

# Homo economicus

(entry prepared for the Handbook of Economics & Ethics, Edited by Peil, Jan and Irene Van Staveren, Edward Elgar Publishing, May 2009)

CARLOS RODRIGUEZ-SICKERT

Pontificia Universidad Católica de Chile  
FACEA, Escuela de Administración

---

*Homo economicus* or *economic man* is an individual that acts so as to maximize his well-being given the constraints he faces. *Homo economicus* is the prevalent model of human behavior among economists, and has also permeated other social sciences through so called rational choice theory.

## Historical Antecedents

The birth of *homo economicus* as we understand the concept nowadays can be found in the work of John Stuart Mill. The concept was developed in his *Essays on some Unsettled Questions of Political Economy* (1848) and full-fledged in his *Principles of Political Economy* (1848). In the *Essays*, he wrote

[Political economy] does not treat at of the whole of man's nature as modified by the social state, nor of the whole conduct of man in society. It is concerned with him solely as a being who desires to possess wealth, and who is capable of judging of the comparative efficacy of means for obtaining that end. It predicts only such of the phenomena of the social state as take place in consequence of the pursuit of wealth. It makes entire abstraction of every other human passion or motive; except those which may be regarded as perpetually antagonizing principles to the desire of wealth, namely, aversion to labour, and desire of the present enjoyment of costly indulgences [John Stuart Mill, 1844, Essay V, Ch. 3]

Mill's specification embodies a commitment to methodological individualism (ie the basic unit of analysis is the individual and not the social system) as well as to a particular abstraction in relation to human nature. The latter embodies:

- 1) instrumental rationality; and
- 2) material self-interest.

Although *homo economicus* was created by Mill, the term was coined by Mill's adversaries of the historical school (Persky, 1995), so from the outset the term carried a pejorative connotation. Mill's adversaries objected both moral lowliness of his nature and the reductionist character of Mill's approach and the amoral character of the assumed model of human nature. Regarding its reductionist character, J.N. Keynes accused Mill "of mistaking a part for the whole, and imagining political economy to end as well begin with mere abstractions" (J.N. Keynes, 1891, Ch. I).

Regarding the second criticism, J.K. Ingram caricatured *Homo economicus* by demoting him from the genus *Homo* and declared it a "money-making animal" (Ingram 1888, Ch. 6).

Methodological individualism and the assumption of a selfish human nature can be traced back to the so called "selfish" school, whose most important representatives were Hobbes and Mandeville. The social implications of such tendency in human beings radically differed in these authors, though. For Hobbes, as he argued on his *Leviathan* (1651), selfish individuals in the absence of an entity which monopolizes power would be stuck in a war of all against all.

In Mandeville's *Fable of the Bees* (1705), on the other hand, he argues that self-love can produce desirable outcomes. In this sense, Mandeville

becomes an antecedent of Adam Smith's *Inquiry into the Nature and Causes of the Wealth of Nations* (1776) where out of the orientation of the individual courses of action on self-interest coordination becomes an emergent property at the societal level.

It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest [Adam Smith, 1776, Book I, Ch. II, Section 1]

Indeed, it is not the novelty (Mill himself situates his work on the *Principles* as a continuation of Smith's treatise) but the abstractness of Mill's *homo economicus* what made it one of the most influential paradigms in modern social sciences.<sup>1</sup> It allowed the formalization of economics half a century later, though the inevitable compromise on generality gave grounds for lasting criticisms.

## **The Model**

In the late 19th century, Marginalists like Jevons, Walras and Menger formalized the ideas of Mill into a set of axioms. Axiomatization guaranteed the internal coherence of economic assumptions and allowed the use of mathematics to deduce testable implications from those assumptions. Formally (as presented in modern microeconomic textbooks), the agent's choice problem involves: a set of possible actions  $A$  a set of possible states of the world  $S$ ; a set of consequences  $C$ . Under incomplete information, an action's consequence will depend on which state of the world turns out to apply, i.e., for each pair of actions and state of the world there is one consequence. Under complete information, there

---

<sup>1</sup> In Mill's interpretation of Smith's work "self-interest" lost its social aspect, explicitly discussed in the first part of the trilogy on moral philosophy, intended by Smith: *The Theory of Moral Sentiments*. See also section: "Discussion".

is a one-to-one relation between actions and consequences.

As economics developed as a discipline, an increasing emphasis on scarcity as the condition faced by *homo economicus* has prevailed. In his widely accepted definition of the domain of study of the discipline, Lionel Robbins captures this trend.

[Economics is] the science which describes human behavior as a relationship between [given] ends and scarce means which have alternative uses. [Lionel Robbins, 1935, Ch.1].

In the axiomatization outlined above, scarcity is expressed by the finiteness of the space of consequences i.e., the actions the *homo economicus* can choose, do not exhaust the possible experiences in the world he can envisage.

Each individual has a set of preferences  $P$  defined over “consequences”  $C$  and rationality is taken to require that those preferences exhibit the following structure:

- 1) Completeness. All consequences can be ranked in an order of preference, i.e, given any two alternative consequence  $c^0$  and  $c^1 \in C$  the individual either prefers  $c^0$  to  $c^1$  ,  $c^1$  to  $c^0$  , or is indifferent between them
- 2) Transitivity. The order of preference is consistent, i.e., given any three alternative experiences  $c^0$  ,  $c^1$  and  $c^2$  , if  $c^0$  is preferred to  $c^1$  and  $c^1$  to  $c^2$  then  $c^0$  must be preferred to  $c^2$ .

Choice is assumed to be the outcome of rational deliberation. Namely, the decision maker has in mind a preference relation  $P$  on the choice set he

faces and, given any choice problem, he will choose an action which leads to an optimal consequence according to  $P$ .

An order of preferences  $P$  admits a utility representation if there exists a function  $u: C \rightarrow \mathbb{R}$  such that if  $c^0$  is preferred to  $c^1$ , then  $u(c^0) > u(c^1)$  and vice-versa. In an environment characterized by scarcity, an agent will maximize utility subject to the constraints he faces. Under incomplete information, following Von-Neumann & Morgenstern (1944)'s approach, an agent will attach subjective probabilities to each consequence for each possible action and choose the action which maximizes expected utility.

What is essential to the concept of rational choice (both under complete and incomplete information) is that it is consistent, but the model is silent regarding the particular content of the preferences so described. Thus, rational choice theory becomes a theory emptied of almost any substance, in which the concept of rational behavior becomes an unfalsifiable tautology --for every form of behavior one could backward engineer a maximization story. The way in which economists have overcome this problem and thus be able to build specific (and thus falsifiable) predictions is by working on the basis of three auxiliary hypotheses in relation to preferences: stability over time; exogeneity of preferences<sup>2</sup>, (eg independent of the institutional environment in which choice is exercised; An appropriate framework to test the theory is provided by Samuelson (1938)'s revealed preferences axioms.

Although it is not constitutive of the notion of rationality outlaid before, for the neoclassical economist practitioner the *homo economicus* has become *de facto* a selfish creature (as in Mill's original abstraction).. Self-regarding preferences implies that the arguments of the individual's utility function

---

<sup>2</sup>Stigler and Becker homologate preferences to the Rocky Mountains by asserting that both are there, will be there next year, and are the same for all men (Stigler & Becker) .

do not include other individuals' utility levels<sup>3</sup>. Assuming self-regarding preferences and that scarcity is fully reflected in market and shadow prices (for non-market interactions), the task of the adherents to the *homo economicus* model becomes to: "[search] often long and frustratingly, for the subtle forms that prices and income take in explaining differences among men and periods." (Stigler and Becker, 1997: 76)<sup>4</sup>.

On the basis of the commandment stated above, the Homo economicus not only has prevailed within the discipline, but permeated other spheres of social action (e.g., crime, voting behavior, educational choices, etc.).

## Criticism

Both, internal criticisms (anomalies identified by experimental economists) and external criticisms coming from other social sciences have focused on the rationality assumption itself and also on the auxiliary hypothesis (best to distinguish) of self-regardness, stability and exogeneity of preferences. Here we focus on these latter hypotheses. For cognitive anomalies, see Kahneman and Tversky, 1979 on prospect theory; Schelling (1984) on inter-temporal inconsistency; and Simon (1976) on bounded rationality).

### Self-regardingness

The self-regarding assumption has been challenged by experimental economists. Forsythe et al. (1994) show that, in a dictator game in which subjects freely choose how to divide a fixed amount of money between

---

<sup>3</sup> Becker (1981) does not neglect the possibility of other-regarding preferences, but confines them to the family.

<sup>4</sup> This "logic of price" requires an additional assumption on the structure of preferences, namely that the preference relation  $P$  is convex. This assumption is usually interpreted in terms of diminishing marginal rates of substitution: in order to compensate for successive losses of one commodity, increasingly larger amounts of an alternative commodity are required.

themselves and an anonymous receptor, these subjects are willing to give away a significant proportion of the pie. Simple models of unconditional altruism (e.g., Collard, 1978) can explain such behavior<sup>5</sup>.

In collective action settings (e.g., public good games or common pool resources games), agents have expressed more complex forms of non-selfish behavior. Specifically, a significant proportion of agents are willing to adhere to a cooperative norm conditionally (on others' cooperative behavior) even when there are no future incentives involved (Fischbacher, Gächter and Fehr 2004). Furthermore, some agents are willing to spend resources sanctioning uncooperative agents (Ostrom et al., 1992; Fehr and Gächter, 2000, 2002). Bowles and Gintis (2002) refer to this form of behavior --cooperate and sanction those who do not, even if it is against their self-interest-- as strong reciprocity; and highlight its central role within community governance mechanisms. It is noteworthy that although high frequencies of strong reciprocators might generate an environment in which cooperation becomes incentive-compatible for a self-regarding agent, costly sanctioning does not<sup>6</sup>. Hence, cooperation in these settings can only be explained beyond the *homo economicus* paradigm. Among the social preferences models, which account for the behavior of strong reciprocators, one can find those based on distributional preferences (Fehr and Schmidt (1999) and Bolton and Ockenfels (2000)); and those based on intentions (Rabin 2003, Dufwenberg and Kirchsteiger, 2004)<sup>7</sup>.

---

<sup>5</sup> Altruism in this context is an observable feature of individual behavior. Allowing internal benefits to be legitimate disqualifiers of genuine altruistic behavior, through the invocation of unobservable benefits, neglects the possibility of classifying with certainty any kind of apparent altruistic behavior, making the concept of altruism itself useless. In other words, altruism is defined behaviourally.

<sup>6</sup> Boyd et al. (2003) and Bowles et al. (2004) explain how such pro-social norms might have evolved under the environmental conditions, which characterized the early stages of human history.

<sup>7</sup> Within models of distributional preferences, besides their own material payoffs, agents also care about the distribution of payoffs. Alternatively, within intention-based models, people are willing to sacrifice resources in order to sanction/reward those people according to the way in which they evaluate their actions towards them. Falk et al. (2003) report experimental evidence for the operation of both types of inclinations.

Sometimes anything that passes for moral inclinations is referred to as irrational. Amartya Sen (1977) has criticized the tendency to link rationality and self-regarding preferences reminding us that it is consistency and transitivity that define rationality and not self-love. Providing an experimental basis for this argument, Andreoni and Miller (2002) show that altruistic behavior can also be described as rational. In their experimental setting, they study how donating behavior in a dictator game changes when the price of altruism varies. They show that most agents satisfy the Generalized Axiom of Revealed Preferences (GARP)<sup>8</sup>.

Exogeneity:

The exogeneity of preferences allows the possibility that preferences are determined by the social/institutional environment in which the individual is immersed. Cooperative dispositions depend not only on agents' intrinsic dispositions (i.e., on their preferences) but on the structure of incentives they face. External enforcement rules might change the intrinsic moral dispositions of –as well as the behavior- an agent. In this scheme, as Sam Bowles warns:

[if] preferences are affected by the policies or institutional arrangements we study, we can neither accurately predict nor coherently evaluate the likely consequences of new policies or institutions without taking account of preference endogeneity [Sam Bowles, 1998, p. 77]

Changes in the structure of incentives associated with a new enforcement structure might eventually erode agents' moral dispositions in that particular context (see Falk and Kosfeld 2006, Gneezy and Rusticini 2000)

---

<sup>8</sup> Satisfying GARP (developed by Afriat, 1967) is a necessary and sufficient condition for choices to be produced by a rational agent facing a linear budget constraint (Varian, 1982).

or may, conversely, trigger the internalization of a social norm (Rodriguez-Sickert et al. 2007).

Henrich et al (2001) provide consistent experimental evidence on deviations from the *homo economicus* model across fifteen traditional societies. Furthermore, they report important differences in both cooperative and fairness dispositions among the members of these societies. By dismissing any specific historical process which might explain current differences in preferences, the *homo economicus* model itself cannot account for these differences. Boyd and Richerson (2005) claim that variation can be maintained by cultural transmission (in the form of conformist social learning). In this line of research, Henrich & Boyd (2001) and Guzmán et al. (2007) show how conformist transmission and strong reciprocity can coevolve.

## **Discussion**

It is important to stress that the criticisms described above have to be understood in opposition to the prevailing interpretation of the fathers of the model, Mill and Smith, to the extent that these authors were fully aware of the complexity of human nature.

In many of his examples, Mill applies a cultural transmission argument: the notion that more or less rational choices made by one generation predispose the tastes of subsequent generations to reinforce similar choice.

Adam Smith, on the other hand, who is mainly known to economists for defending the virtues of self-love, describes human nature in the *Theory of Moral Sentiments* (1759) as possessing moral dispositions opposed to selfish

behavior<sup>9</sup>. He opens his book with the following quote:

How selfish soever man may be supposed, there are evidently some principles in his nature which interest him in the fortune of others, and render their happiness necessary to him though he derives nothing from it except the pleasure of feeling it. [Smith, 1759, part I, section I, chapter 1].

Furthermore, within Smith's framework, moral behavior plays a key role in the adequate functioning of markets. This synergic interaction between markets and the community's moral structure is at the core of a more complex conception of social cohesion where morality, in addition to the state and the markets, becomes a third vertex of social coordination.

Richard Thaler (2000), a prominent behavioral economist, on the light of the accumulated experimental evidence (?), suggests a convergence route between the future of economic research and the original conception of the model. He optimistically claims that *homo economicus* will evolve into *homo sapiens*, or, more simply put, economics will become more related to human behavior. Indeed, it should be in attention to the trade-off between parsimony and the predictive power of this developing process.

## References

ANDREONI, J. & J. MILLER (2002). "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism", *Econometrica* 70(2):737-53.

---

<sup>9</sup> In the History of Economic Thought the apparent contradiction between the two major pieces of work by Smith is known as *Das Adam Smith Problem*. See Sen (1995) for a discussion of this less well-known side of Smithian Economics and its importance in understanding the way in which decentralised economies function.

- AFRIAT, S.N. (1967). "The Construction of a Utility Function from Expenditure Data", *International Economic Review*, 8, 67-77
- BECKER, G. (1981). "Altruism in the Family and Selfishness in the Market Place", *Economica*, 48(189): 1-15.
- BOLTON, G. AND A. OCKENFELS (2000). "ERC: A Theory of Equity, Reciprocity, and Competition", *American Economic Review* 90(1):166-93.
- BOWLES, S. (1998). "Endogenous preferences: the cultural consequences of markets and other economic institutions". *Journal of Economic Literature*, 36(1): 75-111.
- BOWLES, S., AND H. GINTIS (2002). Social capital and community governance. *Economic Journal*, 112(483): 419-36.
- BOWLES, S., CHOI, J. K., & A. HOPFENSITZ (2003). "The coevolution of individual behaviors and social institutions", *Journal of Theoretical Biology*, 223: 135- 47.
- BOYD, R., GINTIS, H., BOWLES, S., & RICHERSON, P. (2003). "The evolution of altruistic punishment". *Proceedings of the National Academy of Science* 100: 3631-35.
- BOYD, R., & RICHERSON, P. J. (2005). "Solving the puzzle of human cooperation" In S. LEVINSON (Ed.), *Evolution and culture* (105–132), Cambridge, MA, MIT Press.
- DUFWENBERG, M. AND G. KIRCHSTEIGER (2004). "A theory of sequential reciprocity", *Games and Economic Behavior*, 47: 268-298.
- FALK, A., U. FISCHBACHER AND E. FEHR (2003). "On the nature of fair behavior", *Economic Enquiry* 41(1): 20-26.
- FEHR, E., AND S. GAECHTER (2000). "Cooperation and punishment in public good experiments", *American Economic Review* 90(4): 980-94.
- FISCHBACHER, U., S. GAECHTER AND E. FEHR (2004). "Are people conditionally cooperative? Evidence from a public goods experiment", *Economic Letters* 71(3): 397-404.

- COLLARD, D. (1978) *Altruism and economy: a study in non-selfish economics*, Oxford: Martin Robertson.
- FALK, A. & M. KOSFELD (2006). "Distrust: The Hidden Cost of Control", *American Economic Review* 96 (5): 1611-30.
- FEHR, E., & S. GAECHTER (2000). "Cooperation and punishment in public good games", *American Economic Review* 90, 980– 994.
- FEHR, E., & S. GAECHTER. (2002). "Altruistic punishment in humans", *Nature* 415: 137– 140.
- FEHR, E. AND K. SCHMIDT (1999). "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics* 118, 817-868.
- FORSYTHE R., J.L. HOROWITZ, N.E. SAVIN & M. SEFTON (1994) "Fairness in simple bargaining experiments", *Games and Economic Behavior* 7(3): 346- 80.
- GNEEZY, U. & A. RUSTICHINI (2000). "A fine is a price", *Journal of Legal Studies* 29(1):1-17.
- GUZMAN, R., C. RODRIGUEZ-SICKERT & R. ROWTHORN (2007). "When in Rome, do as the Romans: The coevolution of punishment, conformism and cooperation", *Evolution and Human Behavior* 28 (2): 112-117.
- HENRICH, J., & BOYD, R. (2001). "Why people punish defectors: Weak conformist transmission can stabilise costly enforcement of norms in cooperative dilemmas.", *Journal of Theoretical Biology*, 208: 79– 89.
- HENRICH, J., R. BOYD, S. BOWLES, C. CAMERER, E. FEHR, H. GINTIS & R. MCELREATH (2001) "Cooperation, Reciprocity and Punishment in Fifteen Small-scale Societies", *American Economic Review* 91 (2): 73-8.
- HOBBS, T. (1651). *Leviathan*.
- INGRAM, J.K (1888). *A History of Political Economy*
- KAHNEMAN, D., & TVERSKY, A. (1979). "Prospect theory: An analysis of decisions under risk." *Econometrica*, 47, 313-27
- KEYNES, J.K. (1891). *The Scope and Method of Political Economy*

- MANDEVILLE, B. (1705). *Fable of the Bees*
- MILL, J.S. (1844) *Essays on some Unsettled Questions of Political Economy.*
- MILL, J.S.(1848) *Principles of Political Economy with some of their Applications to Social Philosophy*
- VON NEUMANN, J. & O. MORGENSTERN (1944) *Theory of Games and Economic Behavior.*
- PERSKY, J. (1995) "The Ethology of Homo Economicus" *The Journal of Economic Perspectives*, 9(2): 221-31.
- RABIN, M. (1993) "Incorporating Fairness into Game Theory and Economics", *American Economic Review* 83: 1281-1302.
- ROBBINS, L. (1935) *An Essay on the Nature and Significance of Economic Science*, 2nd. ed. Macmillan, London.
- SAMUELSON, P. (1938) "A Note on the Pure Theory of Consumer's Behaviour", *Economica* 1: 61-71
- SEN, A. (1977). "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory," *Philosophy and Public Affairs* 6: 317-44.
- SEN, A. (1995). "Moral Codes and Economic Success" in Brittan, S. & A. HAMLIN (eds.) *Market capitalism and moral values*. Ashgate, Brookfield, 23-34.
- SHELLING, T. (1984). "Self-Command in Practice, in Policy, and in a Theory of Rational Choice," *American Economic Review*, 74(2):1-11.
- SIMON, H.,(1976). *Administrative Behavior*, The Free Press, New York.
- SMITH, A. (1759). *The Theory of Moral Sentiments.*
- SMITH, A. (1776). *Inquiry into the Nature and Causes of the Wealth of Nation*
- STIGLER, G. & G. BECKER (1977) "De Gustibus Non Est Disputandum" *The American Economic Review*, 67(2):76-90.
- OSTROM, E., WALKER, J., & GARDNER, R. (1992). Covenants with and without a sword: self-governance is possible. *The American Political Science Review*, 86, 404–417.

THALER, R. (2000) "From Homo Economicus to Homo Sapiens", *The Journal of Economic Perspectives*, 14(1):133-41.

VARIAN, H. (1982). "The Nonparametric Approach to Demand Analysis" *Econometrica* 50(4): 945-74.