

Thoughts on innovations in social science for SFI meeting August 15-17, 2003

Colin F. Camerer, Caltech
camerer@hss.caltech.edu

These are some sketchy thoughts about big areas of research in social sciences. I've tried to avoid the usual bias toward one's own work (I am interested in behavioral economics, by the way, and especially using experimental data to craft formal, parsimonious models of "behavioral game theory"). Also I think a central theme (taking an economist-centric view) is that economics is too musclebound on theory and desperately in need of facts, and of fresh constructs which require economists to think about how obviously important constructs—like norms and culture—fit into the economic framework, or don't. Economics students graduate like body builders who have huge muscles, but are unable to move furniture efficiently. Other social scientists know a lot of facts, often have challenging constructs and use a broader range of methods (surveys—which are pooh-pooed in economics unless conducted by the government :); experiments, computational methods to "prove by simulated regularity" properties of systems too complex to prove theorems about).

HumanAgent-based modelling: The study of complex agent-based systems seems at best limited by, and at worse quite handicapped by, the goal of seeing how the most complex system-wide behavior can be derived from interaction among the simplest possible agents. In biology this is probably the right move—assume birds have very simple heuristics and see whether amazing flock movement can be derived from those heuristics among interacting agents. And in economics the get-more-from-less approach has certainly proved challenging and provocative too (e.g., Axtell-Epstein "sugarscape" and the Arthur, Lane et al SFI paper on stock markets).

I think there are two problems with this approach. One is that it is impossible to prove that just because limited agent structure A leads, with interaction among agents using rules A, to complex behavior C, that A is *necessarily* the source or cause of C. That is, proof by simulation is good for showing *sufficiency* of A, but not its *necessity*. How could one show by simulation alone that A, and only A, caused C?

One solution is to use other sources of data to choose the agent rules A. (That is, rather than trying to maximize the gap in complexity between the inputs A and the outputs C, use other data to choose the "right" A.) I call this HumanAgent-based modelling (just as a gentle reminder that the usual agent-based modelling is not deeply rooted in cognitive details). This creates a huge opportunity which, to my (admittedly limited) knowledge, has not been exploited: Study an economy of agents with plausible decision rules (rooted in cognitive or neural detail) who interact. E.g. equip agents with say, memory of typical cognitive structure (limited working memory that holds information which is later transferred to long-term memory with structure), tastes (or the capacity to learn tastes) and scarce attention. Have the agents interact. See what happens. The leading edge of theoretical behavioral economics is moving in this direction but even the simplest models are so complex that to prove theorems about the resulting behavior C requires focussing

on very narrow domains. (I am thinking of the Bernheim-Rangel models of hot-cold cognition in addiction and savings; these beautiful models are great but are restricted in domain and may not spread widely because the modelling skill required is daunting.)

The whole point of the computational approach is to substitute away from the ability of human brains to prove impressive theorems, and toward the ability of computers to simulate numerically very complex systems. So why not unleash the computers' ability on the problem of starting with cognitively-plausible agent rules A and seeing what happens when they interact? A cognitively-enriched version of Sugarscape seems like an obvious first step.

Sociology and economics (identity, norms, networks): Compared to the influence of psychology on economics, the influence of sociology has been minimal. A major obstacle to intellectual gains from trade seems to be the fact that economists and psychologists share a commitment to methodological individualism—the idea that theories of aggregate behavior should be built up from individual-level foundations. Sociologists tend to start from the idea that properties of a group or society are basic and are not necessarily derived from individual-level considerations (Mark Granovetter once expressed puzzlement to me that agents in the economic model are so “undersocialized”. Point taken. At the same time, sociologists tend to perhaps sloppily use terms like “constraint”—as in social constraint—which is perhaps better construed as a taste for conformity which people will not indulge if the price of nonconforming is right.) Overcoming this deep modelling barrier seems like a ripe area for progress. An imperialist view is that the sociological-level constructs *can* be derived from individual-level microfoundations, if somebody just tried harder.

One way to do it is through the concept of identity, which social psychologists are starting to explore (Eldar Shafir has a nice paper on this), as well as economists (Akerlof and Kranton, Q J Econ 200?). For an economist a natural way to include identity is to assume that a person has several identities and each has a corresponding set of preferences. Environmental cues etc. could activate different identities. My colleague Matt Jackson and Roland Fryer have a beautiful paper on “categorical cognition” which is at the edge of this idea. They assume that people have a limited number of categories into which they can place objects (or people) and show that minorities will be lumped together in coarse categories (“girl bands”, “African-American students”). Their idea is important for microfoundations of discrimination in labor markets and surely has many other uses.

Another sociology-economics bridge is the concept of a norm. To a game theorist a norm is a conventional way of playing a game that (likely) has multiple equilibria. E.g. one group could have a norm of cooperation in a repeated PD (which is an equilibrium under suitable conditions); another has a norm of noncooperation. (This is evident in developing economies where there often patterns of corruption, that are expected and accepted, or patterns of honesty.) The idea of a norm as a belief which is self-fulfilled poses a natural question: What incentives keep the norm together? How do deviators get punished, in order to enforce the norm?

A broader definition is that a norm is an expected pattern of behavior which is punished when it is violated. This might be an equilibrium or might not. Again, the point is just that the game-theoretic concept of equilibrium imposes some intellectual discipline—if a norm is not an equilibrium, how does it adhere?

Networks are another area of potential common ground. To economists, linking in a network is a choice of how to allocate scarce resources (who to talk to, trade with, hang out with); so the natural question is how networks emerge endogenously. (And interestingly, there will often be multiple equilibria in these networks.) My guess is that great advances in modelling technology, and a longer tradition of cleverness conveyed in graduate school, could be a big help to sociologists. In turn, sociologists know a lot about general network patterns—raw facts—that could contribute to disciplining theories of equilibrium selection of network structures.

Neuroeconomics: An area I have been thinking about is how neuroscientific details of brain function could influence economics (see the long and short versions of our papers on my website <http://www.hss.caltech.edu/~camerer>). I think there are two big, easy ideas here: Brain details suggest that behavior results from the interaction of multiple systems, like automatic responses which can be overridden by more controlled processes (at cognitive expense, probably scarce activation in anterior cingulate); or emotions and cognition. To an economist this is heretical because preference is just a utility function over goods—a “one-system” view. But methodologically, economics is superb at understanding what happens when multiple systems interact (e.g. supply-demand, general equilibrium). So I don’t see a huge obstacle, in principle, to adapting these multiple-system approaches to multiple brain systems to create a precise theory of behavior which wipes the “preferences” word off the blackboard and starts over.

Emotions: To an economist, emotions are interesting because they sprawl across modelling categories. Is a quick and dirty fear signal from the amygdala a “belief” about the probability of threat. Uh, could be. Is it a preference? (As when lust influences sexual choices or nausea encodes a dispreference for a food one had a bad experience with.) Uh, that too. Is it a constraint? (As when overwhelming hunger motivates an immediate desire to eat anything in sight). Hmm, could be that too...Social emotions like shame and embarrassment are interesting too because they will depend delicately on what an embarrassed agent knows others know (spinach on your teeth or an unzipped fly aren’t embarrassing until they are pointed out); this invites a game-theoretic analysis in which what an agent knows others know is central.