



TO THINE OWN SELF BE UNTRUE: A DIAGNOSIS OF THE CABLE GUY PARADOX

DARRELL P. ROWBOTTOM AND PETER BAUMANN

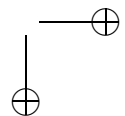
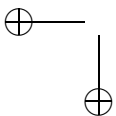
Abstract

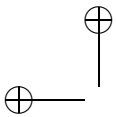
Hájek has recently presented the following paradox. You are certain that a cable guy will visit you tomorrow between 8 a.m. and 4 p.m. but you have no further information about when. And you agree to a bet on whether he will come in the morning interval (8, 12] or in the afternoon interval (12, 4). At first, you have no reason to prefer one possibility rather than the other. But you soon realise that there will definitely be a future time at which you will (rationally) assign higher probability to an afternoon arrival than a morning one, due to time elapsing. You are also sure there may not be a future time at which you will (rationally) assign a higher probability to a morning arrival than an afternoon one. It would therefore appear that you ought to bet on an afternoon arrival.

The paradox is based on the apparent incompatibility of the principle of expected utility and principles of diachronic rationality which are *prima facie* plausible. Hájek concludes that the latter are false, but doesn't provide a clear diagnosis as to why. We endeavour to further our understanding of the paradox by providing such a diagnosis.

1. *The Cable Guy Paradox and Hájek's Attempted Solution*

Hájek (2005) presents the following paradox. You are certain that a cable guy will visit you tomorrow between 8 a.m. and 4 p.m. but you have no further information about when. Nevertheless, you agree to a bet on whether he will come in the morning interval (8, 12] or in the afternoon interval (12, 4); you will win a given sum if you are correct, and lose the same sum if you are wrong. At first you have no reason to prefer one possibility rather than the other, so assign each an equal probability. (More specifically, you opt for a uniform probability distribution, over the 8 a.m.–4 p.m. period,





for the cable guy’s arrival time.¹ You also assume that an arrival at noon on the dot is impossible.) But then you realise that there will certainly be a future time at which you will (rationally) assign higher probability to an afternoon arrival than a morning one, because some time will elapse in the morning interval before the cable guy arrives. You also realise there may not be a future time when you will assign higher probability to a morning arrival than an afternoon one.

The paradox consists in the incompatibility of the following two, *prima facie* plausible, principles of rationality:

Maximise Expected Utility (MEU): you should act so as to maximise your expected utility, if you can. In the event that there is more than one action which maximises expected utility, each is equally as acceptable.

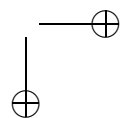
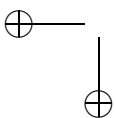
*Avoid Certain Frustration*² (ACF): when you have a choice between two options, you should not choose one of these rather than the other if you are certain that you will subsequently (rationally) wish you had selected the other one, unless this is true of both options.

The expected utility of a bet on MORNING is equal to the expected utility of a bet on AFTERNOON, initially, because the (subjective) probability of MORNING is equal to that of AFTERNOON and the stake will — you are sure — be the same no matter which option you choose. (It is assumed that the only relevant factor, in so far as utility is concerned, is money.) According to MEU, it is rationally permissible to bet either way.

However, you are certain that you will rationally wish you’d bet differently tomorrow if you bet on MORNING, but also that you may never rationally wish you’d bet differently if you bet on AFTERNOON. This is because (you recognise that) some time must elapse in the morning interval, before the cable guy arrives, even if it is only a tiny amount. It’s worth making it explicit at this point, as Hájek does not, that you expect such a change in preference because you are sure that you have adopted a Bayesian strategy

¹ It is not assumed that you are compelled to opt for such a distribution, on pain of irrationality, as objective Bayesians such as Jaynes (2003) and Williamson (2005) might suggest.

² Hájek (2005, p. 116) also discusses a related principle, namely *Avoid Self-Undermining Choices* (ASU): ‘Suppose you now have a choice between two options. You should not make a self-undermining choice if you can avoid doing so.’ Since Hájek’s discussion focuses on ACF, however, we also prefer to do so. We do not believe, and Hájek does not suggest, that considering ASU rather than ACF generates a fundamentally different paradox.



for updating, and will follow it through.³ (Else, how could you be sure how — or even *that* — your assignment will change before the cable guy arrives?) According to ACF, you are therefore in a situation where you should bet on AFTERNOON rather than MORNING.

Note that talk of ‘wishes’ is potentially misleading, because it might be impossible to make any conscious change due to practical (or psychological) constraints. The cable guy might arrive just a picosecond after eight, for instance! Yet the point remains that as far as you’re concerned, the expected utility of a bet on AFTERNOON will, *for (a rational) you at some future point*, be greater than the expected utility of a bet on MORNING. (We can add that you’re sure that if you *were* to adopt a Bayesian strategy for updating, then you *would* assign a higher probability to an AFTERNOON arrival than a MORNING arrival at some point.) So you’re sure that if you were asked which way you wanted to bet at that point, then you’d have to bet on AFTERNOON in order to obey MEU.

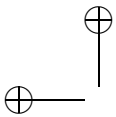
Hájek (2005, p. 118) concludes that ACF is false, but offers only the suggestion that this is because you cannot, in the cable guy scenario, ‘specify appropriately the [future rational] self... at issue’. That is to say, you cannot be sure at which particular point(s) — e.g. 9:15 a.m. or 8:00:01 a.m. — you will be in a scenario where you will prefer a bet on AFTERNOON. Hájek’s reasoning is only correct, however, because if you were certain you could identify such a future self then you would also be certain (if rational) that you could identify a time period after eight during which the cable guy would not arrive. To be sure that you’ll be ‘frustrated’ at 8:30 a.m. if you bet on MORNING, for example, is also to be sure that the cable guy will not arrive until after 8:30 a.m. That is, on the assumption that you are certain that you’ll discover that the cable guy arrives as soon as he does.⁴

Since this is equivalent to becoming certain *before you bet* that the cable guy won’t arrive until after 8:30 a.m. — even for reasons not relating to a consideration of future rational selves, e.g. due to testimony from a trusted friend — we think that Hájek’s brief diagnosis of the problem with ACF is incomplete at best.⁵ In what follows, we shall propose a different diagnosis.

³Note that you need only be convinced that you will update in such a way. Note also that you might be convinced you will update in an Objective Bayesian manner, rather than by Bayesian conditionalisation. See Williamson (Forthcoming).

⁴Let p be ‘I will know that the cable guy has arrived as soon as he does, and thereafter’, q be ‘I will not know at time t that the cable guy has arrived’, and r be ‘The cable guy will not arrive before t ’. It is clear that r is entailed by p and q .

⁵Note that Hájek (2005, p. 118) also admits ‘[T]his much of a paradox still remains: It is rational for you to choose the morning interval, knowing full well that there will be *some* future self of yours, very much like you now but better informed, who will wish you hadn’t.’



We will proceed by considering some variations on the original cable guy scenario.

2. *Three Illuminating Examples*

In the original cable guy scenario, which we will henceforth refer to as (A), you expect to know the outcome as soon as the cable guy arrives; the time of discovery (TD) will be identical to the time of arrival (TA). But what if this isn't the case? Consider the following variation on (A):

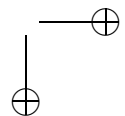
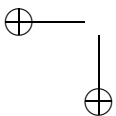
(B) You will not be at home tomorrow — but will leave keys to your abode with your neighbour — so will not discover the cable guy's arrival time until later. Here's how. At 7:30 a.m. on the day after tomorrow, on your return, your neighbour will hand you an alarm clock which will sound precisely 24 hours after the cable guy's arrival. You are sure about all this today. Hence, you are certain that you will later rationally wish to have chosen differently if you bet on MORNING but not certain that you will later rationally wish to have chosen differently if you bet on AFTERNOON.

We contend that (B) is like (A) in all relevant respects. Because we are only dealing with subjective probabilities, it shouldn't matter whether you are sure you will prefer one option rather than the other before the cable guy arrives rather than after. All that matters is that you are sure, and that you aren't sure that you will later prefer the other option. In a nutshell, whether $TD=TA$ is irrelevant for diachronic rationality.

This result is not terribly surprising. But now let us look at another variation on (A), which is much more interesting:

(C) The circumstances are the same as in (B) except that the alarm clock will run backwards. It will start at 4 p.m. and run towards 8 a.m., sounding when it shows the time that the cable guy arrived the day before. You are sure about all this today. You are therefore certain that you will rationally think you should have chosen differently on the day after tomorrow if you bet on AFTERNOON, but not certain that you will ever rationally think you should have chosen differently if you bet on MORNING!

In (B) you expect the evidence to be presented — as in (A) — according to the direction of time; that what happens at earlier times will be presented at earlier times, and so on. In (C), however, you expect the evidence to be



presented 'backwards'; that what happened at earlier times will be presented at later times, and *vice versa*. Now according to ACF, you should bet on AFTERNOON in (A) and (B) but on MORNING in (C). So your conviction about the order in which the evidence will be presented is crucial.

Our intuition, however, is that your conviction about the order in which the evidence will be presented *should not* have an effect on whether you bet on MORNING rather than AFTERNOON. Consider now the following:

(D) You find yourself in a cable guy scenario where the bet will be made by your friend, against you. After he has made his bet, however, you will then be free to choose in which order the evidence will be presented.⁶

In (D), should you choose to have the evidence presented 'forwards' (e.g. as in (B)) if your friend bets on MORNING, but 'backwards' (e.g. as in (C)) if he bets on AFTERNOON? Do you really think that doing this would improve your chances of winning? We guess (and hope) not. You can force him into a situation where he will face certain frustration (and you will avoid it) if you want to, but how will this prevent him from winning (and therefore prevent you from losing)?

If you now share our intuition, perhaps you will also agree with the following rough articulation of why it seems right: the order in which the evidence is presented can't possibly have any *effect* on, or even indicate the truth about, when the cable guy will actually arrive. So (D) forms the basis for an argument that the order of the evidence should not generally matter, and derivatively an argument against ACF. In what follows we will try to flesh this out.

3. The Problem with Avoid Certain Frustration

Above, we have seen that (one's conviction about) the order in which the evidence will be presented, in the cable guy scenario, is crucial in determining how one should act according to ACF. Now we want to explain precisely why the order shouldn't be a consideration in determining rational action.

Consider scenario (A), i.e. the original cable guy scenario, again:

- Let those possible worlds in which MORNING is true be called *M-worlds*.

⁶We can add that your friend is sure that you will choose the order, so need not consider ACF himself, if desired.

- Let those possible worlds in which AFTERNOON is true be called *A-worlds*.
- Now consider only those worlds in which you bet on MORNING.
- In *M-worlds*, the change of mind will be 'in vain' (IV), because MORNING is true.
- In *A-worlds*, the change of mind will *not* be 'in vain' (\sim IV), because MORNING is false.

Your subjective probability that you will have a future rational change of mind is unity. Your subjective probability that said change of mind will be IV is equal to your subjective probability that it will be \sim IV, and you recognise that these options are exhaustive and mutually exclusive. So $P(\text{Rational Change of Mind will be IV}) = P(\text{Rational Change of Mind will be } \sim\text{IV}) = 0.5$. As far as you're concerned, how you will rationally change your mind (in this instance) does not indicate whether MORNING is true or false.

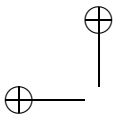
To be sure, you are not certain that you will have any future rational change of mind if you bet on AFTERNOON; to this extent, there is a clear asymmetry between a bet on MORNING and a bet on AFTERNOON. However, this does not help if what matters is whether your change of mind will be IV or \sim IV.⁷ And if all you are interested in is winning the bet, rather than 'the disutility of the pain of regretting your choice' (Hájek 2005, p. 114), then surely you should only be interested in whether or not your change of mind *will* be IV. It is easy to see why. It will be IV if and only if MORNING is true, and \sim IV if and only if MORNING is false!⁸

In support, consider also the following case where you are sure that your Bayesian updating will be interrupted. You are certain today that AFTERNOON is false. However, you are also certain that by tomorrow morning you will have forgotten this crucial fact (through no fault of your own), and will also fail to realise that you have forgotten such a fact. You are thus certain that you will have a rational difference of opinion the next morning if you bet on MORNING, and that you may not have such a difference of opinion if you bet on AFTERNOON, but you are also certain that this will be IV. It is clear, therefore, that your expected future rational change in mind should not matter. It should not matter simply because it is believed in advance, with certainty, to be IV.⁹

⁷Note that if you have no idea whether the change of mind will be IV or \sim IV, you have no reason to prevent it either. It could be a good thing, and it could be a bad thing.

⁸There is thus an asymmetry between what matters for the application of MEU and for the application of ACF.

⁹It is easy to see that this will be also be a concern even if the future rational change in mind is held only to be highly likely to be IV.



This example is imperfect, because it is possible to suggest that it is crucial, for the paradox to retain its force, for you to be sure that your future self will be updating on the basis of new evidence with *the same background information*. (There is also a considerable literature on the Sleeping Beauty problem, where forgetting is central; see, e.g., Hitchcock 2004 and Bradley & Leitgeb 2006.¹⁰ We do not want to run roughshod over this.) In short, the apparent force of ACF in the cable guy scenario may derive from a consideration such as ‘favour the rational self with the greater amount of relevant information’. Yet it remains the case that there need not, in any given case, be a link between obtaining more relevant information and making the correct decision. In fact, there are well-known problems with weight of evidence considerations. As Keynes (1921, p. 76) noted:

Weight cannot... be explained in terms of probability. An argument of high weight is not “more likely to be right” than one of low weight... Nor is an argument of high weight one in which the probable error is small.

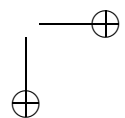
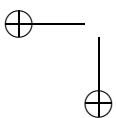
So while it is true that you will have more relevant information after 8 a.m., in the original cable guy scenario, it does not follow that this will enable you to make a ‘better’ decision in some objective sense.¹¹ What’s more, you might surely be able to recognise this at the time of the bet. And perhaps this explains, in part if not in whole, why the order in which the evidence is presented does not matter. Changing the order that the evidence is presented, in scenario (D), is *only* guaranteed to allow one to avoid a change of mind. But such a change of mind might be IV.

4. *Two Further Considerations*

We are not convinced that this is the whole story, however, because two further matters may be relevant in understanding how the cable guy paradox arises. The first is that you expect (with certainty) ‘settling evidence’ — that is to say, to *discover to your final satisfaction* when the cable guy arrives — at some stage. The second is that you do not expect to receive relevant

¹⁰These build upon other scenarios where Dutch Book arguments are shown to be problematic. See, for instance, Hájek (2005b) and Rowbottom (2007).

¹¹This may be unfavourable for Bayesianism *tout court*, but we are not Bayesians. See, for instance, Rowbottom (2008).



information which will bear on the *aleatory* probability of the cable guy's arrival in the morning or afternoon. We will deal with these issues in turn.

First, compare the cable guy scenario with one where you are trying to establish the probability that a given coin lands on heads when flipped. Initially you assume that heads and tails outcomes are equipossible, exhaustive, mutually exclusive, and independent; $P(\text{heads})=P(\text{tails})=0.5$. Then you learn that a future rational self will, *as a result of conditionalising on the evidence after some unspecified number of flips*, believe that $P(\text{heads})=0.6$. Should you opt for $P(\text{heads})=0.6$ (and $P(\text{tails})=0.4$), here and now? While we are not convinced that the answer lies in the affirmative, we think it is plausible that it does. Only after an infinite number of flips, which you can be reasonably assured you will never experience, would you gain 'settling evidence' — would you reach a stage at which you would not anticipate any (serious) possibility of further evidence that would make you change your mind. In example (A), however, you are sure of receiving 'settling evidence' when the cable guy arrives in your house. In examples (B), (C) and (D), this will be when the alarm sounds.¹²

Second, note that the sort of evidence you expect to be presented with — information about some period in which the cable guy doesn't arrive — will not, as far as you're concerned, inform you about the *aleatory* probability of MORNING and/or AFTERNOON. But if you expected to learn that he had a higher *propensity* of arriving in the afternoon than in the morning, e.g. that he would be driving and there would be a serious traffic jam in the morning, then changing your preference *here and now* would seem advisable.¹³ It is easy to motivate this view by considering what would happen if one were put through such a scenario many times in a row. In the long-run, adopting a strategy of agreeing with your future self should increase your number of wins (and definitely will in the limit). Nothing similar is the case in the standard cable guy paradox.

ACKNOWLEDGEMENTS

We are grateful to audience members at the Joint Session of the Aristotelian Society and the Mind Association 2007 (hosted by the University of Bristol) and especially Alan Hájek for comments on early versions of this paper. We

¹² An alternative way to understand the difference, here, is that one would be estimating a propensity by observing a number of repeat flips. Nothing similar is happening in (A)–(D). This leads in to the second point.

¹³ As in the original scenario, you need not know what specific piece of evidence you will receive. You need only know it will indicate that one of the options has a higher propensity than the other, and which option.

should also like to thank anonymous referees for comments on the submitted version.

Darrell P. Rowbottom
Faculty of Philosophy
University of Oxford
10 Merton Street
Oxford OX1 4JJ, UK

E-mail: Darrell.Rowbottom@philosophy.ox.ac.uk

Peter Baumann
Department of Philosophy
Swarthmore College
500 College Avenue
Swarthmore, PA 19081-1397, USA

E-mail: pbauman1@swarthmore.edu

REFERENCES

- Bradley, D. & Leitgeb, H. 2006: 'When Betting Odds and Credences Come Apart: More Worries for Dutch Book Arguments', *Analysis* 66, 119–127.
- Hájek, A. 2005: 'The Cable Guy Paradox', *Analysis* 65, 112–119.
- Hájek, A. 2005b: 'Scotching Dutch Books?', *Philosophical Perspectives* 19, 139–151.
- Hitchcock, C. 2004: 'Beauty and the Bets', *Synthese* 139, 405–420.
- Jaynes, E.T. 2003: *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press).
- Keynes, J.M. 1921: *A Treatise on Probability*. London: Macmillan.
- Rowbottom, D.P. 2007: 'The Insufficiency of the Dutch Book Argument', *Studia Logica* 87, 65–71.
- Rowbottom, D.P. 2008: 'Intersubjective Corroboration', *Studies in History and Philosophy of Science* 39, 124–132.
- Williamson, J. 2005: *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford: Oxford University Press.
- Williamson, J. Forthcoming: 'Objective Bayesianism, Bayesian Conditionalisation, and Voluntarism', *Synthese*.