Preferences, Beliefs and Equilibrium: What Have Experiments Taught Us?

by Daniel Friedman
September 18, 2008

1. Introduction.

The target article, "Theory and Experiment: What are the Questions?" is the latest of Vernon Smith's many attempts to get economists to behave like scientists. We economists have always had our theories, and Smith enjoins us to use them to predict what will happen in new environments. He has shown us, via numerous examples, how to construct enlightening environments in the laboratory. But how should we respond when the prediction works even better than expected, or when the prediction fails badly? The target article sharpens these questions, and lists some commonly accepted auxiliary premises that deserve further scrutiny.

As a student, I learned that the key primitives of microeconomic theory include preferences and beliefs at the individual level, and equilibrium at the aggregate level. In the following response to Smith's target article, I will focus on what experiments can teach us about these primitives and about the theoretical models that we construct from them.

2. Preferences.

Smith (1976) summarizes induced value theory, a cornerstone of experimental economics. Among other things, he lists sufficient conditions ("precepts") for controlling the preferences of human subjects. The basic idea is to get subjects to focus on their own personal cash payoffs, and to induce value on intrinsically worthless objects by redeeming them for cash according to a utility function chosen by the experimenter. For example, by paying a subject $0.50 for each red-black pair of cards the subject turns in, the experimenter induces Leontieff preferences over red cards and black cards.

Induced values, in conjunction with stationary repetition and the continuous double auction (CDA) market institution, work frighteningly well in implementing competitive equilibrium. Along with many other investigators, I personally have induced supply and demand schedules by having subjects draw costs and redemption values from a hat. Nobody knows the actual supply and demand at the time trading occurs, yet transaction prices and trading volume still converge reliably. After trading is all over, the subjects turn in their cost and value cards, get paid in cash (or, in classroom demonstrations, in small grade point bonuses), and I construct the actual supply and demand schedules and compute competitive equilibrium. It predicts quite accurately the converged prices and volumes, even when there are relatively few buyers and sellers (say, three or four of each). Nor does it take long to converge—three periods is usually enough.

The problem is that the prediction is not supposed to work such a setting. I was taught that competitive equilibrium holds only asymptotically as the number of buyers and sellers gets large (so that each of them has a negligibly small market share), and that public knowledge of supply and demand was somehow helpful. Thus the predictive ability of competitive equilibrium is a "scientific mystery" (Smith, 1982) that several game theorists have tried to solve.

The target article generously mentions my own 1984 article and its "no-congestion" assumption (which now would be called "renegotiation proofness"), but neglects more recent progress. My papers with Ostroy (1995), with Cason (1996, 1999) and with Zhang (2007), and work by other researchers cited in those papers, go a long way towards explaining the convergence process across periods.[1] Recent work by Gjerstad (2007) and others sheds new light on the process within each period.[2] Thus we have a pretty example of scientific progress, in which theory suggests experiment, the experimental results catalyze better theories, and the cycle of theory and experiment continues.

But induced values, together with other standard procedures, do not always support classic theoretical predictions. For example, induce constant marginal cost, have subjects choose quantities sequentially each period, and compute price, revenues and profit as in the textbook

[1] To summarize very briefly, traders in simple static theoretical models have the incentive to understate their willingness to transact, and those understatements lead to inefficient outcomes in thin markets. However, stationary repetition in the CDA teaches traders *not* to understate inefficiently: they learn to shade bids and asks towards earlier transaction prices, but not beyond. The learning process leads to a Nash equilibrium that implements a competitive equilibrium outcome.

[2] Again, a brief summary is that each mutually beneficial trade pushes the allocation deeper into the Edgeworth lens, and more closely aligns' traders marginal rates of substitution, narrowing the band (or cone) of plausible prices. The process ends only when the contract curve is attained and mutual gains are exhausted.

Stackelberg oligopoly model. As noted in Huck (2004), the laboratory results are far more egalitarian than predicted. Or have three or four subjects choose quantities simultaneously as in the Cournot model for 25 periods, and then force two of them to "merge," making a single quantity decision and splitting the profits evenly for another 25 periods. The second half results are far less symmetric across firms (but more egalitarian across subjects) than predicted by Cournot-Nash equilibrium. Run various kinds of duopolies and you are likely to see implicit collusion. On the other hand, give Cournot competitors additional feedback on all other subjects' choices and profits, and you probably will see higher outputs and lower profits than predicted by Cournot-Nash equilibrium.

To me, such results suggest a failure of the dominance precept. Perhaps subjects are not entirely focused on their own cash payoff, but instead are also concerned with the payoffs of other subjects. They may try to reward "generous" behavior and try to punish behavior perceived as "greedy." Or, anticipating such behavior by other subjects, they may make choices different than the (stage game Cournot-Nash) equilibrium. The target article considers other examples in which standard predictions fail, and conjectures that using "other peoples' money" may have an important effect. One could test such conjectures in oligopoly settings, and see whether outcomes are closer to Nash equilibrium when (a) the subjects' own money is at stake, or (b) the conversion rate (from points to cash) for other subjects is obscured, or (c) other subjects are replaced by robots. Again, a path of scientific progress seems open.

Things become more complicated when one turns from inducing preferences to eliciting homegrown preferences. Long the province of experimental psychology, homegrown preferences over income distribution have taken center stage in experimental economics in recent years. Results typically are quite muddy, and depend sensitively on context. Unfortunately, existing theory provides little guidance on how and why context matters, a point made in the target article (Assumption 6) and in the response by Gintis (2009).

Similar issues arise in eliciting time preferences, but the problem can perhaps be seen most clearly in the literature on eliciting risk preferences. For about 60 years, mainstream economists have modeled risky choice as maximizing the expectation of a Bernoulli function given a probability distribution (possibly subjective) over cash outcomes (e.g., Mas-Collel et al, 1995, Chapter 6). The Bernoulli (or utility) function, of course, is not observable and may vary arbitrarily across individuals, but in principle it can be inferred from the data, especially if one posits a functional form such as constant absolute risk aversion. In a 2004 manuscript with Shyam Sunder, I argue that the empirical work to date is unkind to theory, and there is no persuasive evidence that the typical person has anything like a stable Bernoulli function (or, for that matter, a well-defined value function as in Prospect Theory.) As one moves from task to task (and sometimes even from one elicitation session to the next), one sees little stability in individual subjects' elicited risk preferences.

This raises a more radical question: do homegrown preferences actually exist? In some sense, all preferences are induced--we care about money mainly for the goods and services it can buy, but we care about specific goods and services mainly to the extent they satisfy generalized desires for comfort and status, but these desires presumably are grounded in biological and psychological imperatives, and so on to the vanishing point.

From the perspective of revealed preference theory, the real question is this: at what level can one demonstrate regularity in choice? Few individuals seem to make transparently consistent choices over income distributions, for example.[3] Choices seem to depend on what the other people did to produce the available distributions, and also depend on the individual's social connections to the others. Gintis (2009) suggests that social norms are key, but until we have an operational theory on the formation and transmission of social norms, it not clear how to follow up the suggestion.

My recent work with Cox and Gjerstad (2007) and Cox and Sadiraj (2008) takes a conservative approach, and fills the gap with the neoclassical idea of state-contingent preferences (e.g., Mas-Colell et al. (1995, Chapter 6). Some aspects of status relationships and reciprocity can be captured in tractable models of "emotional" states. The scientific downside, of course, is that even rather parsimonious models of state-contingent preferences have more free parameters than

[3] Many subjects are consistently selfish in laboratory tasks but, as explained below, this doesn't imply consistent selfishness outside the lab, e.g., with their own families.

non-contingent models. Still, in tests so far, the models seem to explain available data and to have some predictive power.

One last point. Recently it has become popular to conclude from laboratory experiments that people fall into certain types, e.g., x% of US female undergraduates have selfish preferences, or y% are extremely risk averse. The problem is that a subject's observed responses might not reflect her intrinsic personal characteristics as much they reflect the way the subject parses the task—in the language of the target article, the subject's set point depends on her personal autobiographical knowledge filtered by the task description. Isaac and James (2000) find a strong *negative* correlation across individuals between risk aversion elicited in a first price auction and that elicited in a standard Becker-DeGroot-Marschak procedure. As for elicited social preferences, a devout Mormon subject may always tithe his earned income, but may keep all the money in a laboratory dictator game because he perceives the game as a zero-sum contest. Conversely, an intrinsically greedy subject may give away some of the money under the conjecture that that will improve her opportunities in a follow up session. The lesson is clear: it is dangerous to type-cast subjects according to their responses in a laboratory task.

3. Beliefs.

Individual beliefs are another key primitive of microeconomics. According to the orthodox doctrine--Bayesian decision theory--each individual's current beliefs can be summarized as a probability distribution called her prior. When new information arrives, the individual expresses it in terms of a likelihood function that, combined with her prior according to Bayes' Rule, yields a posterior probability distribution that represents her updated beliefs.

Is this model descriptive of the actual process people use? Probably not. Most economists, if pressed, might argue that the model describes the way beliefs *should* evolve, and that it sometimes helps predict the choices actual people make. Gilboa et al. (2008) argue that the Bayesian model is too restrictive in that beliefs seldom are specific enough to define a unique probability distribution, and at the same time, the model is too general in that it offers no insight into how prior beliefs are formed. Most experimentalists would agree, though not all will endorse the authors' proposed remedies: a multiple prior model and a similarity-weighted relative frequency model.

The target article suggests that the actual process is more like trial-and-error: "…people explore an environment by making an accustomed move, observe the consequence, then make … adjustment[s] over time… ." Selten (2004) is somewhat more specific in his exposition of direction learning theory: in adjusting an ordered choice variable in discrete time, people choose higher (lower) values when last period a higher (lower) value would have increased payoff. Oprea et al. (2008) and Cason and I (1999), among others, have estimated parametric versions of direction learning, and found that it nicely explains some apparently anomalous behavior in complicated decision tasks (e.g., exercising deferral options) and games (e.g., bidding in thin call markets[4]).

But there is more to learning than parameter adjustment. Beliefs affect the way we perceive the world, and they respond to all sorts of experience, not just accumulating data. Most of us are in the teaching profession, and offer explicit instruction in classroom lectures and text readings. Don't these activities change students' beliefs? At least once in a while?

Moving towards more researchable questions, one might ask: when a person learns something in a particular context, whether in the laboratory or classroom or field, when does the lesson transfer to a new but relevant context? For example, if a person learns to solve a 100 door Monty Hall problem, when does that help him solve the original 3 door problem (Page, 1998)? Despite precedents going back to mathematical psychology of the 1960s, research on such questions is still in its infancy.

There is a related and even more fundamental question. How can we model and investigate the "aha" experience, when new insight transforms current beliefs? It turns out to be difficult even to model the prior state of "unawareness" within a Bayesian framework (e.g., Heifetz et al., 2008).

[4] The anomaly is that bids and asks in call markets are more nearly competitive (i.e., unaggressive or not understated) than in Bayesian Nash equilibrium. Estimates of the directional learning model explain this as an asymmetric reaction to ex post errors: people respond far more strongly when they miss a trade by bidding too aggressively than when they could have increased profit by bidding more aggressively. Hence they learn not to understate willingness to transact beyond recent clearing prices.

One can model awareness of unawareness by using a lattice of possible worlds, or simply by augmenting the usual set of states with a refinable residual state. Rumsfeldian unawareness of unawareness is even harder to model, and more pernicious. Experiments so far are very few. Smith's recent work on the emergence of trade (e.g., Crockett, Smith and Wilson, 2009) is perhaps the best example so far. One hopes that unawareness theory and experiment will soon begin to catalyze each other.

4. Equilibrium.

Economists usually are more interested in the behavior of multi-agent economic systems than in the behavior of individual agents. We obtain system-level predictions by imposing some sort of mutual consistency condition—competitive equilibrium, or Nash equilibrium, or perfect Bayes equilibrium, or perhaps correlated equilibrium. This theoretical approach seems odd to physicists, chemists and biologists, for whom equilibrium means a rest point that may (or may not) be approximated over time by a dynamical process. Experimental economists tacitly recognize the natural science perspective when our tests of theoretical predictions focus on the later periods of a laboratory session or on experienced subjects.

From this dynamical perspective, it is natural to raise questions about backwards induction, history independence, and many of the other auxiliary premises discussed in the target article. Vernon Smith once criticized predictions based on perfect Bayesian equilibrium (PBE) on the grounds that simply assuming PBE "pushes the real action off stage" (Friedman and Rust, 1993, p. 12). As noted earlier, experimentalists have made progress in understanding competitive market outcomes by studying the dynamic process, not by simply fitting alternative equilibrium concepts. As experimental economists (or as economic scientists), we should often put the dynamical process on center stage. When we better understand the dynamic process, we will be better able to predict when standard equilibrium concepts will describe converged behavior and when they will not.

At the very least, we should be able to distinguish short run predictions from long run. Competitive equilibrium is considered long run when there is no profitable entry or exit or adjustment of variables that are considered fixed in the short run. In a recent paper with Nirvikar

Singh (2007), we propose an analogous notion for Bayesian games. Bayes equilibrium (perfect or merely Nash) takes as given a particular set of "types" and a particular distribution of players over types. This is fine for the short run, but over the longer run the types might evolve, and switching (or learning or demographic changes) could alter the player distribution. If so, in the long run, only the types with highest payoff should survive. One can imagine running laboratory experiments that allow migration of types (or learning or switching) to test this notion of evolutionary perfect Bayes equilibrium, as well as experiments that exogenously fix the type distribution and test the short run concepts of Bayes equilibrium.

5. Concluding remarks.

I've learned a lot by reading Smith's methodological articles, including Smith (1991) as well as those mentioned above. The target article reprises some old issues, raises some new ones and, as with Rorschach's inkblots, elicits our thoughts. It got me thinking more about induced preferences versus elicited homegrown preferences, about the nature of learning, and about the foundations of economic equilibrium. I'm sure that further thoughts on these and other matters will bubble up in readers minds as they think about the target article and the response articles.

Smith opens his target article with a quote from philosopher David Hume on moral judgments, and uses it to question what we can infer from experiments. Let me close my response with a reminder that experiments can be enjoyable as well as enlightening, using a quote from lyricist Cole Porter:

> …do what all good scientists do:
> Experiment!
> Make it your motto day and night.
> Experiment!
> And it will lead you to the light.
> The apple on the top of the tree
> Is never too high to achieve.
> So take an example from Eve:
> Experiment!
> Be curious
> Though interfering friends may frown.
> Get furious

At each attempt to hold you down.
If this advice you'll only employ
The future can offer you infinite joy
And merriment.
Experiment
And you'll see!

Acknowledgement. I am grateful to Herb Gintis, Ryan Oprea and Huibin Yan for helpful comments on an earlier draft of this paper, and to Vernon Smith for serial inspiration.

Bibliography

Aviad Heifetz, Martin Meier, Burkhard C. Schipper, 2008. “A canonical model for interactive unawareness,” *Games and Economic Behavior* 62:1, 304-324.

Cason, Timothy and Daniel Friedman, 1996. "Price Formation in Double Auction Markets," *Journal of Economic Dynamics and Control* 20:8 1307-1337.

Cason, Timothy and Daniel Friedman, 1999."Learning in Laboratory Markets with Random Supply and Demand," *Experimental Economics* 2:1, 77-98.

Crockett, Sean, Vernon Smith and Bart Wilson, 2009. “Exchange and Specialisation as a Discovery Process,” forthcoming, *Economic Journal.*

Friedman, Daniel and Shyam Sunder, 2004. “Risky Curves: From Unobservable Utility to Observable Opportunities,” draft manuscript, UCSC, June 2004.

Friedman, Daniel and Joseph Ostroy, 1995. "Competitivity in Auction Markets: An Experimental and Theoretical Investigation," *Economic Journal* 105:1, 22-53.

Friedman, Daniel and John Rust, 1993. *The Double Auction Market: Institutions, Theories and Evidence*, Santa Fe Institute Proceedings Volume XIV, Addison-Wesley, 1993.

Gilboa, Itzhak, Andrew W. Postlewaite, and David Schmeidler. 2008. "Probability and Uncertainty in Economic Modeling. " *Journal of Economic Perspectives*, 22(3): 173–88.

Gintis, Herbert, 2009. “Experimental Economics Will Foster a Renaissance of Economic Theory,” this volume.

Gjerstad, Steven, 2007. “Price Dynamics in an Exchange Economy,” manuscript, Purdue University Economics Department, November.

Heifetz, Aviad, Martin Meier, Burkhard C. Schipper, 2008. “A canonical model for interactive unawareness,” *Games and Economic Behavior* 62:1, 304-324.

Huck, Steffen, 2004. “Oligopoly,” chapter 10 in D. Friedman and A. Cassar, eds., *Economics Lab: An Intensive Course in Experimental Economics*, London: Routledge.

Issac, Mark, and James, Duncan 2000. “Just who are you calling risk averse?” *Journal of Risk and Uncertainty*, 20(2), 177–187.

Mas-Colell, A., M. Whinston, and J. Green. *Microeconomic Theory*. New York, NY: Oxford University Press, 1995.

Oprea, Ryan, Daniel Friedman and Steven T. Anderson, 2008. “A Laboratory Investigation of Deferral Options,” UCSC Working Paper #635, March 2007 (revised, June 2008).

Page, Scott E., 1998. “Let’s Make a Deal,” *Economics Letters* 61:2, 175-180.

Porter, Cole, 1933. “Experiment!” the finale of Scene 1, Act 1 of the London Stage Musical *Nymph Errant*.

Selten, Reinhard, 2004. “Learning Direction Theory and Impulse Balance Equilibrium,” chapter 12 in D. Friedman and A. Cassar, eds., *Economics Lab: An Intensive Course in Experimental Economics*, London: Routledge.

Smith (1976). “Experimental Economics: Induced Value Theory,” *American Economic Review* 66:2, 274-79.

Smith, Vernon L. 1982. “Microeconomic Systems as an Experimental Science,” *American Economic Review* 72:6, 923-55.

Smith, Vernon L. 1991. “Rational Choice: The Contrast between Economics and Psychology,” *Journal of Political Economy* 99:4, 877-897.

Smith, Vernon L., 2009. “Theory and Experiment: What are the Questions?” this volume.

Zhan, Wenjie and Daniel Friedman, 2007. “Markups in Double Auction Markets,” *Journal of Economic Dynamics and Control* 31:9, 2984-3005.