

# Kenyon College

## Digital Kenyon: Research, Scholarship, and Creative Exchange

---

Faculty Publications

Philosophy

---

7-1994

### Expected utility and constrained maximization: Problems of compatibility

Hans Lottenbach  
lottenbachh@kenyon.edu

Follow this and additional works at: [https://digital.kenyon.edu/philosophy\\_publications](https://digital.kenyon.edu/philosophy_publications)



Part of the [Philosophy Commons](#)

---

#### Recommended Citation

"Expected Utility and Constrained Maximization: Problems of Compatibility," *Erkenntnis* 41 (1994), 37-48.

This Article is brought to you for free and open access by the Philosophy at Digital Kenyon: Research, Scholarship, and Creative Exchange. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Digital Kenyon: Research, Scholarship, and Creative Exchange. For more information, please contact [noltj@kenyon.edu](mailto:noltj@kenyon.edu).

## EXPECTED UTILITY AND CONSTRAINED MAXIMIZATION: PROBLEMS OF COMPATIBILITY

**ABSTRACT.** In recent attempts at deriving morality from rationality expected utility theory has played a major role. In the most prominent such attempt, Gauthier's *Morals by Agreement*, a mode of maximizing utility called *constrained maximization* is defended. I want to show that constrained maximization or any similar proposal cannot be coherently supported by expected utility theory. First, I point to an important implication of that theory. Second, I discuss the question of what the place of constrained maximization in utility theory might be. Third, I argue that no matter how we answer this question, expected utility theory cannot provide the reason why a moral disposition like constrained maximization is to be preferred to its rivals.

Expected utility theory has recently figured prominently in moral philosophy. Contractarian moral theories in particular have been formulated in the language of utility maximization. The most important work in this vein is David Gauthier's *Morals by Agreement*.<sup>1</sup> Gauthier defends there a mode of maximizing one's utility he calls *constrained maximization*. His main project is to justify the constraints of morality as rational constraints on utility maximization.

Critics of *Morals by Agreement* have raised two main types of objections to Gauthier's account of utility maximization as constrained maximization: According to the first line of criticism, constrained maximization involves an unintelligible notion of choosing a disposition to choose.<sup>2</sup> In particular, it has been argued that since the disposition of constrained maximization requires a commitment to refrain from some future maximizing choices, it cannot be accommodated by utility theory.<sup>3</sup> In the second type of objections critics accept the notion of a choice of dispositions, but claim that constrained maximization is not the utility maximizing disposition because one could do even better by acting on other dispositions.<sup>4</sup>

I want to show here that even if the notion of choosing a disposition is made intelligible, constrained maximization (or any such proposal for a maximizing disposition) cannot be defended by expected utility theory. First, I point to an important, but often overlooked, implication of that theory. Second, I discuss the question of what the place of

constrained maximization in utility theory might be. Third, I argue that no matter how we answer this question, expected utility theory cannot provide the reason why constrained maximization is to be preferred to its rivals.

## 1.

Gauthier characterizes a constrained maximizer as follows:

- (i) someone who is conditionally disposed to base her actions on a joint strategy or practice should the utility she expects were everyone so to base his actions be no less than what she would expect were everyone to employ individual strategies;
- (ii) someone who actually acts on this conditional disposition should her expected utility be greater than what she would expect were everyone to employ individual strategies.<sup>5</sup>

Gauthier's thesis is that rationality demands constrained and not straightforward maximization (i.e. maximization over one's individual actions). The core of chapter VI of *Morals by Agreement* is an argument intended to establish that constrained maximizers do better in terms of utility than straightforward maximizers. If practical rationality is identified with utility maximization, this would demonstrate the superiority of constrained maximization. I think this argument cannot get off the ground.

A constrained maximizer cooperates in one-shot Prisoner's Dilemmas if he is certain of the other player's cooperation or thinks it is sufficiently probable. Cooperation in interactive situations with the structure of such a Prisoner's Dilemma seems to allow constrained maximizers to reap benefits unavailable to straightforward maximizers. Gauthier develops a model that aims to specify plausibly realistic conditions under which the expected utility of a constrained maximizer is greater than that of the straightforward maximizer.

The utility theory accepted in *Morals by Agreement* is a version (Harsanyi's) of the familiar approach by von Neumann and Morgenstern.<sup>6</sup> This theory uses the following monotonicity axiom:

- (1) For all  $x, y, z$ , and all  $p$  such that  $0 < p \leq 1$ ;  
 $xPy$  iff  $(x, p; z, 1 - p)P(y, p; z, 1 - p)$ .

$P$  stands for the strict preference relation.  $x, y$ , and  $z$  are sure prospects

or outcomes.  $(x, p; z, 1 - p)$  is a risky prospect or lottery that yields outcome  $x$  with probability  $p$  and outcome  $z$  with probability  $1 - p$ . From (1), given the usual definitions of the relations of strict and weak preference, and indifference ( $P$ ,  $R$ , and  $I$ ), we can derive the following simple lemma (the so-called Substitution Principle):

$$(2) \quad xIy \text{ iff } (x, p; z, 1 - p)I(y, p; z, 1 - p)$$

With the help of (2) and under the assumption of commutativity of lotteries<sup>7</sup> it is straightforward to prove that (1) implies<sup>8</sup>

$$(3) \quad \text{If } (xPy \text{ and } zRw) \text{ then } (x, p; z, 1 - p)P(y, p; w, 1 - p).$$

This is a two-component (weak) dominance principle. In a Prisoner's Dilemma  $x$ ,  $y$ ,  $z$ , and  $w$  could stand for the four possible outcomes;  $(x, p; z, 1 - p)$  and  $(y, p; w, 1 - p)$  for the two available choices.

Cooperation in a one-shot Prisoner's Dilemma violates the dominance principle (3). Violating (3) in turn implies violating the monotonicity axiom (1) which is necessary for the very characterization of a maximizer of expected utility. Therefore particular choices of a constrained maximizer in Prisoner's Dilemma-type situations cannot be represented as maximization of expected utility.

## 2.

Gauthier himself points out that

in defending constrained maximization we have implicitly reinterpreted the utility-maximizing conception of rationality. The received interpretation... identifies rationality with utility-maximizing at the level of particular choices. A choice is rational if and only if it maximizes the actor's utility. We identify rationality with utility-maximization at the level of dispositions to choose.<sup>9</sup>

We are told here to compare the expected utilities not of the particular actions but the *dispositions* of constrained and straightforward maximizers. Before addressing how expected utilities of dispositions are to be understood, it is necessary to specify more precisely what the dispositions investigated here are. I will do this here – very roughly – only for cases where agents find themselves in Prisoner's Dilemmas. (In

other circumstances the dispositions of constrained and straightforward maximizers are supposed to be the same.)

She [a constrained maximizer, *H.L.*] must estimate the likelihood that others involved in the prospective practice or interaction will act co-operatively, and calculate, not the utility she would expect were all to co-operate, but the utility she would expect if she cooperates, given her estimate of the degree to which others will co-operate. Only if this exceeds what she would expect from universal non-co-operation, does her conditional disposition to constraint actually manifest itself in a decision to base her actions on the cooperative joint strategy.<sup>10</sup>

Let *C* and *D* stand for the risky prospects “cooperation” and “defection”;  $P_A(s)$  for the given subjective probability of *s* for *A*;  $P_{iA}$  for the probability in the lottery of *A*’s choosing *C* ( $P_{iA}$  for the cooperative outcome,  $1 - P_{iA}$  for the outcome of being exploited) such that *A* is indifferent between this lottery and his lot under universal non-cooperation. *A* has the disposition of a constrained maximizer if and only if he follows the rule

*RCM*: If  $P_A(\text{others' } C) > P_{iA}$ , choose *C*; otherwise *D*.

We can say that *A* is a straightforward maximizer if and only if he follows the rule

*RSM*: Choose the dominant action.<sup>11</sup>

Since in Prisoner’s Dilemmas *D* dominates *C*, *RSM* will, of course, always prescribe *D*.

How are we to ascribe expected utility to dispositions? Taking  $EU_A$  to be *A*’s expected utility it might seem natural to postulate  $EU_A(RSM) = EU_A(D)$ . This would mean, however, that  $EU_A(RSM)$  represents *A*’s choices as conforming to the von Neumann–Morgenstern axioms for particular actions. By the definition of *RCM*  $EU_A(RCM)$  cannot represent von Neumann–Morgenstern behavior. Under the postulate  $EU_A(RSM) = EU_A(D)$ ,<sup>12</sup> this implies that no matter how we now define  $EU_A(RCM)$  we cannot say that *A* maximizes utility by adopting *RCM* rather than *RSM*. We are now no longer comparing utilities within the same interpretation of the theory of utility.

There is an ambiguity in Gauthier’s notion of straightforward maxim-

ization. It appears that in arguing against *RSM* Gauthier is attacking expected utility theory in its received interpretation. What he compares, however, are the utilities of two *dispositions*. The received theory, which – *ex hypothesi* – is not about dispositions, cannot be the target of the argument. The utilities of a straightforward maximizer understood as someone following *RSM* must be distinguished from those of a straightforward maximizer understood as a follower of the received version of utility theory. Expected utilities of dispositions and particular actions are not commensurable.

In order to define utilities for dispositions like *RSM* and *RCM* we could simply apply the structure of von Neumann-Morgenstern theory to dispositions. As we can understand particular actions to be the domain of the theory we can so understand dispositions to choose. It is, however, not immediately obvious why a utility theory for dispositions to choose should be a theory of *expected utility*. The need for a theory of this form can be made clear by the following example. Assume that the only available dispositions are *RSM* and *RCM* and the only relevant states of the world are to find oneself generally among followers of *RSM* or *RCM*. We can designate these two states by MEET *RSM* and MEET *RCM*. We get the following decision matrix:

	MEET <i>RSM</i>	MEET <i>RCM</i>
CHOOSE <i>RSM</i>	<i>x</i>	<i>y</i>
CHOOSE <i>RCM</i>	<i>z</i>	<i>w</i>

A preference ranking over the four possible outcomes might be  $wPyPxPz$ . (Concerning  $wPy$  assume – with Gauthier – that a straightforward maximizer will generally not be able to exploit followers of *RCM*.) Given these parameters of the decision situation we have to take into account – in some way – the probabilities of the states of the world in order to determine the rational choice. This is exactly where expected utility theory does its job. (Simple principles of dominance will not do here.)

Accepting von Neumann–Morgenstern utility theory for dispositions to choose means accepting its axioms (suitably interpreted as characterizing preferences over lotteries with outcomes as results of choices about dispositions). These include the monotonicity axiom. But now, as the theory must be *general*, we can easily specify Prisoner’s Dilemma problems for this version of expected utility theory. Let *RUC* be the

disposition of unconditional cooperation in Prisoner's Dilemmas involving individual actions. We get:

	MEET <i>RUC</i>	MEET <i>RSM</i>
CHOOSE <i>RUC</i>	<i>a</i>	<i>b</i>
CHOOSE <i>RSM</i>	<i>c</i>	<i>d</i>

Assume everybody's preference ranking is *cPaPdPb*. Our theory tells us to choose *RSM* although we all prefer outcome *a* to *d*. Thus it seems that expected utility theory for dispositions to choose faces a problem of the same kind it was supposed to solve.

Gauthier will object that the above Prisoner's Dilemma depends on an unrealistic choice situation: Only two dispositions are available and only two somewhat peculiar states of the world are relevant. Furthermore, he might argue that a Prisoner's Dilemma could not arise if the choice were between *RCM* and *RSM*: It is part of the definition of a constrained maximizer that he will not leave himself open to exploitation. Thus, in our first example of a decision matrix we would not have to face the Prisoner's Dilemma ranking *yPwPxPz*. We could always rely on *wPy*. But this argument depends on defeasible assumptions about the beliefs of the decision-makers. Notice that from the fact that two constrained maximizers meet it does not follow that they will cooperate. If, for example, the first believes that the second believes him to be a straightforward maximizer, he will not cooperate.<sup>13</sup> If agents expect such problems of recognition to be pervasive, they might very well form the ranking *yRwPxPz* or *yPwPxPz*. This could again seem a far-fetched possibility if strategic considerations are taken into account. In repeated interactions, for instance, recognition problems can be expected to vanish. But if we appeal to such long-run considerations, it is no longer clear why we need constrained maximization in order to allow for cooperative behavior and its benefits. The conditions under which we can exclude the Prisoner's Dilemma rankings (i.e. the conditions under which choosing *RCM* is compatible with expected utility theory) seem to be just those under which reciprocally cooperative strategies are likely to be in Nash equilibrium.<sup>14</sup> That a disposition to choose like constrained maximization is preferable to such strategies now needs to be defended. Thus, the argument intended to show that in the choice between *RCM* and *RSM* a Prisoner's Dilemma is not to be expected seems to undermine the rationale for choosing dispositions rather than conditional strategies.<sup>15</sup>

## 3.

Commenting on his argument Gauthier remarks:

Our argument identifies practical rationality with utility-maximization at the level of dispositions to choose, and carries through the implications of that identification in assessing the rationality of particular choices.<sup>16</sup>

The claim here is that if adopting *RCM* maximizes utility, it is rational to make the particular choices prescribed by the rule. For instance, it is rational to cooperate in certain single-shot Prisoner's Dilemmas. This, however, will not convince someone who does not identify practical rationality with utility maximization at the level of dispositions.<sup>17</sup> That the utility of being disposed to act on *RCM* is greater than that of following some non-cooperative rule does not show that utility maximizing ought to be done at this level. The question of why one should not maximize over particular actions (following expected utility theory in its traditional interpretation) is still open.

It is tempting simply to respond that someone who maximizes over dispositions will do better than an old-fashioned maximizer. But better in what terms? Gauthier is committed to explaining this in terms of utility: Maximizing at the level of dispositions is utility maximizing.<sup>18</sup> But in this claim it is not clear what "utility" means. If it refers to the utility of particular actions, the claim is false. In a Prisoner's Dilemma the action prescribed by *RCM* does not maximize expected utility. If "utility" means the utility of dispositions to choose, the claim is true but tautologous.

As measures of preferences the utilities of particular choices and of dispositions to choose cannot be compared. Comparisons would be meaningful only if the choice set for which the preference relation is defined included both particular choices and dispositions. To allow such a set would make the interpretation of expected utility theory very difficult. It does not seem to make sense to postulate that each individual action (understood as a risky prospect or lottery) is either weakly preferred or dispreferred (or both) to each disposition to choose (also understood as a lottery). In the context of decision theory, what could it mean to prefer a disposition to a particular action? Even if there were no such problems of interpretation, we could ask why we should interpret the theory in this way. Once we allow particular actions and

dispositions in the choice set, there seems to be no reason to stop here. We could also include more or less comprehensive life-plans. The relation between such life-plans and dispositions would be similar to that between dispositions and actions. For my purposes here it is enough to point out that the more is included in the choice set, the less convincing becomes Gauthier's own argument for his favored disposition.<sup>19</sup>

It might clarify matters briefly to consider a similar and well-discussed problem in utilitarianism: Should utilitarians adopt an act- or rule-utilitarian decision procedure? In many utilitarian theories utility is not just a measure of subjective preference. In these theories utility measures interpersonally comparable individual happiness or well-being. On the basis of such a concept of utility the question whether utilitarians should try to maximize social utility at the level of particular choices or dispositions or rules to choose has, at least in principle, a straightforward answer: Adopt the decision procedure which in fact maximizes social utility. The best arguments for rule utilitarianism have tried to show that rule- but not act-utilitarians are able to solve various social coordination problems. If these arguments succeed they do so because solving coordination problems increases social utility. In the debate between proponents of act- and rule-utilitarianism both parties argue in terms of the same concept of utility and share the same goal: to establish what maximizes social utility. In expected utility theory, however, the seemingly analogous problem is in fact very different. The argument for the rationality of a disposition like *RCM* uses an interpretation of the theory unlike the traditional one. Therefore, the proponent of maximizing expected utility at the level of dispositions cannot argue that his approach better realizes the goal both he and the traditional maximizer share. Under the particular-choice-interpretation expected utility theory is not self-undermining in the way act-utilitarianism is if the arguments for rule-utilitarianism are sound.

#### 4.

The argument of this paper does not depend on the specific features of constrained maximization. As mentioned above, some critics of *Morals by Agreement* have proposed dispositions to choose that they claim would improve on constrained maximization. My main point applies to these alternative proposals as well. Gauthier himself has recently discussed constraints on maximization in terms of *plans*. His justification

of the rationality of plans that impose constraints on particular choices is summarized by this:

To suppose, as the orthodox position does, that the imposition of such a constraint is impossible, is to view rationality as in some ways a hindrance to, and not an instrument for, *maximizing one's utility*. The orthodox position treats rationality as self-undermining in situations with prisoner's dilemma-type structures. My alternative account of rational feasibility removes this incoherence.<sup>20</sup>

*No* von Neumann–Morgenstern-type utility theory can be coherent without excluding cooperation in Prisoner's Dilemmas. To shift the interpretation of the theory from particular choices to dispositions or plans does not remove an incoherence. Utility theory for dispositions or plans does not improve on utility theory for particular choices *in the latter's own terms*. Furthermore, at the level of decisions over dispositions or plans expected utility maximizers will again face Prisoner's Dilemmas and their theory will prescribe non-cooperation.

Finally, remember that constrained maximization is introduced by Gauthier as a central part in a subjectivist defense of morality. The project is to show the constraints of morality (or at least some of them) to be justified as utility-maximizing. Fundamental to this project are, first, a subjectivist theory of value identifying value with utility and, second, a maximizing conception of rationality. It was my aim to establish that if the utility concept used here is that of *expected utility theory*, this defense of morality is incoherent. To the (already long) list of problems for justificatory projects of this type in ethics it might be added that we can get no help from expected utility theory on the problem of the level at which we ought to maximize. As much as we may wish Prisoner's Dilemmas to go away, once we do adopt a level of maximization of expected utility we cannot be better – or more moral – utility maximizers by violating the axioms that define our very notion of utility.

#### APPENDIX

It seems that the importance of the fact that the monotonicity axiom implies the dominance principle is underappreciated in the philosoph-

ical literature because the proof of it is never made explicit.<sup>21</sup> So here is a proof. Our claim is:

For all  $x, y, z, w$ , and all  $p$  such that  $0 < p \leq 1$ :

If  $\{xPy \text{ iff } (x, p; z, 1 - p)P(y, p; z, 1 - p)\}$  then:

(a) If  $(xPy \text{ and } zIw)$  then  $(x, p; z, 1 - p)P(y, p; w, 1 - p)$ .

(b) If  $(xPy \text{ and } zPw)$  then  $(x, p; z, 1 - p)P(y, p; w, 1 - p)$ .

*Proof.* For proving (a) assume  $xPy$  and  $zIw$ . By hypothesis we have:  $xPy \text{ iff } (x, p; z, 1 - p)P(y, p; z, 1 - p)$ . Thus we get  $(x, p; z, 1 - p)P(y, p; z, 1 - p)$ . By hypothesis we have:  $zPw \text{ iff } (z, 1 - p; y, p)P(w, 1 - p; y, p)$ . By commutativity of lotteries we get:  $zPw \text{ iff } (y, p; z, 1 - p)P(y, p; w, 1 - p)$ . By the substitution principle and commutativity:  $zIw \text{ iff } (y, p; z, 1 - p)I(y, p; w, 1 - p)$ . Thus we obtain  $(y, p; z, 1 - p)I(y, p; w, 1 - p)$ . This gives us  $(x, p; z, 1 - p)P(y, p; z, 1 - p)I(y, p; w, 1 - p)$ . Thus  $(x, p; z, 1 - p)P(y, p; w, 1 - p)$ . This completes the proof. The proof for (b) is similar.<sup>22</sup>

#### NOTES

<sup>1</sup> Gauthier (1986).

<sup>2</sup> See, for example, Nida-Rümelin (1993), p. 56f.

<sup>3</sup> See, for example, McClennen (1988), pp. 104–108; Binmore (1993), pp. 136–140; Smith (1991), pp. 244–249.

<sup>4</sup> See, for example, Buchanan (1988), pp. 81–83; Danielson (1991).

<sup>5</sup> Gauthier (1986), p. 167. I have omitted the clause about narrow compliance. (See also Gauthier (1986), pp. 178–179.) It is not relevant for my argument.

<sup>6</sup> Chapter II of Gauthier (1986) is an informal exposition and defense of expected utility theory. For Harsanyi's formal account see Harsanyi (1977), ch. 3.

<sup>7</sup> Commutativity of lotteries asserts that the order of the outcomes in a lottery does not matter as long as the probability of each prize remains the same. That is,  $(x, p; y, 1 - p) = (y, 1 - p; x, p)$ .

<sup>8</sup> For the proof see Appendix, pp. 45f.

<sup>9</sup> Gauthier (1986), p. 182.

<sup>10</sup> Gauthier (1986), p. 169.

<sup>11</sup> Remember that here we define constrained and straightforward maximization only for situations with the structure of a Prisoner's Dilemma.

<sup>12</sup> This is how Gauthier ascribes expected utility to RSM in his provisional Argument (1) (Gauthier (1986), p. 171). In his considered argument (Gauthier (1986), pp. 175–176) the utilities calculated are what constrained and straightforward maximizers obtain in typical Prisoner's Dilemma-type situations, i.e. the utilities of *particular actions*. My

claim is that this is illegitimate since in these situations the particular actions of constrained maximizers have no (von Neumann–Morgenstern) expected utilities.

<sup>13</sup> For a discussion of how constrained maximizers can fail to cooperate even if they are transparent to each other, see Smith (1991).

<sup>14</sup> On conditions for Nash equilibria for cooperative strategies in repeated Prisoner's Dilemmas see, for instance, Taylor (1987).

<sup>15</sup> Gauthier distinguishes the disposition of constrained maximization from conditionally cooperative strategies like Tit-for-tat. (Gauthier (1986), pp. 169–170, fn. 19.) The general problem touched upon here – too large to be properly addressed in this paper – is the relation between maximizing expected utility and choosing (Nash) equilibrium strategies, or – in Gauthier's terminology – the relation between parametric and strategic choice. Showing that a disposition to choose is utility maximizing is different from showing that a combination of strategies is an equilibrium. At least in games with multiple equilibria it is quite unclear in what sense choosing an equilibrium strategy maximizes expected utility. This could be the reason why Gauthier's argument for the rationality of constrained maximization is parametric: "We may . . . employ the device of parametric choice among dispositions to choose to show that in strategic contexts the disposition to make constrained choices, rather than straightforwardly maximizing choices, is utility maximizing" (Gauthier (1986), p. 183). See also Hegselmann (1989) (p. 155f.), where Gauthier is criticized for representing the choice situation as parametric and not strategic.

<sup>16</sup> Gauthier (1986), p. 187.

<sup>17</sup> It will certainly not convince someone who maintains that "the primary object of choice is the singular action" (Nida-Rümelin (1993), p. 56). If rational choice is only about singular actions, Nida-Rümelin is right in claiming that constrained maximization and maximization of expected utility cannot be reconciled. Nida-Rümelin objects to the application of utility theory to dispositions because he interprets the choice of a disposition as the imposition of a limit on the freedom of choice of a rational person. This would commit Gauthier to "an inadequate theory of moral agency" (Nida-Rümelin (1993), p. 73). Gauthier, however, rejects this interpretation of the choice of a disposition. (See Gauthier (1993), p. 187.)

<sup>18</sup> It is important to keep in mind that as a general subjectivist theory of doing better or worse in circumstances involving risk expected utility theory has few comparably worked-out and elegant competitors. Intuitions about the relative merits of choices not backed by a utility theory must remain unsatisfactory. This, of course, explains the appeal of von Neumann-Morgenstern theory to contractarian moral philosophers.

<sup>19</sup> Remember that in the argument of chapter VI of Gauthier (1986) compares the utilities of only two dispositions.

<sup>20</sup> Gauthier (1990), p. 122 (my emphasis).

<sup>21</sup> To attribute expected utilities to cooperative dispositions like constrained maximization or unconditional cooperation on the basis of the expected utilities of the *particular actions* implied by these dispositions is simply incoherent. Dominated cooperative actions in Prisoner's Dilemma-type situations cannot have (von Neumann-Morgenstern) expected utilities. For a recent instance of this confusion, see Messerly (1992).

<sup>22</sup> Thanks are due to Philip Kremer, two anonymous referees of this journal and – especially – Sergio Tenenbaum for their suggestions on earlier versions of this paper.

## REFERENCES

- Binmore, K.: 1993, 'Bargaining and Morality', in D. Gauthier and R. Sugden (eds.), *Rationality, Justice and the Social Contract*, The University of Michigan Press, Ann Arbor, pp. 131–156.
- Buchanan, J. F.: 1988, 'The Gauthier Enterprise', in E. F. Paul, F. D. Miller, J. Paul, and J. Ahrens (eds.), *The New Social Contract*, Basil Blackwell, Oxford, pp. 75–94.
- Danielson, P.: 1991, 'Closing the Compliance Dilemma: How it's Rational to be Moral in a Lamarckian World', in P. Vallentyne (ed.), *Contractarianism and Rational Choice*, Cambridge University Press, Cambridge, pp. 291–322.
- Gauthier, D.: 1986, *Morals by Agreement*, Clarendon Press, Oxford.
- Gauthier, D.: 1990, 'Economic Man and the Rational Reasoner', in J. H. Nichols and C. Wright (eds.), *From Political Economy to Economics and Back?*, ICS Press, San Francisco, pp. 105–131.
- Gauthier, D.: 1993, 'Uniting Separate Persons', in D. Gauthier and R. Sugden (eds.), *Rationality, Justice and the Social Contract*, The University of Michigan Press, Ann Arbor, pp. 176–192.
- Harsanyi, J. C.: 1977, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, Cambridge.
- Hegselmann, R.: 1989, 'Rational Egoism, Mutual Advantage and Morality – A Review-Discussion of D. Gauthier: *Morals by Agreement*', *Erkenntnis* 31, 143–159.
- McClennen, E. F.: 1988, 'Constrained Maximization and Resolute Choice', in E. F. Paul, F. D. Miller, J. Paul and J. Ahrens (eds.), *The New Social Contract*, Basil Blackwell, Oxford, 95–118.
- Messerly, J.: 1992, 'The Omission of Unconditional Cooperators: A Critique of Gauthier's Argument for Constrained Maximization', *Philosophical Studies* 67, 19–39.
- Nida-Rümelin, J.: 1993, 'Practical Reason, Collective Rationality and Contractarianism', in D. Gauthier and R. Sugden (eds.), *Rationality, Justice and the Social Contract*, The University of Michigan Press, Ann Arbor, pp. 53–74.
- Smith, H.: 1991, 'Deriving Morality from Rationality', in P. Vallentyne, (ed.): *Contractarianism and Rational Choice*, Cambridge University Press, New York, pp. 228–253.
- Taylor, M.: 1987, *The Possibility of Cooperation*, Cambridge University Press, Cambridge.

Manuscript submitted January 25, 1993

Final version received August 9, 1993

5046 Hobart 11  
Pittsburgh, PA 15217  
U.S.A.