

PHD THESIS SUMMARY:
Reconciling normative and behavioural economics

GUILHEM LECOUTEUX

PhD in economics, May 2015

École Polytechnique

Economic analysis is traditionally separated into two different branches: positive economics seeks to explain and describe how economic agents behave, while normative economics aims at evaluating economic outcomes, policies, and institutions. Both approaches attribute to economic agents coherent preferences, i.e., preferences that are consistent and context-independent. Individual behaviour is described by assuming that agents behave as if seeking to satisfy their preferences, and economic outcomes are desirable to the extent that individual preferences are satisfied. Conventional normative economics therefore takes the satisfaction of individual preferences as the normative criterion: although this approach seems perfectly justifiable when individual preferences are coherent, it is less obvious that this criterion remains valid when individual preferences are incoherent, as suggested by the experimental findings of behavioural economics (it is indeed unclear that it is desirable to give to someone what she prefers *now*, knowing that she is likely to change her mind in other circumstances). The problem of how to reconcile normative and behavioural economics (the ‘reconciliation problem’, McQuillin and Sugden 2012) is that economists need to develop an alternative normative criterion to the standard preference satisfaction paradigm. The main approach suggested up to now—*behavioural welfare economics* (henceforth ‘BWE’)—consists in assuming that the revealed preferences of the agents, when they are incoherent, are a combination of *true preferences* (the preferences the individual would have revealed, were she perfectly rational) and *errors*. The satisfaction of the underlying true preferences is then taken as the normative criterion (see for instance Thaler and Sunstein’s (2008) *libertarian paternalism* and the use of *nudges*).

I argue in the first part of my thesis that this notion of true preferences is psychologically and philosophically problematic:

behavioural welfare economics indeed represents the individual as if it was an ‘inner rational agent trapped in an outer psychological shell’ (Infante, et al. 2015a, 2015b), since psychology is only conceived as a process that is likely to interfere with a latent ‘rational’ mode of reasoning. This mode of reasoning should be able to generate a unique coherent preference ordering (the agent’s *true preferences*), if the agent’s reasoning was not influenced by irrelevant psychological factors (chapter 1). I then trace the historical origins of the notion of true preferences, so as to understand why many behavioural economists accepted this notion as the primitive of their models. I argue that the model of the inner rational agent was already implicit in neoclassical economics from Pareto on, although it was historically designed as a model of representative agent, and therefore could not be used to infer any conclusion about individual welfare (chapter 2). I then highlight the methodological continuity between Pareto’s work and BWE: BWE indeed models the individual as the combination of a *homo economicus* and a *homo psychologicus*, but shares Pareto’s methodological difficulties in defining unambiguously the preferences of the *homo economicus* (chapter 3). I illustrate the methodological and philosophical difficulties of this reductionist approach in the case of intertemporal choices (Lecouteux 2015). I conclude the first part by questioning the ethical claim of BWE, according to which what matters is the satisfaction of one’s true preferences, i.e., of the preferences of one’s *homo economicus*. I suggest that the reason why economists intend to give to people what they ‘truly want’ is the result of the third-person perspective they endorse when providing normative assessments. This perspective may however offer a biased diagnosis of the normative issues faced by boundedly rational individuals: rather than focusing on apparent errors of reasoning, I indeed argue that the normative challenge raised by behavioural economics is that our behaviour may be influenced by reasons we are either not aware of (e.g., framing effects) or that we do not accept (e.g., addiction). Behavioural anomalies may therefore matter since our *autonomy* can be violated: I therefore defend a normative criterion in terms of individual autonomy, according to which it is the ability to choose and accept one’s preferences that matters (chapter 4).

The object of the second part of my thesis is then to provide the basis of a model of preferences that does not rely on a primitive in terms of ‘true preferences’. Rather than considering that the individuals progressively *discover* some latent coherent preferences, I argue that the

individuals progressively *form* their own preferences. I show that strategic interactions characterised by a strategic substitutability give an incentive to the players to adopt more aggressive preferences, while games with strategic complementarities tend to generate cooperative preferences. This result highlights the potential impact of public policies on individual preferences: by changing the strategic nature of a game, the government may also change the preferences of the individuals (chapter 5). The general model of preferences I develop in chapter 6 relies on Bacharach's (2006) variable frame theory and on the theory of team reasoning. I assume that the set of states of the world cannot be unambiguously described, i.e., that there are several equally valid perceptions of the same state of the world. I show that individuals are likely to adopt what Bacharach calls a 'we-frame', i.e., to consider themselves as the members of a group and to be actuated by the group objective. Given their initial individualistic perception of the game, players can therefore choose their *intentions*. Players strategizing with a we-frame are actuated by the collective intention of satisfying the interest of each member of the team, and are therefore *team reasoning*. I argue that collective intentionality can be formally represented as a choice of collective preferences among team reasoners: I show that team reasoners are likely to choose aggressive preferences with outsiders in games with strategic substitutes, while they tend to adopt cooperative preferences with outsiders in games with strategic complementarities. By identifying themselves as the members of a specific group, the individuals build their own preferences and identity through the choice of the preferences of their group: incoherent and non-selfish preferences are therefore not deviations from some 'true' underlying preferences, but the evolving preferences of a socially-embedded agent. I finally show that team reasoning can be interpreted as an ecologically rational heuristic, and more specifically that team reasoning almost systematically outperforms payoff maximising behaviours (chapter 7).

The main conclusion of this thesis is that behavioural findings do not tell us that humans are 'faulty econs', and that it falls to the agents—and not to normative economists—to decide whether the incoherence of their preferences matter. Normative economists should instead focus on the process of preference formation, and provide policy recommendations so as to promote individual autonomy, i.e., so as to ensure that the agents have the opportunity to form their own preferences.

REFERENCES

- Bacharach, Michael. 2006. *Beyond individual choice: teams and frames in game theory*. Natalie Gold, and Robert Sugden (eds.). Princeton (NJ): Princeton University Press.
- Infante, Gerardo, Guilhem Lecouteux, and Robert Sugden. 2015a. Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. Forthcoming in *Journal of Economic Methodology*.
- Infante, Gerardo, Guilhem Lecouteux, and Robert Sugden. 2015b. On the econ within: a reply to Daniel Hausman. Forthcoming in *Journal of Economic Methodology*.
- Lecouteux, Guilhem. 2015. In search of lost nudges. *The Review of Philosophy and Psychology*, 6 (3): 397-408.
- McQuillin, Ben, and Robert Sugden. 2012. Reconciling normative and behavioural economics: the problems to be solved. *Social Choice and Welfare*, 38 (4): 553-567.
- Thaler, Richard, and Cass Sunstein. 2008. *Nudge: improving decisions about health, wealth, and happiness*. New Haven (CT): Yale University Press.

Guilhem Lecouteux obtained a PhD in economics from École Polytechnique in 2015, after having graduated from the École Normale Supérieure de Cachan in economics and management. He is currently a teaching assistant at the University of Bristol. His primary research interest is on the problem of how to reconcile normative and behavioural economics, i.e., the normative implications of behavioural findings such as preference incoherence. His research combines game theory, economic methodology, history of economic thought, and philosophy.

Contact e-mail: <guilhem.lecouteux@bristol.ac.uk>