# How Do Economists Think About Rationality?

Tyler Cowen
Department of Economics
George Mason University
Fairfax, VA 22030
tcowen@gmu.edu

November 24, 2001

The author wishes to thank Bryan Caplan, Robin Hanson, David Schmidtz, and Alex Tabarrok for useful comments.

I. Introduction

It is commonly believed that the rationality postulate, whatever that may mean, stands at the core of economic theory and much of social science research. Many economists go so far as to define their science in terms of the rationality assumption. I therefore try to spell out how working economists approach rationality, with the goal of explaining the practice of economic science to philosophers and other social scientists.

I do not survey the enormous literature on the methodology of economics.[1] Much of this literature focuses on philosophy of science rationales (or lack thereof) for rationality assumptions. The writers debate instrumentalism, the use of rhetoric to discuss rationality, whether assumptions need be realistic, and whether economic propositions are, or should be, falsifiable, to name a few of the better-known issues. These debates have generated insight, but taken alone they give a misleading picture of what economists do. Often they focus on economics as a whole, or one or two fields, rather than on the increasingly diverse ways that contemporary economists conduct their research.[2]

Philosophers, on the other hand, commonly believe that economic logic focuses on instrumental rationality, as exemplified by a Humean ends-means logic. That is, economics focuses on how to use means to achieve given ends, but cannot judge the quality or rationality of those ends. Philosophers have put forth alternative notions of rationality, including "practical reasoning," procedural rationality (do our mental processes for forming values make sense?), and expressive rationality (do we have the right ends or values?). From a philosophic perspective, economic rationality is only one

---

[1] For relevant treatments and surveys, see, for instance, Bruce J. Caldwell, *Beyond Positivism: Economic Methodology in the Twentieth Century*, Routledge, 1994, and Daniel M. Hausman, *The Inexact and Separate Science of Economics*, Cambridge University Press, 1992.

[2] Melvin Reder, *Economics: The Culture of a Controversial Science*, University of Chicago Press, 1999, lays more stress on practice than do most writers on economic method.

very small part of rationality, and for this reason economics appears radically incomplete as a "final theory of the world," whatever its other virtues may be.[3]

I approach the rationality postulate from a differing perspective. In particular, I stress that there is no single, monolithic economic method or approach to rationality. Labor economists, finance theorists, experimental economists, and macroeconomists, among others, all think of rationality, and use the rationality postulate, in different ways. I explicate modern economic method by searching out and identifying the differences across fields, rather than forcing everything into an account of the underlying unities.

Within economics, competing notions of the rationality postulate vie for graduate students, external funding, journal space, Nobel Prizes, public attention, policy influence, and other rewards. A particular approach to the rationality postulate proves useful to the extent that it helps a new field get off the ground, generate useful results, help formulate policy, show intuitive resonance, command attention, and so on. We should think of this competitive process as fundamental to contemporary economics, more fundamental than any single account of rationality that might be provided.

We frequently observe some rationality concepts displacing others. The rational expectations assumption gained in prestige in the 1970s and 1980s and now holds a secure foothold. In more recent times it has been challenged by behavioral assumptions and quasi-rationality. In the last twenty years empirical labor economics and experimental economics have risen greatly in status. Computational economics can be seen as knocking at the door, though it remains inaccessible to most economists, and in the eyes of many outsiders it has not generated new and unique results. Most generally, the number of available rationality concepts has multiplied since the development of

---

[3] For a presentation of this view, see Shaun Hargreaves Heap, *Rationality in Economics*, Basil Blackwell, 1989. The moral general philosophic literature on rationality is vast. For some recent perspectives, see David Schmidtz, *Rational Choice and Moral Agency*, Princeton University Press, Robert Audi, *The Architecture of Reason: The Structure and Substance of Rationality*, Oxford University Press, 2001, and Elijah Millgram, *Varieties of Practical Reasoning*, MIT Press, 2001.

basic consumer theory early in the twentieth century, which reflects both scientific advance and a more intense competitive process.

Starting from these points, I seek to integrate the perspectives of the practitioner and the methodologist, and to explain to non-economists how economists use the rationality postulate. Section II presents some basic ways of thinking about rationality postulates and what role they might play in economics. Section III considers how contemporary economists actually use rationality postulates in differing fields of economics. Section IV considers the implications of this diversity, and for how we should understand criticisms of economics.

## II. How might economists be thinking about rationality?

We see at least five basic ways of interpreting rationality postulates. I call these description, transitivity only, tautology, normative, and pragmatic. In section III below I will illustrate each with particular applications. Note that the five categories may overlap to some degree.

### 1. Description

In this view the rationality postulate describes individual behavior and has definite empirical implications. These implications are in principle falsifiable. For the time being, and perhaps forever, the rationality postulate best describes economic behavior. Note that this view may attach varying substantive meanings to the empirical content of the rationality concept.[4]

### 2. Transitivity only

[4] See, for instance, James C. Cox, "Testing the Utility Hypothesis," Economic Journal, 107 (1997): 1054-1078

Some economists believe that rationality requires only the postulate of transitivity of preference. Transitivity, in this context, stipulates that if A is preferred to B, and B is preferred to C, then A is preferred to C. In other words, preferences can be represented by a global rank ordering. Typically it is left open whether transitivity must always be satisfied, or if we require only that transitivity is true "often enough." Note that the transitivity only view can be a special case of the first "descriptive" view, at least if we interpret transitivity as an empirical property rather than a logical one.

3. Tautology

Any and all behavior can be described as rational, provided we are willing to manipulate the theory enough to avoid any testable implications. The rationality postulate therefore involves no substantive commitment to any empirical claims. A strong version of this view postulates that rationality is a useful tautology, while a weaker view postulates that it is an arbitrary use of terms offering no particular advantage.

4. Pragmatic, or useful organizing category

This view rejects foundationalist approaches to rationality. We do not necessarily know exactly what the rationality postulate does or means. Nonetheless economists who use the rationality postulate come up with better work and better ideas than those who do not. It is a useful heuristic for the economist. The rationality postulate is part of a research strategy for generating new ideas, regardless of its descriptive or logical status.

5. Normative

Individuals are not always rational, but rationality is an ideal that we should strive to achieve. Economic theory can be used to improve the quality of decision-making. Note that this approach can be combined with any number of more substantive commitments to what rationality means.

III. Rationality postulates in various economic fields

With these basic ideas in place, let us now consider some of the roles they play, in various combinations and permutations, in economic research. I consider several fields in the discussion that follows, including macroeconomics, theory of the consumer, expected utility theory, finance, empirical labor economics, and experimental economics. I do not mean to suggest that every practitioner in a cited area holds the same views. Rather I consider a "typical" use of the rationality concept, as it might be found in the specified field or subfield of economics. Again, I am referring primarily to research practice rather than to the methodological literature.

a. Theory of the consumer under certainty

The theory of consumer behavior under certainty provides the first few chapters of most microeconomics texts. It concerns how to represent preference orderings, relative price effects, income effects, indifference curves and so on. These results stand at the base of many other propositions in economic theory.[5]

Strictly ordinal utility theory, under conditions of perfect certainty, typically treats the rationality postulate as a tautology. For any observed market behavior, we can try to redescribe that behavior ex post in terms of an underlying ordinal utility function. Virtually nothing could refute the hypothesis that individual behavior can be described in such terms. The theory of "revealed preference" tells us that observed market demands can be retranslated into underlying utility functions under a wide variety of assumptions. Over time economists have worked hard at relaxing the various assumptions behind consumer theory. Originally assumptions such as continuity and divisibility were seen as important, but now consumer theories have been developed that dispense with those assumptions, while still delivering the standard results of neoclassical theory. Even observed "intransitivities" can be redescribed as "changing preferences over time." Note

[5] The reader who wishes further detail should consult Hal Varian, *Microeconomic Analysis*, W.W. Norton & Company, 1992.

that this basic theory does not assume that individuals are selfish but accommodates altruistic motives regularly.[6]

b. Theory of the consumer under uncertainty

The theory of choice under uncertainty started with von Neumann-Morgenstern subjective expected utility theory and since expanded to cover various modifications of that basic approach. For our purposes here, I will speak simply of the von Neumann-Morgenstern approach as a more general appellation, even though many modern researchers reject the original axiomatic formulation of the theory.

In contrast to the theory under certainty, economists typically do not treat their theories of expected utility as tautologies. To the extent the empirical tests reject the axioms, it is considered grounds for rejecting the theory. Those who work with the narrower versions of the von Neumann-Morgenstern approach typically admit some degree of falsification has occurred. They do not defend the theory on tautological grounds, but rather question whether some equally tractable alternative has been developed, and whether the extant alternatives in fact yield better predictions across the board.

Several axioms of the von Neumann-Morgenstern approach are subject to falsification. Most notorious in this regard in the so-called "independence axiom." The independence axiom typically it specifies that an additional percentage chance of some outcome has the same utility value to an individual, whether it is added to a thirty, forty, or fifty percent probability of the outcome in question. In other words, mutually exclusive world-states ("what I could have had") should not influence the value of what I have, according to this postulate.

---

[6] On revealed preference, see Paul Samuelson, "A Note on the Pure Theory of Consumer's Behavior," Econometrica, 5, (1938): 353-354. On relaxing assumptions behind existence proofs, see, for instance, Andreu Mas-Colell, "The Price Equilibrium Existence Problem in Topological Vector Lattices," Econometrica, 54, (1986): 1039-1054.

This axiom is contradicted by the evidence, whether we look at questionnaires or experiments with real dollar prizes. Notice, for instance, that the phenomenon of "nerves" often falsifies the independence axiom. If I move from a .99 chance of a good outcome to a 1.0 chance of that same outcome (certainty), I can stop worrying about what will happen. This extra one percent, in this context, may be worth more than moving from a .01 chance to a .02 chance. For some individuals, the move from .00 to .01, or to some smaller positive increment, is especially important. This can be taken to represent the value of hope, and may explain why so many people play the lottery. Similarly, regret may falsify the independence axiom. If we feel bad about what we could have had, but did not get, the values of mutually exclusive world states will influence each other.[7]

Note that even here, we could try to define the expected utility axioms as tautologies. Assume, for instance, that an individual values the "move from .00 to .01 chance of becoming a millionaire" more than the "move from .90 to .91 chance of becoming a millionaire." This would appear to falsify the independence axiom. Yet such behavior is consistent with the independence axiom <u>if</u> we specify the relevant outcomes differently. Perhaps "the chance of becoming a millionaire" is not the relevant outcome. Instead, imagine the relevant outcome is "the chance of becoming a millionaire, plus the hope that is enjoyed in the process of waiting to discover one's fate." The moves from .00 to .01 and from .90 to .91 thus represent different outcomes, and valuing them differently can be fully rational. Under this approach no observed behavior could refute the independence axiom.

Taking this logic further, any apparent violation of the independence axiom could be taken to mean that we had not specified the appropriate outcomes or world-states

---

[7] The Allais paradox provides the best known counter to the independence axiom. In the Allais paradox, individuals are first asked to rank gamble A against gamble B. If the individual chooses A, the independence axiom implies that same person should prefer C over D. Yet we systematically observe individuals preferring A over B, and D over C. See Mark J. Machina, "Expected Utility Without the Independence Axiom," Econometrica, 50, (1982): 277-324.

correctly. Under this maneuver, however, the definition of an outcome varies with the probability of that outcome. For technical reasons, this interdependence of outcomes and probabilities would make expected utility theory intractable, since an outcome could never be defined as separate from the probabilities of that outcome. Common judgment holds that expected utility theory would cease to be a useful tautology if it were treated in this fashion.

Note that the "tautologizing moves" in choice theory under certainty are seen as less destructive of tractability. They may decrease the usefulness of the theory, but they pose no immediate technical problem comparable to the intermingling of outcomes and probabilities, as we find in the theory of choice under uncertainty. For this reason economists treat consumer behavior under uncertainty as having falsifiable axioms to a greater extent than does the theory of choice under certainty.

Some choice theorists have attempted to reconstruct expected utility theory without the independence axiom.[8] In his theory expected utility is linear in the probabilities only locally, not globally. Yet even this theory admits of possible empirical refutation. It implies particular attitudes towards how gambles are resolved over time, and how individual preferences will change with the temporal resolution of uncertainty. While these propositions are considered difficult to test, they are not tautologies either.

For these reasons, few economists interpret expected utility theory as a tautology. Instead, expected utility theory is considered as either a testable hypothesis, a normative standard, or as a useful analytical category, depending on the field of investigation. When economists do choice theory, especially in a laboratory setting, the expected utility approach serves as a testable hypothesis, to be either supported or falsified by the data.

For more general theoretical purposes, the expected utility hypothesis is a useful building block for presenting some larger idea. In labor economics, for instance, it is commonly postulated that workers choose some degree of shirking, depending on their chance of

[8] See Machina, Ibid.

being caught and fired. Expected utility theory, especially in its simplest forms, is used to represent this problem. Most of the economists who use expected utility theory in this way do not believe it is descriptively true. Nonetheless they hold it to be the most tractable and convenient approach at hand. They regard the theory as both falsifiable, and false, yet useful nonetheless. The empirical failings of expected utility theory are considered inessential to the basics of work/shirking decisions.

Expected utility theory also can be used normatively. The experimental literature now shows that people do not satisfy the independence axiom across probabilities. Nonetheless a normative theorist still might believe that people ought to satisfy such axioms. The entire field of risk analysis tries to help people make better decisions under uncertainty. Risk analysis has been used to advise the Pentagon, to help adjudicate lawsuits, to make securities markets more efficient, and to help companies make decisions about how to invest and when to buy insurance. Normative risk analysis forms a substantial part of applied economics, often in the context of consulting, even when it does not show up in academic research more narrowly defined.

c. Macroeconomics

A large body of macroeconomics uses the assumption of rational expectations, henceforth RE. While a majority of working macroeconomists do not accept the empirical validity of RE, most of the important work in macroeconomics over the last thirty years has used the RE assumption. Rationality, for macroeconomists, refers primarily to rationality of expectations, rather than to some property of preferences.

RE has been defined in several regards, which may or may not boil down to the same thing. Under one account, RE means that individuals understand the "true model" of the economy; in other words, trading individuals have the understanding of a fully accomplished macroeconomist. Under another account, individual forecasts of economic variables are correct on average. This can mean either that the errors of an individual average out to zero over time, or that at any point in time, individual forecasts are

scattered around the true variable but with a correct mean. Finally, RE may mean that errors are serially uncorrelated over time. That is, if I guess too high one period, that has no predictive power for whether I guess too high or too low the next period. Again, this proposition can hold for either individuals or groups.[9]

Economists put these assumptions into macroeconomic models for several reasons. First, some economists believe those assumptions are roughly true. A more common view is that they provide a kind of modeling discipline. The view is commonly voiced that "errors can be used to explain anything." By forcing the theorist not to rely too heavily on errors, the RE assumption makes it harder to come up with ad hoc models. Finally, the RE assumption may be useful as a foil, to see what a world without systematic errors would look like. By comparing this idealized picture to the real world, we may get a better sense of whether systematic errors are central to the phenomenon of business cycles and other economic problems.

The RE assumption has received a wide variety of tests. When we look at questionnaires about expectations, RE commonly fails to predict measured expectations. Similarly, RE fails tests in the laboratory setting. Some RE predictions, however, are commonly (though not always) validated at the macroeconomic level. It appears that the money supply does not affect real output, once we take interest rates into account, and that budget deficits do not cause real interest rates to rise.[10] Both of these predictions follow from some standard RE models, though of course the studies test several hypotheses jointly, rather than just RE taken alone.

Rationality, in the form of RE, is considered testable, both in principle and in reality. Failing the tests lowers the status of the RE assumption, without making it worthless

---

[9] For a discussion of different forms of rational expectations, see Tyler Cowen, *Risk and Business Cycles: New and Old Austrian Perspectives*, Routledge, 1997, pp.8-9, which also provides more detailed citations to the literature.

[10] See, respectively, Robert B. Litterman, and Laurence Weiss, "Money, Real Interest Rates, and Output: A Reinterpretation of Postwar U.S. Data," Econometrica, 53 (1985): 129-156, and Paul Evans, "Interest Rates and Expected Future Budget Deficits in the United States," Journal of Political Economy, 95, (1987): 34-58.

altogether. RE thus has descriptive, pragmatic, and normative components, to refer back to the categories outlined above.

d. Theory of finance

Financial economics has one of the most extreme methods in economic theory, and increasingly one of the most prestigious. Finance concerns the pricing of market securities, the determinants of market returns, the operating of trading systems, the valuation of corporations, and the financial policies of corporations, among other topics. Specialists in finance can command very high salaries in the private sector and have helped design many financial markets and instruments. To many economists, this ability to "meet a market test" suggests that financial economists are doing something right.

Depending on one's interpretation, the theory of finance makes either minimal or extreme assumptions about rationality. Let us consider the efficient markets hypothesis (EMH), which holds the status of a central core for finance, though without commanding universal assent. Like most economic claims, EMH comes in many forms, some weaker, others stronger. The weaker versions typically claim that deliberate stock picking does not on average outperform selecting stocks randomly, such as by throwing darts at the financial page. The market already incorporates information about the value of companies into the stock prices, and no one individual can beat this information, other than by random luck, or perhaps by outright insider trading.

Note that the weak version of EMH requires few assumptions about rationality. Many market participants may be grossly irrational or systematically biased in a variety of ways. It must be the case, however, that their irrationalities are unpredictable to the remaining rational investors. If the irrationalities were predictable, rational investors could make systematic extra-normal profits with some trading rule. The data, however, suggest that it is very hard for rational investors to outperform the market averages. This suggests that extant irrationalities are either very small, or very hard to predict, two very

different conclusions. The commitment that one of these conclusions must be true does not involve much of a substantive position on the rationality front.

The stronger forms of EMH claim that market prices accurately reflect the fundamental values of corporations and thus cannot be improved upon. This does involve a differing and arguably stronger commitment to a notion of rationality.

Strong EMH still allows that most individuals may be irrational, regardless of how we define that concept. These individuals could literally be behaving on a random basis, or perhaps even deliberately counter to standard rationality assumptions. It is assumed, however, that at least one individual does have rational information about how much stocks are worth. Furthermore, and most importantly, it is assumed that capital markets are perfect or nearly perfect. With perfect capital markets, the one rational individual will overwhelm the influence of the irrational on stock prices. If the stock ought to be worth $30 a share, but irrational "noise traders" push it down to $20 a share, the person who knows better will keep on buying shares until the price has risen to $30. With perfect capital markets, there is no limit to this arbitrage process. Even if the person who knows better has limited wealth, he or she can borrow against the value of the shares and continue to buy, making money in the process and pushing the share price to its proper value.

So the assumptions about rationality in strong EMH are tricky. Only one person need be rational, but through perfect capital markets, this one person will have decisive weight on market prices. As noted above, this can be taken as either an extreme or modest assumption. While no one believes that capital markets are literally perfect, they may be "perfect enough" to allow the rational investors to prevail.

"Behavioral finance" is currently a fad in financial theory, and in the eyes of many it may become the new mainstream. Behavioral finance typically weakens rationality assumptions, usually with a view towards explaining "market anomalies." Almost always these models assume imperfect capital markets, to prevent a small number of

rational investors from dwarfing the influence of behavioral factors. Robert J. Shiller claims that investors overreact to very small pieces of information, causing virtually irrelevant news to have a large impact on market prices. Other economists argue that some fund managers "churn" their portfolios, and trade for no good reason, simply to give their employers the impression that they are working hard. It appears that during the Internet stock boom, simply having the suffix "dot com" in the firm's name added value on share markets, and that after the bust it subtracted value.[11]

Behavioral models use looser notions of rationality than does EMH. Rarely do behavioral models postulate outright irrationality, rather the term "quasi-rationality" is popular in the literature. Most frequently, a behavioral model introduces only a single deviation from classical rationality postulates. The assumption of imperfect capital markets then creates the possibility that this quasi-rationality will have a real impact on market phenomena.

The debates between the behavioral theories and EMH now form the central dispute in modern financial theory. In essence, one vision of rationality -- the rational overwhelm the influence of the irrational through perfect capital markets -- is pitted against another vision -- imperfect capital markets give real influence to quasi-rationality. These differing approaches to rationality, combined with assumptions about capital markets, are considered to be eminently testable.

e. Empirical labor economics

Empirical labor economics has been one of the most successful areas in neoclassical economics over the last fifteen years. Largely through careful use of panel data sets, economists have generated much knowledge about wages, labor supply decisions, the effects of the welfare state, and the nature of unemployment, among other topics.[12]

[11] See Robert J. Shiller, *Irrational Exuberance*, Broadway Books, 2000. Richard H. Thaler, *Quasi Rational Economics*, Russell Sage Foundation, 1994, presents many behavioral themes.

[12] By its very nature empirical labor economics cannot be presented in any canonical form, nor is it typically found in book length treatments. The best source are the articles

Empirical labor economists tend to be among the least theoretical of economists, and tend to be the most likely to defer to the data. Labor economics offers few (if any) a priori propositions. Even traditional supply and demand patterns are difficult to forecast in advance. We cannot say a priori whether an increase in wages causes people to work more (being paid more money per hour, the substitution effect) or to work less (they have more money and can now consume leisure, the income effect). The operation of the law of demand is contingent in similar fashion. Under some circumstances, a rise in real wages can increase the demand for labor, rather than lower it (as standard demand theory would otherwise predict). A hypothesis known as "efficiency wage" theory stresses how higher wages can make workers more productive, either by making them happier, or by making them more afraid to shirk. Employers therefore may therefore demand more labor, the higher the wage, at least along some margins.

Given the difficulty of establishing clear theoretical predictions, labor economists tend to be among the most positivistic of economists. They consider the test to be everything, and theory to tell us little. A typical piece of labor economics writes down some utility function for laborers, and thus assumes a kind of rationality. This utility function, however, usually has few testable implications on its own, other than implying that individuals prefer income and leisure. This is more of a concrete statement about preferences rather than about rationality. In other cases labor economists may use subjective expected utility theory, for reasons of convenience, but without possessing any deep commitment to the underlying theoretical apparatus. Therefore we can think of labor economics as involving minimum assumptions about rationality, and as relying almost entirely on the data for its substantive hypotheses. Theory certainly helps labor economists interpret the data, but this theory is remarkably thin with regard to its assumptions about rationality.

f. Game theory

---

published in the American Economic Review in the last fifteen years, under the editorship of Orley Ashenfelter, himself an empirical labor economist.

In game theory, notions of rationality are highly specific and economists debate the propriety of one notion against another. Such debates are central to the theoretical enterprise.

To see how this works, let us step back for just a moment for some context. In most (non-game-theoretic) economic models, individuals are assumed to play what are called "dominant strategies." A dominant strategy unambiguously maximizes the individual's return, such as when an individual maximizes utility or a firm maximizes profit. Exercising a dominant option of this kind is uncontroversial in economic theory, even though at a deeper level economists may disagree on what an institution seeks to maximize (is it profit, revenue, prestige?, etc.). Once the maximand is given, individuals will pursue the dominant strategy. In other words, more is preferred to less.

When we move to game theory, and the world of strategic interdependence, dominant strategies frequently do not exist. The payoff of a given strategy depends on what the individual expects others to do. Often no single strategy yields higher returns for every possible response from others. So getting a good return depends on making accurate predictions about the behavior of others; of course these same others are trying to make accurate predictions about their opponents as well.

At this point game theory resorts to equilibrium concepts to resolve games. The most famous of these is the Nash equilibrium, which describes a series of moves as an equilibrium only when neither player has an incentive to deviate with his or her strategy, taking the strategy of the others as given. Nash equilibrium thus represents one attempt of game theory to model the notion of rationality in a game. In more complicated games, game theorists resort to such notions as "subgame perfection," "time consistency," "perfect Bayesian equilibrium," and many others. In each case, these concepts attempt to define rationality for the purpose of "solving" the game or games. Note also that the solution concept can be viewed in positive terms, in normative terms, or more abstractly as possessing certain formal properties, without necessarily being either predictive or normative.

Debates about rationality arise for two reasons. First, many common solution concepts are sometimes implausible. Second a solution concept may imply multiple equilibria.

To start with plausibility, Nash equilibria appear unlikely in many contexts. If we consider a multi-period Prisoners' Dilemma, played ten times, the Nash concept usually predicts defection in the first round.[13] Yet even most game theorists do not play this strategy of immediate defection when asked to play the game for real money. Economists therefore have tried to come up with richer notions of equilibrium that can account for the predictive limitations of the Nash concept.[14]

In many games we find a multiplicity of Nash equilibria, often an infinite or near-infinite number of such equilibria. The simplest way of generating multiple equilibria is to set up coordination game with two possible points of high return. Very large numbers of equilibria arise most easily when we consider trigger and response strategies. For instance, if an individual plays "if you defect, I will defect for seven successive periods with $p = 0.9$," it may also be an equilibrium strategy to play the same but with $p = 0.90001$, with $p = 0.90002$, and so on. Many games have an infinite number of possible equilibria.

When multiple Nash equilibria exist, the game-theoretic notion of rationality does not, taken alone, yield unique predictions. Economists then have developed "refinement concepts," which seek to discriminate amongst the various Nash equilibria (or other multiple equilibria, derived from other solution concepts) and elevate one as the uniquely rational outcome. Some of these refinements are highly complex, and consist of classificatory criteria that have no clear intuitive referents. Other refinement concepts deliberately introduce bounded rationality and ask which equilibria could be "computed" by persons in the real world. These debates have given rise to a voluminous and complex literature, and there is no general agreement on what is the best refinement concept.

---

[13] The so-called "folk theorem," which allows for cooperation, kicks in only when the number of rounds is very large and the discount rate is small.

[14] These attempts have spawned a huge literature. See, for instance, John C. Harsanyi and Reinhard Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, 1988.

Game theory has shown economists that the concept of rationality is more problematic than they had previously believed. What is rational depends not only on the objective features of the problem but also depends on what actors believe. This short discussion has only scratched the surface of how beliefs may imply very complex solutions, or multiple solutions. Sometimes the relevant beliefs, for instance, are beliefs about the out-of-equilibrium behavior of other agents. These beliefs are very hard to model, or it is very hard to find agreement among theorists as to how they should be modeled.

In sum, game theorists spend much of their time trying to figure out what rationality means. They are virtually unique amongst economists in this regard. Game theory from twenty years ago pitted various concepts of rationality against each other in purely theoretical terms. Empirical results had some feedback into this process, such as when economists reject Nash equilibrium for some of its counterintuitive predictions, but it remains striking how much of the early literature does not refer to any empirical tests. This enterprise has now become much more empirical, and more closely tied to both computational science and experimental economics.

g. Experimental economics

Experimental economists test economic propositions in a controlled laboratory setting with real dollar prizes. By using controlled experiments in this regard, experimental economics comes closer to the methods of many of the natural sciences.[15]

Experimental economics typically treats all propositions about rationality as up for grabs. The ability to perform controlled experiments relieves experimental economics from the necessity to start with unquestioned assumptions. Virtually every economic assumption about individual behavior has been subject to test in a laboratory setting. Furthermore, virtually all of these assumptions have been falsified.

---

[15] On experimental economics, see, for instance, John H. Kagel and Alvin E. Roth, editors, *The Handbook of Experimental Economics*, Princeton University Press, 1995.

Some of the most radical experiments appear to have falsified the assumption of transitivity. Individuals can be induced to show preferences that cycle indefinitely. Other experiments have focused on what is called "preference reversal." These experiments ask individuals whether they prefer prize x or prize y. Shortly thereafter, these same individuals are asked whether they prefer some probability of x or the same probability of y. The same individuals who prefer x to y often prefer the probability of y to the probability of x. Preference reversal, like violations of transitivity, suggest there is no well-defined preference ordering. They appear to undercut any well-specified notion of instrumental rationality.[16]

Experimental economics, however, does not necessarily imply nihilistic conclusions. Much of the literature shows how markets can operate even when individuals are not rational as in standard theory. To provide a simple example, laboratory markets rapidly reach a coherent equilibrium, even when the traders fail numerous other tests of individual rationality, including transitivity. These results contradict mainstream analysis, which treats individual rationality as an underpinning of market operation. The results imply that markets have some means of working even when rationality is not present, which implies that standard theory, in addition to its failings on the rationality question, does not provide a very good account of markets.

Twenty or thirty years ago experimental economics was considered to be a highly speculative method of investigation. Many economists thought it was outright dubious. Today experimental methods are used commonly at the highest levels of the profession. Most issues of the American Economic Review, the profession's flagship journal, contain at least one experimental piece.

Experimental economics is also tied closely to game theory, though the two branches of economics often stand in conflict. Experimentalists are well aware of the failings of game theory, and hope to fill in the empty boxes with results from the laboratory. The

[16] See David M. Grether and Charles R. Plott, "Economic Theory of Choice and the Preference Reversal Phenomenon," The American Economic Review, 69, (1979): 623-638.

experiments typically find more cooperation than game-theoretic accounts would have predicted. Experimentalists wish to introduce a desire for cooperation into economics, whereas the game theorists are more skeptical about the lab results, often stressing the differences between the experiments and real world conditions.

Of all the branches of economics, experimental economics probably goes furthest in the empirical direction, with regards to rationality. Experimental economics regards all or most stipulations of the rationality postulate as empirical in nature and as potentially falsifiable. The surprising lesson is that what we assume about rationality does not seem to matter for many economic contexts. Markets economize on rationality, whatever that concept may mean, more than neoclassical economics usually implies.

h. Economic imperialism

Some economists argue that the economic method should be extended to many or all of the other social sciences, such as political science, sociology and anthropology. This movement has already made strong progress in political science, some inroads in sociology, and is starting to influence anthropology.

Gary Becker's work on the economics of the family is a prototypical example of economic imperialism. Becker has sought to bring the understanding of the family under the rubric of economic reasoning. For each family member he postulates a utility function and constraints, and then uses various models to examine family interactions. Using this method, Becker and other researchers have studied the effects of familial division of labor, familial altruism, government policies towards children, and government welfare policies.[17]

---

[17] See Gary S. Becker, *A Treatise on the Family*, Harvard University Press, 1993. For a more general treatment of economic imperialism, see Gerard Radnitzky and Peter Bernholz, editors, *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics*, Paragon House, 1987.

Becker's work has proven among the most controversial applications of rationality postulates. Applying rationality postulates to the family differs from most other parts of economics. Most significantly, the economist cannot resort to the usual assumption of profit-maximization, as is applied to business firms. The objective function of the family is less well-specified, and is harder to link to any empirically measurable variable. In addition to its economic functions, the family provides love, affection, the chance to raise children, and many other personal goals. Furthermore, while families engage in many implicit forms of trading, they eschew explicit trading for many choices. Most families do not bid as to who gets to watch their favorite television program on a given evening. Economic models are well-suited to handle trade, but have a harder time defining the operation of non-trade allocative mechanisms.

I am not aware of a philosophically sophisticated defense of economic imperialism (advocates tend to believe in the primacy of practice), but the following defense might be offered in response to these problems. Rational choice explanations, in the realm of economic imperialism, are most defensible when we think of them as complements to alternative approaches, rather than substitutes. This will weaken some of the more extreme pretensions of the imperialistic approach, but also place it on sounder footing.

We start with the notion of explanation in general, however that term might be understood. Trying to explain family outcomes is bound to be problematic, for the reasons specified above. The vagueness of family "objectives" is not just a problem for economic modes of reasoning, it renders all explanation more difficult. Given these problems, however, rational choice approaches do make a particular claim about how to approach explanation. Specifically, when offering an explanation, it is useful to break that explanation into two parts: claims made about preferences and claims made about constraints. Once this breakdown is made, we have a rational choice attempt at explanation, motivated by the pragmatic grounds discussed earlier in this paper. The rational choice approach may help us generate new ideas about the family, and new ways of testing old ideas.

If we translate this method into the context of the family, the rational choice theorist believes that this parsing will yield useful results. Insight may arise through (at least) three mechanisms. First, the categories may allow us to more easily observe connections between apparently unrelated events. Some of the problems of the family, for instance, such as "moral hazard," may resemble the problems of the welfare state. Second, the categories may make causal chains easier to isolate. An understanding of division of labor in the household, for instance, may help us understand when dowry payments are likely to be positive or negative. Third, the categories may make it easier to test hypotheses against the data, however problematic this enterprise may be. Andrei Shleifer, for instance, considers whether older individuals use their bequests to manipulate the behavior of the young; his basic insights came out of a rational choice model.[18]

i. Fixed preferences?

Some researchers treat the constant preference assumption as part of the core of rationality and the economic approach, especially as applied to other disciplines. Economists typically assume that preferences are constant over time. George Stigler and Gary Becker go further and insist this practice defines the economic or rational choice method. While most economists reject this extreme view, in practice economists remain skeptical of explanations that invoke changing preferences. It is commonly argued that changing preferences can be used to "explain anything."[19]

This view is not a rationality postulate in the full-blown sense of the word. Nonetheless, in the eyes of some, it is an important auxiliary hypothesis for what it means to apply a theory of rationality. In this view, adherence to a rationality-based approach must rule

---

[18] On these points, see Theodore C. Bergstrom, "A Fresh Look at the Rotten Kind Theorem -- and Other Household Mysteries," Journal of Political Economy, 93 (1985): 1045-1076, and Douglas B. Bernhaim, Andrei Shleifer, and Lawrence H. Summers, "The Strategic Bequest Motive," Journal of Political Economy, 34 (1985): 1045-1076.
[19] See George S. Stigler and Gary J. Becker, "De Gustibus Non Est Disputandum," American Economic Review, 67, (1977): 76-90.

out changing preferences, since a rational choice model -- if based on the idea of given preferences -- could not explain how preferences could ever change.

Economists are more likely to accept explanations based on varying preferences across individuals than explanations based on changing preferences for a single individual. Principal-agent models, for instance, commonly postulate one person who is risk-neutral and another who is risk-averse. Similarly, it is obvious that some people like classical music, others like rock and roll, and so on. Even in these cases, however, Stigler and Becker stake out an extreme position. They suggest that the idea of differing human capital endowments be used in lieu of the concept of changing preferences. In their terminology, all individuals have the same tastes for music, but some individuals have better complementary endowments in their human capital for enjoying classical music. Most other economists see this difference as a semantic one rather than a real one. Thus they will accept varying preferences, across different individuals, but they are reluctant to accept changing preferences over time for a single individual.

Even changing preferences are making inroads in the research community. Fischer Black invoked changing preferences as a fundamental cause of business cycles. Contemporary consumer theory considers individuals with non-hyperbolic discounting, in other words individuals who seek to postpone their pains and commitments as they draw near. The Christmas boom in economic activity also seems hard to explain without invoking changing preferences across the seasons. The same might be said for fashion cycles, or increasing demand for safety as most individuals, especially males, age. Note that a purely economic approach would imply that the young would take the fewest risks, since they have the most years left to lose. The data show convincingly that young males take the most risks, and take correspondingly fewer risks as they age.[20]

The stable preference assumption may have been an unfortunate addition to the basic economic method. Economists typically reject preference-based explanations on the

[20] On some of these ideas, see Fischer Black, *Exploring General Equilibrium*, MIT Press, 1995, and Robert Barsky and Jeffrey Miron, "The Seasonal Cycle and the Business Cycle," Journal of Political Economy, 97 (1989): 503-534.

grounds that "anything can be explained by changing preferences." Furthermore changes in constraints are seen as potentially measurable, and thus offering potentially testable hypotheses, whereas changes in tastes are seen as harder to pin down. These criticisms, however, are misleading for at least two reasons. First, we often have actual data about preferences and preference changes, against which preference-based explanations can be tested. Second, many economic phenomena cannot be explained very easily on the preference side, no matter how much wiggling and squirming is done. An increase in the real wage rate, for instance, will cause many entrepreneurs to substitute machinery for labor. This phenomenon can be explained easily on the constraint side, but not very easily on the preference side; examples of this kind could be multiplied. Invoking changing preferences is therefore not a universal license for abuse.[21]

j. Satisficing

The early work of Herbert Simon, Richard Cyert, and James March led to the idea of "satisficing." Satisficing refers to the idea that individuals do not seek the very best outcome, but rather they stop once they find an outcome that is "good enough." The concept of satisficing came originally from the realization that most maximizing problems are extremely complex and often lead to simple "rule of thumb" solutions.

Many economists do not regard the satisficing postulate as an independent alternative to a rationality assumption. Consider a simple example of what might be called satisficing. An individual is searching in a grocery store for a good piece of fruit, say a cherry. Most people will not spend their day searching through the entire bin of cherries. Rather at some point they will be content to pick a cherry that is "good enough." It is easy to see how this example can be reinterpreted as an application of rationality. The searching individual will face an opportunity cost of time, and the benefits of getting a good cherry are only so large. The individual will make some calculation, either explicit or implicit,

[21] Bryan Caplan, "Stigler-Becker versus Myers-Briggs: Why Preference-Based Explanations Are Scientifically Meaningful and Empirically Important," working paper, George Mason University, 2001, argues that economists should be willing to consider varying tastes across individuals, and sometimes changing tastes over time.

about how much time is worth devoting to the cherry search. It is well known that an optimal strategy often consists of a "stopping rule." That is, the individual should stop searching and pick the best option available, after some finite period of time.[22]

Defenders of the satisficing approach do not regard this criticism as decisive. In their view satisficing has more predictive power, even if it does not have well-defined microfoundations. Some taxi drivers, for instance, simply may stop taking fares once they have accumulated a certain amount of money for the evening.

The economists who oppose the satisficing model typically point out that we can always define such behavior as rational in one set of terms or another. We could say, for instance, that an "income effect" has kicked in. Once the driver has a certain amount of money, he may prefer to consume leisure instead of working. Defenders of satisficing, however, may still claim superior predictive power for their approach. Some taxi drivers may still behave in closer accord with the satisficing model. The mainstream model implies that high income or wealth -- regardless of its source -- causes the taxi driver to work less. But if the taxi drivers receive more money from some other source, perhaps a gift, they often drive their cab just as much. So perhaps it is not the "income effect" (or whatever other mechanism might be postulated) at work. The taxi drivers may be making context-laden decisions that are best described by the satisficing construct. Many drivers might have an inclination to stop, once they have accumulated a certain quantity of fares on a given evening. They feel they have done "a good night's work" and feel no further need to bring home money that evening.

k. Computation

The last two decades have brought increasing interest in computational models in economics. Computational models have a number of roots, but in part they attempt to

---

[22] For a philosophic examination of satisficing, see Michael Slote, *Beyond Optimizing*, Harvard University Press, 1989. Michael Slote, "Satisficing and Optimality," Ethics, 109 (1998): 67-93, provides a critique of the philosophic underpinnings of many satisficing models.

improve on the satisficing concept. Satisficing approaches typically postulate an absolute stopping rule at some point. Computational approaches attempt to derive when an individual will stop calculating, depending on the complexity of the problem at hand. So we can think of computational approaches as providing a kind of microfoundations for satisficing. Under these microfoundations, however, we do not always generate a fixed stopping rule, as under satisficing.

Rather than allowing all solutions to a problem, the computational research agenda focuses on which of these solutions might be computable to rational agents with limited abilities. Kenneth Binmore has applied this point to the choice among solution concepts in game theory (see the discussion of game theory above), and Thomas Sargent has tried to integrate the idea of computability into monetary theory and macroeconomics. Sargent has also worked to integrate the idea of economic computation with the theory of artificial intelligence.[23]

Computational theories, like experimental economics, have not arisen in a vacuum. In part they are designed to address some of the weaknesses in previous accounts of rationality. For instance, Binmore's computational approach tries to resolve the problems of game theory, as discussed above. Sargent's computational work attempts to work around the limits of rational expectations macroeconomics.

Note that the work on computation moves closer to what philosophers sometimes call procedural rationality. Rationality now becomes defined in terms of some computational algorithm.

Most work in computational economics is new, highly mathematical, still resists easy summary, and is relatively inaccessible to non-experts (which includes this author).

---

[23] See Kenneth Binmore, "Modeling Rational Players, Part I," Economics and Philosophy, 3, (1987): 179-214, and Kenneth Binmore, "Modeling Rational Rational Players, Part II," Economics and Philosophy 4, (1988): 9-55, and Albert Marcet and Thomas J. Sargent, "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," Journal of Political Economy, 97 (1989): 1306-1322.

Nonetheless it is easy to see how the emphasis on computability puts rationality assumptions back on center stage, and further breaks down the idea of a monolithic approach to rationality. The choice of computational algorithm is not given a priori, but is continually up for grabs. Furthermore the choice of algorithm will go a long way to determining the results of the model. Given that the algorithm suddenly is rationality, computational economics forces economists to debate which assumptions about procedural rationality are reasonable or useful ones.

The mainstream criticism of computational models, of course, falls right out of these issues. Critics believe that computational models can generate just about "any" result, depending on the assumptions about what is computable. This would move economics away from being a unified science. Furthermore it is not clear how we should evaluate the reasonableness of one set of assumptions about computability as opposed to another set. We might consider whether the assumptions yield plausible results, but if we already know what a plausible result consists of, it is not clear why we need computational theories of rationality.

To make matters even more difficult, human beings appear to have vastly different computational abilities in different activities and different spheres of life. To give a simple example, most individuals have relatively good abilities to remember faces, voices, and to sort through various aspects of interpersonal relationships. It has proven very hard to get computers to do these same tasks. In more abstract, less personal contexts, individuals do not have nearly the same abilities. Hardly anyone can perform complex long division in his or her head. This point implies that computational theories of rationality will probably fail to settle on a single, simple account of how computational rationality proceeds.

Currently computational theories of rationality are an open box. It is not clear how they will develop. Nonetheless they illustrate the high and growing diversity of the rationality concept within economics.

## IV. Concluding remarks

Our examination of how economists use the rationality postulate suggests diversity more than anything else. There is no single set of assumptions that economists make about rationality, and no single set of attitudes about the role of rationality assumptions in economic theory and practice. Rather we have seen the rationality postulate, and attitudes toward it, evolve in different directions in different fields of economics. Some fields view the rationality postulate quite broadly, while other fields identify it with some very particular assumptions (macroeconomics). Some fields regard rationality as a tautology (consumer behavior under certainty), while others regard it is potentially testable (consumer behavior under uncertainty). Game theory places great importance on the rationality concept, as does computational economics, whereas empirical labor economics and experimental economics imply that the rationality concept does not matter so much.

We therefore need an account of the rationality postulate that can explain its varied uses across fields and subfields of economics. To return back to our list in section II, the first three views (tautology, transitivity only, and descriptive) must stand as incomplete. They may account for how rationality is used in particular instances, but without providing an account of rationality in economics more generally. The most serious contender thus remains the pragmatic view that rationality, and various rationality assumptions, serves as useful organizing categories. Given the plurality of investigative methods in economics, this will imply that particular rationality assumptions take numerous forms.

This look at the rationality postulate may not satisfy those critics who, in any case, do not like the content of economic theory. But at the very least the rationality postulate is a less vulnerable target of criticism than is usually thought. I do not mean to suggest that economic uses of the rationality postulate cannot be improved upon. I have, however, attempted to revise the picture of the rationality postulate as an easy whipping boy for critics. Economic practice, in this regard, is more sophisticated than the views of any single economist. This is an appropriate result for a science that emphasizes invisible

hand mechanisms and the limitations of central planning. I see many things wrong with contemporary economics, but the rationality postulate is not near the top of my list. At the very least, I hope this paper induces some of the critics of economics to be blunter and more specific about their objections.

I believe that economists hold a default view about their own use of rationality concepts, although this view is rarely if ever articulated. They see the rationality concepts in use as the result of a competitive process, and thus the best available practice, at least given our current state of scientific information. A rationality concept, in this regard, is like a market price or a management practice. Not every price or practice is optimal when compared to full information, but the competitive process nonetheless gives us as good a menu of choice as we are likely to have. It provides the best available options, relative to the imperfections in our information. I believe this optimistic view, whether right or wrong, is the underlying reason why practicing economists pay so little heed to methodology.[24]

This underlying model of economic science has a surface simplicity. If we refer back to rationality postulates as "inputs" into scientific research, a competitive process (under certain conditions) allocates inputs efficiently. Economists, who tend to accept efficiency as a relevant standard, therefore believe that this competitive process yields an approximation of good science, and this of course involves the rationality postulate in its diverse forms.

I do not raise this possibility to evaluate it, but rather to note that any effective criticism of economics must start with the institutions that produce (and evaluate) economics. Methodological criticisms alone, especially if they focus on rationality, are unlikely to be very persuasive. In the meantime we should deemphasize monolithic attempts to

---

[24] Sherwin Rosen, "Austrian and Neoclassical Economics: Any Gains From Trade?" The Journal of Economic Perspectives, 11, (1997): 139-152 comes closest to stating this view. See Leland B. Yeager, "Austrian Economics, Neoclassicism, and the Market Test," The Journal of Economic Perspecitves, 11, (1997): 153-165 for a response.

characterize (or criticize) the method of economics and recognize the strongly plural character of research practice.