Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 112-120.
http://ejpe.org/pdf/5-1-br-5.pdf

Review of Mark D. White's *Kantian ethics and economics: autonomy, dignity, and character*. Stanford University Press, 2011, 288pp.

NICOLAS GRAVEL
Aix-Marseille University

This interesting book presents itself as an introduction to economists of the moral philosophy of Immanuel Kant. I believe however that this self-presentation is not very reflective of the actual content of the book. Indeed, the book spends most of its 190 pages (excluding notes and references) in criticizing the ethical foundations of normative economics from a Kantian perspective. The presentation of Kant's moral philosophy *sensu stricto*, while present, serves mainly as a reference point from which the author develops his critique of the ethical foundations of economics. Hence, this book is more a Kantian critique of the ethical foundations of modern mainstream economics than an introduction to Kantian moral philosophy.

The ethical foundations of economics that are the object of the "Kantian" critique of the book are, essentially, those generally known under the heading of "welfarism"—using the term coined by Amartya Sen (1979)—and which White refers to as "utilitarianism". Welfarism is a rather coherent approach to social ethics. It asserts that the distribution of individual welfare levels is the *only* information that is relevant to evaluating the "ethical goodness" of social arrangements. Individual attributes such as rights, freedom, discrimination, and the like, are only important for welfarism insofar as they affect individuals' welfare. Similarly, welfarism attaches no intrinsic importance to the procedure by which a particular distribution of individual welfare levels is brought about. Welfarism is thus a *consequentialist* approach to social ethics, and this feature is at the very heart of the critique developed in the book:

> In this chapter—and, more broadly, this entire book—I propose to construct an economic model of decision-making based on nonconsequentialist ethics, specifically the moral theory of Immanuel Kant, in which the nature of actions themselves, rather than their consequences, determines their moral worth (p. 16).

Given the strong anti-welfarist stance taken by the author, it is somewhat surprising that the book makes no reference to the vast non-welfarist literature that has developed in the last forty years or so in economics and philosophy. I am referring here especially to the literature on equality of opportunity and responsibility (see, e.g., Rawls 1971; Arneson 1989; Arneson 1990; Dworkin 1981a; Dworkin 1981b; Roemer 1998; and the survey Fleurbaey 2008); on freedom of choice (see, e.g., Sen 1988; Sen 1990; Sen 1991; Pattanaik and Xu 1990; Arrow 1995; Sugden 1998; and the survey Barberà, et al. 2004); or on the modelling of individual rights (see, e.g., Sen 1970; Gibbard 1974; Gärdenfors 1981; Deb, et al. 1997; or Peleg 1998). The absence of reference to this literature is all the more surprising as several contributors to it—including Rawls himself (1971) but also Kolm (1997) and Van Hees (2003)—have made explicit reference to Kantian moral philosophy in developing their own non-welfarist approach to social ethics.

Be that as it may, the reference point from which the author develops his critique of the ethical foundations of mainstream economics is a quite sophisticated (if not abstract) view of Kant's conception of morality. This view, presented in the first chapter of the book, follows Kant's writings closely. It is founded on the notion of the autonomy of the human person, described as her "capacity [...] to make choices according to laws that she sets for herself, without undue influence from either external pressures or internal desires" (p. 19). According to White, autonomy—which distinguishes humans from animals or mere "biological machines" (p. 84)—imbues persons with 'dignity': an "unconditional and incomparable worth" that as such "demands respect from all persons" (p. 21). The autonomous agent imposes on herself a moral law or a maxim. In order to be moral, such a maxim must satisfy the "universalization" test provided by Kant's first formulation of the categorical imperative: "Act only according to that maxim whereby you can at the same time will that it should become a universal law". The maxim must also satisfy the second formulation of Kant's categorical imperative—which White seems to prefer: "act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means" (p. 26). Duties are maxims adopted by an autonomous agent which satisfy the two versions of the categorical imperative. A perfect duty "permits no exception in the interest of inclination" while an imperfect duty is one which "the agent has some latitude in executing, both in degree and method" (p. 29). For instance the well-known—in Kantian writings—maxim "do not lie" is a perfect duty, while the maxim "do not be indifferent to others" is an imperfect one. White recognizes the possibility that various duties (perfect or imperfect) may contradict each other. For instance, the duty "do not be indifferent to your friends" may conflict with the duty "do not lie" for a person whose lie (say to a possible murderer of her friend) may save the life of this friend. White calls the ability of a person to solve the "problem of conflicting obligations or rules" 'judgement' (p. 34), but does not provide much insight as to how the exercise of this judgement may work in practical situations.

I mentioned that I find this view of individual morality rather abstract. Are there really such moral rules "which permit no exception in the interest of inclination"? I cannot think of any. Even the most basic—and widely discussed—Kantian duty "do not lie" seems to suffer several exceptions. After all, who can seriously maintain that it was the "duty" of partisans arrested by the Nazis in World War II to reveal truthfully to their enemies everything they knew about their comrades? If duties permit exceptions, perhaps in order to make them consistent with other duties, this means that moral maxims cannot be unconditional. The important question therefore becomes identifying the circumstances that could allow a moral individual to deviate from his/her maxim. I cannot see how this question can be answered once and for all, independently from the circumstances and the "inclinations" of the agent who is supposed to adopt the moral behaviour. For instance, in the specific context of prisoners' dilemma games, discussed in chapter 1, several authors, including myself (see, for instance, Bilodeau and Gravel 2004), have suggested that an interpretation of the first formulation of the categorical imperative would impose cooperative behaviour on moral agents. Indeed, any agent placed in an interactive context in which all agents have preferences like those that give rise to prisoners' dilemma situations would prefer that the maxim "cooperate", rather than the maxim "defect", becomes a universal law. But clearly this particular use of the universalization test depends upon the inclinations of the participants of the interaction, as would, I believe, most ethical rules of conduct.

The second chapter of the book discusses the difficulty that Kantian persons may encounter in practical situations when the maxim resulting from the exercise of their judgement, in the sense just given to this word, conflicts with their spontaneous inclination or preference. The author insists that the conflict between the aggregate duty resulting from the exercise of judgement and that person's inclination is of a different nature than the conflict between ethical duties solved by judgement. The latter conflict is managed by the autonomous person out of rational deliberation. The former conflict is solved every day by any person who must confront his or her spontaneous inclination with the general 'duty' that she imposes on herself.

White discusses at length the various forms of "weakness of will" that a person may experience in her day to day life, and sketches a "probabilistic" model of weakness of will. In this model, an individual is depicted as having two preferences: one resulting from duty, and one resulting from inclination. In any choice situation, the individual will choose her inclination-most preferred option with a certain probability p, and her duty-most preferred option with the probability $1\text{-}p$. In White's view, the probability p can be seen as a measure of the weakness of the will of the individual. The author calls "character" the ability of a person to solve the conflict between the pursuit of spontaneous inclination and the call of duty. The author illustrates the conflict by discussing at length the phenomenon of procrastination.

While the issue discussed in this chapter—that of a conflict between a "superior" goal decided in an *ex ante* situation by the individual and the spontaneous inclination that may, at any single moment, distract the individual from the pursuit of her goal—is interesting and important, it is somewhat surprising that the author discusses it without any reference to the economic literature on the subject. After all, what the author is describing in this chapter is akin to a problem of time inconsistency. A plan of action—say to eat vegetarian food—is judged optimal once decided initially (perhaps out of a thoughtful Kantian judgement) but, when appetite preceding lunch time manifests itself, the initially recommended plan of action no longer seems appealing, especially in view (or in smell) of the hamburgers served in a neighbouring restaurant. Economic theory has provided several models of this situation that could have been, I believe, usefully connected to the discussion of this chapter. For example, Gul and Pesendorfer's

(2001) article on temptation and self-control seems particularly relevant to this issue.

The third chapter discusses some issues concerning the "social embeddedness" of the ideal Kantian person depicted in the two preceding chapters. It argues that the Kantian approach is founded on an atomistic view of the individual whose "character is self-created and maintained by her choices and actions" (p. 88). The chapter, whose main ideas are essentially those developed in the excellent book by Christine M. Korsgaard (2009), also discusses various views of individual identity, including those of John B. Davis (2003). The picture that one gets from reading this chapter is that the ideal Kantian person is indeed minimally socially embedded. The only "duty" that Kantian morality imposes on a socially embedded person is respect for the dignity of other persons. But it is unclear what this "respect" is supposed to mean in practical situations of real (social) life. White argues, repeatedly, that the working of the market—voluntary transactions between free private owners of resources—satisfies this requirement of mutual respect for individual dignity. White even writes that "the market represents a kingdom of ends which is limited but nonetheless complete within its scope—ensuring the maximal freedom *from interference* consistent with the same freedom for all" (p. 117). But in order to be convincing on this, White should have discussed examples where free exchange between owners may not appear at first glance to be spontaneously compatible with mutual respect. A good example of this is prostitution. Is the person who satisfies their sexual desire by using the body of a prostitute really treating the prostitute "as an end and not simply as a means to an end?" It would have been interesting to read the author's thoughts on this.

The last two chapters of the book present the author's rather violent critique of the economic approach to law and, more generally, normative economics. The starting point of the critique is an unashamed economic libertarianism according to which only minimal intervention by the state is compatible with the Kantian ideal of the "kingdom of the ends". I must confess that I do not find the author very convincing on this front.

Consider for instance his discussion of (negative) externalities, based on a distinction between "wrongful" externalities and those which are not "wrongful". I find the distinction rather blurred. The author for instance considers "crime" to be a "wrongful" activity, while the action of driving a car is not "wrongful". It is hard to understand the basis of this distinction. A city like London, for instance, has imposed significant congestion pricing on private vehicles to protect the inhabitants of the city against the harms of air pollution (among other things). Hence, at peak hours, driving a car can indeed be "wrongful", perhaps even more "wrongful" than some other activities that are considered criminal in the U.S. (e.g., low-level pick pocketing, marijuana dealing, and so on). What is the criterion that would make one consider that selling a joint of marijuana is more "wrongful" than driving a car at peak hours in central London? The author does not provide any clue about what this criterion might be, or how a Kantian ethics can help one in identifying it.

White frequently criticizes the economic approach to crime and regulation for advocating the "management" or "optimization" of the level of criminal activity instead of its "elimination" (p. 127). But again, I do not see how one can avoid "management" of the human activities that are harmful to others, be they considered "criminal" or not. The amount of material and human resources that would be needed to "eliminate" homicides—assuming it were even possible—is just too great to make it a worthwhile objective. There seems therefore to be no alternative than to reduce the number of homicides to a level that is deemed appropriate, given the resources that we wish to devote to this endeavour. As the Nobel laureate James Mirrlees wrote more than twenty year ago: "A good way of governing is to agree upon objectives, discover what is possible, and optimize" (Mirrlees 1986). It is hard to see how one can oppose this basic wisdom.

The author also spends some time discussing the Kaldor-Hicks potential efficiency criteria, as well as the book by Kaplow and Shavell (2002). The discussion of Kaldor-Hicks is rather surprising, as most economists have been convinced that these criteria—proposed in the forties, and since shown to suffer from severe logical inconsistencies—are flawed (see, e.g., Bossert 1996; or Gravel 2001). It is therefore rather misplaced to base one's critique of normative economics on the Kaldor-Hicks criteria, or on the defence of welfarism provided by Kaplow and Shavell (2002).

The last chapter is concerned with challenging the ethical plausibility of the Pareto principle, according to which unanimous consent by the persons concerned is a sufficient condition for recommending a policy or a reform. Here again I found the argument very difficult to follow. The author emphasizes on many occasions that the working of the market instantiates something close to the Kantian ideal of the kingdom of ends. Yet a distinctive feature of the working of the market is that it respects the Pareto principle: if there is a transaction between two parties, it is because both parties consent to it. In this sense, markets bring about what economists call "Pareto-improvements". Given this, it is very difficult to understand the author's arguments against Pareto improvements, some of which are very surprising. For instance:

> At the same time, persons who are not affected by the policy must be denied any right of consent, for policy decisions based on their input would again violate the dignity of those who are truly affected by the policy and deserve the right of consent (p. 179).

This sentence is strange. Of course people who are not affected by a policy are not affected by it. There is therefore no need to "deny them any right of consent". Sentences like this abound in the chapter, and give to the reader the impression that the author's artillery barrage against the Pareto principle largely stems from a basic misunderstanding.

Overall, the main criticism I would make of this book is of the gap and disconnection that seems to exist between the interesting and reasonably deep philosophical account of Kantian morality provided in the three first chapters and the rather confusing and dogmatic criticism of normative economics that is developed in the two last chapters. As I wrote above, a better starting point for criticizing the consequentialist and welfarist underpinnings of mainstream normative economics would have incorporated the extensive non-welfarist literature.

REFERENCES

Arneson, Richard J. 1989. Equality and equal opportunity for welfare. *Philosophical Studies*, 56 (1): 77-93.

Arneson, Richard J. 1990. Liberalism, distributive subjectivism, and equal opportunity for welfare. *Philosophy & Public Affairs*, 19 (2): 158-194.

Arrow, Kenneth J. 1995. A note on freedom and flexibility. In *Choice, welfare, and development: a festschrift in honour of Amartya K. Sen*, eds. Kaushik Basu, Prasanta K. Pattanaik, and Kōtarō Suzumura. Oxford: Oxford University Press, 7-16.

Barberà, Salvador, Walter Bossert, and Prasanta K. Pattanaik. 2004. Ranking sets of objects. In *Handbook of utility theory, vol. 2: extensions*, eds. Salvador Barberà, Peter J. Hammond, and Christian Seidl. Dordrecht: Kluwer, 893-977.

Bilodeau, Marc, and Nicolas Gravel. 2004. Voluntary provision of a public good and individual morality. *Journal of Public Economics*, 88 (3-4): 645-666.

Bossert, Walter. 1996. The Kaldor compensation test and rational choice. *Journal of Public Economics*, 59 (2): 265-276.

Davis, John B. 2003. *The theory of individual in economics: identity and values*. London: Routledge.

Deb, Rajat, Prasanta K. Pattanaik, and Laura Razzolini. 1997. Game forms, rights and the efficiency of social outcomes. *Journal of Economic Theory*, 72 (1): 74-95.

Dworkin, Ronald. 1981a. What is equality? Part 1: equality of welfare. *Philosophy & Public Affairs*, 10 (3): 185-246.

Dworkin, Ronald. 1981b. What is equality? Part 2: equality of resources. *Philosophy & Public Affairs*, 10 (4): 283-345.

Fleurbaey, Marc. 2008. *Fairness, responsibility, and welfare*. London: Oxford University Press.

Gärdenfors, Peter. 1981. Rights, games, and social choice. *Noûs*, 15 (3): 341-356.

Gibbard, Allan. 1974. A Pareto-consistent liberal claim. *Journal of Economic Theory*, 7 (4): 388-410.

Gravel, Nicolas. 2001. On the difficulty of combining actual and potential criteria for an increase in social welfare. *Economic Theory*, 17 (1): 163-180.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. Temptation and self-control. *Econometrica*, 69 (6): 1403-1435.

Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus welfare.* Cambridge (MA): Harvard University Press.

Kolm, Serge-Christophe. 1997. The values of liberty. In *Freedom in economics: new perspective in normative analysis*, eds. Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy. London: Routledge, 25-46.

Korsgaard, Christine M. 2009. *Self-constitution: agency, identity and integrity.* Oxford: Oxford University Press.

Mirrlees, James A. 1986. The theory of optimal income taxation. In *Handbook of mathematical economics*, eds. Kenneth J. Arrow, and Michael D. Intriligator. Amsterdam: North Holland, 1197-1249.

Pattanaik, Prasanta K., and Yongsheng Xu. 1990. On ranking opportunity sets in terms of freedom of choice. *Recherches Économiques de Louvain/Louvain Economic Review*, 56 (3-4): 383-390.

Peleg, Bezalel. 1998. Effectivity functions, games and rights. In *Freedom in economics: new perspective in normative analysis*, eds. Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy. London: Routledge, 116-132.

Rawls, John. 1971. *A theory of justice.* Cambridge (MA): Berknap Press of Harvard University Press.

Roemer, John E. 1998. *Equality of opportunity.* Cambridge (MA): Harvard University Press.

Sen, Amartya K. 1970. The impossibility of a Paretian liberal. *Journal of Political Economy*, 78 (1): 152-157.

Sen, Amartya K. 1979. Utilitarianism and welfarism. *Journal of Philosophy*, 76 (9): 463-489.

Sen, Amartya K. 1988. Freedom of choice: concept and content. *European Economic Review*, 32 (2-3): 269-294.

Sen, Amartya K. 1990. Welfare, freedom, and social choice: a reply. *Recherches Économiques de Louvain/Louvain Economic Review*, 56 (3-4): 451-485.

Sen, Amartya K. 1991. Welfare, preferences, and freedom. *Journal of Econometrics*, 50 (1-2): 15-29.

Sugden, Robert. 1998. The metric of opportunity. *Economics and Philosophy*, 14 (2): 307-337.

Van Hees, Martin. 2003. Acting autonomously versus not acting heteronomously. *Theory and Decision*, 54 (4): 337-355.

Nicolas Gravel is professor of economics at Aix-Marseille University and director of the research unit GREQAM. His research interests include normative and public economics and the intersection of moral theory and economics.

Contact e-mail: <nicolas.gravel@univ-amu.fr>