Tough Decisions, Pragmatic Encroachment, and Closure

Anon

2022-08-23

One of the original motivations for an interest-relative theory of knowledge (IRT) was that it avoided the closure failures that dogged rivals. But it turns out that there is a closure problem for natural versions of IRT. Alex Zweber (2016) and, separately, Charity Anderson and John Hawthorne (2019), showed that, on pain of generating closure violations, the following principle cannot be the only way interests enter into our theory of knowledge.

Conditional Preferences If S knows that p, and is trying to decide between X and Y, then her preferences over X and Y are the same unconditionally as they are conditional on p.

They show that this adding this principle, and no other principle, to a natural theory of knowledge leads to an absurd situation where S knows $p \land q$ but does not know p. They go on to argue that the natural ways for versions of IRT to avoid this trap will lead to implausible amounts of scepticism. The point of this paper is to respond to this challenge, and develop a somewhat novel version of IRT that avoids the problem.

The short version of what I'm going to say is that if the defender of IRT wants to maintain **Conditional Preferences**, as they should, they have to give up on the idea that rational choice involves strictly maximising expected utility. Expected utility maximisation only makes sense as a theory of rational choice if the chooser faces zero deliberation costs. Real choosers are not in that situation, and must be sensitive to the costs of choosing one or other means of deliberation. I'm going to defend a particular account of how to make choices when deliberation costs matter, but my defence of IRT won't entirely turn on the details. What I will argue is that for non-ideal choosers, i.e., real humans, adding **Conditional Preferences** to the theory of knowledge does not lead to implausible results.

1 An Example

Let's start with an example from a great thinker. It will require a little exegesis, but that's not unusual when using classic texts.

> Well Frankie Lee and Judas Priest
> They were the best of friends
> So when Frankie Lee needed money one day
> Judas quickly pulled out a roll of tens
> And placed them on the footstool
> Just above the potted plain
> Saying "Take your pick, Frankie boy,
> My loss will be your gain."
> > "The Ballad of Frankie Lee and Judas Priest", 1968.
> > Lyrics from Bob Dylan (2016, 225)

On a common reading of this, Judas Priest isn't just asking Frankie Lee how much money he wants to take, but which invidual notes. Let's simplify, and say that it is common ground that Frankie Lee should only take $10, so the choice he has is which of the individual notes he will take. This will be enough to set up the puzzle.

Assume something else that isn't in the text, but which isn't an implausible addition to the story. The world Frankie Lee and Judas Priest live in is not completely free of counterfeit notes. And it would be bad for Frankie Lee to take a counterfeit note. It won't matter just how common these notes are, or how bad it would be. But our puzzle will be most vivid if each of these are relatively small quantities. So there aren't that many counterfeit notes in circulation, and the (expected) disutility to Frankie Lee of having one of them is not great. There is some chance that he will get in trouble, but the chance isn't high, and the trouble isn't any worse than he's suffered before. Still, other things exactly equal, Frankie Lee would prefer a genuine note to a counterfeit one.

Now for some terminology to help us state the problem Frankie Lee is in. Assume there are k notes on the footstool. Call them $n_1, \ldots, n_k$. Let c_i be the proposition that note n_i is counterfeit, and its negation g_i be that it is genuine. Let g, without a subscript, be the conjunction $g_1 \wedge \cdots \wedge g_n$; i.e., the proposition that all the notes are genuine. Let t_i be the act of taking note n_i. Let U be Frankie Lee's utility function, and Cr his credence function.

In our first version of the example, we'll make two more assumptions. Apart from the issue of whether the note is real or counterfeit, Frankie Lee is indifferent between the notes, so for some $h, l, U(t_i | g_i) = h$ and $U(t_i | c_i) = l$ for all i, with of course $h > l$. And Frankie Lee

thinks each of the banknotes is equally likely to be genuine, so for some p, $Cr(g_i) = p$ for all i. (And the probability of any of them being a counterfeit is independent of the probability of any of the others being counterfeit.)

That's enough to get us three puzzles for the form of IRT that just uses **Conditional Preferences**. I'm going to refer to this form of IRT a lot, so let's give it the memorable moniker IRT-CP. So IRT-CP is what you get by taking a standard theory of knowledge, adding **Conditional Preferences** as a further constraint on knowledge, assuming that rational choosers are expcted utility maximisers, and stopping there. I don't know that anyone endorses IRT-CP, but it's a good theory to have on the table. It says a number of implausible things about Frankie Lee, and the big challenge, as I see it, is to craft a version of IRT that doesn't fall into the same traps.

First, Frankie Lee doesn't know of any note that it is genuine. As things stand, Frankie is indifferent between t_i and t_j for any i, j. But conditional on g_i, Frankie prefers t_i to t_j. Right now, the expected utility of taking either i or j is $ph + (1 - p)l$. If Frankie Lee conditionalises on g_i, then the utility of t_j doesn't change, but the utility of t_i now becomes h, and that's higher than $ph + (1 - p)l$. Since IRT-CP says that one doesn't know p if conditionalising on p changes one's preferences over pragmatically salient options, and t_i and t_j are really salient to Frankie Lee, it follows that he doesn't know g_i. And i was arbitrary in this proof, so he doesn't know of any of the notes that they are genuine. That's not very intuitive, but worse is to follow.

Second, Frankie Lee does know that all the notes are genuine, although he doesn't know of any note that it is genuine. Conditional on g, Frankie Lee's preferences are the same as they are unconditionally. He used to be indifferent between the notes; after conditionalising he is still indifferent. So the one principle that IRT-CP adds to a standard theory of knowledge does not rule out that Frankie Lee knows g. So he knows g; but doesn't know any of its constituent conjuncts. This is a very unappealing option.

Change the example a bit, so we can generate the third problem. Keep that the probabilities of each note being genuine are equal and independent. But assume that the notes are laid out in a line, and Frankie Lee is at one end of that line. So to get a note that is further away from him, he has to reach further. And this has an ever so small disutility. Let r_i be the disutility of reaching for note i. And assume this value increases as i increases, but is always smaller than $(1 - p)(h - l)$. That last quantity is important, because it is the difference between the utility of taking an arbitrary note (with no penalty for the cost of reaching for it), and the utility of taking a genuine banknote.

If all these assumptions are added, Frankie Lee knows one more thing. He knows g_1. That's because as things stand, he prefers t_1 to the other options. Conditional on g_i for any

$i \geq 2$, he prefers t_i to t_1. But if $i > 2$, conditionalising on g_i changes Frankie's preferences, so he doesn't know g_i.

This third puzzle is striking for two reasons. One is that it involves a change of strict preferences. Unconditionally, Frankie strictly prefers t_1 to t_i; conditional on g_i he strictly prefers t_i to t_1. When I first saw these puzzles, I thought we could possibly get around them by restricting attention to cases where conditionalisation changes a strict preference. This example shows that way of rescuing IRT-CP won't work. And the other reason is that it heightens the implausibility of the sceptical result that Frankie doesn't know g_i. It's one thing to say that the weird situation that Judas Priest puts Frankie Lee makes Frankie Lee lose a lot of knowledge he ordinarily has. That's just IRT in action. It's another to say that within this very situation, Frankie Lee knows of some notes that they are genuine but does not know that others are genuine, even though his evidence for the genuineness of each note is the same.

So we have three puzzles to try to solve, if we want to defend anything like IRT-CP.

1. In the case where Frankie Lee has no reason to choose one note rather than another, he doesn't know of any note that it is genuine. And this is surprisingly sceptical.
2. In the case where he has a weak reason to choose one note, he knows that note is genuine, but not the others. This retains the surprisingly sceptical consequence of the first puzzle, and adds a surprising asymmetry.
3. In both cases, there seems to be a really bad closure failure, with Frankie Lee knowing that all the notes are genuine, but not knowing of all or most individual notes that they are genuine.

Before we leave Frankie Lee for a while, let's note one variation on the case that somewhat helps IRT. Imagine that the country they are in has just reached the level of technological sophistication where it can mass produce plastic banknotes. And, as Frankie Lee knows, no one in the country has yet figured out how to produce forgeries of plastic banknotes that are even remotely plausible. Finally, assume that one of the notes, lucky n_8, is one of the new plastic notes, while the others are the old paper notes. If Frankie Lee cares about counterfeit avoidance at all, he should take n_8. And he should do so because it definitely isn't a counterfeit, while each of the others might be. So in that case, Frankie Lee doesn't know that the notes other than n_8 are genuine, at least if whatever might be false isn't known.

Now we have a case where IRT-CP gives the right answers for the right reasons. A theory that disagrees with IRT-CP about this case has to either (a) deny this intuition that the uniquely rational choice for Frankie Lee is n_8, or (b) say that Frankie Lee should choose n_8 because the other choices are too risky, even though he knows the risk in question will not

4

eventuate. Neither option is particularly appealing, at least if one is unhappy with making Moore-paradoxical assertions, so this is a good case for IRT-CP. Or, at least, for something like it.

2 Responding to the Challenge, Quickly

The second half of this paper is going to get into the weeds a bit about how choices do and should get made in cases like Frankie Lee's. So before we do that, I am going to outline how my version of IRT, which of course differs from IRT-CP, handles these cases.

Let's start with closure, and assume that Frankie Lee doesn't know of any note that it is genuine. And assume that's because the conditional utility of a salient act is importantly different, conditional on that note being genuine, to what its unconditional utility. Now we can avoid the closure problem by stressing that what matters is not that the conditional and unconditional questions end up with the same verdict, but that the process of getting to that verdict is the same. This is why, if Frankie Lee doesn't know of any note that it's genuine, he also doesn't know g. Right now, when choosing a note (and trying to maximise expected utility), he should be indifferent because the risk that any note is counterfeit, given his evidence, is more or less the same as the risk that any other note is counterfeit. When he is choosing conditional on g, he doesn't have to attend to risks, or his evidence, or anything that might be more or less equal to anything else. He just takes it as fixed, for purposes of answering the question of what to choose conditional on g, that the notes are genuine. He ends up in the same place both times, indifference between the notes, but he gets there via different pathways. That's enough to defeat knowledge that g.

In English, saying that two questions are answered the same way is ambiguous. It might mean that we end up in the same place when answering the two questions. Or it might mean that we get to that place the same way. There are any number of examples of this. The questions *What is three plus two*, and *How many Platonic solids are there*, get answered the same way in the first sense, but not the second sense. **Conditional Preference** stresses that certain conditional and unconditional questions get answered the same way in this first sense. My version of IRT says that what matters is that these conditional and unconditional questions get answered the same way in the second sense.

That deals with the closure problem satisfactorily, but it makes the sceptical problems worse. To solve those problems we need to rethink our theory of decision. I added, almost as an aside, an assumption in the earlier discussion that Frankie Lee was trying to maximise expected utility. That's a mistake I think; he shouldn't do that. In a lot of cases like Frankie Lee's, the rational thing to do is to simply ignore the possibility that the notes are counterfeit.

This will sometimes lead to taking a sub-optimal choice. But given that choice has computational (and sometimes investigative) costs, it is worth risking a sub-optimal choice if the reward is a cheaper decision procedure. And in some cases like this, the best decision procedure will ignore certain risks.

If Frankie Lee ignores the risk that the notes are counterfeit, then the argument that he doesn't know g_1, g_2, etc., doesn't get off the ground. Given that he's ignoring the risk that the notes are counterfeit, conditionalising on them not being counterfeit changes precisely nothing. So there is no pragmatic argument that he does not know they are genuine.

This approach will avoid the sceptical problems if, but only if, this kind of 'ignoring' is rational and widespread. I aim to make a case that it is. But first I want to make things if anything worse for IRT, by stressing how quotidian examples with the structure of Frankie Lee's are. This will prevent me from being able to dismiss the example as a theorist's fantasy, but will ultimately help see why ignoring in this sense is so prevalent, so important, and even so rational.

3 Back to Earth

The Frankie Lee and Judas Priest case is weird. Who offers someone money, then asks them to pick which note to take? And intuitions about such weird cases cases are sometimes deprecated. Perhaps the contrivance doesn't reveal deep problems with a philosophical theory, but merely a quirk of our intuitions. I am not going to take a stand on any big questions about the epistemic significance of intuitions about weird cases here. Rather, I'm going to note that cases with the same structure as the story of Frankie Lee and Judas Priest are incredibly common in the real world. Thinking about the real world examples can show us how pressing the problems are, and as I just hinted, eventually show us a way out of those problems.

So let's leave Frankie Lee for now, just above the potted plain, and think about a new character. We will call this one David, and he is buying a few groceries on the way home from work. In particular, he has to buy a can of chickpeas, a bottle of milk, and a carton of eggs. To make life easy, we'll assume each of these cost the same amount: $5.[1] None of these purchases is entirely risk free. Canned goods are pretty safe, but sometimes they go bad. Milk is normally removed from sale when it goes sour, but not always. And eggs can crack, either in transit or just on the shelf. In David's world, just like ours, each of these risks

---

[1]If that sounds implausible to you, make the can/bottle/carton a different size, or change the currency to some other dollars than the one you're instinctively using. But I think this example works tolerably well when understand as involving, for example, East Carribean dollars.

is greater than the one that came before.

David has a favorite brand of chickpeas, of milk, and of eggs. And he knows where in the store they are located. So his shopping is pretty easy. But it isn't completely straightforward. First he gets the chickpeas. And that's simple; he grabs the nearest can, and unless it is badly dented, or leaking, he puts in in his basket. Next he goes onto the milk. The milk bottles have sell-by dates printed in big letters on the front. [2] And David checks that he isn't picking up one that is about to expire. His store has been known to have adjacent bottles of milk with sell-by dates 10 days apart, so it's worth checking. But as long as the date is far enough in the future, he takes it and moves on. Finally, he comes to the eggs. (Nothing so alike as eggs, he always thinks to himself.) Here he has to do a little more work. He takes the first carton, opens it to see there are no cracks on the top of the eggs, and, finding none, puts that in his basket too. He knows some of his friends do more than this; flipping the carton over to check for cracks underneath. But the one time he tried that, the eggs ended up on the floor. And he knows some of his friends do less; just picking up the carton by the underside, and only checking for cracks if the underside is sticky where the eggs have leaked. He thinks that makes sense too, but he is a little paranoid, and likes visual confirmation of what he's getting. All done, he heads to the checkout, pays his $15, and goes home.

The choice David faces when getting the chickpeas is like the choice Frankie Lee faces. In a normal store, it will be more like the version where Frankie Lee has to reach further for some notes than others, but sometimes there will be multiple cans equidistant from David. More normally though, some of the cans will be towards the front, and others towards the back, and it will be easier to grab one of the ones from the front. That's why it is weird to get one from the back; reaching incurs costs without any particular payoff.

But in a deeper sense, in all three cases, the choice David faces is something like the choice Frankie Lee faced. He has to choose from among a bunch of very similar seeming options. In at least the chickpeas example, there is something you'd want to say that he knows: canned goods sold at reputable stores are safe. But the arguments above seem to show that David does not know this, at least if IRT-CP is true. Indeed, the argument so far relies just on **Conditional Preferences** being part of the correct theory of IRT, and that's a much weaker claim than that IRT-CP is true. Assuming there is some probability of the chickpeas not being safe, and the costs of reaching for some other can are low enough, David is in exactly the same situation as Frankie Lee. Right now, he maximises utility by taking the front-most can. But conditional on one of the other cans being safe, he maximises utility by taking it. So he does not know of any of the other cans that they are safe.

Frankie Lee's situation is weird. Who lays out some ten dollar bills and asks you to pick

---

[2]This is true of milk in Australian supermarkets, but not, typically, in American supermarkets.

one? (Judas Priest, I guess.) But David's situation is not weird. Looking at a fully stocked shelf of industrially produced food, and needing to pick one can out of an array of similar items, is a very common experience. If a theory of knowledge yields bizarre verdicts about a case like this, it is no defence at all to say the situation is too obscure. In this modern world, it's an everyday occurrence.

4   I Have Questions

So far in this paper I've mostly assumed that these two questions are equivalent:

1. Which option has highest expected utility?
2. What to do?

In doing this, I've faithfully reproduced the arguments of some critics of IRT. But those critics were hardly being unfair to proponents of IRT in treating these questions alike. They are explicitly treated as being interchangable in some defences of IRT, such as Weatherson (2005). But defenders of IRT need not, and should not, have equated these questions. The beginning of a solution to the problems raised by Frankie Lee is to separate the questions out. I already mentioned one respect in which these questions differ back in section 2. I'll rehearse that difference, briefly mention a second difference, then spend some time on a third difference.

The point I made much of back in section 2 was that someone might know the utility facts, but not know what to do. When Frankie sits down, with his fingers to his chin, and tries to decide which of the tens to take, it's possible he knows that they each have the same utility. But he still has to pick one, and with his head spinning he can't decide which one to take. In cases like these answering questions about utility comparisons won't settle questions about what to do.[3]

A second reason for not treating the questions alike is that to treat them alike assumes away something that should not be assumed away. It simply assumes that risk-sensitive theories of choice, as defended by Quiggin (1982) and Buchak (2013), are mistaken. We probably shouldn't simply assume that. It turns out the difference between expected utility theory and these heterodox alternatives isn't particularly relevant to Frankie's or David's choices, so I'll leave this aside for the rest of the paper.

---

[3] James Joyce (2018) suggests the following terminology. If Frankie is rational, then utility considerations settle questions about what to *choose*, but not questions about what to *pick* in the case of a tie. I haven't quite followed that terminology; I've let Frankie pick and choose more freely than that. But I'm following Joyce in stressing this conceptual distinction.

The third way in which the equivalence is wrong takes a little longer to set up. The short version is that rational people are satisficers, and for a satisficer you can answer the question *What to do* without taking a stand on questions about relative utility. The longer version is set out in the next section.

5   You'll Never Be Satisfied (If You Try to Maximise)

The standard model of practical rationality that we use in philosophy is that of expected utility maximisation. But there are both theoretical and experimental reasons to think that this is not the right model for choices such as that faced by Frankie or David. Maximising expected utility is resource intensive, especially in contexts like a modern supermarket, and the returns on this resource expenditure are unimpressive. What people mostly do, and what they should do, is choose in a way that is sensitive to the costs of adopting one or other choice procedure.

There are two annoying terminological issues around here that I mostly want to set aside, but need to briefly address in order to forestall confusion.

I'm going to assume maximising expected utility means taking the option with the highest expected utility given facts that are readily available. So if one simply doesn't process a relevant but observationally obvious fact, that can lead to an irrational choice. I might alternatively have said that the choice was rational (given the facts the chooser was aware of), but the observational process was irrational. But I suspect that terminology would just add needless complication. I'm going to come back to another point that is partially terminological, and partially substantive. That's whether we should identify the choice consequentialists recommend in virtue of the fact that it maximises expected utility with one of the options (in the ordinary sense of option), or something antecedent.

And I'm going to call any search procedure that is sensitive to resource considerations a satisficing procedure. This isn't an uncommon usage. Charles Manski (2017) uses the term this way, and notes that it has rarely been defined more precisely than that. But it isn't the only way that it is used. Mauro Papi (2013) uses the term to exclusively mean that the chooser has a 'reservation level', and they choose the first option that crosses it. This kind of meaning will be something that becomes important again in a bit. And Chris Tucker (2016), following a long tradition in philosophy of religion, uses it to mean any choice procedure that does not optimize. Elena Reutskaja et al (2011) contrast a 'hybrid' model that is sensitive to resource constraints with a 'satisficing' model that has a fixed reservation level. They end up offering reasons to think ordinary people do (and perhaps should) adopt this hybrid model. So though they don't call this a satisficing approach, it just is a version of what

9

Manski calls satisficing. Andrew Caplin et al (2011), on the other hand, describe a very similar model to Reutskaja et al's hybrid model - one where agents try to find something above a reservation level but the reservation level is sensitive to search costs - as a form of satisficing. So the terminology around here is a mess. I propose to use Manski's terminology: agents satisfice if they choose in a way that is sensitive to resource constraints. Ideally they would maximise, subject to constraints, but saying just what this comes to runs into obvious regress problems (Savage 1967). Let's set aside this theoretical point for a little, and go back to David and the chickpeas.

When David is facing the shelf of chickpeas, he can rationally take any one of them - apart perhaps from ones that are seriously damaged. How can expected utility theory capture that fact? It says that more than one choice is permissible only if the choices are equal in expected utility. So the different cans are equal in expected utility. But on reflection, this is an implausible claim. Some of the cans are ever so slightly easier to reach. Some of the cans will have ever so slight damage - a tiny dint here, a small tear in the label there - that just might indicate a more serious flaw. Of course, these small damages are almost always irrelevant, but as long as the probability that they indicate damage is positive, it breaks the equality of the expected utility of the cans. Even if there is no visible damage, some of the labels will be ever so slightly more faded, which indicates that the cans are older, which ever so slightly increases the probability that the goods will go bad before David gets to use them. Of course in reality this won't matter more than one time in a million, but one in a million chances matter if you are asking whether two expected utilities are strictly equal.

The common thread to the last paragraph is that these objects on the shelves are almost duplicates, but the most careful quality control doesn't produce consumer goods that are actual duplicates. This is particularly true in Frankie Lee's choice situation. If all the notes he looks at are really duplicates, down to the serial numbers, he should run away. There are always some differences. It is unlikely that these differences make precisely zero difference to the expected utility of each choice. And even if they do, discovering that is hard work.

So it seems likely that, according to the expected utility model, it isn't true that David could permissibly take any can of chickpeas that is easily reachable and not obviously flawed. Even if that is true, it is extremely unlikely that David could know it to be true. But one thing we know about situations like David's is that any one of the (easily reached, not clearly flawed) cans can be permissibly chosen, and David can easily know that. So the expected utility model, as I've so far described it, is false.

I'll return in the next section to the question of whether this is a problem for theories of decision based around expected utility maximisation broadly, or whether it is just a problem for the particular way I've spelled out the expected utility theory. But for now I want to run

through two more arguments against the idea that supermarket shoppers like David should be maximising expected utility (so understood).

In all but a vanishingly small class of cases, the different cans will not have the same expected utility. Indeed, that they have the same expected utility is a measure zero event. One way to note that expected utility maximisation can't be the right theory of choice-worthiness is that cases where multiple cans are equally choice-worthy is not a measure zero event; it's the standard case. And figuring out which can has the highest expected utility is a going to be work. It's possible in principle, I suppose, that someone could be skilled at it, in the sense that they could instinctively pick out the can whose shape, label fading, etc., reveal it to have the highest expected utility. Such a skill seems likely to be rare - though I'll come back to this point below when considering some other skills that are probably less rare. For most people, maximising expected utility will not be something that can be done through skill alone; it will take effort. And this effort will be costly, and almost certainly not worth it. Although one of the cans will be ever so fractionally higher in expected utility than the others, the cost of finding out which can this is will be greater than the difference in expected utility of the cans. So aiming to maximise expected utility will have the perverse effect of reducing one's overall utility, in a predictable way.

The costs of trying to maximise expected utility go beyond the costs of engaging in search and computation. There is evidence that people who employ maximising strategies in consumer search end up worse off than those who don't. Schwartz et al (2002) reported that consumers could be divided in 'satisficers' and 'maximizers'. And once this division is made, it turns out that the maximizers are less happy with individual choices, and with their life in general. This finding has been extended to work on career choice (Iyengar, Wells, and Schwartz 2006) where the maximisers end up with higher salaries but less job satisfaction, and to friend choice (Newman et al. 2018), where again the maximizers seem to end up less satisfied.

There is evidence here that maximizing is bad at what it sets out to achieve. But there are both empirical and theoretical reasons to be cautious about accepting these results at face value. Whether maximizers are worse off seems to be tied up to the 'paradox of choice' (Schwartz 2004), the idea that sometimes giving people even more choices makes them less happy with their outcome, because they are more prone to regret. But it is unclear whether such a paradox exists. One meta-analysis (Scheibehenne, Greifeneder, and Todd 2010) did not show the effect existing at all, though a later meta-analysis finds a significant mediated effect (Chernev, Böckenholt, and Goodman 2015). But it could also be that the result is a feature of an idiosyncratic way of carving up the maximizers from the satisficers. Another way of dividing them up produces no effect at all (Diab, Gillespie, and Highhouse 2008).

The theoretical reasons relate to Newcomb's problem. Even if we knew that maximizers were less satisfied with how things are going than satisficers, it isn't obvious that any one person would be better off switching. They might be like a two-boxer who would get nothing if they took one-box. There is a little evidence in Iyengar, Wells, and Schwartz (2006) that this isn't quite what is happening, but the overall situation is unclear.

The upshot of all this, I think, is that there are potentially two kinds of cost of engaging in certain kind of search and choice procedures. Some procedures are more costly to implement than others: they take more time, or more energy, or even more money. But further, some procedures might have a hedonic cost that extends beyond the time that the procedure is implemented. There is no theoretical or empirical guarantee that choosing widget W by procedure P1 will produce the same amount of happiness as choosing widget W by procedure P2. And especially for choices that are intended to produce happiness, this kind of factor should matter to us. In short, there are many more ways to assess a consumer choice procedure than the quality of the products in ends up choosing. And this will be the key to our resolution of the puzzles about closure.

6   Deliberation Costs and Infinite Regresses

The picture I've been building towards in the last section, on which rational people don't maximise but choose arbitrarily between choices that are close enough, is not a new one. As I mentioned, the experimental work in Reutskaja (2011) suggests this is how people do reason. But the idea that people should reason this way goes back much further. It is often traced back to a footnote of Knight's. Here is the text that provides the context for the note.

> Let us take Marshall's example of a boy gathering and eating berries … We can hardly suppose that the boy goes through such mental operations as drawing curves or making estimates of utility and disutility scales. What he does, in so far as he deliberates between the alternatives at all*, is to consider together with reference to successive amounts of his "commodity," the utility of each increment against its "cost in effort," and evaluate the net result as either positive or negative (Knight 1921, 66–67)

And the footnote attached to 'at all' says this

> Which, to be sure, is not very far. Nor is this any criticism of the boy. Quite the contrary! It is evident that the rational thing to do is to be irrational, where deliberation and estimation cost more than they are worth. That this is very

often true, and that men still oftener (perhaps) behave as if it were, does not vitiate economic reasoning to the extent that might be supposed. For these irrationalities (whether rational or irrational!) tend to offset each other. (Knight 1921, 67fn1)

Knight doesn't really give an argument for the claim that these effects will offset. And as John Conlisk (1996) shows in his fantastic survey of the late 20th century literature on bounded rationality, it very often isn't true. Especially in game theoretic contexts, the thought that other players might think that "deliberation and estimation cost more than they are worth" can have striking consequences. But our aim here is not to think about economic theorising, but about the nature of rationality.

There is something paradoxical, almost incoherent, about Knight's formulation. If it is "rational to be irrational", then being "irrational" can't really be irrational. There are two natural ways to get out of this paradox. One, loosely following David Christensen (2007), would be to say that "Murphy's Law" applies here. Whatever one does will be irrational in some sense. But still some actions are less irrational than others, and the least irrational will be to decline to engage in deliberation that costs more than it is worth. I suspect what Knight had in mind though was something different (if not obviously better). He is using 'rational' as more or less a rigid designator of the the property of choosing as a Marshallian maximiser does. And what he means here is that the disposition to not choose in that way will be, in the long run, the disposition with maximal returns.

This latter idea is what motivates the thought that rational agents will take what Conlisk calls "deliberation costs" into account. And Conlisk thinks that this is what rational agents will do. But he also raises a problem for this view, and indeed offers one of the clearest (and most widely cited) statements of this problem.

However, we quickly collide with a perplexing obstacle. Suppose that we first formulate a decision problem as a conventional optimization based on the assumption of unbounded rationality and thus on the assumption of zero deliberation cost. Suppose we then recognize that deliberation cost is positive; so we fold this further cost into the original problem. The difficulty is that the augmented optimization problem will itself be costly to analyze; and this new deliberation cost will be neglected. We can then formulate a third problem which includes the cost of solving the second, and then a fourth problem, and so on. We quickly find ourselves in an infinite and seemingly intractable regress. In rough notation, let P denote the initial problem, and let $F(.)$ denote

the operation of folding deliberation cost into a problem. Then the regress of problems is $P, F(P), F^2(P), ...$ (Conlisk 1996, 687)

Conlisk's own solution to this problem is not particularly satisfying. He notes that once we get to F^3 and F^4, the problems are 'overly convoluted' and seem to be safely ignored. This isn't enough for two reasons. First, even a problem that is convoluted to state can have serious consequences when we think about solving it. (What would *Econometrica* publish if this weren't true?) Second, as is often noted, $F^2(P)$ might be a harder problem to solve than P, so simply stopping the regress there and treating the rational agent as solving this problem seems to be an unmotivated choice.

As Conlisk notes, this problem has a long history, and is often used to dismiss the idea that folding deliberation costs into our model of the optimising agent is a good idea. I use 'dismiss' advisedly here. As he also notes, there is very little *discussion* of this infinite regress problem in the literature before 1996. The same remains true after 1996. What is done is that instead people appeal to the regress in a sentence or two to set aside approaches that incorporate deliberation cost in the way that Conlisk suggests.

Up to around the time of Conlisk's article, the infinite regress problem was often appealed to by people arguing that we should, in effect, ignore deliberation costs. After his article, the appeals to the regress comes from a different direction. It is usually from theorists arguing that deliberation costs are real, but the regress means it will be impossible to consistently incorporate them into a model of an optimizing agent. So we should instead rely on experimental techniques to see how people actually handle deliberation costs; the theory of optimization has reached its limit. This kind of move is found in writers as diverse as Gigerenzer and Selton (2001), Odell (2002), Pingle (2006), Mangan et al (2010), Ogaki and Tanaka (2017) and Chakravarti (2017). And proponents of taking deliberation costs seriously within broadly optimizing approaches, like Miles Kimball (2015), say that solving the regress problem is the biggest barrier to having such an approach taken seriously by economists.

And it really matters for the story of this book that there is a solution to the infinite regress problem within a broadly optimizing framework. More precisely, we need a solution to the regress problem that does not defeat knowledge. At least some of the time, the fact that a belief was formed by a rationally problematic procedure means that the belief is not a piece of knowledge. As we might say, the irrationality of the procedure is a defeater of the claim to knowledge. But perhaps if the procedure is optimal (even if not rational) that defeats the defeater. 'Optimal' here need not mean rationally optimal; it means optimal given the computational limitations on the agent. But now I've said enough to suggest that the regress problem will arise.

Here's how I plan to solve the regress problem. What matters for optimality is that the thinker is following the procedure that is the optimal solution to $F(P)$. It doesn't matter that they compute that it is the optimal solution, or even that they are following it because it is the optimal solution. Because there is this external, success oriented condition, there isn't any extra problem that arises by 'folding in' computational costs. The thinker doesn't have to compute whether the optimal solution to $F(P)$ is indeed optimal; they just have to follow it.

So this is a solution to the regress problem. Externalism about optimality of processes suggests that all that matters is that agents follow optimal processes, not that they compute that these are optimal. And so there is no computational regress problem, since we don't do any computation at the process choosing stage, so there are no computational costs to minimise. And there is no hedonic regress problem, because what makes a process optimal is that it leads to choices that are satisfying when chosen by that very process. If all this is true, the problem is solved.

This solution to the regress problem is easy to state, but a little harder to motivate. There are two big questions to answer before we can say it is really motivated.

1. Why should we allow this kind of unreflective rule-following in our solution to the regress?
2. Why should we think that $F(P)$ is the point where this consideration kicks in, as opposed to P, or anything else?

There are a few ways to answer 1. One motivation traces back to the work by the artificial intelligence researcher Stuart Russell (1997). (Although really it starts with the philosophers Russell cites as inspiration, such as Cherniak (1986) and Harman (1973).) He stresses that we should think about the problem from the outside, as it were, not from inside the agent's perspective. How would we program a machine that we knew would have to face the world with various limitations? We will give it rules to follow, but we won't necessarily give it the desire (or even the capacity) to follow those rules self-consciously. That might be useful some of the time - though really what's more useful is knowing the limitations of the rules. And that can be done without following the rules as such. It just requires good dispositions to complicate the rules one is following in cases where such complication will be justified.

Another motivation is right there in the quote from Knight that set this literature going. Most writers quote the footnote, where Knight suggests it might be rational to be irrational. But look back at what he's saying in the text. The point is that it can be perfectly rational to use considerations other than drawing curves and making utility scales. What one has to

do is follow internal rules that (non-accidentally) track what one would do if one was a self-consciously perfect Marshallian agent. That's what I'm saying too, though I'm saying it one level up.

Finally, there is the simple point that on pain of regress any set of rules whatsoever must say that there are some rules that are simply followed. This is one of the less controversial conclusions of the debates about rule-following that were started by Wittgenstein (1953). That we must at some stage simply follow rules, not follow them in virtue of following another rule, say the rule to compute how to follow the first rule and act accordingly, is an inevitable consequence of thinking that finite creatures can be rule followers.

So question 1 is not really a big problem. But question 2 is more serious. Why $F(P)$, and why not something else? The short answer will be that any reason to think that rational actors maximize *expected* utility, as opposed to actual utility, will also be a reason to think that they solve $F(P)$ and not P. The longer answer is a bit more roundabout, but it helps us to see what a solution to $F(P)$ will look like.

Start by stepping back and thinking about why we cared about *expected* utility in the first place. Why not just say that the best thing to do is to produce the best outcome, and be done with it? Well, we don't say that because we take it as a fixed point of our inquiry that agents are informationally limited, and that the best thing to do is what is best given that limitation. Given some plausible assumptions, the best thing for the informationally limited agent to do would be to maximise expected utility. This is a second-best option, but the best is unavailable given the limitations that we are treating as unavoidable.

But agents are not just informationally limited, they are computationally limited too. And we could treat that as the core limitation to be modelled. As Conlisk says, it is "entertaining to imagine" theorists who worked in just this way (Conlisk 1996, 691). So let's imagine we meet some Martian economists, and they take computational, and not informational, limitations as the core constraint on rational choosers. So in their models, every agent has all the information relevant to their choice, but can't always compute what to do with that information.

Conlisk doesn't spell out the details of this thought experiment, and it's a little tricky to say exactly how it should work. (I'm indebted here to Harvey Lederman.) After all, you might think that 'information' should include things like information about the results of various computations, or about what would be best to do given their information. So how can we make sense of a being that is computationally but not informationally limited?

Here's one way to make sense of what Conlisk's Martians might be like. Assume that the Martians are very strict positivists. (This isn't going to make them optimal social scientists, but presumably we never thought they were.) So the truths can be divided up into observa-

tion sentences, and things derived from observation sentences by definition and deduction. And in their model, every agent knows every true observation sentence - including those about observations that have not yet been made. But they don't know all the results of deriving further truths from the observation sentences by definition and deduction. So such an agent might know precisely all the points she has to drive to today, and know the cost of travelling between any two points, but not know the optimal route to take on her travels. That last claim won't be 'information' in the relevant sense since it is not an observation sentence. Indeed, we've assumed already that she knows a lot beyond mere observations - but this is the kind of thing she might not know.

The point is not that the Martian economists think that every agent knows every observation sentence, any more than human economists think that every agent has a solution to every travelling salesman problem in their back pocket. Rather, it's that they that this is a good modeling assumption. Conlisk has some fun imagining what Martian economists who make this modeling assumption might say in defence of their practice. They might disparage their colleagues who take informational limitations seriously as introducing ad hoc stipulations into theory. They might argue that informational limitations are bound to cancel out, or be eliminated by competition. They might argue that apparent informational limitations are really just computational ones, or at least can be modelled as computational ones. (Here it might be helpful to think of the Martian economists as positivists, and in particular as positivists who think that the notion of observation sentence is flexible enough to behave differently in different theoretical contexts.) And so on, replicating almost every complaint that human economists have ever made about theorists who want to take computational limitations seriously.

What Conlisk doesn't add is that they might suggest that there is a regress worry for any attempt to add informational constraints. Imagine that inside one of these models, an agent is deciding what to have for dinner. Let Q be the initial optimisation problem as the Martians see it. That is, Q is the problem of finding the best outcome, the best dinner, given full knowledge of the situation, but the actual computational limitations of the agent. Then we suggest that we should also account for the informational limitations. Let's see if this will work, they say. Let I be the function that transforms a problem into one that is sensitive to the informational limitations of the agent. But if we're really sensitive to informational limitations, we should note that $I(Q)$ is also a problem the agent has to solve under conditions of less than full information.[4] So the informationally challenged agent will have to solve not

---

[4] At this point the Martians might note that while they are grateful that Williamson (2000) has highlighted problems with the KK principle, and these problems show some of the reasons for wanting to idealise away from informational limitations, they aren't in fact relying on Williamson's work. All they need is that agents

just $I(Q)$, but $I^2(Q)$, and $I^3(Q)$ and so on.[5]

Orthodox defenders of (human versions of) rational choice theory have to think this is a bad argument. And I think most of them will agree with roughly the solution I'm adopting. The right problem to solve is $I(Q)$, on a model where Q is in fact the problem of choosing the objectively best option. Put in philosophers' terms, we should think of Q as rigidly, and transparently, designating the problem the agent is facing. So $I(Q)$ is not the problem of doing what's best given how little one knows about both the world and one's place in it. Rather, it's the problem of how to do the best one can in this very situation, given one's ignorance about the world. Even if one doesn't know precisely the situation one is in, and one doesn't know what utility function one has, or for that matter what knowledge one has, one should maximise expected utility given actual expectations and actual utility. The problem to solve is $I(Q)$, not $I^2(Q)$.

But the bigger thing to say is that neither we nor the Martians really started with the right original problem. The original problem, O, is the problem of choosing the objectively best option; i.e., choosing what to have for dinner. The humans start by considering the problem $I(O)$, i.e., P, and then debate whether we should stick with that problem, or move to $F(I(O))$. The Martians start by considering the problem $F(O)$, i.e., Q, then debate whether we should stick with that or move to $I(F(O))$. And the answer in both cases is that we should move.

Given the plausible commutativity principle that introducing two limitations to theorising has the same effect whichever order we introduce them, $I(F(O)) = F(I(O))$. That is, $F(P) = I(Q)$. And that's the problem that we should think the rational agent is solving.

But why solve that, rather than something more or less close to O? Well, think about what we say about an agent in a Jackson case who tries to solve O not $I(O)$. (A Jackson case, in this sense, is a case where the choice with highest expected value is known to not have the highest objective value. So trying to get the highest objective value will mean definitely not maximising expected value.) We think it will be sheer luck if they succeed. We think in the long run they will almost certainly do worse than if they tried to solve $I(O)$. And in the rare case where they do better, we think it isn't a credit to them, but to their luck. In cases where the well-being of others is involved, we think aiming for the solution to O involves needless, and often immoral, risk-taking.

The Martains can quite rightly say the same things about why $F(O)$ is a more theoreti-

---

do not exactly what they know. And that will be true as long as the correct epistemic logic is weaker than S5. And that will be true as long as someone somewhere has a false belief. And it would just be weird, they think, to care about informational limitations but want to idealise away from the existence of false beliefs.

[5] At this point, some of the Martians note that the existence of Elster (1979) restored their faith in humanity.

cally interesting problem than *O*. Assume we are in a situation where *F*(*O*) is known to differ from *O*. For example, imagine the decision maker will get a reward if they announce the correct answer to whether a particular sentence is a truth-functional tautology, and they are allowed to pay a small fee to use a computer that can decide whether any given sentence is a tautology. The solution to *O* is to announce the correct answer, whatever it is. The solution to *F*(*O*) is to pay to use the computer. And the Martians might point out that in the long run, solving *F*(*O*) will yield better results. That if the agent does solve problems like *O* correctly, even in the long run, this will just mean they were lucky not rational. That if the reward is that a third party does not suffer, then it is immorally reckless to not solve *F*(*O*), i.e., to not consult the computer. And in general, whatever we can say that motivated "Rational Choice Theory", as opposed to "Choose the Best Choice Theory", they can say too.

Both the human and the Martian arguments look good to me. We should add in both computational and informational limitations into our model of the ideal agent. But note something else that comes from thinking about these Jackson cases. In solving a limitation sensitive problem, we aren't trying to approximate a solution to the limination insensitive problem. This is part of why the regress can stop here. To solve *F*(*X*), we don't have to solve *X*, and then see how close the various computationally feasible solutions get to this solution. That's true in general because of Jackson cases, but it's especially true when *X* is itself a complex problem. In trying to solve *F*(*I*(*O*)), i.e., *I*(*F*(*O*)), we aren't trying to maximise expected value, and then approximate that solution given computational limitations. Nor are we trying to be optimal by Martian standards (i.e., solve *F*(*O*)), then approximate that given informational limitations. We're just trying to get as good an outcome as we can, given our limitations. Doing that does not require solving any iterated problem about how well we can solve *F*(*I*(*O*)) given various limitations, any more than rationally picking berries requires drawing Marshallian curves.

So that's the solution to the regress. It is legitimate to think that there is a rule that rational creatures follow immediately, on pain of thinking that all theories of rationality imply regresses. And thinking about the contingency of how Rational Choice Theory got to be the way it is suggests that the solution to what Conlisk calls *F*(*P*), or what I've called *F*(*I*(*O*)), will be that point.

What might that stopping point look like in practice? In his discussion of the regress, Miles Kimball (2015) suggests a few options. I want to focus on two of them.

> Least transgressive are models in which an agent sits down once in a long
> while to think very carefully about how carefully to think about decisions
> of a frequently encountered type. For example, it is not impossible that
> someone might spend one afternoon considering how much time to spend

> on each of many grocery-shopping trips in comparison shopping. In this
> type of modelling, the infrequent computations of how carefully to think
> about repeated types of decisions could be approximated as if there were
> no computational cost, even though the context of the problem implies that
> those computational costs are strictly positive. (Kimball 2015, 174)

And that's obviously relevant to David in the supermarket. He could, in principle, spend one Saturday afternoon thinking about how carefully to check each of the items in the supermarket before putting it in his shopping cart. And then in future trips, he could just carry out this plan. This isn't terrible, but I don't think it's optimal. For one thing, there are much better things to do with Saturday afternoons. For another, it suggests we are back in the business of equating solving $F(P)$ with approximately solving P. And that's a mistake. Better to just say that David is rational if he just does the things that he would do were he to waste a Saturday afternoon this way, and then plan it out. And that thought leads to Kimball's more radical suggestion for how to avoid the regress,

> [M]odelling economic actors as doing constrained optimization in relation to
> a simpler economic model than the model treated as true in the analysis. This
> simpler economic model treated as true by the agent can be called a "folk the-
> ory". (Kimball 2015, 175)

It's this last idea I plan to explore in more detail. (It has some similarities to the discussion of small worlds in Joyce (1999, 70–77).) The short version is that David can, and should, have a little toy model of the supermarket in his head, and should optimize relative to that model. The model will be false, and David will know it is false. And that won't matter, as long as David treats the model the right way.

7 Ignorance is Bliss

There are a lot of things that could have gone wrong with a can of chickpeas. They could have gone bad inside the can. They could have been contaminated, either deliberately or through carelessness. They could have been sitting around so long they have expired. All these things are, at least logically, possible.

But these possibilities, while serious, have two quite distinctive features. One is that they are very rare. In some cases they may have never happened. (I've never heard of someone deliberately contaminating canned chickpeas, though other grocery products like strawberries have been contamination targets.) The other is that there are few easy

ways to tell whether they are actualised. You can scan each of the cans for an expiry date, but it is really uncommon that this is relevant, and it takes work since the expiry dates are normally written in such small type. If a can is really badly dented, I guess that weakens the metal and raises ever so slightly the prospect of unintentional contamination. But it's common to have shelves full of cans that have no dents, or at most very minor ones.

Given these two facts - the rarity of the problems and the difficulty in getting evidence that significantly shifts the probability that this is one of the (rare) problems - the rational thing to do is choose in a way that is insensitive to whether those problems are actualised. Or, perhaps more cautiously, one should be vigilant, in the sense of Sperber et al (2010), to some of these problems, and ignore the rest. But being vigilant about a problem means, I take it, being willing to consider it if and only if you get evidence that it is worth considering. In the short run, you still ignore the potential problem.

And to ignore a potential problem is to choose in a way that is insensitive to evidence for the problem. That makes sense for both the banknotes and the chickpeas, because engaging in a choice procedure that is sensitive to the probability of the problem will, in the long run, make you worse off.

In Kimball's terms, the rational shopper will have a toy model of the supermarket in which all cans of chickpeas that aren't obviously damaged are safe to eat. This will be a defeasible model, but on a typical grocery trip, it won't be defeated. (In Joyce's terms, the small worlds the shopper uses in setting up the decision problem they face will all be ones in which the chickpeas are safe.)

But if that's the toy model the shopper has, then conditionalising on the fact that a particular can is safe doesn't change anything. Assuming that the can was not obviously damaged, the toy model says that it is fine to eat. So there is no argument from interest-relativity that the shopper does not know this can is fine to eat. So we've avoided the worst of the sceptical challenges. By similar reasoning, Frankie Lee knows all of the banknotes are genuine.

This paper started with a problem for pragmatic theories of knowledge like IRT. They seemed to lead to rampant scepticism in cases like Frankie Lee's. But the solution to this problem was, ultimately, more pragmatism. The rational chooser won't typically use a model where the probability that a particular note is a forgery, or a particular can is contaminated, is 0.99999. Using a model like that, rather than a model where the probability is 1, is more trouble than it's worth, especially since there is no actionable difference between the models. At least, it's usually more trouble than it's worth. If passing a forged banknote will lead to torture, then maybe you use the more complex model. If there is a something to do about differences between nearly 1 and 1, as in the case when some banknotes are

plastic, or in the case where you can look inside the egg carton, then you use the more complex model. But usually one should, and does, use the simpler model.

So David does know that the chickpeas are safe. He believes this on the basis of evidence that is connected in the right way to the truth of the proposition that the chickpeas are safe. There is a potential pragmatic defeater from the fact that **Conditional Preference** seems to rule out this knowledge. But there is a pragmatic defeater of that pragmatic defeater. The thought behind **Conditional Preference** assumes that David will be insensitive to deliberation costs when choosing. He shouldn't be, on practical grounds. He should use a toy model that says all safe looking cans are safe. And once he uses that toy model, there is no pragmatic defeat of his well-supported, well-grounded true belief. He knows the chickpeas are safe.

But he doesn't know the eggs aren't cracked. The toy model that says all eggs on the shelf are uncracked is a bad toy model to use. It isn't bad because it's wrong. It's bad because there is a model that will yield better long run results even once we account for its complexity. That's the model that says that only eggs that have been visually inspected are certain to be uncracked; all other eggs are at best probably uncracked. So David doesn't know the eggs aren't cracked.

And note this would be true even if improvements in the supply chain made the probability of cracked eggs much lower than it is today. What matters in the canned goods case is not just that the risk of contamination is low, it's also that there isn't anything to do about it. As long as it remains easy to flip the lid of egg cartons to check whether they are cracked, it will be hard to know without flipping they aren't cracked.

Note here that the version of IRT I'm defending is like IRT-CP, and unlike most forms of IRT in the literature, i thinking that there can be pragmatic encroachment in low stakes context. When buying eggs, the stakes are not zero - buying cracked eggs wastes money and that's why David should check. But it isn't a 'high stakes' case in anything like the sense that phrase is used. The stakes are exactly the same as in the chickpeas case. What matters is not the cost of being wrong about an assumption, but rather the relative cost of being wrong compared to the probability that one is wrong and the cost of checking.

The milk case is only slightly more complicated. At least in some places, the expiry date for milk is written in very large print on the front of the bottle. In those cases, it is worth checking that you aren't buying milk that expires tomorrow. So before you check, you don't know that the milk you pick up doesn't expire tomorrow. (And, like in the eggs case, that's true even if the shop very very rarely sells milk that close to the expiry date.) But there is no way to check whether a particular unexpired milk has gone bad. You can't easily open a milk bottle in the supermarket and smell it, for example. So that's the kind of rare and

uncheckable problem that the sensible chooser will ignore. Their toy model will include that in a well functioning store, all milk that is well away from the expiry date is safe. So once they've checked the expiry date, they know it is safe (assuming it is safe).

And in the normal case, Frankie Lee knows that the notes aren't forgeries. His toy model of the currency, like ours, should be that all bank notes are genuine unless there is a clear sign that they are not.[6] So we have a solution from within IRT to both the closure problems and the sceptical problems.

References

Anderson, Charity, and John Hawthorne. 2019. "Knowledge, Practical Adequacy, and Stakes." *Oxford Studies in Epistemology* 6: 234–57.

Buchak, Lara. 2013. *Risk and Rationality*. Oxford: Oxford University Press.

Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and Satisficing." *American Economic Review* 101 (7): 2899–2922. https://doi.org/10.1257/aer.101.7.2899.

Chakravarti, Ashok. 2017. "Imperfect Information and Opportunism." *Journal of Economic Issues* 51 (4): 1114–36. https://doi.org/10.1080/00213624.2017.1391594.

Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman. 2015. "Choice Overload: A Conceptual Review and Meta-Analysis." *Journal of Consumer Psychology* 25 (2): 333–58. https://doi.org/10.1016/j.jcps.2014.08.002.

Cherniak, Christopher. 1986. *Minimal Rationality*. Cambridge, MA: MIT Press.

Christensen, David. 2007. "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." *Oxford Studies in Epistemology* 2: 3–31.

Conlisk, John. 1996. "Why Bounded Rationality?" *Journal of Economic Literature* 34 (2): 669–700.

Diab, Dalia L., Michael A. Gillespie, and Scott Highhouse. 2008. "Are Maximizers Really Unhappy? The Measurement of Maximizing Tendency." *Judgment and Decision Making* 3 (5): 364–70.

Dylan, Bob. 2016. *The Lyrics: 1961-2012*. New York: Simon & Schuster.

Elster, Jon. 1979. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press.

Gigerenzer, Gerd, and Reinhard Selten. 2001. *Bounded Rationality: The Adaptive Toolbox.*

---

[6]Or at least a sign. Arguably, the fact that a note is a high value note that someone is trying to use in the betting ring half an hour before the Melbourne Cup is in itself a sign that it is not genuine. A sceptical theory that says no one there knows whether they are passing on forged bank notes is not a problematic sceptical theory.

Cambridge, MA: MIT Press.

Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.

Iyengar, Sheena S., Rachael E. Wells, and Barry Schwartz. 2006. "Doing Better but Feeling Worse: Looking for the 'Best' Job Undermines Satisfaction." *Psychological Science* 17 (2): 143–50. https://doi.org/10.1111/j.1467-9280.2006.01677.x.

Joyce, James. 2018. "Deliberation and Stability in Newcomb Problems and Pseudo-Newcomb Problems." In *Newcomb's Problem*, edited by Arif Ahmed, 138–59. Cambridge: Cambridge University Press.

Joyce, James M. 1999. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Kimball, Miles. 2015. "Cognitive Economics." *The Japanese Economic Review* 66 (2): 167–81. https://doi.org/10.1111/jere.12070.

Knight, Frank. 1921. *Risk, Uncertainty and Profit*. Chicago: University of Chicago Press.

Mangan, Jean, Amanda Hughes, and Kim Slack. 2010. "Student Finance, Information and Decision Making." *Higher Education* 60 (5): 459–72. https://doi.org/10.1007/s10734-010-9309-7.

Manski, Charles F. 2017. "Optimize, Satisfice, or Choose Without Deliberation? A Simple Minimax-Regret Assessment." *Theory and Decision* 83 (2): 155–73. https://doi.org/10.1007/s11238-017-9592-1.

Newman, David B., Joanna Schug, Masaki Yuki, Junko Yamada, and John B. Nezlek. 2018. "The Negative Consequences of Maximizing in Friendship Selection." *Journal of Personality and Social Psychology* 114 (5): 804–24. https://doi.org/10.1037/pspp0000141.

Odell, John S. 2002. "Bounded Rationality and World Political Economy." In *Governing the World's Money*, edited by David M. Andrews, C. Randall Henning, and Louis W. Pauly, 168–93. Ithaca: Cornell University Press.

Ogaki, Masao, and Saori C. Tanaka. 2017. *Behavioral Economics: Toward a New Economics by Integration with Traditional Economics*. Singapore: Springer.

Papi, Mario. 2013. "Satisficing and Maximizing Consumers in a Monopolistic Screening Model." *Mathematical Social Sciences* 66 (3): 385–89. https://doi.org/10.1016/j.mathsocsci.2013.08.005.

Pingle, Mark. 2006. "Deliberation Cost as a Foundation for Behavioral Economics." In *In Handbook of Contemporary Behavioral Economics: Foundations and Developments*, edited by Morris Altman, 340–55. New York: Routledge.

Quiggin, John. 1982. "A Theory of Anticipated Utility." *Journal of Economic Behavior & Organization* 3 (4): 323–43. https://doi.org/10.1016/0167-2681(82)90008-7.

Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel. 2011. "Search

Dynamics in Consumer Choice Under Time Pressure: An Eye-Tracking Study." *American Economic Review* 101 (2): 900–926. https://doi.org/10.1257/aer.101.2.900.

Russell, Stuart J. 1997. "Rationality and Intelligence." *Artificial Intelligence* 94 (1-2): 57–77. https://doi.org/10.1016/S0004-3702(97)00026-X.

Savage, Leonard. 1967. "Difficulties in the Theory of Personal Probability." *Philosophy of Science* 34 (4): 305–10. https://doi.org/10.1086/288168.

Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd. 2010. "Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload." *Journal of Consumer Research* 37 (3): 409–25. https://doi.org/10.1086/651235.

Schwartz, Barry. 2004. *The Paradox of Choice: Why More Is Less*. New York: Harper Collins.

Schwartz, Barry, Andrew Ward, John Monterosso, Sonja Lyubomirsky, Katherine White, and Darrin R. Lehman. 2002. "Maximizing Versus Satisficing: Happiness Is a Matter of Choice." *Journal of Personality and Social Psychology* 83 (5): 1178–97. https://doi.org/10.1037/0022-3514.83.5.1178.

Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, Gloria Origgi, and Deirdre Wilson. 2010. "Epistemic Vigilance." *Mind and Language* 25 (4): 359–93. https://doi.org/10.1111/j.1468-0017.2010.01394.x.

Tucker, Chris. 2016. "Satisficing and Motivated Submaximization (in the Philosophy of Religion)." *Philosophy and Phenomenological Research* 93 (1): 127–43. https://doi.org/10.1111/phpr.12191.

Weatherson, Brian. 2005. "Can We Do Without Pragmatic Encroachment?" *Philosophical Perspectives* 19 (1): 417–43. https://doi.org/10.1111/j.1520-8583.2005.00068.x.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. London: Macmillan.

Zweber, Adam. 2016. "Fallibilism, Closure, and Pragmatic Encroachment." *Philosophical Studies* 173 (10): 2745–57. https://doi.org/10.1007/s11098-016-0631-5.