The (Human) Nature of Economic Choice

Eli Spiegelman[*†]

Abstract

The standard view of the economic actor is dismally restrictive. The notorious *Homo economicus* is motivated only by material self-interest. Economists recognize that this restriction is unnecessary from a theoretical point of view, and also empirically untenable. However, research on broadening the vision has so far been piecemeal. This paper describes the nature of choice in theory, proposes the term "normative incentive" to describe these broader interests, and proposes an analytical classification to organize the general literature around the concept of an interaction.

1 Introduction

Economics has a famously impoverished view of human nature. Just what makes the science seem so thoroughly dismal may be the view it takes of the human subjects of its investigations: the notorious *Homo economicus*[1]. In its most primal version, this strange species is portrayed as something of a comic book supervillain: the brilliant sociopath, with infinite intellectual capacity aimed only at satisfaction of his own desires. Although it is nearly universally accepted that this caricature is not just an illegitimate portrait of human experience, but also an unnecessarily simplistic interpretation of the theory which bore it, economics lacks a *coherent* vision of the richer aspects of human nature.

The goal of this paper is an attempt to motivate and sketch the rough outlines of a more unified vision[2]. The point of departure is a surprisingly basic proposition: *When human actors are aware that their decisions are interrelated with those of other human actors, they trade material costs and benefits off against normative incentives.* A normative incentive is a component of the

[*]PhD. Candidate in economics, Université du Québec à Montréal. My research focus on the effect of personal communication on economic choice. Contact: spiegele@vaniercollege.qc.ca

[†]The author thanks Claude Fluet for guidance and support, as well as Raúl López-Pérez and participants at several conferences for informal discussion.

[1]Joseph Persky (1995) has noted that from its inception, this term has been one of criticism. It was apparently originally coined to differentiate the agents of John Stuart Mill's (1836) economic analysis from actual human beings. Nevertheless, it is representative of the heuristic used in "workaday" economics. I will therefore often refer to this H. economicus model as the "standard" model of economics.

[2]This is far too broad a task for the size of this paper. Interested readers may contact the author for an expanded version, which constitutes the introduction to a Ph.D. thesis.

decision-making process that makes some alternative more appealing from a moral, personal or socially constructed perspective. If, in the standard model, human nature is to calculate and carry out the lowest-cost method of getting what you want, then in modest contrast I distinguish what people *want* to do from what they feel they *should* do.

The first goal of the paper is to establish that both of these inclinations have a place within the framework of rational action, dictated by something called a "preference relation." The habit of focusing on *Homo economicus* has generated the illusion that this is difficult, and that should be dispelled. Following that, I will try to elaborate a comprehensive classification that differentiates, I hope usefully, between different kinds of normative incentive. This is illustrated with non-exhaustive examples from the literatures that apply to each classification.

2 Rationality and Ethics

The view of rationality adopted in this paper can be succinctly summarized: A *person* is rational if she makes rational *choices*. A choice is rational if the chooser can *acceptably* answer the question "Why did you make that choice?" Acceptable answers are those which are logically unfalsifiable given the person's beliefs. That is, if the choice is P, then using beliefs as premises[3], one cannot construct an argument with *not*-P as a consequence. One could then ask if the beliefs were rational, by the same procedure. This leads to a chain of rationality. Such a chain may be circular, but usually there will be some non-rational end to it.

Now any meaningful choice, rational or otherwise, requires *mutually exclusive* alternatives from which to choose. *Rational* choice requires a *criterion* or unified standard by which one can discriminate between the mutually exclusive alternatives (Machan, 1995). The economic jargon for this criterion is *preferences*. In order to guarantee some set of rational choices, preferences must satisfy three conditions: they must be complete, transitive and independent of irrelevant. These conditions guarantee some "most preferred" subset of the available options , and it becomes quite direct to define any element of that subset as a rational choice.

While admirable in its generality, however, this theory has almost no descriptive content. Utterly absent is an *explanation* of why any given option is preferred to another. To join the chorus of citations of Milton Friedman:

> "The great Saints of history have served their 'private interest' just as the most money grubbing miser has served his interest. The *private interest* is whatever it is that drives an individual."

This has led many to argue that rationality is *empirically* indeterminate, even vacuous. However, the argument hangs crucially on the assumption that we

[3]Thus a person cannot claim to act rationally, based on premises she does not believe to be true. For instance, an atheist cannot rationally claim that she prays because God might answer her prayers, but a theist can.

don't know others' preferences. This assumption has a certain plausibility in some cases. But if, as Hayek (1943) argues, different people's preferences are in fact broadly similar, then we can, by considering why we *ourselves* would perform a given action, and assessing its rationality as a path towards one of our goals, at the same time form a good impression of its rationality for others. This is supported in psychology by the so-called theory of mind (see (Astington and Dack, 2008). It should be noted that a decent term to describe the broad similarity of preferences would be "human nature". Indeed, the hypothesis that there is such a thing as human nature contains a strong suggestion that this kind of similarity exists.

A closely related issue concerns the relative merit of different preference orderings: which preferences are preferred? This is the domain of ethics: the study of what should be valued, and it is highly relevant to a discussion of normative incentives. The claim that preferences are ultimately based on some arational belief implies, on the one hand, that there is (in general) no *rational* way for any individual to directly compare her own goals with those of another person. Thus, as a matter of *logic* I concur with von Mises that:

> "Applied to the ultimate ends of action, the terms rational and irrational are inappropriate and meaningless. [...] No man is qualified to declare what would make another man happier or less discontented."(Mises, 1996, 21)[4]

On the other hand, this is a distinctly unpromising ethical stance. It suggests, for instance, that the person who has a preference for lies is not comparable to the person who has a preference for honesty on any (rational) ethical grounds. The solution to the dilemma is that in *fact* (as opposed to logic), just as we do have some insight into other people's preferences, so we have some common ideas about what kinds of things *should* be valued. This is to say: although we cannot as individuals judge any preferences rational or irrational, it does not follow that as a society we do not have some over-arching meta-preferences, defined over the actions of each individual. In this case, as a society, we would "prefer" individuals to have preferences which are more in line with these meta-preferences.

This is, indeed, the essence of morality. Moral precepts are behavioral rules which can be logically determined from social (or perhaps universal?) meta-preferences. *All such ethical preferences constitute normative incentives.* However, the reverse need not hold. For instance, evolutionary mechanisms might well favor a taste for conformity or for position within a group. These make one think one "should," for instance, wear a tie to work, simply because others do. There is no particular ethical precept behind this kind of conformity, and yet it represents a normative incentive inasmuch as it is experienced as an "obligation" rather than a desire.

Economists, and those steeped in the mythology of the Invisible Hand, may at this point be objecting that normative incentives cannot explain *why* peo-

[4] This has been called the "pig in muck" doctrine since it implies that some people may find happiness in alternatives that others find abhorrent.

ple feel these obligations. Is it not, they may be thinking, more *fundamentally* true that those who fail to conform are punished in some way, and it is actually fear of that (material) punishment, and not a normative incentive, which drives adherence? This is the argument of the "folk theorem". If interactions are repeated indefinitely, then individually costly, cooperative (in the sense of aggregate welfare-enhancing) actions can be supported as a Nash equilibrium in a game defined purely over material benefits if deviations are punished. In this light, adding "social" preferences to the model seems simply an unparsimonious way of expressing behavior that could just as well be explained with a more "fundamental" model of pure self-interest and strategic reputation building.

Aside from the argument above, that there is nothing in the pure theory of economic choice which suggests that self-interest should be more "fundamental" than social-interest, there are three main reasons why I think the real answer to the economists' rhetorical question must be "no". First, it cannot explain the abundant experimental evidence that, *as long as they are aware that their actions are related to those of another human being*, people behave socially even in anonymous, one-shot encounters, where such strategic reputation building is useless. Second, the folk theorem requires behavior that raises expected benefits. Normative incentives, however, may as well lead to surplus-destroying actions as to surplus-enhancing ones. For instance, social pressure may lead one group of teens to study diligently, while it leads another to smoke and skip class. Third, and not least in my view, the folk theorem denies the introspective validity of normative incentives. It is a fact of human experience that we feel we *ought to do* some things, and *ought not to do* others, despite what we otherwise might prefer. People *do* experience the private and normative incentives distinctly[5].

3 A Framework for Analysis: The moments of an interaction

I will term the basic unit, or context in which choices are made, an *interaction*. This concept is very similar to that of a *game*, as used in economics, but with some structure added. The structure I add divides the payoff-relevant components of the game into three essentially distinct phases, or "moments". First, the individuals in the interaction all have various *intentions* before the interaction begins. Once the interaction has begun, some individuals make choices: they act, and those *actions* constitute the second moment of the interaction. Finally, the choices made by all the acting individuals come together to produce an *outcome* for each individual[6].

[5]This is not to deny that non-normative incentives also bind, and can also contribute to explanations of social behavior. Folk theorem arguments of mutual threat-based détente, information cascades of "rational herding" theories and the heuristics of bounded rationality are all certainly important *additional* considerations.

[6]In standard game theory, all payoff-relevant features are collected in the "payoff" at the end of the game. This has caused much unfortunate confusion, as researchers, perhaps thinking of *Homo economicus*, took this payoff to be approximable by money, and found – much to their surprise – that all of the conditions for rational choice listed above seemed to be violated.

Let us begin with the end: the outcome. An outcome is a vector of material and informational consequences of an interaction. In general, a person may have preferences over all the different possible vectors. *H. economicus* is concerned with just one element of this vector: his own material consequence. By contrast, models of "benevolence", which have a long pedigree (Edgeworth, 1881, 16), allow people to care about others' material payoffs, as well. People may also care about the relative size of different material payoffs, as in inequity aversion models (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999), or the rank-based or positional utility(Clark and Oswald, 1998; Frank, 1985). The importance of informational consequences – which can be interpreted as esteem or stigma – has if anything a better pedigree than benevolence (Marshall, 1985; Velben, 1973), although these were not formalized until somewhat later (Becker, 1974; Schelling, 1978).

All of these can be interpreted as normative incentives. Their distinguishing feature is that they are also *consequentialist* in the strong sense that the process by which a result occurs does not directly matter to the evaluation it receives. However, experimental evidence tends to agree with intuitive experience than such strongly consequentalist models are inherently insufficient to completely predict people's preferences in an interaction. People don't only care *what* happens, they also, for various non-instrumental reasons, care *how* and *for what reason* it happens. That is, they may have preference over changes in the two other moments of an interaction; the acts themselves, and the intentions of the actors leading into the interaction.

Act-based normative incentives can stem from "pure" aversion to the act in question, from the actor's interpretation of other people's *prior* expectations (guilt aversion) or from observers' approval or disapproval of the act itself, in which case, as in reputational concerns, the *posterior beliefs* of other people are the source of the utility. The difference between stigma and disapproval is that for disapproval to be operative, the source of the approval must be the act itself, and not the resultant inference. The admonition to "love the sinner, hate the sin" reflects the difference.

The final moment of the interaction is the intention of the interacting people going in. Informally, there can be no question that intentions matter. To give a few examples:

- Someone cuts ahead of you in line at the cinema. Your reaction will be different, depending on whether you think the person didn't see you, or whether they intentionally ignored you.
- A new acquaintance doesn't return a phone call. Are they busy, or are they avoiding you?
- One of the key requirements in labor negotiations is often that the parties feel they are negotiating "in good faith". This comes down to whether they really intend to find middle ground, or just to push through their preconceived expectations.

A significant difference between intention-based incentives and act- or outcome-based incentives is that while outcomes are observable more or less by definition, and actions may well be observable, unless they are hidden by the actor, intentions are, as a general rule, unobservable, and so will usually have to be inferred. This inference can be effected mostly through observable signs, i.e., through actions or outcomes. As a result, models of intention-based normative incentives must specify how people use observable outcomes and/or actions to infer intentions. Indeed, they must specify exactly what an "intention" is. In general, it seems that a person's intention is closely related to the goal they are trying to attain. In other words, to some extent people's *intentions* may be the same thing as their *preferences*. Following this line of thought, a model of intention-based normative incentives would involve people who intrinsically cared what other people's preferences were (Levine, 1998; Rotemberg, 2008). Notice that this means that the perceived intention is essentially same thing as a reputation, a posterior belief about an unobservable type. The model nevertheless is *not* an outcome-based reputational model, because the player whose reputation is established does not get any direct benefit from it. The reputation provides them with instrumental benefit, because the normative incentive acting on the others (who assess the reputation) leads them to act favorably towards those whom they perceive to have good intentions, and unfavorably towards those whom they perceive to be spiteful.

4 The human nature of economics

The models described above are all normative because they explicitly include considerations of interests that go beyond personal preferences. Outcome-based normative incentives dictate how we "should" respond to different posterior beliefs about payoffs or unobservable types, or what kinds of posterior beliefs we "should" try to instill. Act-based normative incentives dictate how we "should" respond to certain actions, or what actions we "should" take. Finally, intention-based incentives dictate how people "should" think of or act towards each other, and how we "should" respond when they do or don't.

The classification system proposed is very general, encompassing many conceptually different normative incentives. However, I believe that, as the discussion above on the distinction between intention-based reciprocity and reputation shows, it can be useful in clarifying differences between concepts. This is an important task both for theorists who want to build clean models, and for empirical researchers who want to test them. As such it is proposed as an analytical tool rather than as some foundationally true characteristic of the world. There are other classification systems that could also partition normative incentives more or less independently. Indeed, the system here will probably prove much more useful when applied in tandem with others. Those, however, fall outside the scope of this paper.

References

Astington, J. W. and L. A. Dack (2008). Theory of mind. In M. M. Haith and J. B. Benson (Eds.), *Encyclopedia of Infant and Early Childhood Development*, pp. 343–356. Oxford: Academic Press.

Becker, G. (1974). A theory of social interactions. *Journal of Political Economy 82*, 1063–1093.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity and competition. *American Economic Review 90*, 166–193.

Clark, A. E. and A. J. Oswald (1998). Comparison-concave utility and following behavior in social and economic settings. *Journal of Public Economics 70*, 133–155.

Edgeworth, F. Y. (1881). *Mathematical Psychics*. London: Kegan Paul and Co.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness competition and cooperation. *The Quarterly Journal of Economics 114*, 817–868.

Frank, R. (1985). *Choosing the Right Pond: Human behavior and the question for status*. New York: Oxford University Press.

Hayek, F. A. v. (1943). Scientism and the study of society. *Economica 10*, 34–63.

Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics 1*, 593–622.

Machan, T. (1995). Reason in economics versus ethics. *International Journal of Social Economics 22*, 19–38.

Marshall, A. (1985). *Principles of Economics*. Philadelphia: Porcupine Press.

Mill, J. S. (1836). On the definition of political economy, and on the method of investigation proper to it. *London and Westminster Review October*.

Mises, L. v. (1996). *Human Action*. Irvington: Foundation for Economic Education.

Persky, J. (1995). Retrospectives: The ethology of homo economicus. *The Journal of Economic Perspectives 9*, 221–231.

Rotemberg, J. J. (2008). Minimally acceptable altruism. *Journal of Economic Behavior and Organization 66*, 457–476.

Schelling, T. (1978). *Micromotives and Macrobehavior*. New York: Norton.

Velben, T. (1973). *Theory of the Leisure Class: An Economic Study of Institutions*. Boston: Houghton Mifflin.