

A Money-Pump for Acyclic Intransitive Preferences

Johan E. Gustafsson*

The standard argument for the claim that rational preferences are transitive is the pragmatic money-pump argument. However, a money-pump only exploits agents with cyclic strict preferences. In order to pump agents who violate transitivity but without a cycle of strict preferences, one needs to somehow induce such a cycle. Methods for inducing cycles of strict preferences from non-cyclic violations of transitivity have been proposed in the literature, based either on offering the agent small monetary transaction premiums or on multi-dimensional preferences. This paper argues that previous proposals have been flawed and presents a new approach based on the dominance principle.

1. Introduction

The pragmatic money-pump argument is a standard argument for the claim that rational preferences have to be transitive.¹ Pragmatic arguments for rationality requirements aim to show that agents who violate a certain requirement can be made to act to their guaranteed disadvantage, which is taken to be a sign of irrationality.² Two transitivity properties for preferences can be stated as follows:³

PP-transitivity: $\forall x \forall y \forall z ((xPy \ \& \ yPz) \supset xPz)$

PI-transitivity: $\forall x \forall y \forall z ((xPy \ \& \ yIz) \supset xPz)$

The money-pump argument is often presented in an overly simplified form. For example, Amos Tversky writes:

Transitivity, however, is one of the basic and the most compelling principles of rational behaviour. For if one violates transitivity, it is a well-known conclusion that he is acting, in effect, as a “money-pump.” Suppose an individual prefers y to x , z to y , and x to z . It

* I would be grateful for any thoughts or comments on this paper, which can be sent to me at johan.eric.gustafsson@gmail.com.

¹ See, e.g. Davidson et al. (1955, 146) and Raiffa (1968, 78).

² For an overview, see Rabinowicz (2008).

³ Together, *PP*- and *PI*-transitivity imply transitivity of weak preference and transitivity of indifference, given completeness of weak preference. See Sen (1970, 18–19).

is reasonable to assume that he is willing to pay a sum of money to replace x by y . Similarly, he should be willing to pay some amount of money to replace y by z and still a third amount to replace z by x . Thus, he ends up with the alternative he started with but with less money. (Tversky, 1969, 45)

A weak point in Tversky's way of formulating the argument is that it only shows how to pump agents with cyclic strict preferences for money. In order to complete [p. 252] the argument we also need to show how to turn agents with the following acyclic types of preferences into money-pumps:⁴

PPI-preferences: $aPb \ \& \ bPc \ \& \ aIc$

PII-preferences: $aPb \ \& \ bIc \ \& \ aIc$

In order for the money-pump argument for transitive preferences to work it must be shown that agents with *PPI*- or *PII*-preferences are rationally committed to act as money-pumps. However, the method used to pump the agent in Tversky's example will not work for agents with *PPI*- or *PII*-preferences. Since an agent with these preferences does not prefer c over a , he need not be rationally committed to swap from a to c . So, an implicit assumption of Tversky's method is that a cycle of strict preferences can be induced from *PPI*- or *PII*-preferences. This paper explores approaches to how this can be done.

One might object that there is no need to make the agent rationally *committed*, given his non-transitive preferences, to act as a money-pump; it would be enough to make the agent rationally *permitted*, given the non-transitive preferences, to act as a money-pump. An agent with *PPI*- or *PII*-preferences would not act contrary to his preferences if he voluntarily swapped a for c and c for b and then paid a small amount to swap b for a . Thus, we have a non-forcing money-pump, one in which the agent who is acting consistently with his preferences can be, but is not rationally committed to be, exploited.⁵ The trouble with non-forcing money-pumps is that it does not follow from the consistency of having some preferences and behaving as a money-pump that it is rationally permitted, given these preferences, to behave like a money-pump. So although the agent may go along with the pump without choosing an option over a preferred alternative, the pump may still be blocked by some other rationality requirement on how to act given one's preferences.

⁴ There has been much debate on whether the money-pump works even on agents with cyclic strict preferences. See, e.g. Schick (1986), McClennen (1990), Rabinowicz (2000), and Sobel (2001).

⁵ For an example of a non-forcing money-pump, see Lehrer and Wagner (1985, 249–250). For the distinction between forcing and non-forcing pumps, see Gustafsson and Espinoza (2010).

If we grant the possibility of incomparable alternatives there are three further cases to consider, where ‘#’ denotes incomparability:⁶

PP#-preferences: $aPb \ \& \ bPc \ \& \ a\#c$

PI#-preferences: $aPb \ \& \ bIc \ \& \ a\#c$

IP#-preferences: $aIb \ \& \ bPc \ \& \ a\#c$

[p. 253] However, none of the proposals in the literature nor the one defended in this paper is able to induce cyclic preferences from agents with *PP#*-, *PI#*-, or *IP#*-preferences. So if we grant the possibility of incomparability, we need to supplement the money-pump argument with an argument for that rational preferences are complete.⁷ In the following I will take comparability for granted.

2. The small-bonus approach

The usual way to amend the method to work also for agents with *PII*-preferences is to offer the agent a very small sum of money if he is willing to take c instead of a and then another sum if he is willing to take b instead of c , in order to make the agent swap. As long as the agent is willing to pay more than these sums for the swap from b to a , he is still acting as a money-pump. In the case of *PPI*-preferences one only needs to offer a small sum to make the agent swap from a to c .⁸

However, the small-bonus approach seems to be begging the question. On this approach it is necessary that a small sum of money will make the agent with *PPI*- or *PII*-preferences swap a for c . Let c^+ be this improved alternative of c with an added small sum of money. The agent is indifferent between a and c and it seems very plausible that the agent should prefer c^+ over c . The crucial premise is that the agent given these preferences is rationally committed to prefer c^+ over a . That is, the small-bonus approach needs the following premise:⁹

PI⁺-transitivity: $\forall x \forall y ((x^+Px \ \& \ xIy) \supset x^+Py)$, where x^+ is x with an additional sum of money.

⁶ *IP#*-preferences do not violate *PP*- or *PI*-transitivity but they violate *IP*-transitivity, i.e. the condition $\forall x \forall y \forall z ((xIy \ \& \ yPz) \supset xPz)$. *IP*-transitivity does not follow from the combination of *PP*- and *PI*-transitivity without completeness.

⁷ For examples of pragmatic arguments that rational preferences are complete, see Broome (1999, 156) and Peterson (2007). However, the pumps they propose are of the weak non-forcing type and suffer the same problem as the non-forcing pump discussed above.

⁸ For examples of this approach, see, e.g. McClennen (1990, 90–91), Hansson (1993, 478), and Rabinowicz (2008, 149–150, n. 7).

⁹ Strictly, a principle like $\forall x \forall y (xIy \supset x^+Py)$ would be enough. But since an underlying assumption in any money-pump argument is that any option with an addition of money is strictly preferred to the same option without the added money (why else would the loss of money be irrational?), this alternative principle would not be weaker.

Without PI^+ -transitivity we cannot conclude that agents with PPI - or PII -preferences are in general rationally committed to swap a for c^+ . But PI^+ -transitivity just seems to be a special case of PI -transitivity. Since PI^+ -transitivity is part of an argument that is supposed to establish that PI - and PP -transitivity are rationally required, any support for PI^+ -transitivity from PI - or PP -transitivity would beg the question. The problem then is that there does not seem to be any plausible support for the claim that violating preferences should be considered irrational that is independent of PI - or PP -transitivity.

[p. 254] Furthermore, many of the proposed counterexamples to transitivity contradict PI^+ -transitivity. For example, let d be a trip to Florida, let e be a trip to California, and let e^+ be the trip to California with an additional \$1.¹⁰ It seems at least *prima facie* plausible that it is rationally permitted to be indifferent between d and e and between d and e^+ while still preferring e^+ to e . It seems question-begging to argue that such counterexamples are irrational by help of PI^+ -transitivity since someone who finds them plausible is not likely to find the premise PI^+ -transitivity plausible.

One might object that one could provide support for PI^+ -transitivity by showing that violating preferences are irrational since an agent with these preferences could be used as a money-pump. But this would just lead to a regress as we, in order for this new money-pump to work, would still need to make the agent swap a for c , and thus, we would still have the same problem.

Nevertheless, one might object that the small-bonus approach does not really need PI^+ -transitivity. The agent can simply prefer c^+ to a . The problem with this reasoning is that it at most shows that an agent with either PPI - or PII -preferences who also prefers c^+ to a is rationally committed to behave as a money-pump. It does not show that an agent is rationally committed to behave as a money-pump just by having PPI - or PII -preferences, which is needed in order to show that any agent who violates transitivity is rationally committed to behave as a money-pump.

3. The multi-dimensional approach

George F. Schumm (1987, 436) has proposed another method for converting violations of the transitivity of indifference into cycles of strict preference.¹¹ He presents an example where an agent has PII -type pref-

¹⁰ This example seems to be due to Restle (1961, 62–63). A similar example is ascribed to Armstrong (1939) by Lehrer and Wagner (1985, 255). Nevertheless, the example is lacking in Armstrong (1939). For a similar conclusion on the origin of this example, see Hansson (2001, 332).

¹¹ For an earlier similar proposal, see Ng (1977, 52).

erences over three independent alternative sets: three red balls over which he holds r_1Pr_2 & r_2Ir_3 & r_3Ir_1 , three green balls over which he holds g_1Ig_2 & g_2Pg_3 & g_3Ig_1 , and three blue balls over which he holds b_1Ib_2 & b_2Ib_3 & b_3Pb_1 . He is presented three boxes, each containing a red ball r_i , a green ball g_i , and a blue ball b_i . Schumm then argues that the agent's preferences over these combined options should be cyclic:¹²

$$\begin{aligned} &< r_1, g_1, b_1 > P < r_2, g_2, b_2 > \\ &\& < r_2, g_2, b_2 > P < r_3, g_3, b_3 > \\ &\& < r_3, g_3, b_3 > P < r_1, g_1, b_1 > \end{aligned}$$

[p. 255] This conclusion presupposes that there are no interactions between the different types of objects, for example, the balls of different colours in one box do not look especially good or bad together. Also, if objects of some type in one bundle are preferred over the objects of the same type in another bundle and all other types of objects are equi-preferred between the bundles, then the first bundle should be preferred over the other. A virtue of Schumm's example is that it does not rely on any monetary offers to the agent. But as a general method for converting violations of the transitivity of indifference into cycles of strict preference it has a serious drawback. The approach does not work if there are just one or two sets of alternatives over which the agent has *PII*-type preferences. We need to be able to induce cyclic strict preferences from agents with *PII*-type preferences over just one set of alternatives.

4. The dominance approach

I propose that the problems that face the small-bonus approach and the multi-dimensional approach can be overcome if we just add as a premise that the dominance principle is a rational requirement:

Dominance: If there is a partition of states of the world such that it is independent¹³ 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' is not weakly preferred to the outcome of L'' , then L' is strictly preferred to L'' .¹⁴

¹² Of course the same move could be made with an agent who has the *PPI*-type preferences r_1Pr_2 & r_2Pr_3 & r_3Ir_1 , g_1Ig_2 & g_2Pg_3 & g_3Pg_1 , and b_1Pb_2 & b_2Ib_3 & b_3Pb_1 .

¹³ The discovery of Newcomb's problem has led to a dispute on whether this independence should be causal or evidential. See Nozick (1969), Gibbard and Harper (1978, 146–151), and Joyce (1999, 150–151). Since it does not affect my argument I will not take a stand on this issue.

¹⁴ See, e.g. Savage (1951, 58), Milnor (1954, 55), Luce and Raiffa (1957, 287), Nozick (1969, 118), Jeffrey (1983, 9), and McClennen (1990, 48).

The dominance principle is less controversial than PI^+ -transitivity. Note also that the usual counterexamples to transitivity are not counterexamples to dominance. Therefore, unlike PI^+ -transitivity, the dominance principle does not in itself beg the question against the usual counterexamples to transitivity. So, it should not beg the question as a premise in an argument for transitivity.

Suppose an agent has PPI - or PII -preferences, that is, $aPb \& bPc \& aIc$ or $aPb \& bIc \& aIc$. We then construct the following lotteries:

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

[p. 256] The states are chosen so that they are jointly exhaustive, incompatible in pairs, positively probable, and independent of the lotteries. So, for example, L_1 is a lottery with a probability $P(S_1)$ of prize a , $P(S_2)$ of b , and $P(S_3)$ of c . The dominance principle implies that the agent is rationally required to have the following cyclic strict preferences over the lotteries:

$$L_1PL_2 \& L_2PL_3 \& L_3PL_1.$$

Like the multi-dimensional approach, this approach does not rely on any monetary offers to the agent. Note also that this approach, unlike the multi-dimensional approach, can induce cyclic strict preferences from single violations of transitivity.

So, to show that any violation of transitivity that satisfies completeness can be turned into money-pumps involves two steps. First we will induce a cycle of strict preferences with help of the dominance principle. Then we pump the cyclic strict preferences by the standard methods; for example, the one Tversky sketched. Of course, adding the dominance principle as a premise comes at a price. But as I have argued, the price is relatively low and the pump should still be worth it since the dominance principle is a much less contested rationality requirement than transitivity.

References

- Armstrong, W. E.: 1939, "The Determinateness of the Utility Function", *The Economic Journal* 49(195), pp. 453–467.
- Broome, J.: 1999, *Ethics Out of Economics*, Cambridge: Cambridge University Press.
- Davidson, D., J. C. C. McKinsey, and P. Suppes: 1955, "Outlines of a Formal Theory of Value, I", *Philosophy of Science* 22(2), pp. 140–160.

- Gibbard, A. and W. L. Harper: 1978, "Counterfactuals and Two Kinds of Expected Utility", in: C. A. H. Hooker, J. J. Leach, and E. F. McClennen (eds.): *Foundations and Applications of Decision Theory*, Vol. I, Dordrecht: D. Reidel, pp. 125–162.
- Gustafsson, J. E. and N. Espinoza: 2010, "Conflicting Reasons in the Small-Improvement Argument", *The Philosophical Quarterly* 60(241), pp. 754–763.
- Hansson, S. O.: 1993, "Money-Pumps, Self-Torturers and the Demons of Real Life", *Australasian Journal of Philosophy* 71(4), pp. 476–485.
- Hansson, S. O.: 2001, "Preference logic", in: D. M. Gabbay and F. Guenther (eds.): *Handbook of Philosophical Logic*, Vol. 4, 2nd edn., Dordrecht: Kluwer, pp. 319–394.
- Jeffrey, R. C.: 1983, *The Logic of Decision*, 2nd edn., Chicago: University of Chicago Press.
- Joyce, J. M.: 1999, *The Foundations of Causal Decision Theory*, Cambridge: Cambridge University Press.
- Lehrer, K. and C. Wagner: 1985, "Intransitive Indifference: The Semi-order Problem", *Synthese* 65(2), pp. 249–256.
- Luce, R. D. and H. Raiffa: 1957, *Games and Decisions*, New York: Wiley.
- McClennen, E. F.: 1990, *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge: Cambridge University Press.
- Milnor, J.: 1954, "Games Against Nature", in: R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.): *Decision Processes*, New York: Wiley, pp. 49–59.
- Ng, Y.-K.: 1977, "Sub-semiorder: A Model of Multidimensional Choice with Preference Intransitivity", *Journal of Mathematical Psychology* 16(1), pp. 51–59.
- Nozick, R.: 1969, "Newcomb's Problem and Two Principles of Choice", in: N. Rescher (ed.): *Essays in Honor of Carl G. Hempel*, New York: Humanities Press, pp. 114–146.
- Peterson, M.: 2007, "Parity, Clumpiness and Rational Choice", *Utilitas* 19(4), pp. 505–513.
- Rabinowicz, W.: 2000, "Money Pump with Foresight", in: M. J. Almeida (ed.): *Imperceptible Harms and Benefits*, Dordrecht: Kluwer, pp. 123–154.
- Rabinowicz, W.: 2008, "Pragmatic Arguments for Rationality Constraints", in: M. C. Galavotti, R. Scazzieri, and P. Suppes (eds.): *Reasoning, Rationality, and Probability*, Stanford: CSLI Publications, pp. 139–163.
- Raiffa, H.: 1968, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*, Reading, Massachusetts: Addison-Wesley.
- Restle, F.: 1961, *Psychology of Judgment and Choice: A Theoretical Essay*, New York: Wiley.
- Savage, L. J.: 1951, "The Theory of Statistical Decision", *Journal of the*

- American Statistical Association* 46(253), pp. 55–67.
- Schick, F.: 1986, “Dutch Bookies and Money Pumps”, *The Journal of Philosophy* 83(2), pp. 112–119.
- Schumm, G. F.: 1987, “Transitivity, Preference and Indifference”, *Philosophical Studies* 52(3), pp. 435–437.
- Sen, A. K.: 1970, *Collective Choice and Social Welfare*, San Francisco: Holden-Day.
- Sobel, J. H.: 2001, “Money Pumps”, *Philosophy of Science* 68(2), pp. 242–257.
- Tversky, A.: 1969, “Intransitivity of Preferences”, *Psychological Review* 76(1), pp. 31–48.

Thanks to John Cantwell, Sven Ove Hansson, Martin Peterson, and three anonymous referees for *dialectica*.