradox" to refer to the attempts by some writers to create a paradox out of a more or less simple decision problem.

In this chapter, I will argue that Priest and Restall ([2003]) have not managed to sort out the TEP decision problem, and that neither Form and Mechanism 1–3 nor particular forms of reasoning and mechanisms, in general, are really relevant for the decision-maker. Finally, I will give an argument such that the premises imply the conclusion that the decision-maker should be indifferent between a_1 and a_2. It is clear that anything except indifference has no rational ground: "arguments attempting to justify some difference between switching and keeping get no grip" (Priest & Restall, [2003], p. 10). Nevertheless, a sound argument justifying indifference between switching and keeping has not yet been propounded.

4.2 Close but no cigar

The unsolved decision problem
As I see it, Priest and Restall's ([2003]) argument for solving the *decision problem* runs as follows:

(P1′)   The decision-maker acts in accordance with the *Bayesian Principle*, i.e. he chooses an act of maximum estimated desirability.
(P2′)   The decision-maker wishes to maximise his money. Only monetary values are relevant for his preferences.[7]
(P3′)   The decision-maker knows that his decision situation is a Two Envelopes situation.
(P4′)   The decision-maker knows that, for each Two Envelopes situation, there are k different mechanisms or classes of mechanisms—mechanisms by which the money can end up in the envelopes A and B—and k corresponding forms of reasoning such that (for $k = 3$)
(P5′)   If, and only if, Mechanism i was employed, Form i of reasoning is correct (for $i=1, …, k$).
(P6′)   The decision-maker knows what Form i of reasoning says and which conclusion it contains (i.e. choose either a_1 or a_2 or a_3).
(P7′)   The decision-maker does not know which mechanism was in fact employed.
(P8′)   The principle of indifference holds.
(C1′)   The decision-maker should ascribe the probability $1/k$ to the event that Mechanism i was employed.

---

[7] That is to say, the shape or the colour of the two envelopes (etc.) is irrelevant for decision-making.

(C2′)   The decision-maker should be indifferent between a_1, a_2 and a_3.

(C3′)   The decision-maker should pick a mixed strategy such that each main strategy (i.e. a_1, a_2 and a_3) is played with the probability of 1/3, which allows for the decision-maker to randomly select a pure strategy (i.e. a_1, a_2).

The crucial point in this argument is that k equals 3 and that it equals a *finitely* large number, respectively. The validity of this assumption in (P4′) is doubtful. Doubts can be expressed in the following questions:

It is beyond dispute that there are only and not more than two or three, respectively, possible strategies available to the agent [a_1 and a_2 (and a_3 consisting of infinitely many *mixed* strategies)]; but are there only three possible processes by which the money can end up in the envelopes? Do Mechanism 1–3 stand for classes of mechanisms which contain *all* possible mechanisms? What do Priest and Restall (2003) say exactly in this respect?

Unfortunately, Priest and Restall (2003) make no clear statement. They simply speak of three *examples* for three different types of processes by which the conditions of TEP can be implemented (see above). They do not say whether Mechanism 1–3 correspond to different classes of mechanisms or whether there are procedures of a different sort for putting the money into the envelopes (and if so, how many there are).

In fact, one can imagine other mechanisms (and corresponding forms of reasoning) which are *different* from Priest and Restall's (2003) Mechanism 1–3 (Form 1–3) and which, therefore, cannot be assigned to one of them. I will argue that their argument does not have to work for $k>3$. Here is an example for an additional mechanism and form of reasoning, respectively:

Mechanism 4: Either envelope A or B is selected first, depending on the toss of a fair coin. A number z, chosen arbitrarily, is put in *that* envelope. Either $2z$ or $z/2$ is then put in the other envelope, depending on the toss of a fair coin.

The corresponding form of reasoning is as follows:

Form 4: Let z be the amount of money put into the envelope selected first *by the game master*, i.e. "z" stands for "the sum of money put into the envelope selected first *by the game master*". There are four possibilities, which can be depicted in this manner (Matrix 4):

	A selected first (by the game master) and $C(A) < C(B)$	A selected first (by the game master) and $C(A) > C(B)$	B selected first (by the game master) and $C(A) < C(B)$	B selected first (by the game master) and $C(A) > C(B)$
A	z	z	$z/2$	$2z$
B	$2z$	$z/2$	z	z

By the principle of indifference, the probability of each possibility is ¼. The expected value of a_1 is

$$1/4 \times z + 1/4 \times z + 1/4 \times z/2 + 1/4 \times 2z = 9z/8$$

The expected value of a_2 is

$$1/4 \times 2z + 1/4 \times z/2 + 1/4 \times z + 1/4 \times z = 9z/8$$

Conclusion: Be indifferent between a_1 and a_2, choose a_3: $p \times a_1 + (1-p) \times a_2$ (for $p \in [0, 1]$).

Thus, there is at least one mechanism that does not belong to Form 1–3 (and vice versa) and there is at least one form of reasoning that does not belong to Mechanism 1–3 (and vice versa). Priest and Restall (2003) disregard such an equally good alternative to Form/Mechanism 1–3.

Since the decision-maker knows that the probability that Mechanism 4 was deployed is greater than zero (once he has seen this other possibility), he might be inclined to conclude that he should no longer be indifferent between a_1, a_2 and a_3. To reach this conclusion, he might reason as follows:

(Mechanism 1 was employed)$\rightarrow$(a_3 is the thing to do).
(Mechanism 2 was employed)$\rightarrow$(a_2 is the thing to do).
(Mechanism 3 was employed)$\rightarrow$(a_1 is the thing to do).
(Mechanism 4 was employed)$\rightarrow$(a_3 is the thing to do).

By the principle of indifference, the probability that Mechanism i was employed is 1/4, for $i=1, 2, \ldots, 4$. Therefore,

in 1 out of 4 cases a_1 is the thing to do;
in 1 out of 4 cases a_2 is the thing to do;
in 2 out of 4 cases a_3 is the thing to do.

Conclusion: The decision-maker should still be indifferent between a_1 and a_2, but the argument is slightly different since (P4′) was challenged.

Moreover, as an outlook, (C2′) does not have to follow from (C1′) and the foregoing premises if $k>3$. In general, if the number of forms of reasoning where a_1 is recommended was unequal to the number of forms where a_2 is recommended, Priest and Restall's (2003) argument would not only be incorrect but lead to an unacceptable conclusion. Apart from the one example given (Mechanism 4), how many other mechanisms could be added to the list of possible processes by which the money can end up in A and B?

To put it in a nutshell, there may be infinitely many. That there are indefinitely many different ways of filling the two envelopes can be proved, and proof can be found in the Appendix.

But if an infinite number of possibilities are all equally likely, the chance of any one outcome must be zero. Then every outcome has a zero chance, and this is nonsense. Taken in (standard) Cantorian terms, if ∞ is the smallest Cantorian set-sizing infinity, then $1/\infty$ is not defined. Also the limit of $1/k$ as k increases without limit— $\lim_{k \to \infty} 1/k$ —is 0.[8]

---

[8] Some decision theorists therefore view 'infinite values' as the stuff of a "monstrous hypothesis" (Nalebuff 1989, p. 176, 178) that should be banned from decision theory (cf. Jeffrey 1983, p. 153–4).
 For the controversial debate on the extension of decision theory to infinite quantities cf. McClennen (1994), Sorensen (1994), Sobel (1996).

Hence, (P4') is false and it would presumably be true if it said that $k=\infty$. But then (C1') would be at least problematic—the probability that Mechanism i was employed would be $1/\infty=0$ which does not make much sense.

Moreover, the introduction of *mechanisms* in the problem solving for TEP and Priest and Restall's (2003) choice of words,[9] respectively, is inappropriate. Because the conditions of TEP might, of course, be satisfied without a plan or design. Perhaps it is just a coincidence that there are two envelopes, one of which contains twice as much as the other. John's mother informs him that his Uncle Ali and his Uncle Baba have each sent John an envelope with a sum of money for his birthday and that it happens that one sum is twice as large as the other.

4.3 Solving the two envelopes decision problem

Admittedly, one might draw the lesson from this objection to Priest and Restall's (2003) argument (for solving the Two Envelopes decision problem) that abandoning appeal to probability altogether, in favour of some other principle of decision-making, could be promising. But in this case, one would still not have an argument justifying indifference between switching and keeping. As it turns out, it is not necessary to give up probabilities for this purpose.

My argument for solving the Two Envelopes decision problem runs as follows:

(P1*)   The decision-maker acts in accordance with the *Bayesian Principle*, i.e. he chooses an act of maximum estimated desirability.

(P2*)   The decision-maker wishes to maximise his money. Only monetary values are relevant for his preferences.

(P3*)   The decision-maker knows that the conditions of TEP, i.e. (C1)–(C5), are satisfied in his decision situation and that they can be fulfilled in various ways.

(P4*)   The decision-maker does not know whether the expected value of a_1 is greater than, less than, equal to the expected value of a_2 (otherwise there would not be a decision problem). But he does know that either EV $(a_1) >$ EV (a_2) or EV $(a_1) <$ EV (a_2) or EV $(a_1) =$ EV (a_2) is the case.

(C1*)   The decision-maker knows that he could have selected the *other* envelope *in the first place* just as well, he has no decision criterion—he chose the envelope he actually chose, to which we refer by the term "*A*", for no particular reason (see (C3)).[10]

(C2*)   The decision-maker knows that in each Two Envelopes situation envelope A would be envelope B and envelope B would be A if he had decided differently in the first place.

---

[9] That is to say, they refer to circumstances of Two Envelopes situations or to distributions or to manners in which the conditions of TEP can be implemented by using the word "mechanism".

[10] (C3): The decision-maker randomly selected an envelope—this envelope is called "*A*", the other envelope is called "*B*" –, without knowing which envelope contains the larger sum and which the smaller. This argument requires that "*A*" and "*B*" function as definite descriptions whereas it seems to be a more plausible interpretation to read "*A*" and "*B*" as rigid designators when the different mechanisms are described.

(C3*)   The decision-maker knows that in each Two Envelopes situation where EV $(a_1) >$ EV (a_2) holds, EV $(a_1) <$ EV (a_2) would hold if A was B and B was A. He also knows that in each Two Envelopes situation where EV $(a_1) <$ EV (a_2) [EV $(a_1) =$ EV (a_2)] holds, EV $(a_1) >$ EV (a_2) [EV $(a_1) =$ EV (a_2)] would hold if A was B and B was A.

(C4*)   There are *pairs* of Two Envelopes situations: For each Two Envelopes situation, implemented in a certain way, there is another Two Envelopes situation, implemented in the same way, but where the expected values of the alternatives are just interchanged (because the decision-maker had decided differently in the first place).

(C5*)   It does not follow that the decision-maker should think that indefinitely many Two Envelopes situations, implemented in different ways, occur with equal probability; i.e. for different pairs different probabilities are possible. But his degree of belief that his decision situation is a situation where EV $(a_1) >$ EV (a_2) holds should be equal to the degree that it is a situation where EV $(a_1) <$ EV (a_2) is the case; i.e. he should be *pairwise indifferent*.

(C6*)   The decision-maker should be indifferent between a_1 and a_2.

(C7*)   The decision-maker should play a_3.

The moral of the story is simple: The agent can safely ignore Form 1, 2, 3, 4 etc. of reasoning and Mechanism 1, 2, 3, 4 etc. He does not have to think about various arguments for solving TEP or manners in which the money was selected and deposited into the two envelopes because, on the one hand, the particular forms of reasoning are only useful in the right context whereas the agent does not have the relevant information on the context of his situation; and on the other hand, Priest and Restall's (2003) argument for solving the decision problem where forms of reasoning and mechanisms play an important role turned out to be untenable.

Instead, I suggested an argument with uncontroversial premises which justifies the view that the information given in the original characterisation of the Two Envelopes scenario seems to give us no good reason to switch. There is therefore nothing to break the *symmetry*—for every Two Envelopes situation we can point to its dual where the opposite choice is recommended –, and so to give ground for anything other than indifference.

This is the simple solution to the well-known Two Envelopes Problem!

None of the authors have any competing interests in the manuscript.

Appendix

Constructive Approach to proving the following claim

Claim: There are indefinitely many different ways of filling the two envelopes A and B.

Let M_1 and M_2 be two mechanisms of filling the envelopes A and B. We define a new mechanism M' as follows.

Proposition: There is no $s \in \mathbb{N}$ such that the number of mechanisms to fill envelopes A and B is smaller than or equal to s.

Proof (Reductio ad absurdum). : Assume that s exists as above. We know that $s \geq 4$, because of the four mechanisms stated earlier. Let M_1 and M_2 be two different mechanisms for filling the envelopes as above. We create a new mechanism M_3 as follows: A fair coin is tossed. If the outcome is heads, then the envelopes will be filled using method M_1. If the outcome is tails, then M_2 is used. M_3 is different from M_2 and M_1. We can then define M_4 using M_3 and M_2 as before until we obtain $s+1$ different mechanisms for filling the envelopes.

References

Albers, C. J., Kooi, B. P., & Schaafsma, W. (2005). Trying to resolve the two-envelope problem. *Synthese, 145*, 89–109.

Arntzenius, F., & McCarthy, D. (1997). The two envelope paradox and infinite expectations. *Analysis, 57*, 42–50.

Broome, J. (1995). The two-envelope paradox. *Analysis, 55*, 6–11.

Castell, P., & Batens, D. (1994). The two-envelope paradox: the infinite case. *Analysis, 54*, 46–49.

Chalmers, D. (2002). The St. Petersburg two-envelope paradox. *Analysis, 62*, 155–157.

Chase, J. (2002). The non-probabilistic two envelope paradox. *Analysis, 62*, 157–160.

Chen, J. (2007). The puzzle of the two-envelope puzzle—A logical approach. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1132506

Christensen, R., & Utts, J. (1992). Bayesian resolution of the "exchange paradox." *The American Statistician, 46*, 274–276.

Clark, M., & Shackel, N. (2000). The two-envelope paradox. *Mind, 109*, 415–442.

Dietrich, F., & List, C. (2005). The two-envelope paradox: An axiomatic approach. *Mind, 114*, 239–248.

Heijenoort, J. V. (1967). Logical paradoxes. In P. Edwards (Ed.), *The encyclopedia of philosophy* (Vol. 5, pp. 45–51). Macmillan.

Jackson, F., Menzies, P., & Oppy, G. (1994). The two envelope "paradox." *Analysis, 54*, 43–45.

Jeffrey, R. C. (1983). *The logic of decision*. University of Chicago Press.

Katz, B. D., & Olin, D. (2007). A tale of two envelopes. *Mind, 116*, 903–926.

Linzer, E. (1994). The two envelope paradox. *The American Mathematical Monthly, 101*, 417–419.

McClennen, E. F. (1994). Pascal's Wager and finite decision theory. In J. Jordan (Ed.), *Gambling on God: Essays on Pascal's Wager* (pp. 115–137). Rowman and Littlefield Publishers Inc.

McGrew, T., Shier, D., & Silverstein, H. (1997). The two-envelope paradox resolved. *Analysis, 57*, 28–33.

Nalebuff, B. (1989). Puzzles: The other Person's envelope is always Greener. *Journal of Economic Perspectives, 3*, 171–181.

Neumann, J. V., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Norton, J. D. (1998). When the sum of our expectations fails us: The exchange paradox. *Pacific Philosophical Quarterly, 79*, 34–58.

Priest, G., & Restall, G. (2003). Envelopes and indifference. http://consequently.org/papers/envelopes.pdf (07/01/2012).

Smullyan, R. (1992). *Satan, cantor, and infinity: And other mind-boggling puzzles*. Alfred A. Knopf.

Sobel, J. H. (1994). Two envelopes. *Theory and Decision, 36*, 69–96.

Sobel, J. H. (1996). Pascalian Wagers. *Synthese, 108*, 11–61.

Sorensen, R. (1994). Infinite decision theory. In J. Jordan (Ed.), *Gambling on God: Essays on Pascal's Wager* (pp. 139–159). Rowman and Littlefield Publishers Inc.

Zabell, S. (1988a). Loss and gain: The exchange paradox. In: J. M. Bernardo, M. H. DeGroot, D. V. Lindley, & A. F. M. Smith (Eds.), Bayesian Statistics 3, Proceedings of the Third Valencia International Meeting, Oxford, pp. 233–36.

Zabell, S. (1988b). Symmetry and its discontents. In: B. Skyrms, W. Harper (Eds.) Causation, Chance, and Credence: Proceedings from the Irvine Conference on Probability and Causation (Vol. 1, pp. 155–90), Dordrecht.