A Universal Money Pump for the Myopic, Naive, and Minimally Sophisticated

Johan E. Gustafsson*

Draft: April 17, 2022 at 9:58 p.m.

ABSTRACT. Suppose that you prefer A to B, B to C, and C to A. The standard argument that such cyclic preferences are irrational is the money-pump argument. This argument can be based on a number of different money pumps that vary in what needs to be assumed about the agent. The Standard Money Pump works for myopic and naive agents, but not for sophisticated agents who use backward induction. The Upfront Money Pump works for sophisticated agents, but not for myopic or naive agents. In this paper, I present a new money pump, the Universal Money Pump, that works for myopic, naive agents, and sophisticated agents. Moreover, the Universal Money Pump (just like the Upfront Money Pump) also works for minimally sophisticated agents who do not assume that they will choose rationally at nodes that can only be reached by irrational choices.

Suppose that you prefer A to B, B to C, and C to A. Letting '$X \succ Y$' denote that X is (strictly) preferred to Y, we can represent your preferences as follows:

(1) $A \succ B \succ C \succ A$.

These preferences are cyclic. More specifically, these preferences violate Three-Step Acyclicity:[1]

> *Three-Step Acyclicity* If $X \succ Y \succ Z$, then it is not the case that $Z \succ X$.

The standard argument that Three-Step Acyclicity is a requirement of rationality is the money-pump argument. A *money-pump argument* for an

* 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] Samuelson 1947, p. 151 and Sen 1977, p. 62.

alleged requirement of rationality (such as Three-Step Acyclicity) is an argument that otherwise rational agents who violate the requirement would in some possible situation end up paying for something they could have kept for free even though they knew in advance what decision problem they were facing.[2]

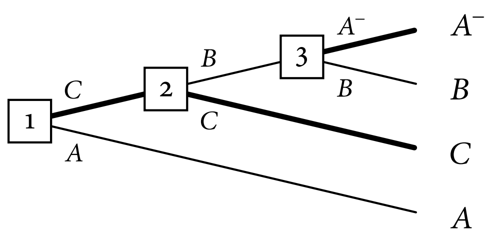
The standard version runs as follows: Suppose that, initially, you can walk away with A—that is, you end up with A if you turn down all trades. First, you are offered a trade from A to C. Since you prefer C to A, you accept this trade. Then you are offered a trade from C to B. Since you prefer B to C, you also accept this trade. Finally, you are offered a trade from C to A for a small payment. Since you prefer A to B with some margin, there should be some small amount of money that you would be willing to pay to get A instead of B. Likewise, since you prefer C to A, you should prefer C without any payment to A with a payment. Accordingly, there is a soured version A^- of A such that

(2) $A \succ A^-$, and $C \succ A^- \succ B$.

So we let the third offer be an offer to trade from B to A^-. Since you prefer A^- to B, you accept this final offer and end up with A^- even though you could have walked away with A (that is, you could have keep A without paying anything).[3]

We can diagram this set-up with a decision tree:[4]

The Standard Money Pump



$A \succ B \succ C \succ A \succ A^-$, and $C \succ A^- \succ B$.

Here, the squares represent the choice nodes where you are offered the trades. Accepting a trade corresponds to going up at a node, and turning

[2] Gustafsson forthcoming.

[3] Davidson et al. 1955, p. 146, Edwards et al. 1965, p. 273, and Pratt et al. 1965, ch. 2, p. 10.

[4] Rabinowicz 1995, p. 393.

a trade down corresponds to going down.[5] The outcome on the upper right of each square is what you get if you accept the trade. The outcome on the lower right of each square is what you get if you turn the trade down.

The standard version of the money-pump argument, as presented earlier, assumes that the agent is myopic. *Myopic* agents make their choices under the assumption that they will walk way from all future trades. If we treat the walk-away option (that is, what you get if you turn down all future trades) as what you currently possess, then being myopic can be thought of as choosing between possessions without taking future choices into account.[6]

The Standard Money Pump, however, also works for naive agents. *Naive* agents (i) consider the outcomes of all available plans and assess which of these outcomes are choice-worthy in a choice between all of them and (ii) choose in accordance with a plan to end up with one of these choice-worthy outcomes, without taking into consideration whether they would later depart from that plan.[7] To be naive in the Standard Money Pump, we need to make a choice between all potential outcomes—that is, between A, A^{-}, B, and C. But, given the preferences in (1) and (2), we can't rely on maximization—that is, the following rule:[8]

> *The Maximization Rule* It is rationally permitted to choose an outcome X if and only if there is no feasible outcome Y such that $Y \succ X$.

The trouble is that, for each potential outcome, there is another that's preferred to it. So, rather than maximization, we adopt the following rule:[9]

> *The Uncovered-Choice Rule* It is rationally permitted to choose a outcome X if and only if there is no feasible outcome Y such that $Y \succ X$ and, for all feasible outcomes Z, it holds that $Y \succ Z$ if $X \succ Z$.

Choosing between all potential outcomes at node 1 with the Uncovered-Choice Rule, you only deem A, B, and C as choice-worthy. So it's rationally permitted both to accept and to turn down the trade at node 1. So you

[5] Following Rabinowicz 2008, p. 152.

[6] See Dow 1984, p. 96 and McClennen 1990, pp. 11–12.

[7] Pollak 1968, pp. 202–203 and Hammond 1976, p. 162.

[8] Uzawa 1956, p. 37.

[9] Schwartz 1990, p. 21; see also Miller 1980, pp. 72–74.

can rationally accept the trade at node 1. Then, at node 2, the potential outcomes are just A^-, B, and C. Choosing between them with the Uncovered-Choice Rule, you deem all of them as choice-worthy. Hence it's rationally permitted both to accept and to turn down the trade at node 2. So you can rationally accept the trade at node 2. At node 3, the potential outcomes are A^- and B. Choosing between them with the Uncovered-Choice Rule, you only deem A^- as choice-worthy. So you accept the trade at node 3. Hence, with rationally permitted choices, you end up with A^- even though you could have walked away with A.

The Standard Money Pump for naive choosers is an example of a permitting, non-forcing money pump. A money pump is *forcing* if and only if the agent is rationally required, at each step, to go along with the exploitation. A money pump is *permitting* if and only if, at each step, the agent is rationally permitted to go along with the exploitation. A money pump is *non-prohibiting* if and only if, at each step, the agent is not rationally prohibited from go along with the exploitation. Finally, a money pump is *non-forcing* if and only if it is non-prohibiting and, at some step, the agent is not rationally required to go along with the exploitation.[10]

While the money pumps we shall deploy against other kinds of agents will be forcing, we can't do better than permitting money pumps when we deal with naive agents. The trouble, roughly, is that—as long as all of A, A^-, B, and C are potential outcomes—the plans leading to A (the walk-away outcome) will be rationally permitted. And, as long as A is a potential outcome, it cannot be rationally required to follow a plan that potentially leads to A^- and not potentially to A. So, when A is a potential outcome, the only way it can be rationally required to follow a plan that potentially leads to A^- and not potentially to A is when, in addition, not all of A, B, and C are potential outcomes. But, if A, B, and C aren't all potential outcomes, we can't exploit the fact that the agent has cyclic preferences over these outcomes.

So far, we have only covered myopic and naive agents. Some agents, however, are neither myopic nor naive. A problem with the Standard Money Pump is that cyclic preferrers can avoid being money pumped if they are sophisticated, rather than myopic or naive. *Sophisticated* agents make their choices using backward induction based on what they predict they will choose in the future, assuming that they choose rationally at all

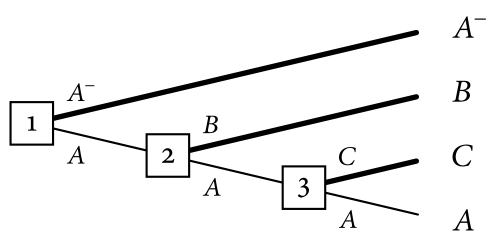
[10] Gustafsson and Espinoza 2010, pp. 761–762 and Gustafsson forthcoming.

future choice nodes.[11]

To see that sophisticated agents with the preferences in (1) and (2) avoid the Standard Money Pump, suppose that you are sophisticated with those preferences. At node 3, you would accept the trade from B to A^-, since you prefer A^- to B. Taking this prediction into account, the choice at node 2 is effectively between A^- (accepting the trade) and C (turning it down). Since you prefer C to A^-, you would turn down the trade at node 2. Taking this prediction into account, the choice at node 1 is effectively between C (accepting the trade) and A (turning it down). Since you prefer C to A, you would accept the trade at node 1. So you end up with C and avoid being money pumped. (The choices that are recommend by backward induction—assuming rational choices at future nodes—are marked with thicker lines in the decision tree.)

But there are money pumps that work for sophisticated agents. Consider the following decision problem:[12]

The Upfront Money Pump



$$A \succ B \succ C \succ A \succ A^-, \text{ and } C \succ A^- \succ B.$$

To see that sophisticated agents with the preferences in (1) and (2) cannot avoid exploitation in the Upfront Money Pump, suppose that you are sophisticated with those preferences. At node 3, you would accept the trade from A to C, since you prefer C to A. Taking this prediction into account, the choice at node 2 is effectively between B (accepting the trade) and C (turning it down). Since you prefer B to C, you would accept the trade at node 2. Taking this prediction into account, the choice at node 1 is effectively between A^- (accepting the trade) and B (turning it down). Since

[11] von Neumann and Morgenstern 1944, pp. 116–117, Pollak 1968, p. 203, and Hammond 1976, p. 162.

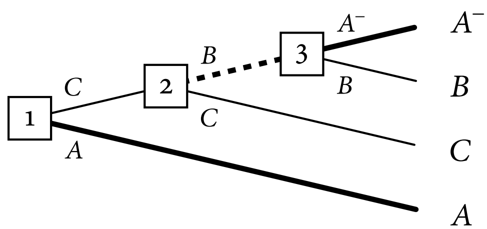
[12] Gustafsson and Rabinowicz 2020, p. 583. For a similar construction in voting theory, see Moulin 1983, pp. 96–97. Rabinowicz (2000, p. 141) presents another money pump for sophisticated agents, but—unlike the Upfront Money Pump—it does not work for minimally sophisticated agents (defined later).

you prefer A^- to B, you accept the trade at node 1. Hence you end up with A^- even though you could have walked away with A.

The standard form of backward induction on which sophisticated agents rely can be challenged, however. It is based on the dubious assumption that the agent would choose rationally even at future choice nodes that can only be reached by irrational choices.[13] There is, however, a weaker form of sophistication which doesn't rely on this assumption. *Minimally sophisticated* agents choose using backward induction based on what they predict they will choose in the future assuming that, at nodes that can be reached without making any irrational choices, they retain (i) their rationality and (ii) their trust in their rationality at nodes that can be reached without making any irrational choices.[14]

In some decision problems, minimally sophisticated agents need not make the same choices as sophisticated agents with the same preferences. Consider, once more, the Standard Money Pump. Suppose that you predict that you would, irrationally, accept the trade at node 2. We mark this predicted irrational choice with a thick dashed line in the decision tree:

The Standard Money Pump (with a predicted irrational choice)



$A \succ B \succ C \succ A \succ A^-$, and $C \succ A^- \succ B$.

Given the prediction that you would accept the trades at nodes 2 and 3, the choice at node 1 is effectively between A^- (accepting the trade) and A (turning it down). Since you prefer A to A^-, you turn down the trade at node 1. In fact, given your prediction, it would be irrational to accept the trade at node 1. Hence the predicted irrational choice at node 2 would follow an initial irrational choice at node 1. So a minimally sophisticated agent can consistently turn down the initial trade in the Standard Money Pump.

[13] Binmore 1987, pp. 196–200, Bicchieri 1988, pp. 334–335, and Pettit and Sugden 1989, pp. 171–174.

[14] Rabinowicz 1998, pp. 108–109 and Gustafsson and Rabinowicz 2020, p. 583.

In some decision problems, however, minimally sophisticated agents must make the same choices as sophisticated agents. One such decision problem is the Upfront Money Pump. We will show this by a proof by contradiction.

Given that you're minimally sophisticated, we note that you choose using backward induction based on what you predict you will choose in the future assuming that, at nodes that can be reached without making any irrational choices, you retain (i) your rationality and (ii) your trust in your rationality at nodes that can be reached without making any irrational choices.

Firstly, we assume (for proof by contradiction) that node 3 in the Upfront Money Pump can be reached without making any irrational choices. Then, at all choice nodes, you retain your rationality and your trust in your rationality at these nodes. Accordingly, you would accept the trade from A to C at node 3, since you prefer C to A. But, if so, the choice at node 2 is effectively a choice between B (accepting the trade) and C (turning it down). Since you prefer B to C, the choice to turn down the trade at node 2 was irrational, which contradicts our assumption that node 3 can be reached without making any irrational choices.

Secondly, we assume (for proof by contradiction) that node 2 can be reached without making any irrational choices. Then, at nodes 1 and 2, you retain your rationality and your trust in your rationality at these nodes. Since we have already shown that node 3 cannot be reach without making irrational choices, it follows that it's irrational to turn down the trade at node 2. So you would accept the trade at node 2. But, if so, the choice at node 1 is effectively between A^- (accepting the trade) and B (turning it down). Since you prefer A^- to B, the choice to turn down the trade at node 1 was irrational, which contradicts our assumption that node 2 can be reached without making any irrational choices.

Hence it's irrational to turn down the trade at node 1. So you accept the initial trade and end up with A^-, even though you could have walked away with A.[15]

So neither sophisticated nor minimally sophisticated agents avoid exploitation in the Upfront Money Pump. How do myopic and naive agents fare?

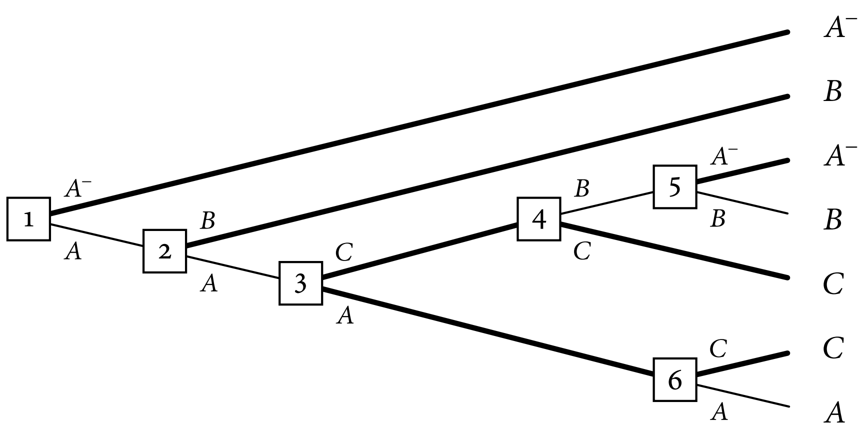
[15] Gustafsson and Rabinowicz 2020, p. 585, which adapts the argument from Broome and Rabinowicz 1999, pp. 240–242. See also Rabinowicz 1998, pp. 108–109 and Aumann 1998, p. 103.

If you are myopic, you turn down the trade from A to A^- at node 1, since you prefer A to A^-. Then, at node 2, you turn down the trade from A to B, since you prefer B to A. Finally, at node 3, you accept the trade from A to C, since you prefer C to A. Hence you end up with C and avoid exploitation.

If you're a naive agent following the Uncovered-Choice Rule at node 1, you only deem A, B, and C as choice-worthy among the potential outcomes A, A^-, B, and C. So you turn down the trade at node 1. Then, at node 2, you deem all the potential outcomes A, B, and C as choice-worthy. So it's rationally permitted both to accept the trade and to turn it down. If you accept the trade at node 2, you end up with B and avoid exploitation. If you turn it down, you will accept the trade at node 3, since you only deem C as choice-worthy among the potential outcomes A and C. And then you end up with C and, likewise, avoid exploitation.

Hence neither myopic nor naive agents are open to exploitation in the Upfront Money Pump. There is, however, a universal money-pump that works for myopic, naive, and (minimally) sophisticated agents. Consider the following decision problem:

The Universal Money Pump



$$A \succ B \succ C \succ A \succ A^-, \text{and } C \succ A^- \succ B.$$

To see that myopic agents with the preferences in (1) and (2) get money pumped in the Universal Money Pump, suppose that you are myopic with those preferences. At node 1, you turn down the trade from A to A^-, since you prefer A to A^-. At node 2, you turn down the trade from A to B, since you prefer A to B. At node 3, you accept the trade from A to C, since you prefer C to A. At node 4, you accept the trade from C to B, since you prefer B to C. Finally, at node 5, you accept the trade from B to A^-, since you prefer A^- to B. Hence you end up with A^- even though you could have walked away with A.

To see that naive agents with the preferences in (1) and (2) may be money pumped in the Universal Money Pump, suppose that you are naive with those preferences. If you're a naive agent who follows the Uncovered-Choice Rule at node 1, you only deem A, B, and C as choice-worthy among the potential outcomes A, A^-, B, and C. So you turn down the trade at node 1. At node 2, the potential outcomes are the same, so it's rationally permitted to turn down the trade (and permitted to accept it). If you turn the trade down, you reach node 3, where the potential outcomes are still the same. So it's rationally permitted to accept the trade (and permitted to turn it down). If you accept, you reach node 4, where the potential outcomes are A^-, B, and C. You deem all of them as choice-worthy, so it's rationally permitted to accept the trade at node 4 (and permitted to turn it down). If you accept, you reach node 5, where the potential outcomes are A^- and B. Since you only deem A^- as choice-worthy, you accept the trade. Hence you end up with A^- even though you could have walked away with A.

To see that sophisticated agents with the preferences in (1) and (2) get money pumped in the Universal Money Pump, suppose that you are sophisticated with those preferences. At node 5, you would accept the trade from B to A^-, since you prefer A^- to B. Taking this into account, the choice at node 4 is effectively between A^- (accepting the trade) and C (turning it down). Since you prefer C to A^-, you would turn down the trade at node 4. Note, next, that you would accept the trade from A to C at node 6, since you prefer C to A. Taking these predictions into account, your choice at node 3 is effectively a choice between C (accepting the trade) and C (turning it down). So, if you were to reach node 3, you would end up with C. Taking this into account, the choice at node 2 is effectively between B (accepting the trade) and C (turning it down). Since you prefer B to C, you would accept the trade at node 2. Finally, taking this prediction into account, your choice at node 1 is effectively between A^- (accepting the trade) and B (turning it down). Since you prefer A^- to B, you accept the initial trade and end up with A^- even though you could have walked away with A.

To see that minimally sophisticated agents with the preferences in (1) and (2) get money pumped in the Universal Money Pump, suppose that you are minimally sophisticated with those preferences. As before, we note that, being minimally sophisticated, you choose using backward induction based on what you predict you will choose in the future assuming that, at nodes that can be reached without making any irrational choices, you retain (i) your rationality and (ii) your trust in your rationality at nodes that can be reached without making any irrational choices.

Firstly, we assume (for proof by contradiction) that node 5 can be reached without making any irrational choices. Then, at each choice node, you retain your rationality and your trust in your rationality at these nodes. Hence you would accept the trade from B to A^- at node 5, since you prefer A^- to B. But then the choice to accept the trade at node 4 was irrational, since it is effectively a choice of A^- over C, even though you prefer C to A^-. So node 5 can only be reached by making some irrational choices, which contradicts our assumption.

Secondly, we assume (for proof by contradiction) that nodes 4 and 6 can be reached without making any irrational choices. Then—at nodes 1, 2, 3, 4, and 6—you retain your rationality and your trust in your rationality at these nodes. And, since we have already shown that node 5 cannot be reached without making irrational choices, we find that it must be irrational to accept the trade at node 4. Hence you would turn down the trade at node 4. And, since you prefer C to A, you would accept the trade at node 6. Then the choice at node 3 is effectively a choice where you end up with C regardless of whether you accept or turn down the trade. But then the choice to turn down the trade at node 2 was irrational, since it is effectively a choice of C over B, even though you prefer B to C. So nodes 4 and 6 can only be reached by making some irrational choices, which contradicts our assumption.

Thirdly, we assume (for proof by contradiction) that node 4 can be reached without making any irrational choices. Then—at nodes 1, 2, 3, and 4—you retain your rationality and your trust in your rationality at these nodes. And, since we have shown that node 5 can't be reached without making irrational choices, it must be irrational to accept the trade at node 4. So you would turn down the trade at node 4. Since we have shown that nodes 4 and 6 cannot both be reached without making any irrational choices, we find that 6 can't be reached without making any irrational choices. Hence it's irrational to turn down the trade at node 3, and so you would accept that trade. But then the choice to turn down the trade at node 2 was irrational, since it is effectively a choice of C over B, even though you prefer B to C. So node 4 can only be reached by making some irrational choices, which contradicts our assumption.

Fourthly, we assume (for proof by contradiction) that node 3 can be reached without making any irrational choices. Then—at nodes 1, 2, and 3—you retain your rationality and your trust in your rationality at these nodes. And, since we've shown that node 4 can't be reached without making irrational choices, we find that it must be irrational to accept the trade at node 3. So you would turn down the trade at node 3. And node 6 can be reached without making any irrational choices. Hence, at node 6, you retain your rationality. And so you would accept the trade from *A* to *C* at node 6, since you prefer *C* to *A*. But then the choice to turn down the trade at node 2 was irrational, since it is effectively a choice of *C* over *B*, even though you prefer *B* to *C*. So node 3 can only be reached by making some irrational choices, which contradicts our assumption.

Finally, we assume (for proof by contradiction) that node 2 can be reached without making any irrational choices. Then, at nodes 1 and 2, you retain your rationality and your trust in your rationality at these nodes. And, since we've shown that node 3 can't be reached without making irrational choices, it must be irrational to turn down the trade at node 2. Hence you would accept the trade at node 2. But then the choice to turn down the trade at node 1 was irrational, since it is effectively a choice of *B* over A^-, even though you prefer A^- to *B*. So node 2 can only be reached by making some irrational choices, which contradicts our assumption.

Hence node 2 can't be reached without making an irrational choice. So, at node 1, it is irrational to turn down the trade. So you go up at node 1 and end up with A^-, even though you could have walked away with *A*.

Hence we have a money pump that works for agents who violate Three-Step Acyclicity regardless of whether they are myopic, naive, or (minimally) sophisticated.[16]

References

Aumann, Robert J. (1998) 'On the Centipede Game', *Games and Economic Behavior* 23 (1): 97–105.

[16] A notable limitation of the Universal Money Pump is that, unlike the Standard Money Pump and the Upfront Money Pump for myopic and sophisticated agents, there's no straightforward way to extend it to cover preference cycles over more than three outcomes—while retaining it universality. So it can't be used to defend Acyclicity in general for myopic, naive, and minimally sophisticated agents; it only works for Three-Step Acyclicity.

Bicchieri, Christina (1988) 'Backward Induction without Common Knowledge', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1988 (2): 329–343.

Binmore, Ken (1987) 'Modeling Rational Players: Part I', *Economics and Philosophy* 3 (2): 179–214.

Broome, John and Wlodek Rabinowicz (1999) 'Backwards Induction in the Centipede Game', *Analysis* 59 (4): 237–242.

Davidson, Donald, J. C. C. McKinsey, and Patrick Suppes (1955) 'Outlines of a Formal Theory of Value, I', *Philosophy of Science* 22 (2): 140–160.

Dow, Gregory K. (1984) 'Myopia, Amnesia, and Consistent Intertemporal Choice', *Mathematical Social Sciences* 8 (2): 95–109.

Edwards, Ward, Harold Lindman, and Lawrence D. Phillips (1965) 'Emerging Technologies for Making Decisions', in Theodore M. Newcomb, ed., *New Directions in Psychology II*, pp. 259–325, New York: Holt, Rinehart and Winston.

Gustafsson, Johan E. (forthcoming) *Money-Pump Arguments*, Cambridge: Cambridge University Press.

Gustafsson, Johan E. and Nicolas Espinoza (2010) 'Conflicting Reasons in the Small-Improvement Argument', *The Philosophical Quarterly* 60 (241): 754–763.

Gustafsson, Johan E. and Wlodek Rabinowicz (2020) 'A Simpler, More Compelling Money Pump with Foresight', *The Journal of Philosophy* 117 (10): 578–589.

Hammond, Peter J. (1976) 'Changing Tastes and Coherent Dynamic Choice', *The Review of Economic Studies* 43 (1): 159–173.

McClennen, Edward F. (1990) *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge: Cambridge University Press.

Miller, Nicholas R. (1980) 'A New Solution Set for Tournaments and Majority Voting: Further Graph-Theoretical Approaches to the Theory of Voting', *American Journal of Political Science* 24 (1): 68–96.

Moulin, H. (1983) *The Strategy of Social Choice*, Amsterdam: North-Holland.

Pettit, Philip and Robert Sugden (1989) 'The Backward Induction Paradox', *The Journal of Philosophy* 86 (4): 169–182.

Pollak, R. A. (1968) 'Consistent Planning', *The Review of Economic Studies* 35 (2): 201–208.

Pratt, John W., Howard Raiffa, and Robert Schlaifer (1965) *Introduction to Statistical Decision Theory*, New York: McGraw-Hill, preliminary edn.

Rabinowicz, Wlodek (1995) 'To Have One's Cake and Eat It, Too: Sequential Choice and Expected-Utility Violations', *The Journal of Philosophy* 92 (11): 586–620.

— (1998) 'Grappling with the Centipede: Defence of Backward Induction for BI-Terminating Games', *Economics and Philosophy* 14 (1): 95–126.

— (2000) 'Money Pump with Foresight', in Michael J. Almeida, ed., *Imperceptible Harms and Benefits*, pp. 123–154, Dordrecht: Kluwer.

— (2008) 'Pragmatic Arguments for Rationality Constraints', in Maria Carla Galavotti, Roberto Scazzieri, and Patrick Suppes, eds., *Reasoning, Rationality, and Probability*, pp. 139–163, Stanford, CA: CSLI.

Samuelson, Paul Anthony (1947) *Foundations of Economic Analysis*, Cambridge, MA: Harvard University Press.

Schwartz, Thomas (1990) 'Cyclic Tournaments and Cooperative Majority Voting: A Solution', *Social Choice and Welfare* 7 (1): 19–29.

Sen, Amartya K. (1977) 'Social Choice Theory: A Re-Examination', *Econometrica* 45 (1): 53–89.

Uzawa, Hirofumi (1956) 'Note on Preference and Axioms of Choice', *Annals of the Institute of Statistical Mathematics* 8 (1): 35–40.

von Neumann, John and Oskar Morgenstern (1944) *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.