The Extended Money Pump Argument

Philippe Mongin
CNRS & HEC Paris
mongin@greg-hec.com

June 2012

The classic money pump argument (henceforth MPA) allegedly shows that an agent with cyclical strict preferences is open to monetary exploitation, which allegedly establishes that he cannot be rational. Whatever the merits of the two steps of this reasoning (both have been disputed), the MPA falls short of defending the transitivity of preference, since it does not cover the other kinds of intransitivities. In the most interesting attempt to date, Gustafsson (2010) extends the MPA to these cases. Essentially, he substitutes triples of alternatives on which non-cyclical intransitivities occur with triples of lotteries, to which he applies a dominance principle from the theory of decision under uncertainty; the result is a cycle of *strict* preferences among the new objects, i.e., one that is open to monetary exploitation according to the usual argument. This ingenious reduction strategy has a cost that we think is not entirely well assessed. The dominance principle involves a new rationality consideration that Gustafsson takes to be less demanding that the kind of transitivity he tries to justify. Not questioning the principle itself, we point out two problems in applying it. One has to do with the assignment of probabilities, which cannot be decided at will. The other, deeper one is that there does not always exist a relevant triple of lotteries for every non-cyclical intransitive triple of alternatives that may occur; this follows from the complication of *event-dependence* that plagues decision theory since its beginnings in Ramsey and Savage. The upshot is the extended MPA in terms of uncertainty, though by no means invalid, is restricted in scope. Schumm's (1987) initial extended MPA, which involves a multi-attribute rather than an uncertainty framework, was objected to by Hansson (1993) and Gustafsson essentially on the same grounds. So the present note can be interpreted as carrying the critical argument one step further along the same line.

There are several kinds of non-cyclical intransitivies, but we concentrate on one of them:

(IIP) $aIb \& bIc \& aPc$.

We use a, b, c for alternatives and P, I, R for strict preference, indifference, and weak preference, respectively, among alternatives. The (IIP) type is widely evidenced in the psychological literature, which often explains it in terms of perceptive thresholds. A normative assessment of (IIP) might have to pay attention to this explanation, but we do not pursue the point here (see, e.g., Mongin, 2001).

With (PII), a relevant triple of lotteries L_1, L_2, L_3 is given by the rows of the following matrix:

	S_1	S_2	S_3
L_1	a	b	c
L_2	b	c	a
L_3	c	a	b

in which the columns S_1, S_2, S_3 stand for states of the worlds that are fixed in certain way. In Gustafsson's words, "the states are chosen so that they are jointly exhaustive, incompatible in pairs, positively probable, and independent of the lotteries" (2010, p. 256).

Lotteries are new objects calling for new preference relations, which we denote by $\succ, \sim, \succeq$ for strict preference, indifference, and weak preference, respectively. The reduction of the initial intransitivity to the standard MPA is achieved by applying a dominance principle that Gustafsson states thus:

(DP) "If there is a partition of states of the world such that it is independent of lotteries L' and L'' and relative to it, there is at least one positively probable state where the outcome of L' is strictly preferred to the outcome of L'' and no state where the outcome of $L\prime$ is not weakly preferred to the outcome of L'', then $L\prime$ is strictly preferred to L''" (ibid.).

In view of (DP) and the (IIP) assumption, strict preferences hold between successive rows and between the last and first ones, i.e., $L_3 \succ L_2 \succ L_1 \succ L_3$. QED.

Now to the objections, starting with the milder one. (DP) is a version of what decision theorists describe as the *strong* dominance principle, and to clarify it by a contrast, we state the *weak* dominance principle in the same words:

(DP') "If there is a partition of states of the world such that it is independent of lotteries L' and L'' and relative to it, and the outcome of L' is strictly preferred to the outcome of L'' for every state, then L' is strictly preferred to L''".

This simpler principle is also logically weaker because of the tacit assumption that there is at least one positively probable state. Although (DP) is fairly widely accepted, as references to Savage and others show, only (DP') is universally endorsed. Objections to (DP) arise in game-theoretic contexts where players applying it recursively sometimes end up discarding very natural equilibria. The extended MPA in terms of uncertainty needs (DP) precisely, but appears to be immune to these objections. Gustafsson's slightly obscure expression, "a partition of states of the world that is independent of lotteries", is explicitly intended to exclude a Newcomb-like situation in which the agent's preference for a lottery would influence the probability that a certain state results. We may interpret it as suggesting, more generally, that there is no strategic interaction between the theorist who fixes the states and the agent who evaluates the lotteries.

Thus, the ideal experiment belongs to the realm of individual decision theory, not game theory. But in this event, contrary to the suggestion in the first quote, it is not for the theorist to *choose* the states so that they are positively probable. Whether they are or not is a property of the agent's subjective beliefs, and these may attribute zero probability to one or two among S_1, S_2, S_3, granting that not all three can be in this case. This would block the application of (DP) once or twice, and once is enough to distroy the cycle. Assume specifically that the agent believe S_1 to be absolutely improbable; then he ends up with $L_3 \succ L_2 \succ L_1$, a well behaved ordering.

This is only the beginning of an objection since, generally speaking, the ideal experiment does not belong to the (von Neumann-Morgenstern) framework of lotteries with preassigned probabilities, but to the (Ramsey-Savage) framework of acting or betting on states that have non-apparent, purely subjective probabilities. Perhaps Gustafsson has in mind some physical device, like a roulette, by which S_1, S_2, S_3 could be endowed with universally agreed probabilities. However, the theorist may well claim that his roulette operates in such and such way, and the agent noneless disbelieve him, having his own probability assignments.

Granting that the Ramsey-Savage framework of subjective uncertainty is the appropriate one, we must clean the terminology somewhat. We should speak not only of *acts* (Savage's term) or *bets* (Ramsey's) instead of lotteries for L_1, L_2, L_3, but also of *events* for S_1, S_2, S_3 instead of states. In the framework now considered, states of the world are what they are given the relevant uncertainties, and surely not a matter of *choice* by anybody. The ideal experiment is best reformulated by saying that the theorist fixes a partition of events on the preexisting structure of states, and that he defines bets (let us settle for Ramsey's term) based on this partition, which he submits to the agent's preferences.

But now comes a major problem: these preferences may be event-dependent. Suppose the theorist has a roulette with a disk divided into into three differently coloured areas, A_1, A_2, A_3, and decides that event S_i occurs if the pointer stops on A_i. But suppose also that the agent is superstitious about colours. Oddly enough, he does not evaluate a, b, c in the same way whether the pointer stops on red A_1 or blue A_2 or green A_3. Does the argument based on (DP) carry through to this agent? The answer depends on his precisely defined preferences. Take the following ones:

- on S_1, aIb, bIc, aPc (the initial intransitivity)
- on S_2, aIb, bIc, cPa (a different (IIP)-intransitivity)
- on S_3, aIb, bIc, aPc (the initial intransitivity again).

The original bets do not bring about a cycle, but others would, say,

	S_1	S_2	S_3
L'_1	a	a	b
L'_2	b	c	c
L'_3	c	a	a

Indeed, by (DP), $L'_3 \succ L'_2 \succ L'_1 \succ L'_3$, as desired.

Here is now a case where the argument fails:

- on S_1, aIb, bIc, aIc (complete indifference)
- on S_2, aIb, bIc, aIc (complete indifference)
- on S_3, aIb, bIc, aPc (the initial intransitivity).

Such preferences are handled in terms of the *indifference principle* (IP), which is an easy variant of (DP) and (DP'): it says that two bets are indifferent if, for every event in the defining partition, the outcome of the first is indifferent to the outcome of the second. Now, (DP) and (IP) entail that $L_3 \sim L_2 \succ L_1 \sim L_3$, and generally, that no betting scheme is available for the extended MPA. This is because there is only one strict preference among bets if there is only one strict preference among alternatives. The same holds for two strict preferences instead of one, so it is necessary for the argument that there are at least three strict preferences among alternatives. As readily seen, this condition is also sufficient

We have illustrated event-dependence for an agent with odd preferences for coulours, but this was only to follow up the roulette example and dispel any impression that the use of chance devices is sufficient to preclude the difficulty. Here is a psychologically more realistic variant. Suppose the partition of events is determined by the agent's physical ability a year ahead, which he does not know: it is excellent in S_1, mediocre in S_2, and fair in S_3. The agent is a rockclimber and the alternatives are as follows:

a = a climb to Mont-Blanc (4810m)
b = a climb to Mont-Rose (4633m)
c = a climb to Grand Paradis (4061m).

If either S_1 or S_2 occurs, the agent is flatly indifferent between a, b, c (in the former, because each climb is equally dull, in the latter because each is equally horrendous). Only with S_3 does he make a difference; however, he can tell two alternatives from each other only if they are sufficiently apart, so he falls prey to (IIP). This is the earlier case of failure, now dressed up in a more intuitive way.

But if the partition $\{excellent,\ mediocre,\ fair\}$ does not support the extended MPA, why not simply replace it by another that does, say by $\{red,\ blue,\ green\}$ as in the roulette device? (We now assume that the agent has no preferences for colours.) This suggestion seems commonsensical, but actually runs against decision theory, when it is viewed as a theory of rational preference, not as a descriptive theory or as a decision-aid method. States of the world must be construed in an encompassing way. They must refer to every aspect of the world that is uncertain in the agent's eyes. In the particular example, the uncertainty that the colour partition captures is compounded with the preexisting uncertainty connected with the ability partition. Accordingly, there are nine states of the world, each corresponding to an ability-colour pair. If this set of states is divided by $\{red,\ blue,\ green\}$, no definite preferences can be associated with the partition cells, since they contain different states entailing different preferences. In sum, even if the selected events differ from those bringing about preference changes, the complication of event-dependence is still there.

A possible complaint is that the agent declared aIb, bIc, aPc at the outset without explicitly referring to any event at all. We can analyze these preference statements in two ways, each of which corresponds to a possible reading of the expression "preferences under certainty". In one analysis, the agent believes a particular event to be realized, and his initial preferences are conditional on it although it is unstated. Uncertainty occurs only at a later stage. In the rockclimber's example, S_3 would be the initially conditioning event, since it is the only one to support the three preference statements. At the time when he is faced with the ideal experiment, the rockclimber would have become uncertain between S_1, S_2 and S_3. In another, deeper and more challenging analysis, there is uncertainty throughout, "certainty" is a misnomer, and the statements aIb, bIc, aPc are in fact *averages* of more basic preference statements that may be event-dependent, the averages being taken in accordance with some theory of decision under uncertainty with event-dependent preferences. Glossing over the technical reconstruction, we may simply say that, when issuing aIb, bIc, aPc, the rockclimber is uncertain about S_1, S_2, S_3 and takes this uncertainty duly into account. In the latter theoretical option as well as in the former, event-dependence is logically compatible with the beginning step of the argument.

Digressing briefly, we would like to emphasize that this complication - also referred to as *state-dependence* - has attracted surprisingly little attention from philosophers, whereas it is central to current decision theory (see Karni, 1985, and Drèze, 1987, for path-breaking treatments; Karni and Mongin, 2000, for more interpretation; Hill, 2009, for a recent technical statement). With event-dependence, it becomes impossible to infer unique subjective probabilities from betting behaviour alone. Ramsey acknowledged the obstacle and simply put it aside. Savage claimed to supersede it by a suitable remodelling of the uncertainty situation, but his move is by now widely rejected as being ad hoc. Philosophers of probability and Bayesian philosophers of science still have to face these foundational problems squarely (see Mongin, 2011, for a plea to this effect). The present note testifies to the philosophical significance of event-dependent preferences in a different way.

Let us summarize our objection to Gustafsson's extension of the MPA by relating it to Schumm's original attempt. There, the agent was faced with nine different objects ranked along three dimensions, and he was intransitive along each dimension in the same IIP way. These three preference relations with the same pattern are necessary for a cycle to emerge, which restricts the scope of Schumm's argument, as Gustafsson (2010, p. 255) rightly emphasizes after Hansson (1993, p. 482). But the positive suggestion of replacing dimensions by uncertain events does not necessarily help. If event-dependence occurs, there can be as many different preference relations as there are events, so that Gustafsson's approach would be no more economical than Schumm's. The difference between the two is that event-dependence is only a *possibility* in Gustafsson, while preferences *need* to be dimension-dependent in Schumm. Whether one version effectively improves on the other hinges on a theoretical assessment of the event-dependence possibility - does it captures the general case of uncertainty or only some cases? Even the low-profile answer to this question means trouble for the new version, since it cannot claim to be universally applicable. Schumm wrongly believed to have shown "that one cannot plausibly abandon the transitivity of indifference without giving up that of preference as well" (1987, p. 437). Gustafsson's more raffined argument has not yet established this bold claim.

REFERENCES

Drèze, J. (1987), *Essays on Economic Decisions Under Uncertainty*, Cambridge, Cambridge University Press.

Karni, E. (1985), *Decision Making Under Uncertainty*, Harvard, Harvard University Press.

Gustafsson, J.E. (2010), "A Money-Pump for Acyclic Intransitive Preferences", *Dialectica*, 64, p. 251–257.

Hansson, S.O. (1993), "Money-Pumps, Self-Torturers and the Demons of Real Life", *Australasian Journal of Philosophy*, 71, p. 476–485.

Hill, B. (2009), "What is State-Dependence?", *Journal of Economic Theory*, 144, p. 1119–1134.

Karni, E. and P. Mongin (2000), "On the Determination of Subjective Probabilities by Choices", *Management Science*, 46, p. 233-248.

Mongin, P. (2001), "Does Optimization Imply Rationality?", *Synthese*, 124, p. 73-111.

Mongin, P. (2011), "What the Bayesian Decision Theorist Could Tell the Bayesian Philosopher", EPSA Lecture, LMPS Conference, Nancy.

Schumm, G.F. (1987), "Transitivity, Preference and Indifference", *Philosophical Studies*, 52, p. 435-437.