*Erasmus Journal for Philosophy and Economics,*
*Volume 13, Issue 2,*
*Winter 2020, pp. 85–97.*
*https://doi.org/10.23941/ejpe.v13i2.502*

# Roemer on the Rationality of Cooperation

PETER VALLENTYNE
*University of Missouri*

In *How We Cooperate: A Theory of Kantian Optimization*, John Roemer (2019), a philosophically informed and highly influential normative economist, builds upon some previous work and defends a Kantian, and hence unorthodox, theory of the rationality of cooperation. More exactly, he addresses the question of when it is rational to cooperate with other agents in *non-cooperative games*. These are games with no possibility of first negotiating an agreement that is externally binding (for example, with externally imposed penalties that are large enough to make it irrational to fail to comply with the agreement).

I shall first summarize the philosophical core of Roemer's project. Then I shall raise some concerns about it.

## I. AN OVERVIEW OF KANTIAN OPTIMIZATION

Throughout, like Roemer (I believe), I focus on game theory as a normative theory of rational choice in the context of interacting agents, rather than as a descriptive (predictive) theory of such choice.

I follow Roemer and make the standard assumption of *idealized* game theory that the following are common knowledge (that is, they are true, each player knows it, each player knows that each player knows it, etc.): (1) that each agent is perfectly rational, (2) what choices each player confronts, and (3) what the payoffs are for each player given the choices for every player. These conditions do not hold in real life: agents are not perfectly rational. They suffer from deductive incompleteness, inconsistency, confusion, and weakness of the will. Moreover, what one agent knows is typically not known by all agents, let alone common knowledge to all. *Realistic* game theory is based on more realistic assumptions. The problem of rational choice in non-cooperative games is difficult, even in the idealized case, and so focusing on it makes sense.

**AUTHOR'S NOTE:** For helpful comments, I am grateful to Akshath Jitendranath, Luc Lauwers, John Roemer, Bertil Tungodden, Alex Voorhoeve, and Paul Weirich.

| | | COLUMN | |
|---|---|---|---|
| | | SILENT | CONFESS |
| ROW | SILENT | (8, 8) | (1, 9) |
| | CONFESS | (9, 1) | (3, 3) |

**Table 1**: The Prisoner's Dilemma.

Following Roemer, we shall assume that the outcome payoffs in a game represent, for each agent, the *prudential value* (what makes her life go well for her), on some arbitrary scale, of the outcome for that agent.

The standard view of non-cooperative games is that rationality requires, when possible, agents to *Nash optimize,* which means making a choice (adopting a strategy) that is part of a *Nash equilibrium*—a choice *n*-tuple (one choice for each player) such that each player's choice is a best response to the choices of the other players.

To illustrate this, consider, for example, the Prisoner's Dilemma game in Table 1, where (1, 9) designates the payoffs of 1 to the row player (*Row*) and of 9 to the column player (*Column*), respectively.

Here, suppose the police caught two criminals and placed them in isolated cells. Each has two options: to remain silent or to confess. If one confesses and the other does not, then the confessor will go to prison for one year and have a life worth 9 units of value. In this case, the silent agent will be convicted and sent to prison for nine years, and have a life worth 1 unit. If they both remain silent, they will each go to prison for 2 years and have lives worth 8 units of value each. If they both confess, they will each go to prison for 7 years and have lives worth 3 units of value each.

In this Prisoner's Dilemma, the only Nash equilibrium is for each to confess. (*Confess*, *Confess*) is a Nash equilibrium because neither player (on their own) can unilaterally do better by remaining silent. (*Silent*, *Silent*) and (*Silent*, *Confess*) are not Nash equilibria since *Row* would do better by switching to confessing. (*Confess*, *Silent*) (as well as (*Silent*, *Silent*)) is not a Nash equilibrium, since *Column* would do better by switching to confessing. Nash optimizing here requires performing one's choice in the only Nash equilibrium.

Above we focused on a single play of the Prisoner's Dilemma. If the game will be played infinitely many times, or with no known upper bound on the number of times, then rational agents will take into account that

failing to cooperate in one round (by confessing) may have the result that, in the future, other agents won't trust them to cooperate and will thus exclude them from the possibility of cooperation. In such cases, it may be Nash optimal to cooperate, since the long-term benefits of cooperating may exceed the one-round benefits of defecting. If, however, the number of rounds is finite and also common knowledge to the players, then Nash optimization will require defecting on the first round. This is because all know that Nash optimizers will not cooperate on the final round, and so all know that Nash optimizers will not cooperate on the previous round (since there are no last-round benefits), and, by backward induction, all know that Nash optimizers will not cooperate in the first round. So, although indefinite repeated play can open up the possibility of rational cooperation among Nash optimizers, it does not do so when there is a finite upper bound on the number of plays and this is common knowledge. In what follows, we (like Roemer) shall therefore focus on single-play games.

Roemer denies that, in single-play games, rationality precludes cooperation. In the above Prisoner's Dilemma, if each player chooses cooperatively and remains silent, each would have a life worth 8 units, rather than 3 units. How can it be rational for each to confess (as required by Nash optimization) when this leads to an outcome that is worse for each of them compared to each remaining silent?

Roemer's core claim, which he strengthens in various ways, is that, in simple symmetric games, with players one trusts sufficiently to cooperate, rationality requires players to choose their part of a simple Kantian equilibrium when it produces a Pareto optimal outcome.[1] Very crudely, this requires that each agent make a choice that, *if all agents chose in a like manner* (a kind of Kantian universalization), it would have the best consequences for the agent. I will explain this in terms of two-person games.

*Symmetric games* are games in which players are 'identically situated' in the sense that (1) each player has the same possible choices (where choice $a_i$ for player 1 is the 'same choice' as choice $b_i$ for player 2), and (2) the ordinal rank, relative to the feasible outcome, for agent 1 for choice-pair ($a_i$, $b_k$) is the same as the ordinal rank to agent 2 for the choice-pair ($a_k$, $b_i$) (for example, each agent gets their second best feasible outcome).

[1] Initially, I thought that Roemer claimed that rationality requires Kantian optimization (with those one trusts sufficiently to cooperate), whenever the game has a common diagonal (that is, is symmetric when players 'make the same choice'), even if so optimizing is not Pareto optimal. In e-mail correspondence, however, Roemer clarified that his claim is the weaker one made in the text.

The Prisoner's Dilemma displayed above, with staying silent as each agent's first choice, and confessing as each agent's second, is symmetric. For example, (*Confess*, *Silent*) gives payoffs to the two players of (9, 1), and (*Silent*, *Confess*) gives them a payoff of (1, 9).

In a symmetric game, a *simple Kantian equilibrium* is a choice-pair ($a_i$, $b_i$) for which no player gets a greater benefit from any alternative pair ($a_j$, $b_j$). This considers only choice-pairs in which everyone 'does the same thing', understood as playing the *n*-th choice, for some *n*. A simple Kantian equilibrium is a choice-pair in which everyone 'does the same thing' and doing so is at least as good, for each agent, as any other choice-pair in which everyone 'does the same thing'. In the above Prisoner's Dilemma, each player confessing is a choice-pair in which everyone 'does the same thing', as is each player remaining silent. Each player gets a greater benefit from the latter, and hence only that outcome (that is, each remaining silent) is a simple Kantian equilibrium. Kantian optimization thus requires that each remain silent, which gives each 8 units of value, rather than the 3 units that each would receive, if each confessed (as required by Nash optimization).

The definition of symmetric games I gave above assumes that player 1's *n*-th choice can be identified in a non-arbitrary way with player 2's *n*-th choice. This is so, because symmetry requires that, when both players make their *n*-th choice (for any *n*), the outcome has the same ordinal rank for each player relative to the feasible outcomes. In a later section, I will question whether it is plausible to claim that the rationality of a choice depends on such, as I will call it, 'interpersonal identification' of choices. In the present section, however, I will make Roemer's assumption that the specification of a game includes a privileged enumeration of choices for each agent (such that the *n*-th choice of one agent is the same choice as the *n*-th choice of the other agent).

A Nash optimizer (in the two-person case) asks themself "Given the strategy chosen by my opponent, what is the best strategy for me?", whereas the Kantian optimizer asks "[If I trust my opponent] [w]hat is the strategy I would like both of us to play?" (12). That is, the Kantian optimizer, unlike the Nash optimizer, does not treat the choices of the others as given. Roemer makes clear (9, 20) that Kantian optimization requires cooperation only when the agent sufficiently trusts the other(s) to cooperate. Such optimizers are willing to cooperate with those they believe to be cooperators, but not with those they believe to be non-cooperators.

Kantian optimization is not based on an altruistic concern for others (although it is compatible with that). Nor is it based on team reasoning about some collective goal that each player seeks to promote. Instead, it is based on reasoning cooperatively with those disposed to reason cooperatively (39) about how to promote *shared interests* (distinct interests, but ones that can sometimes be jointly promoted).

Roemer establishes that simple Kantian equilibria always exist for symmetric games, and, more generally, for games with a "common diagonal" (23)—a left-to-right downward diagonal (the 'main diagonal') in the respective strategy matrix (that is, choice *n*-tuples where agents make the same choices) such that the ordering of the diagonal entries is the same for all agents. Moreover, if the payoffs for the game are strictly monotone (that is, strictly increasing, or strictly decreasing, in the strategies of others), then simple Kantian equilibria are Pareto optimal (such that no alternative feasible choice *n*-tuple makes some agent better off and no agent worse off).

Where individuals have different preferences over individual payoffs/outcomes, there may be no common diagonal, and simple Kantian equilibria may not exist. Roemer, however, extends Kantian optimization to require, where the payoffs are strictly monotone, either a multiplicative Kantian equilibrium, an additive Kantian equilibrium, or a mixture of the two. (Simple Kantian equilibria are a special case of each.) My critical comments below won't address these; so I won't explain these notions here.

Kantian optimization among people who trust each other to cooperate solves two major problems that confront Nash optimization: (1) the inefficiency of Nash equilibria in the presence of negative externalities (for example, the tragedy of the commons problem), and (2) the inefficiency of Nash equilibria in the presence of positive externalities (for example, the free-rider problem for public goods). In strictly monotonic, symmetric games, Kantian optimization will select the level of negative (or positive) externalities that, if imposed on each player, will leave each as well off as any other universalized level.

In what follows, I will focus on the philosophical foundations of simple Kantian optimization. I shall thus not address the more sophisticated forms of Kantian optimization, nor the many interesting theorems and applications. For simplicity, I focus on two-person games without the possibility of randomization among pure options.

## II. PRACTICAL RATIONALITY (AND ALTRUISM, FAIRNESS, AND MORALITY)

Roemer defends his theory primarily as a normative theory of rational choice (69, 215), although he thinks it has some descriptive/predictive value as well. He sometimes describes his theory as a moral theory (69–70), but it seems clear that he means it to be a theory of practical rationality, where this is the theory of rational choice relative to the values that the agent has (which need not be moral values).

Practical rationality is often thought of as rational choice relative to the agent's *prudential values* (their well-being, narrowly understood), but there is no reason to so limit it. Agents clearly care about their own well-being, but most agents are also partially *altruistic* (care about the well-being of at least some others), care about *cooperative fairness* (for example, the extent to which outcomes approximate the outcome of the negotiation of an externally enforceable agreement), or care about conforming to *moral values* (for example, choosing in morally permissible ways or in morally better ways). To the extent that agents care about such matters (and this may not be rationally required), practical rationality is sensitive to how well a choice promotes them (in addition to the agent's well-being).

Roemer is well aware of this, but, in most of his book, Roemer assumes that agents care only about their own well-being. This is because he believes that *cooperative reasoning*, rather than altruism (or, presumably, valuing cooperative fairness or conformance to moral values), is key to the rationality of cooperation (4). I'm skeptical about this, but I fully agree that any realistic appeal to such valuings will not eliminate the possibility and desirability of cooperative reasoning (since there will still be conflicts between the overall values of different agents).

In most of his book, Roemer assumes that agents care only about their own well-being, but, in chapter 5, Roemer considers the implications of adding some altruism to the values of agents. He there establishes that Kantian equilibria for economies with partially altruistic preferences are observationally equivalent to those in the same economy without altruism.

I agree with Roemer that appealing to altruism, concern for cooperative fairness, or concern for moral values does not eliminate the possibility or desirability of cooperative reasoning.

| | | COLUMN | |
|---|---|---|---|
| | | CHOICE 1 (CONFESS) | CHOICE 2 (SILENT) |
| ROW | CHOICE 1 (SILENT) | (1, 9) | (8, 8) |
| | CHOICE 2 (CONFESS) | (3, 3) | (9, 1) |

**Table 2**: The Prisoner's Dilemma Relabeled.

## III. SOME CRITICISM

I have two criticisms of Kantian optimization. One concerns its presupposition that there is some privileged manner of making *interpersonal identifications of choices* of different agents (different agents performing the same action). The other concerns its general requirement to cooperate with those one trusts sufficiently to cooperate with one.

### *III.I. The Irrelevance of Interpersonal Identifications of Choices*

Consider the reformulation of the Prisoner's Dilemma in Table 2. This is the same as the Prisoner's Dilemma presented earlier (in Table 1), except that (1) the two columns have been permuted, and (2) the labels of the rows and columns have been changed. For standard game theory, these changes in presentation are irrelevant, and it remains true that each confessing is the only Nash equilibrium. For Roemer's theory, however, this relabeling makes a big difference. First, the game is no longer symmetric (because, for example, the payoffs are not the same for both players, when both make choice 1). Second, there is now no simple Kantian equilibrium, since there is no choice such that all weakly prefer everyone making that choice to everyone making some alternative feasible choice (for example, player 1 most prefers all making choice 2, whereas player 2 prefers all making choice 1). Roemer makes no claim about what rationality requires under such conditions.

My objection here is *not* that Roemer's theory is silent in this case. Rather, it is that his theory gives different answers to what rationality requires in the two different ways of presenting the game. In the original presentation, Roemer's theory holds that rationality requires that both agents remain silent (not confess), but, in the second presentation, his theory is silent about what rationality requires. My claim is that they are the same game, and the different presentations change nothing. More exactly, my claim is that Kantian optimization presupposes that there is a privileged way of identifying the choices of different players (for example,

player 1's choice to confess 'is the same choice' as player 2's choice to confess), but such interpersonal identifications of choice are arbitrary and irrelevant to rational choice. Let me explain.

Kantian optimization is well-defined only if each choice of one player can be uniquely identified with a distinct choice of the other player (for example, confess for *Row* is the same choice as confess for *Column*) in a way that is relevant to rational choice. This permits the choices of the two agents to be listed in the same order in the matrix of the game. Thus, the downward common diagonal has the players 'making the same choice' (their *n*-th choice for each *n*). This is why, in the above Prisoner's Dilemma Relabeled, *Row* and *Column* each choosing silent is not a Kantian equilibrium. They are making 'different choices'. *Row* is choosing her first enumerated choice, whereas *Column* is choosing his second enumerated choice. By assumption, a choice by *Row* is 'the same' as a choice by *Column* if and only if they are both the agent's *n*-th choice, for some *n*.

Roemer is quite aware of this issue. He explicitly points out (26–28) how a common diagonal, and hence a simple Kantian equilibrium, may exist for one interpersonal identification of the choices of different agents (one 'common choice space'), but not for another.

This interpersonal identification of choices seems completely arbitrary. Of course, we can identify one person's confessing with another person's confessing, but we can also identify it with the other person remaining silent (for example, if each person is doing what they most want to do, or if each person is doing what their mother told them to do). The question is whether such identifications have any normative significance for rational choice. Nash optimization, like most approaches to rational choice, does not require any specification of which choice by one player 'is the same choice' as that of the other. Roemer's framework, however, assumes that there is a privileged way of doing this (as given by the 'common choice space' from which agents choose).

My point is not that, although there are facts about which choice of one agent is *the same choice* as another agent's, rational choice is insensitive to these facts. It is rather that there are no such (interest-invariant) facts. My pushing the red button and your pushing the red button may be the same choice *relative* to our choices of what colored button to push, but they may be different choices with respect to what shape the button we push is (mine is square and yours is round), with respect to our intentions (for example, I push the button to save my mother's life, whereas you push it in order to kill your father), with respect to the anticipated

consequences, etc. I am not, that is, claiming that facts of whether agents make the same choices are substantively irrelevant (as a permutation invariance condition might require). I am claiming that there are no such facts. As far as I know, no other theory of rational choice appeals to such facts.

For non-simple games, which I have not introduced, Roemer further assumes that the strategies have a 'natural' order, based on the 'effort' involved (for example, labor time). One can indeed hold a *moral* view according to which equal effort should receive equal reward, but I see little reason to think that rationality in non-cooperative games requires this, or that it requires the interpersonal identification of choices with the same effort levels. Indeed, in simple normal-form games (such as those discussed above), effort plays no role, and I see no reason to think that it does in non-simple games in which efforts (of some specified sort) are commonly known. This, however, is a complex issue, and I here set it aside.[2]

Let me propose, without endorsing, a friendly but radical revision to the formulation of Kantian optimization that avoids this problem. Let us note that simple Kantian optimization does not depend on any interpersonal comparisons of value, nor on any cardinal (that is, interval) measurability of the values of choices. It merely requires that, in symmetric games, each agent make a choice that, if all other agents make the same choice, maximizes the payoff for the agent and is Pareto optimal. Thus, Roemer's idea of 'same payoff for the same choice' is that of 'same intrapersonal ordinal rank for same choice'. For example, in the Prisoner's Dilemma (the first version in Table 1, where choice 1 for each agent is staying silent), the four outcomes are (8, 8), (1, 9), (9, 1), and (3, 3), and this is ordinally and non-comparably equivalent to respective outcomes of (3, 3), (1, 4), (4, 1), and (2, 2). For each agent, the ordering of the four outcomes is the same for both sets of numbers. Kantian optimization requires a choice that gives each player their best outcome on the common diagonal (that is, (3, 3) rather than (2, 2)), and that requires each to choose to stay silent. So, Roemer's symmetry condition does not (as Roemer knows) require that players get the same cardinal interpersonally comparable payoff when they make the same choice. It only requires that players

[2] Even for simple games, Roemer assumes that the choices of each agent have a 'natural' order. Although his definitions of simple Kantian equilibrium, symmetry, and Pareto optimality do not depend on this order, his definition of monotonicity does. Given, however, that monotonicity is relevant only as a sufficient condition for Pareto optimality, no natural order is required for simple Kantian optimization.

get the same (non-comparable, intrapersonal) ordinal rank for the same choice (for example, that when all make the same choice, they each get their second best payoffs).

Once, however, we drop the idea of a privileged interpersonal identification of choices, the appeal to the diagonal becomes arbitrary (since the diagonal depends on arbitrary identifications of choices). Instead, we should look at *all possible outcomes*, in terms of each individual's ordinal ranking, focus on those outcomes that, for example, give players the same intrapersonal ordinality (that is, all get their *n*-th best outcome), and require that each do their part in producing the outcome that makes that ordinality as good as possible. (For simplicity, I here ignore how ties in ordinality are handled.) For example, with respect to the intrapersonal ordinality of each outcome, the Prisoner's Dilemma Relabeled game (Table 2) is identical to the original Prisoner's Dilemma game (Table 1). In both cases, the only two outcomes with equal ordinality for the two agents are (*Silent*, *Silent*) with payoffs (3, 3), the second-best outcome for each, and (*Confess*, *Confess*) with payoffs (2, 2), the third-best outcome for each. The revised approach would thus require that each agent be silent. This agrees with Roemer's version where confessing for *Row* is the same action as confessing for *Column*, but, unlike Roemer's version, this result holds no matter how choices are interpersonally identified. Roemer's version, by contrast is silent when confessing for *Row* is the same action as staying silent for *Column* (since there is no common diagonal).

The approach just sketched requires *equality* of intrapersonal ordinal payoff, but sometimes equality will not be possible. Moreover, even when possible, it may be possible to make both players better off than they are under the most ordinally equal outcome. So, a more plausible approach is probably to require *leximinning the intrapersonal ordinal payoff* of outcomes (making the lowest ordinal payoff as high as possible, and then doing the same for the second lowest ordinal ranking, etc.). So, if the outcomes are (1, 1), (2, 2), (2, 4), and (4, 3), this revised version of cooperative optimizing would require each to do their part in bringing about (4, 3). I leave open how things are assessed when there is more than one outcome that leximins intrapersonal ordinal payoff (as in (1, 1), (2, 2), (3, 4), and (4, 3)).[3]

[3] If one allows, as Roemer does not, intrapersonal *cardinality* to be relevant to cooperation, cooperation might require leximinning *intrapersonal relative benefit*, defined, for a given agent, as $(X - Min)/(Max - Min)$, where $X$ is the payoff for the agent for a given outcome, and $Min$ and $Max$ are, respectively, the smallest and the greatest payoffs for the agent in the choice situation. For example, if the five possible outcomes are (0, 100),

Call this (leximin) approach *ordinal cooperation*. Like Roemer's simple Kantian optimization, cooperation is to be understood as conditional on sufficiently trusting the other to cooperate appropriately. Unlike Roemer's simple Kantian optimization, it is insensitive to how choices are interpersonally identified.

Call a joint strategy *an ordinally cooperative equilibrium* just in case it leximins the ordinal rank of the outcome. We can now note the following trivial results, which correspond to Roemer's main results for simple Kantian optimization (23):

(1) Ordinally cooperative equilibria exist for all games (whereas Kantian equilibria do not).
(2) Ordinally cooperative equilibria are always Pareto efficient (whereas Kantian equilibria are not).

Of course, I here leave many key issues unresolved. My key claims are simply that (1) Roemer's appeal to the notion of 'same choice' is arbitrary and irrelevant to rational choice, and (2) one can preserve some of the ideas of Kantian optimization (although perhaps not much) by focusing on something like leximinning ordinal benefit. I don't claim that this is a plausible approach. I merely claim that it is more plausible than Roemer's version of Kantian optimization. Indeed, I shall now suggest that the form of cooperative reasoning in ordinal cooperation (and in Kantian optimization) is too strong. The role of cooperative reasoning, I shall suggest, is limited to selecting among Nash equilibria.

### *III.II. Does Rationality Require (or Even Allow) Cooperation with Cooperators?*

Roemer claims that, for symmetric, single-stage games, where cooperation is Pareto optimal, and one's trust that the other (or others) will cooperate is sufficiently high, rationality requires that one cooperate. I have no new argument against this view. I will simply state a fairly standard objection. It applies not just against Kantian optimization, but also to all theories of rationality that require cooperation with trusted cooperators (such as ordinal cooperation above).

---

(1, 95), (2, 90), (3, 80), and (4, 0), instead of requiring each to do their part to produce the third-best outcome for each (namely, (2, 90)), this approach would require each to do their part to choose (3, 80), since this leximins intrapersonal relative benefits, (0.75, 0.8)—where the five relative-benefit pairs are (0, 1), (0.25, 0.95), (0.5, 0.9), (0.75, 0.8), and (1, 0).

| | | COLUMN | | |
|---|---|---|---|---|
| | | C1 | C2 | C3 |
| | R1 | (7, 5) | (3, 2) | (2, 2) |
| ROW | R2 | (4, 3) | (6, 6) | (2, 3) |
| | R3 | (4, 4) | (3, 4) | (5, 7) |

**Table 3**: A game with three equilibria.

Suppose that we have a symmetric, single-stage, two-person Prisoner's Dilemma (and thus cooperation is Pareto optimal). Suppose that it is common knowledge that in this situation each agent is highly disposed to cooperate with the other. For example, suppose that it is common knowledge that there is only an arbitrarily small (infinitesimal, if you like) probability that a player will not cooperate with the other under the given conditions. Each agent is thus virtually (but not perfectly) certain that the other will cooperate. What does rational choice require? Given that, for each player, non-cooperation has higher payoffs, *no matter what the other player chooses*, it seems clear to me that rational choice requires non-cooperation, even though each agent is virtually (or even absolutely) certain that the other will cooperate.

This is not to say that cooperative reasoning is irrelevant to single-play games. Where there is more than one Nash equilibrium, I find it quite plausible that some form of cooperative reasoning is rationally required to choose among the Nash equilibria. The point is rather that rationality, I claim, limits the role of cooperative reasoning, if any, to the selection of Nash equilibria.

To illustrate this, consider, for example, the game in Table 3.

Here, the three (pure) Nash equilibria are (*R1*, *C1*), (*R2*, *C2*), and (*R3*, *C3*). From these three possibilities, only (*R2*, *C2*) leximins intrapersonal ordinal rank (second choice for each agent), and thus a restricted version of ordinal cooperation selects only this choice-pair as rational.[4] This seems reasonably plausible to me.[5] Of course, it's controversial that cooperative reasoning plays even this weak role in rational choice in non-

[4] Ordinal cooperation applied to select a Nash equilibrium will always select a joint strategy that is Pareto optimal *relative to the set of Nash equilibria*, but the selected joint strategy need not be Pareto optimal relative to the entire feasible set. For example, in the Prisoner's Dilemma, there is only one Nash equilibrium, and it is not Pareto optimal.
[5] In the example, none of the three Nash equilibria risk dominates any other. A more restrictive account of the role of cooperation might restrict it to risk undominated Nash

cooperative games. My claim is only that, even if it does, it does not play the strong role defended by Roemer.

In sum, although I agree with Roemer that cooperative reasoning is sometimes relevant to rational choice in non-cooperative games, I claim that: (1) the relevant cooperative reasoning does not depend on any privileged interpersonal identification of choices, and (2) the role of cooperative reasoning, if any, is limited to selecting among Nash equilibria. Obviously, these are big issues that warrant further analysis.

## IV. CONCLUSION

*How We Cooperate* is well-written, philosophically informed, and informative. My discussion has addressed only the core idea of the book, which is the focus of only a small part of the book (roughly 40 of the 218 pages). There are also many important extensions (with production functions, etc.), and applications. Although many of these technical aspects won't be of interest to many philosophers, the presentation typically includes useful accompanying discussion. Those with strong interests in rational choice theory will definitely profit from reading it.

## REFERENCES

Harsanyi, John C., and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: The MIT Press.

Roemer, John E. 2019. *How We Cooperate: A Theory of Kantian Optimization*. New Haven, CT: Yale University Press.

**Peter Vallentyne**, Florence G. Kline Professor of Philosophy at the University of Missouri, has held visiting professorships at Xiamen University, China, and at Central European University, Hungary. He has published over 100 articles in ethics and political philosophy, and edited or co-edited 16 volumes in those areas. He has been associate editor of the *Journal of the American Philosophical Association*, *Economics & Philosophy*, *Social Choice and Welfare*, *Ethics*, and *Politics, Philosophy, and Economics*.

Contact e-mail: <vallentynep@missouri.edu>

equilibria. For a definition of risk dominance, see, for example, Harsanyi and Selten (1988, 82).