7-29-2015

Adam Smith: Homo Socialis, Yes; Social Preferences, No; Reciprocity Was to Be Explained

Vernon L. Smith
Chapman University, vsmith@chapman.edu

Adam Smith: Homo Socialis, Yes; Social Preferences, No; Reciprocity Was to Be Explained

Adam Smith: *Homo Socialis*, Yes; Social Preferences, No; Reciprocity Was to Be Explained

Vernon L. Smith

Chapman University, USA, vsmith@chapman.edu

ABSTRACT

I argue that the authors accept too casually the neo-classical framework of thought that was incapable of predicting choices in 2-person and other experimental games in the 1980s and 1990ss. The *ex post hoc* hypothesis that social preference can describe homo socialis reduces inevitably to a rescue of neo-classical economics in which Max-U (own payoff, other payoff) substitutes mechanically for Max-U (own payoff) in our personal groupings. This static procedure unnecessarily and inappropriately robs human conduct of its sociality as a process relationship. The model I articulate was masterfully developed by Adam Smith, which back-predicts the results of these earlier small group experiments, and argues the central importance of context—a finding of experimentalists in their attempt to come to terms with the predictive failures of Max-U (own payoff).

The featured paper by Gintis and Helbing (2015) offers much that I agree with:

- The title suggesting that our species is well-described as "*homo socialis.*"

- Rejection of the neo-classical assumption that across the spectrum of actions humans could be modeled choosing to Max-U(.) defined on own outcomes (payoffs) only (McCloskey, 2010).

- Gene-cultural co-evolution.

- That some form of correlated social equilibrium is a better representation of human development and change than traditional neo-classical qua Nash equilibrium models defined on outcomes.

1   The Concern for Behavioral and Empirical Relevance

There are core issues that I want to develop and that will return me to the classical structure of modeling human sociability:

The hypothesis that social preference can describe *homo socialis* reduces inevitably to a rescue of neo-classical economics in which Max-U (own payoff, other payoff) substitutes for Max-U (own payoff) in our personal groupings. But that approach—lingering in the article's rhetoric – doesn't compute from within, and fails to serve our understanding of the dynamics of sociality which is the primary and praiseworthy purpose of the article.

The other-regardingness of conduct in *homo socialis* does not scale up to "payoffs, such as the welfare of others, environmental integrity, fairness, reciprocity, and conformance with social norms." That is an outcome-centered overreach, too authoritarian to be consistent with *homo socialis* as an exploration and discovery process. Non-conforming forms of disagreement are essential sources of adaptation and change in both individual conduct and norm evolution.

The alternative process model I want to articulate was skillfully developed by Adam Smith (1759, 1982), hereafter TMS; what is astonishing is how relevant that depth is in understanding our *homo socialis* conduct and nature, and in anticipating the experimental discoveries that provides the empirical foundation for Gintis and Helbing (2015).[1]

Neo-classical economics under the egis of Jevons in the 1870s wrote under the long shadow of Bentham, launching an irreconcilable diversion from the path articulated in TMS; that diversion became a path dependent lock-in, and it is difficult to jar economic thinking enough to get out of that myopic rut. It might be thought—as in "as if" arguments—that by retreading the utility function to include the elements articulated by Smith, it can embrace the same implications he derived. However, there are two flaws in this methodology:

(1) Because of the neo-classical outcome-payoff focus, Max-U failed to define a natural on-the-ground treatment of rule-governed human sociality as in TMS, and we were ill-prepared to understand the predictive failures in applying the theory to two person experimental games in the 1980s and 1990s. That formulation was wrong-headed, not lacking in its specification

---

[1]An important question is why, if *TMS* is so refreshingly relevant, it never survived to influence the development of socio-economic thought into the 19th and 20th century. That puzzle is partially addressed in Smith (2015).

of utility, although it was correct to pursue better models of human sociability.

(2) Even if a utilitarian route can be shown to possess technical equivalence with the TMS model, it fails to convey an understanding of the mainsprings of human action. Smith fully recognized that the emergent rules of his *homo socialis* model might well be efficient, but warned that this corollary does not enable us to understand the process, whereby these rules came about: "In every part of the universe we observe means adjusted ... to the ends which they are intended to produce ... But though, in accounting for the operations of bodies, we never fail to distinguish in this manner the efficient from the final cause, in accounting for those of the mind we are very apt to confound these two different things with one another. When by natural principles we are led to advance those ends, which a refined and enlightened reason would recommend to us, we are very apt to impute to that reason, as to their efficient cause, the sentiments and actions by which we advance those ends ..." (TMS, 1759, p. 87).

Once the modeling process is understood, there is time enough to ask in what way it might achieve efficiency.

My comments derive from a sympathetic reading of Gintis and Helbing (2015). My argument is not an exercise in the history of economic thought, although that may be of interest. Neither is it a claim that much in TMS was not discovered by sociologists and social psychologists; indeed TMS was a contribution to these fields that seamlessly connected to economics and is fittingly captured in the term humanomics. Rather, the core of my argument is that TMS offers a systematic treatise modeling human sociality that is entirely modern and predictive in its application to our social milieu.

2   Gratitude, Gratuities and Culture: An Example

My wife and I arrived late at our South Korean hotel. I offered tips to the two baggage handlers when we were shown to our hotel room signaling our recognition of and gratitude for their superb and pleasing service. However, the offer was most graciously declined, and in a manner that made it evident that their expectation was that good service be seen as part of their task; any additional payment was neither required nor appropriate. Their propriety of conduct was a rule of good service in welcoming a valued customer; to accept a tip was to disturb the harmony conveyed in that intentional message. Good service was simply an expression of their gratitude for my business; whereas the rule of conduct that I was following presumed to convey my expression of gratitude by explicitly rewarding good service. You only have to experience

that once to begin reassessing your accustomed rules in the light of local custom. It is an adaptation that concerns human rule learning; what Smith meant in referring to the effect on us of others "which always mark when they enter into, and when they disapprove of his sentiments." (TMS, 1759, p. 110)

The mark can be entirely friendly yet subtly disapproving, with the result that your personal rule accommodates a new context, but you are still the same adaptive social person. (At the opposite extreme, try not tipping in an upscale New York restaurant).[2] There has been no change in whatever Bentham and Jevons said, and what behavioral or experimental economists have been saying is the individual's utility function.

2.1   *If It's Not Modeled by Utility, How Might it be Modeled?*

This event is not well-modeled with a utilitarian social preference function activating an action: Max-U(own payoff, other payoff) in place of Max-U(own payoff). No, the event was about discordance in rule space—a negative rather than a positive resonant correlation between a rule followed by the individual and the rule that constitutes the emergent convention or norm. A rule maps context, inclusive of the available set of outcome payoffs, into an action, but the resulting outcome only has meaning in the context (circumstances) that led to the action and is not separable from the context. Equilibrium, if the concept applies, is in rule space and stems from empathy, but more significantly, from mutual empathy as in TMS. Because it would be another 150 years before the word "empathy" entered the English language, Adam Smith, who thoroughly understood this concept as the foundation of his system, emphasized the role of sympathetic "fellow-feeling" in learning the rules we live by. In TMS, Smith brings this home again and again; the word "fellow-feeling" is used 40 times in articulating the mutual "propriety and fitness" between the rules that individuals follow and the co-evolutionary conventions that emerge and change in the culture.

3   How Experimentalists Got Here

Max-U theory has done tolerably well in games against nature, and in impersonal market experiments confined to perishables connecting people through

---

[2]Error is easy because tipping rules are so variable with their context. You don't tip contract Limo service drivers, because it is included in the bill as an 18% service charge—unless you ask him or her to divert and stop at Starbucks, or mail a package; otherwise, if you offer a tip, he/she will remind you that the tip is prepaid; this avoids offending uninformed customers who find that out later.

I once paid a restaurant bill for six people leaving no tip as I was certain they had a policy of adding a service charge. The waiter followed me into the parking lot! "What was wrong with the service?" "Nothing, it was excellent." Whereupon I learned that there was no service charge, and we settled "accounts."

prices, and supported by (perfect) third party enforcement of property rights; Max-U continues to prosper in these two important environments.[3] However, observe that the second environment reduces to a game against nature. Indeed, static game-theoretic solutions reduce to games against an inert nature by each agent choosing to Max-U subject to the constraint that all others choose best response strategies.

This neo-classical perspective, and its limitations, changed decisively in the 1980s and 1990s as experimentalists were drawn into the exploration of two person games such as ultimatum (Güth *et al.*, 1982) and trust games (Berg *et al.*, 1995; McCabe *et al.*, 1996). Traditional Max-U (own payoff) game theoretic models failed by a large margin to predict subject actions. This was particularly evident with designs intended to push the boundary of that failure in favor of Max-U. In McCabe and Smith (2000), Player 1 can opt out, yielding payoffs $(P1, P2) = (\$10, \$10)$ or pass to Player 2 who chooses between $(P1, P2) = (\$15, \$25)$ and $(P1, P2) = (\$0, \$40)$. We thought the prospect of Player 2 taking all the money would greatly discourage cooperative moves, giving self-love its best shot. With stubborn resistance, people still cooperated far "too much" to rescue Max-U (own payoff); it was a fundamentally different world than had become a habitual feature of the way we thought about decision making. Half of the Player 1s passed to their counterpart Player 2 s and 75% of these did not take all the money, choosing ($15,$25) instead. These were undergraduates, but the only change for graduate students was that $(P1, P2) = (\$15, \$25)$ was chosen with frequency 67%.[4] The undergraduate results were robust under replication: Gillies and Rigdon (2008); Cox and Deck (2005).

In response, two main explanations were pursued experimentally: Social Preference (or Pro-social behavior) and Reciprocity. Both were *ex post* rationalizations of results falsifying the original Max-U (own payoffs) model, and failed to provide an account for the results by considerations outside that framework.[5]

---

[3]In asset trading where durable items can be re-traded and agents can switch roles as buyers or sellers depending on price and price changes equilibrium convergence is much more elusive (Dickhaut *et al.*, 2012). Non-durable consumer goods and services account for about 75 percent of gross private product, and is highly stable. The other 25 percent, especially new home construction, accounts for the persistent re-occurrence of economic instability and recession. Ideological devotees of "the market" need to distinguish the two kinds of markets the good and the sometimes ugly, and to seek better property right rules for the ugly if Max-U is to continue to do service.

[4]However, these results did not extend to economics faculty; in a similar game comparing faculty with undergraduates, Coricelli *et al.* (2000) found that undergraduates cooperated more, took less time in choosing, and made more money; hardly irrational.

[5]TMS provides an *ex ante* explanation of the conduct of Person 1 and Person 2 by separately specifying the motivations of both Person 1 and Person 2. The TMS model derives the (back) predicted results from more fundamental considerations related to more comprehensive *ex ante* accounts of human conduct than simply re-parameterizing utility or re-interpreting the results as reciprocal exchange considerations introduced after learning that an alternative was needed for these cases.

Social preference functions copped out in explaining "social" by inserting a utility function that gave you an answer consistent with the data, while preserving neo-classical optimization. Similarly, those of us following the reciprocity explanation lacked a deeper way of understanding why, imagining somehow that the name "reciprocity" was an explanation. Nevertheless, these explorations led to important experimental findings documenting that context and intentions mattered (McCabe *et al.*, 2003; Falk *et al.*, 2007), and the experimental findings were consistent with the sociology/social-psychological literature cited by Gintis and Helbing (2015). Yet TMS provided an alternative modeling perspective that easily accommodated all these findings—a model not so much forgotten as never found by the generations brought up on neo-classical marginalism, which became the dominant tradition after the 1870s.

4   Humans Are Self-Loving, But That Did Not Mean That Motivation in Smith Was Utilitarian

Adam Smith accepted the Stoic axiom that humans are self-loving—more for self is better, less is worse: "Every man, as the Stoics used to say, is first and principally recommended to his own care; and every man is certainly, in every respect, fitter and abler to take care of himself than of any other person." (TMS, 1759, p. 210) And " ... every animal was by nature recommended to its own care, and was endowed with the principle of self-love, that it might endeavour to preserve, not only its existence, but all the different parts of its nature, in the best and most perfect state of which they were capable." (TMS, 1759, p. 272)

Hence, *for Smith, self-love enables each to have common knowledge of what it means for an action to be beneficent or hurtful to others: the action raises or lowers the payoff to non-satiated others. With that common knowledge* axiom, *homo socialis can read particular circumstances and acquire sociability by using its capacity for empathy to learn other-regarding rules of conduct.*

The decisive failure of Bentham, Jevons and the neo-classical paradigm was to suppose that common knowledge of non-satiation implied that all decision, including our most personal human interactions, stemmed from choice in the pursuit of self-love.

5   Motivation Was Social: Desiring Praise, Praise-Worthiness; Dreading Blame, Blame-Worthiness

Man naturally

"... desires, not only praise, but praise-worthiness; or to be that thing which, though it should be praised by nobody, is, however,

the natural and proper object of praise. He dreads, not only blame, but blame-worthiness; or to be that thing which, though it should be blamed by nobody, is, however, the natural and proper object of blame." (TMS, 1859, p. 113–114)

6 Process in Smith

In order to achieve these desires:

"We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it. If, upon placing ourselves in his situation, we thoroughly enter into all the passions and motives which influenced it, we approve of it, by sympathy with the approbation of this supposed equitable judge. If otherwise, we enter into his disapprobation, and condemn it." (TMS, 1759, p. 110) "This is the only looking-glass by which we can, in some measure, with the eyes of other people, scrutinize the propriety of our own conduct." (TMS, 1759, p. 112)

7 Leads to Self-Restraint and the Principles of Self-Command

"Though it may be true, therefore, that every individual, in his own breast, naturally prefers himself to all mankind, yet he dares not look mankind in the face, and avow that he acts according to this principle ... and that how natural soever it may be to him, it must always appear excessive and extravagant to them ... to them he is but one of the multitude in no respect better than any other in it. If he would act so as that the impartial spectator may enter into the principles of his conduct ... he must ... humble the arrogance of his self-love, and bring it down to something which other men can go along with" (TMS, 1759, p. 83).

We are all non-satiated self-lovers, but we learn early that to act always to Max-U, without regard to the sentiments of others, is not the way to achieve harmony with our neighbors.

8 Sociality Takes the Form of Rule Following

Human sociality is expressed in rule following conduct, where a rule involves choosing an action based on the actor's judgment of its propriety, given its context (or circumstances), where context includes the alternative outcome payoffs in the choice set. Since outcomes have utility value it is easy for

economists to reverse this chain and slip mistakenly into thinking that action implies outcome implies utility. That is not the process in TMS. Rather the actor, and other(s) affected by the action, judge its propriety given the context (including payoffs). And the mapping from context into action is a rule. Praise and praise-worthiness in TMS are elements to be weighted in making the judgment.[6] The rule arises because the actors and others necessarily must invoke knowledge of the context to evaluate the meaning of the outcome. That is why I began this essay with the Korean hotel event. I and the baggage-handlers were following discordant rules. It was not that they disliked money or that I was not happy in giving up money in gratitude for their service; we each knew that money was a good to the other. Rather, their rule and mine were not in equilibrium in rule space, and they signaled their disapproval by not accepting my money in that context.

8.1   How Rules Emerge

"Our continual observations upon the conduct of others, insensibly lead us to form to ourselves certain general rules concerning what is fit and proper either to be done or to be avoided. Some of their actions shock all our natural sentiments. We hear everybody about us express the like detestation against them ... It is thus that the general rules of morality are formed. They are ultimately founded upon experience of what, in particular instances, our moral faculties, our natural sense of merit and propriety, approve, or disapprove of. We do not originally approve or condemn particular actions; because, upon examination, they appear to be agreeable or inconsistent with a certain general rule. The general rule, on the contrary, is formed, by finding from experience, that all actions of a certain kind, or circumstanced in a certain manner, are approved or disapproved of." (TMS, 1759, p. 159)

"Insensibly," that is without being aware, we acquire these rules of propriety in *homo socialis* maturation and occasionally must re-learn the rule protocol (as for tipping) when we change cultures. Rules are "normed", but are not followed without uncertainty or interpretive divergence between action signals sent and received. TMS emphasizes this uncertainty, as well as the fitness between the rules followed by the individual and those evolving in the social environment. Thus, in contemporary language, the rules followed by individuals in choosing

---

[6]In Smith and Wilson (2014, footnote 5) we express an individual's (action | propriety) choice criteria in terms of weighting functions multiplying praise and praise-worthiness as indicator $(0, 1)$ measures. The weights depend on context, $C(m1, m2)$, where $(m1, m2)$ are own and other payoff vectors across the set of alternative actions. The meaning of a chosen action as a signal requires knowledge of the alternatives that might have been chosen but were not. If the weights on the praise/praise-worthiness indicator variables are identically zero, only an additive function of $C(m1, m2)$ remains for self-love to prevail, un-modified by sociability.

actions are subject to error and co-evolutionary processes of change.[7] The individual does not change in moving from Japan to Italy, but uncertainty about whether her interpretation of appropriate action accords with local norms is bound to increase and condition a change in choice depending on experience. Rules are perhaps best seen as social-referential, but self-identity is modified by the social. The "individual" harbors a social mind subject to adaptation!

8.2 Why Contextual Circumstances Matter

"The only consequences for which he can be answerable, or by which he can deserve either approbation or disapprobation of any kind, are those which were someway or other intended, or those which, at least, show some agreeable or disagreeable quality in the intention of the heart, from which he acted. To the intention or affection of the heart, therefore, to the propriety or impropriety, to the beneficence or hurtfulness of the design, all praise or blame, all approbation or disapprobation, of any kind, which can justly be bestowed upon any action, must ultimately belong." (TMS, 1759, p. 93)

9 Smith's Propositions On Beneficence and Justice

Smith uses his model to arrive at three propositions:

1. "Actions of a beneficent tendency, which proceed from proper motives, seem alone to require reward; because such alone are the approved objects of gratitude, or excite the sympathetic gratitude of the spectator . . .

2. Actions of a hurtful tendency, which proceed from improper motives, seem alone to deserve punishment; because such alone are the approved objects of resentment, or excite the sympathetic resentment of the spectator . . .

3. "Beneficence is always free, it cannot be extorted by force, the mere want of it exposes to no punishment; because the mere want of beneficence tends to do no real positive evil." (TMS, 1759, p. 78) "Though the breach rules of justice . . . exposes to punishment, the observance of the rules of that virtue seems scarce to deserve any reward." (TMS, 1759, pp. 80–81)

Some of the implications of these propositions for ultimatum and trust games, for the development of property rights and for understanding Smith (1776) are discussed in Smith and Wilson (2014) and Smith (2013).

---

[7]Sympathy was central to Darwin's non-preference idea that groups composed of individuals with strong emotional ties based on sympathy would have fitness advantages over groups deficient in these ties. Likewise, we have noted above that Adam Smith saw the emergent rules as efficient, but that this derivative property does not explain why people followed them.

10   Summary

Adam Smith's model In TMS is neither about outcomes nor equilibrium in outcomes; rather TMS is about rules; equilibria, if they exist, are in rule space. Non-satiation of individuals is an essential axiom in TMS enabling each individual to identify what is beneficent or hurtful to others; this common knowledge of self-love is used not for individual decision making, but to learn self-restraint and follow rules that are in harmony with those of our neighbors. Rules carry context, including payoffs, into actions motivated by the desire for praise and praise-worthiness (or dread of blame and blame-worthiness). Actions of course determine outcomes, but their meaning and grounds for choice is derived inseparably from their context and this is what makes us *homo socialis*. The underlying model in TMS explains (predicts) trustworthiness and trusting actions (as well as resentful and punishing actions) and, *para passu*, reciprocity in personal exchange. *Homo socialis* has no need for the artifice of utilitarian social preferences.

References

Berg, J., J. Dickhaut, and K. McCabe. 1995. "Trust, Reciprocity, and Social History". *Games and Economic Behavior.* 10(1): 122–142.

Coricelli, G., K. McCabe, and V. L. Smith. 2000. "Theory-of-Mind Mechanism in Personal Exchange". In: *Affective Minds.* Ed. by G. Hatano, N. Okada, and H. Tanabe. Amsterdam: Elsevier. 249–259.

Cox, J. C. and C. A. Deck. 2005. "On the Nature of Reciprocal Motives". *Economic Inquiry.* 43(3): 623–635.

Dickhaut, J., S. Lin, and D. Porter. 2012. "Commodity durability, trader specialization, and market performance". *Proceedings National Academy of Science.* 109(5): 1425–1430.

Falk, A., E. Fehr, and U. Fischbache. 2007. "Testing Theories of Fairness—Intentions Matter". *Games and Economic Behavior.* 62(5): 267–303.

Gillies, A. S. and M. L. Rigdon. 2008. "Epistemic Conditions and Social Preferences in Trust Games". *Munich Personal RePEc Archive*, 19th July.

Gintis, H. and D. Helbing. 2015. "Homo Socialis: An Analytical Core for Sociological Theory". *Review of Behavioral Economics.* 2: 1–59.

Güth, W., R. Schmittberger, and B. Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining". *Journal of Economic Behavior and Organization.* 3(4): 367–388.

McCabe, K., S. Rassenti, and V. L. Smith. 1996. "Game Theory and Reciprocity in Some Extensive Form Experimental Games". *Proceedings of National Academy of Sciences.* 93(23): 13421–13428.

McCabe, K., M. L. Rigdon, and V. L. Smith. 2003. "Positive Reciprocity and Intentions in Trust Games". *Journal of Economic Behavior and Organization.* 52(2): 267–275.

McCabe, K. and V. L. Smith. 2000. "A Comparison of Naïve and Sophisticated Subject Behavior with Game Theoretic Predictions". *Proceedings of the National Academy of Arts and Sciences.* 97(7): 3777–3781.

McCloskey, D. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World.* Chicago: The University of Chicago Press.

Smith, A. 1759, 1982. *Theory of Moral Sentiments.* Indianapolis, IN.: Liberty Fund.

Smith, A. 1776, 1981. *An Inquiry into the Nature and Causes of the Wealth of Nations.* Vol. I & II. Indianapolis, IN.: Liberty Fund.

Smith, V. 2013. "Adam Smith: From Propriety and Sentiments to Property and Wealth". *Forum for Social Economics.* 42(4): 283–297.

Smith, V. 2015. "Humanomics of Adam Smith". In: *Adam Smith Review.* 8. Ed. by Fonna Forman. London: Routledge. 161–173.

Smith, V. and B. Wilson. 2014. "Fair and Impartial Spectators in Experimental Economics". *Review of Behavioral Economics.* 1(1-2).