

It's Convenience, Not Ontology: The Modest Pretense of Social Science Modeling

One of the earliest applications of what is known today as rational choice theory was developed to defend faith. "Pascal's Wager" is widely known, and I've often heard versions of it deployed by believers in basic form, if not in name, when explaining why they believe, or in appealing others to believe.

The wager drew on a primitive form of integrating probabilities with rational choice theory, today called expected utility theory. Pascal's original argument in *Pensées* differs in telling ways from the versions commonly deployed today in everyday conversations (which we'll return to below), but the argument as I typically hear it today goes something like this: If God exists and one believes, the believer would experience infinite bliss or happiness for eternity. If God did not exist and one believed, that would only entail finite loss in this life - the wealth that one gave away to charity, or worldly pleasures foregone. And if God existed and one did not believe, then one will experience infinite torment. (Pascal doesn't actually include the possibility of going to hell in his argument, but all of the versions I've heard include it.) Whatever (fixed) probability one assigns to the possibility that God exists, no matter how small, the receipt of infinite gain and the avoidance of infinite loss outweighs any temporal pleasures forgone, and so, the argument goes, one should believe.

Pascal's argument prompted centuries of discussion and criticism; some of those criticisms apply equally well to rational choice theory as deployed in modern social science. Nonetheless, given the popularity of the argument among the faithful, there is perhaps an irony that the broader application of rational choice approaches to study human behavior today now seems more to disturb believers than to comfort them. Steven Green, an economics professor at Baylor University, reported in an essay of his on rational choice theory that most of the Baylor faculty and students attending a lecture by Laurence Iannaccone, an economist who applies rational choice theory to study religion, were troubled by his deployment of the approach "to understand the evolution of the doctrine of hell in Christian thought."

Economists have taken the brunt of criticism for rational choice (or, without derogation, just "rat choice") theory, given the longer period of time, and the great breadth, that rational choice analysis has been applied and developed in economics. Calvin College economist John P. Tiemstra, for example, criticizes mainstream "neoclassical" economists for being committed to the proposition that "all human behavior is driven by the self-interested pursuit of material gain." Aside from the crabbed view of the human person that this assumption suggests, for Tiemstra, the neoclassical account of human behavior fails to capture the breadth of motivations that prompt human behavior. "[S]ince all humans are created in the image of God, and hence are by their very nature religious and moral beings, the neoclassical model fails to capture an essential dimension of human behavior."

The question of the ability of rational-choice models to reflect human behavior accurately aside, the broader criticism is that rational-choice theory warps souls. Tiemstra writes "By teaching people the utilitarian ideology of neoclassical economics, economists encourage the very kind of self-interested, greedy behavior that is inconsistent with the demands of the Christian life . . ." This is a form of concern expressed by Christians and other religious believers regarding social science more generally: the itch of grand theory in the social sciences to get "behind" human behavior ostensibly reduces human behavior and action, such as religious or moral behavior, to mere epiphenomenon. The essence of the human person is reduced from an acting moral agent to no more than a combination of unseen forces at work manifesting themselves in the actions, beliefs, and behavior of people.

For example, in the 1981 film, *The Chosen*, Danny Saunders, son of the leader of a Hasidic congregation in Brooklyn, receives a letter of admission from Columbia University. Danny's younger brother looks to their father and asks, "What's at Columbia?" Reb Saunders, whose voice drips with the bitterness of

losing his son to another god, answers, “*Psychology*.” Critics reject rational choice theory also as a normative failure and as a positive failure.

“Bad science,” the critics maintain, and, even worse, “bad ethics.”

R.R. Reno summarized the fear more recently, writing:

In my twenty years of teaching undergraduates, more often than not it was a class in the social sciences that challenged the faith of students, not a class in biology. . . .

It’s not hard to see why faith finds itself challenged by modern social science. Christianity and Judaism privilege sacred texts as the ultimate horizon for our historical, social, and moral imaginations. By contrast, Freud saw monotheism as emerging out of an Oedipal psychodrama enacted in the distant past. Weber interpreted religion in terms of an oscillation between charisma and institutionalization. These and other theories have their secular critics, but the general structure of explanation was (and remains) constant. As a modern tradition of inquiry, social science wants to get underneath religion, as it were, explaining it in terms of something more fundamental.

. . . Almost all social scientists make this claim, even if only implicitly. Need it be so?

Reno aims to be fair and to leave the answer open with whether it “Need . . . be so” that “almost all” social scientists make this claim. Still, my own observation about what most social scientists do is that they – we – are mainly engaged in building up practical wisdom about human behavior rather than in attempting to build grand theories. This more-modest aspiration in fact does not interest most people as much as head-line grabbing, grandiose (and, usually, untestable) theories.

Most social scientists do not engage in developing “Grand Theory.” Many think pursuit of Grand Theory is not particularly illuminating and, to be honest, not particularly interesting (except perhaps for generating books sales). Often times observers outside of social sciences see only Grand Theory, and so assume that that is what social science is about. Newspaper reporters call me up all the time inviting me to pontificate on this or that. Almost always I respond by saying I don’t know any data that can speak to the question, and tell the reporter that his or her answer to the question would be as good as mine. The reporter invariably responds, “Yeah, but I can’t quote myself.”

Outsiders then get in the habit of thinking of social science in this way. A couple of years ago I stepped into my daughter’s high school IB economics class. The teacher framed the upcoming year of study as something of a grudge match between Keynes versus Hayek. My (unspoken) thoughts were, first, “Surely they must devote some time to microeconomics.” My second thought was, “How boring.” (For the sake of my daughter I didn’t ask the teacher whether he really thought that macroeconomics had not developed since the 1930s.) Aside from the issue of whether the instructor actually understood Keynes or Hayek – or simply used them as placeholders for ideology – in substituting politics for analysis, the instructor was missing the opportunity to motivate the study of economics in its own right – as an area of study in which scholars can marvel at puzzles of human economic behavior, and seek to understand the whys and wherefores.

As a result of the high-level attention accorded to Grand Theory over much-more modest, local claims of middle-range theory, there is an itch by many critics, and, truth be told, by some proponents, to think of the modest epistemological claims of many social science theories, including rational choice theory, as making the larger claims of grand theory. Assumptions are turned from being made for analytical convenience into broader ontological claims about the nature of being human. This need not be so. And

doing so results in missing the ability of social science to add to the sum of practical wisdom about humanity rather than implicating grand questions of existence.

Because I know rational choice theory best – and because it’s a particular lightning rod both in the academy and without, I’ll focus on that.

The problem is this: Many critics (and, to tell the truth, some practitioners), think that rational choice theorists engage in building comprehensive systems of thought seeking to explain, as one critic put it, “all human behavior.” But they don’t. At least they don’t necessarily. Rather, rational choice theory is one tool which scholars can bring to puzzles of human behavior in an attempt to understand and explain those puzzles. In doing so, scholars can build up practical wisdom about behavior, events, and human institutions. Too often, the more-modest task of social science is caught up in criticism aimed, appropriately to my mind, at “grand theories” that are based on little more than speculation and assertion. But to say *something* about the human person we don’t need to say *everything*. Many social scientists focus not on developing grand intellectual schemes but rather on, as it were, more-local matters of human behavior. Simplifying assumptions are helpful, and necessary, to build up intuition about behavior in these more local regions of knowledge. Indeed, this mode of developing practical knowledge is not limited to scholars; it’s applied by people throughout history in trying to understand ordinary matters of everyday life. Social science is no more than a refinement of this habit. To be sure, simplifying assumptions mean that human behavior in an area can be represented by a mathematical function with just a few elements in it, something that may be expressed generally as $U(x,y,z)$. This attracts its own criticism – arguments that the models are unrealistic and reductionistic. And, of course, they *are* unrealistic and reductionistic. But that’s a *virtue*, not a vice.

Critics of rational choice theory often impute too much ontology to what are intended only to be models of human behavior. But to understand why – or how – that is an answer, we need first understand a bit about rational choice theory and how it is deployed in social science today. Understanding, I think, defangs most of the suspicion of rational choice theory and mathematical modeling, and points to how it is that many, even most, forms of doing social science do not actually implicate religious (or non-religious) commitments. Rather, it is one way to build practical wisdom about human interaction.

Defining Rational Choice Theory

First, we need a workable definition of rational choice theory. The essence of rational choice theory is no more than the insight that people respond to incentives in often predictable ways. This is not a modern innovation, and one does not need to be committed to a strong version of utilitarianism to hold it. Much of the analysis of the writers of, say, the *The Federalist* are arguments consistent with a form of “soft” (i.e., non-mathematical) rational choice theory. In *Federalist 51*, for example, James Madison writes that

the great security against a gradual concentration of the several powers in the same department, consists in giving to those who administer each department the necessary constitutional means and personal motives to resist encroachments of the others. The provision for defense must in this, as in all other cases, be made commensurate to the danger of attack. Ambition must be made to counteract ambition.

His discussion of the influence of factions on politics in *Federalist 10* also discusses incentives that persons, and groups of persons, respond to.

Rational choice theory seeks to refine this notion, not because one must posit that people respond always and only to incentives, but because enough of us do so enough of the time that one can gain significance leverage on many types of behavior by inquiring into the incentive structures the persons inhabit.

Part of the difficulty lies with a divergence between the everyday definition of “rationality” and the specialized meaning of “rationality” for rational choice scholars. The notion of “rationality” in rational choice theory is very thin relative to the popular conception of the term. Rationality means little more than that humans pursue goals. More specifically, rationality means that a person can rank alternatives (which includes ranking alternatives as “tied”) and those rankings are transitive. So if I have a choice between apple, blueberry, and cherry pie, I can rank those choices (which includes being indifferent between two or more choices) and my choices are transitive. So if I prefer apple to blueberry, and blueberry to cherry, then I will also prefer apple to cherry.

This notion of rationality is so thin that even persons that most people would think of as “insane” are perfectly rational for the rational choice theorist: A person can believe himself or herself to be Napoleon, but as long as the person can rank outcomes and that ranking is transitive, then that person is rational according to rational choice theory. If a person cannot rank choices at all – meaning that given options A and B, a person cannot say “I prefer A to B, B to A, or I am indifferent between A and B,” or if that person’s choices are intransitive (“I prefer A to B, and I prefer B to C, but I prefer C to A”), it does not challenge the broader usefulness of the theory, but it does mean that the theory doesn’t apply in those instances. To claim that a theory generates useful insights is not to claim that it generates insights universally. Indeed, the question of when individuals act rationally and when they don’t, as well as why they do in certain areas but not in other areas, is itself an interesting intellectual puzzle.

A common mistake of many critics (and some rational choice theorists as well), is to equate rational choice theory with the rational choice models that economists develop to understand specifically economic behavior. Critics object that human behavior cannot, and should not, be reduced to monetary self-interest, and that market exchange is not an appropriate framework to understand lots of human interaction.

They’re right about that. But while economics is the 800-lb gorilla in the rational-choice room, there are other sizeable primates as well, particularly in political science and sociology. Rational choice theory limits human goals neither to cash nor to material or economic goals more generally, and the theory does not require that the analysis think of human interaction only through the lens of market-like exchange.

Scholars in other social sciences like political science, international relations, and sociology apply rational choice models that posit goals other than the pursuit of material or financial gain. In political science, for example, people are often assumed to be motivated by multiple goals including policy goals, winning elections, and retaining office.

There’s a defense for economists focusing on material gain in their models that I’ll discuss below. But even when we get past the idea that rational choice theory necessarily posits cash or material gain as the goal which people pursue, the mathematical expression of the utility function to model human goals in rational choice models provokes a similar criticism: that rational choice models are reductionistic, that is, that they reduce complicated human motivations and lives to something like this: $U(x,y,z)$. Perhaps this is the way that Dr. Strangelove thinks of human beings, the thought goes, but it is bad both for science and for humans to think of human behavior this way.

There’s where critics get it largely wrong.

Reductionistic assumptions are made for convenience, not as ontology. This is a critical point to understand. Secondly, humans necessarily simplify reality when trying to understand parts of it that are salient to them. Let’s consider each of these points in turn.

First, simplifying assumptions are made for convenience – they are made to promote understanding; they are not intended as holistic descriptions of reality. Indeed, it is the very *unrealism* of the models that make them helpful to understand reality. (Note that this is not the same as saying that the mere feature of being unrealistic is what makes modeling useful.)¹

To be sure, abstraction in models can be poorly done – and that’s often not recognized sufficiently in arguments over modeling. But modeling can also be useful. No theory really aims to replicate reality. Indeed, if one wants reality in all its complexity, then all one needs to do is look out the window. But what’s “out there” is intrinsically too complicated to be understood. (I might suggest that’s one reason that art is inescapably necessary for humans.) The point of analysis is to understand reality. And humans do that, inescapably, using words, pictures and illustrations, and sometimes mathematical symbols. The unrealism of models is not an intellectual vice, it is their virtue. As with any abstraction (whether prosaic, mathematical, graphical or some combination), the value of the abstraction hinges on the model’s usefulness. But it is the very *unrealism* of a model, if properly constructed, is what makes it useful.

For an every-day example, consider the instrumental abstraction of a city street map. It is obvious that street maps are highly reductionistic representations of the real topography of a city. Maps utterly distort what is *really* there and leave out numerous details about what a particular area looks like. But it is precisely *because* the map distorts reality – because it abstracts away from a host of details of what is really there – that it becomes useful. A map that attempts to portray the full details of a particular area would be too cluttered to be useful in finding a particular location or too large to be stored in a glove compartment. So it is with the game theoretic models. They seek to abstract away from a host of details that are not relevant to the phenomenon under study. It is the very abstraction that permits us analytically to “hold everything else equal” and to focus on the most salient aspects of the phenomenon. Indeed, in the social sciences where there are additional ethical and practical constraints on experimentation, models are a relatively nonintrusive means to try to get a grasp on understanding what’s happening “out there.”

Of course, everything is not always equal and omitted details can matter. Just as with street maps, there can be better models and worse models for particular purposes. Maps that abstract away from too much detail won’t be useful in finding a particular street or address. Models that abstract away from too much pertinent detail aren’t useful for understanding the phenomenon being studied. The trick, then – which is as much matter of aesthetic taste as it is intellectual judgment – is to develop models that provide just enough detail to be useful for their intended purpose without being so complex as to confuse rather than illuminate. Further, maps can be initially off-putting abstractions from reality for those unfamiliar with them. Many maps use specialized symbols and representations that require some time and effort to learn. Nonetheless, if someone were to point at a map, point out the window and sneer that “that doesn’t look anything like what’s out there.”

Of course, simply because maps can be simplifying abstractions from reality doesn’t mean that we always use them. Sometimes it’s easier to write directions to your house in words; sometimes it’s easier to draw a map. Again, the basis of your choice is instrumental – what best accomplishes the goal of directing someone to your house successfully. So it is with models. There is nothing inherently wrong with a model

¹ Much of this section is drawn from James R. Rogers, “The Gimlet-Eye of Social Science,” in *Religion and the Social Sciences* (R.R. Reno, B. McClay, eds.). Cascade Books (Eugene Oregon: 2015): 88-98, and James R. Rogers, “A Primer on Game Theory,” in *Institutional Games and the Supreme Court* (J. Rogers, R. Flemming, J. Bond, eds). University of Virginia Press (2006).

that is written entirely in words. Sometimes, however, it's more convenient for a given purpose to use symbols and equations. Again, the choice is instrumental. To be sure, as with the stylization of maps, sometimes the abstractions may be initially off putting to a person untrained in what the symbols mean. But as with the increasing usefulness of abstract maps relative to written directions when routes become complicated, attempts to understand more complex phenomena can benefit from the convenience and parsimony of more mathematically abstract models relative to more prosaic models. And, of course, reasonable people can have different personal preferences for written directions relative to mapped-out directions. Reasonable people can also disagree over whether a particular map is detailed enough to serve its purpose (or whether the map contains superfluous and distracting details). So, too, scholars have different preferences over using more prosaic models relative to more mathematical models. Scholars can further disagree whether a specific model, whether prosaic or mathematical, is, on the one hand, sufficiently detailed to help us understand the phenomenon under investigation or, on the other hand, so detailed that the phenomenon is obscured rather than illuminated. To the extent that mathematizing concepts in a model help us effectively to express, manipulate, or reason from those concepts, it's a useful enterprise. To the extent that mathematization doesn't achieve a useful purpose, then there's no reason to mathematize an otherwise prosaic model.

Indeed, there is a temptation to the modeler. In a "medium becomes the message" sort of way, the fun of representing things mathematically can entice the analyst to seek to mathematize when there is no reason to mathematize. One of the first "original" models I developed (and long lost) applied a fixed-point theorem to prove the existence of equilibrium in the use of hot and cold water among separate apartment dwellers who were all attempting to optimize the temperature individually, but whose individual decisions affected the temperature of the water in the other apartments as well. On occasion, this temptation gets the best of the modeler, and arguments that are perfectly sensible in non-mathematical form are nonetheless "mathematized" with the thought that that, by itself, makes them worthy of publication. It doesn't.

Although more controversial, the role of self-interest and methodological individualism in modeling play a similar role as mathematical modeling. These are assumptions made for analytical convenience, not as universalistic claims as to the nature of reality or, more specifically, about the nature of humanity.

The point is that if the assumptions assume away what one wants to understand or explain, then they're poor assumptions. The question is not whether the assumption of self-interested behavior is always true (although more about that in a moment), the question is whether the assumption is correct enough in areas of human behavior of substantive interest so that the analyst gets insight into human behavior in these areas.

We All Make Simplifying Assumptions

We all make simplifying assumptions about human behavior to focus on a puzzle or topic that interests us. In motivating the usefulness of a separation-of-power system, for example, James Madison highlights the role of self-interest as the most pertinent human desire to be understood and at least channeled, if not controlled. In *The Federalist* No. 51, for example, Madison develops a political analogue to Adam Smith's argument that, through markets, people's self-interest leads them as though by an invisible hand to promote the public good. Madison argues that the structure of political institutions in separation-of-power systems channels ambition and self-interest so that they reduce the risk of tyranny rather than invite it:

But the great security against a gradual concentration of the several powers in the same department, consists in giving to those who administer each department the necessary constitutional means and personal motives to resist encroachments of the others. . . .

Ambition must be made to counteract ambition. The interest of the man must be connected with the constitutional rights of the place. . . .

This policy of supplying, by opposite and rival interests, the defect of better motives, might be traced through the whole system of human affairs, private as well as public. We see it particularly displayed in all the subordinate distributions of power, where the constant aim is to divide and arrange the several offices in such a manner as that each may be a check on the other -- that the private interest of every individual may be a sentinel over the public rights. These inventions of prudence cannot be less requisite in the distribution of the supreme powers of the State.

While we can contest whether the separated institutions that share power in the U.S. constitutional system exercise too little or too much check and balance over the other institutions, Madison's argument here seems intuitive and straight forward. With government officials holding the power of the sword, the unchecked government official might use his or her power to pursue private interests rather than the common good. Madison recognizes that not all, or even most, government officials would succumb to the temptation to use their power for personal gain. But he also recognizes that those officials who do want to use their power in that manner constitute a special problem not only for democratic government, but for any form of government. Therefore his discussion at this point (and others throughout *The Federalist*) highlights, or assumes, public officials as self-interested, ambitious people who need to have their power checked.

Madison (and Hamilton, and even oft-forgotten John Jay) do not posit that humans, or even the narrower set of politicians, are actuated *only* by self-interest and ambition. They often discuss how politicians can and do reflect honor and virtue in their work as well. At critical points in developing the case for the separation of powers in the U.S. Constitution, however, they focus on developing a form of government that does not require virtuous people to be in charge to survive, but can survive tolerably well, at least for a time, even when the venal are in charge. For example, Madison argues in *The Federalist #10*,

It is vain to say that enlightened statesmen will be able to adjust these clashing interests and render them all subservient to the public good. Enlightened statesmen will not always be at the helm. Nor, in many cases, can such an adjustment be made at all without taking into view indirect and remote considerations, which will rarely prevail over the immediate interests which one party may find in disregarding the rights of another or the good of a whole.

Indeed, in discussing "the most common and durable source of factions has been the various and unequal distribution of property," Madison invokes an early version of the common assumption in economic models that people pursue wealth at least cost to themselves:

The apportionment of taxes on the various descriptions of property is an act which seems to require the most exact impartiality; yet there is, perhaps, no legislative act in which greater opportunity and temptation are given to a predominant party to trample on the rules of justice. Every shilling with which they overburden the inferior number is a shilling saved to their own pockets.

Madison considers the test case of how the institutions work with ambitious, narrowly self-interested politicians *not* because he thinks it is an accurate, full-fledged anthropology of the nature of humans. For him it is the critical test case, however, and so it suits his analytical purposes as an institutional designer and as a political engineer, to test his design under realistic, albeit extreme conditions. If it survives the

acid test, the argument would go, then it's a fit political design for conditions that would almost always be less extreme.

We could take issue with Madison's risk aversion, asking what cost might there be to insure against the extreme political outcomes he fears. Nonetheless, his approach is not unreasonable. So while not immune to criticism or extension, Madison highlights a distinctive feature of human motivation not because he thinks that is all there is to human motivation, but because he thinks it's a salient feature of human motivation for the cases he's most interested in considering. It's not supposed to be an exhaustively, or even a largely, realistic account of the human person. It is rather an analytical convenience given the topic of focus. While we can be thankful for as often as we have a Cincinnatus in charge, that case is, for Madison, an uninteresting case for the issue of institutional design: "If angels were to govern men," he famously wrote,

No government would be necessary. If Angels were to govern men, neither external nor internal controls on government would be necessary. In framing a government which is to be administered by men over men, the great difficulty lies in this: you must first enable to government to control the governed; and in the next place oblige it to control itself.

Selfishness versus Self-Interest

Rational choice theory receives a great deal of criticism because of its postulate of self-interested behavior. This is understandably, but not accurately, taken as identical to selfish behavior. This distinction is of more than rhetorical interest.

Economics, as usual, is the lightning rod for this criticism. This by itself should give us pause, as many rational choice models in political science posit that political actors pursue policy goals rather than pecuniary goals. (Other times, modelers posit that politicians do or can seek to pursue other, more personal goals, whether financial, or electoral success, or a comfortable office, or power, etc.)

But let's consider the hard case of economic models in which agents are pursuing economic goals. "Self-interested" behavior in this context, which is the assumption, is not the same thing as "selfish" behavior, at least not as commonly understood. An entrepreneur or a laborer may be seeking to maximize profit or wage in order to provide for a family, or even to generate money to give to those in need (Eph 4.28). Economic models are silent about why the actors they model engage in optimizing behavior. To be sure, there may be cases in which the entrepreneur or laborer does not try to maximize short-term profits or wages – the entrepreneur may feel ethically bound to resist price-gouging after a hurricane, the worker may accept a lower wage to allow the business to stay in business. (Although the latter merely exemplifies workers sacrificing short-run interests on behalf of longer-run interests.)

Economists don't deny that these other forms of behavior exist – indeed, sometimes they seek to model that type of behavior itself. Rather, they write their models with a certain case in mind that they wish to explain. In many interesting cases, although no means in all such cases, it's just fine to think of what happens when workers seek to maximize their wages and entrepreneurs seek to maximize their profits. And by "just fine" I mean "useful."

This can be contemplated more concretely by using economist Philip Wicksteed's example of thinking about how the Apostle Paul engaged in his tent-making vocation.

Given that Paul used his profession to avoid burdening churches with supporting him – this was a major "boast" of his, and waving the interesting theological issue of whether Paul's actions here are examples of ethical supererogation, we know consequently that Paul was not giving his labor and his tents away for

free. Paul purchased materials and sold tents. While I anticipate that Paul's tents were a good value for the price, there is no reason to think that, ordinarily, Paul sold his tents at a systematically lower price than most other tent-makers, or that his tents were of systematically higher quality at a given price than other tents. There is no reason to think that much of Paul's ordinary, everyday vocational activity would not accord with economic models that focus on consumption and production. That's fine, because that's the type of behavior that most economists are interested in studying. So they throw away lots of details that affect behavior but do not interest them. They do so because "economic" behavior interests them. If we were interested in understanding other aspects of Paul's behavior, we would want to assume different entries in his utility function.

The point is that the self-interest postulate applies to Paul as well as it does to any other individual to study the type of behavior in which most economists are interested most of the time.

That said, there is the next step: Just as it can be fun to press mathematical modeling for its own sake, some economists and others are interested to press specifically economic forms of rational choice models into other areas not traditionally thought of as economic. This is what prompts complaints of the "imperialism" of the economic perspective, although this is also sometimes conflated with the extension of rational choice models to additional areas of human behavior. One can argue that normal assumptions of distinctly "economic" models of behavior are less useful in understanding some types of "non-economic" behaviors without throwing out the rational choice approach at large. For example, I've typically found analysis based on analogizing political or religious behavior with economics markets not particularly illuminating. As with different types of maps, scholars can, and do, argue over the usefulness of bringing distinctly economic assumptions to the study of behavior usually thought to be non-economic. Whether there are alternative specifications that scholars find more illuminating, or better at predicting the type of behavior that interests analysts. I suggest patience with economists, and let them try, whether mischievously or not, to take their models as far as they can. The proof of the theoretical pudding is in the empirical eating, after all.

In any event, when we take a step back from rational choice models as instantiated in economics, the "self-interest" assumption is little more than an assumption of goal-seeking behavior. Given constraints (both external and internal) people pursue what they desire. To be sure, in cases in which people cannot rank different outcomes as more preferred, less preferred, or indifferently preferred, or when those preferences are not transitive, then the rationality postulates don't apply. And that's just fine. As with any other theory, rational choice theory need not explain everything in order to be helpful in explaining some things.

What's the Purpose?

A central question in social science – by which I mean the study of human interaction with other humans – is the question, "Why can't we all just get along?" Why is cooperation and coordination between people seemingly so easy in some situations, and so difficult in others? Relatedly, when should cooperation be facilitated, and when should it be deterred? While an Augustinian view of human nature may be a sufficient answer to the question, perhaps more surprisingly, it is not necessary to answer the question. Cooperation and coordination between humans can be difficult even when people want to get along together. To put it a different way, even a society of saints would face cooperation and coordination failures, not because they don't intend the best for everyone, but because cooperation and coordination among numerous individuals, however well motivated, can sometimes be inherently difficult. In his poem *The Rock*, T.S. Eliot wrote of those who dream of "systems so perfect that no one will need to be good." The flip side of that is those who dream of people so perfect that no system can help.

A question I often ask my students to write about is to draw on basic game-theoretic models and respond to this claim: “Most social and political problems in the world result because people do not want to cooperate with other people.”

The answer I want from them is: it depends. In the case of the so-called “Battle-of-the Sexes,” in which a husband and a wife are deciding whether to go to a boxing match or a ballet, the couple actually *wants* to be together, but each also has a preferred activity, and so there are multiple outcomes to the game. Despite the desire to be together, the couple may fail to achieve their desired togetherness not because of a lack of desire, but because the situation they face makes it more or less difficult to coordinate their activity without something to help beyond their desire to spend time together.

The problem of cooperative behavior of course does become more acute with a more-realistic anthropology. In contrast, in the well-known “prisoners’ dilemma,” there is a sense in which the players recognize if they cooperate they will be better off – and they desire that outcome relative to where they end up in the game – but because of incentives to exploit other players if they voluntarily behave cooperatively, it is difficult to sustain the cooperative outcome without changing the game.

It is sometimes tempting to root for cooperative outcomes in the abstract, to think that cooperation is always a good thing. But cooperation is not an unalloyed good. Forms of cooperation that, say, produce additional goods, whether material goods like bread, or social or spiritual goods like family life, are generally good. Forms of cooperation in the form of price fixing among businesses, or criminal conspiracies, are generally bad.

Indeed, markets harness the central insight of the prisoners’ dilemma to facilitate the *failure* of cooperation in the form of price and product competition, and channel that cooperative failure to benefit society: as competing businesses play a prisoners’ dilemma against each other in price, the rest of us benefit by enjoying lower prices relative to the businesses achieving the cooperative outcome of oligopolistic pricing.

Critics of rational choice theory like to point out that, in experiments testing game theory models like the prisoners’ dilemma, the results don’t confirm 100% non-cooperation. That’s interesting in itself, and offers its own “why” puzzle. But the critics often want to draw a conclusion bigger than the result: Unless we live in a first-best world where prisoners’ dilemma-like situations are never experienced, then the existence of those incentive structures and the behavior they induce among significant swaths of humanity is grist both for scholars and as a practical matter for folks living together in society. Unless one is far gone in utopian beliefs, or the belief that human nature is perfectible through education (or reeducation), no one seriously believes that voluntary cooperation between individuals can eliminate air pollution, or overfishing, or arguments between roommates over washing dishes. Like the quality of a map, the prisoners’ dilemma and models, some quite sophisticated, highlight important aspects of human interaction for study. That a model doesn’t explain everything is no more relevant than the observation that a world globe doesn’t help us find a restaurant in Manhattan. One need not explain everything to explain something.

This is not to wave off the significance of experimental results. If, in fact, one is interested in understanding why and when human behavior is not accurately predicted by a simple prisoners’-dilemma models, then we can relax assumptions of the model, or incorporate other assumptions, and see where it gets us. If the map you have can’t get you where you want to go, the solution is not to reject “maps” as an entire category, the solution might be to get a better map - a map with the details sufficient to get you where you’re interested in going.

Limitations: changing preferences

None of the above should be taken to suggest that rational choice models can, in principle, explain the totality of human behavior. For example, if a person cannot compare two outcomes or if a person's preferences are not transitive, then the axiomatic primitives for rational choice theory are not met. This is, however, different from a person desiring more information before wanting to make a decision, or the need to make decisions under uncertainty. (Particularly in political science an oft-asserted line of criticism of rational choice theory is that its models necessarily assume "perfect information." This hasn't been true since the 1970s. There has been a veritable revolution in modeling uncertainty and informational asymmetries over the last four decades. Indeed, entire courses in graduate schools in sundry disciplines focus on understanding and predicting human decision-making without perfect information.)

More limiting for modeling human decision making over time using rational choice theory, however, is that, as a practical matter, analysts need take people's preferences as a primitive, that is, as a given. Or, perhaps better, theorists have not generated tractable theories of predictable preference change. There has been some progress on this topic in recent years, but that progress is pretty much limited to modeling predictable preference changes. For example, that one's appreciation of opera, or taste for finer wines, increases with one's experience. It's not simply a matter of information; one's tastes actually change.

It is more difficult to model conversions, or dramatic preference changes.

Again, however, we must be careful as to what we mean. The assumption that preferences do not change means the underlying preferences of a person do not change. It does not refer to a person's revealed preferences. It is possible, indeed, extremely likely, that revealed preferences change with the acquisition of information, experience, learning, or changes in income, costs, and alternatives etc. That I now drink finer wines because I can afford them relative to when I was in college does not reflect a change in my underlying preferences: I have the same preferences, it's just that the income constraint is less binding today than when in college. Similarly, a person's preference that the government adopt one policy rather than another, or a preference for dining at one restaurant over another, may change frequently as a result of changes in the environment, institutions, and constraints that people face. None of this means that there is a change in underlying preferences.

Irrespective of all this, people do have conversion experiences. While this, importantly, includes religious conversions, I mean the term to describe a change in the underlying components of a person's preferences. "Conversion" need not engage any supernatural affirmation, as in the conversion of a rock-ribbed high school Republican into an equally rock-ribbed socialist during a first year at Harvard. This may reflect a change in information, environment, and constraints, but there is no reason to think that revealed preferences, even if initially induced by other factors, do not ultimately become reflected by a change in the person's underlying preferences, at least some of the time.

Rational choice models assume underlying preferences, and cannot generate useful theories of conversion.

This is one difficulty, to return to the introduction, with common versions of Pascal's wager. Often times religiosity is treated merely as the affirmation of one or more additional propositions by the believer relative to the unbeliever: The unbeliever affirms N propositions to be true; the believer affirms $N + 1$ propositions to be true, where the $N+1^{\text{st}}$ proposition is that "God exists." (Believers may indeed affirm more than $N+1$ additional propositions, the thought goes, depending on just how sectarian they are.)

These abstract affirmations, however, reflect a person's preferences and desires; they do not transform those preferences and desires. The preferences of the person after "conversion" is in a continuum with the

person's preferences before conversion. It is, after all, the pre-conversion preferences that, making the gamble, see the ostensible gain to be had from belief.

While in obvious ways the same person continues after conversion as before conversion, some religions also teach the transformation of the pre-conversion person in discontinuity with the post-conversion person. Jesus taught that a person must be "born anew." Paul taught that a Christian is a "new person" (*kainos anthropos*), that the old person has died and a new person, once again created in the image of the Creator, now lives as a part of a new humanity. For Paul, "it is no longer I who live, but Christ lives in me" (Gal 2.20).

Consistent with this, Paul adopts a different mode of admonition than he is often taken to teach. Christians are often taken to admonish one another, and others, to be good: that the point of the faith is to become holy by acting holy. But the way that Paul admonishes his Christian listeners is different than that. Arguing from their transformed natures, he admonishes his audiences to become what they are – more particularly, to become the new *anthropos* that they have been re-created to be. The idea is that the Christian has a new identity in Christ, and with that identity comes new desires and preferences. (Which is not to say that the old desires and preferences simply evaporated.) In the First letter to the Corinthians, Paul writes that the non-converted not only do not understand the things of God, but that the non-converted *cannot* understand the things of God. For the non-converted then, belief cannot simply be a matter of weighing eternal life versus eternal separation from God. In a sense, the question is whether one *wants* to spend eternity in God's presence or not. As Jesus put it, do we think what God offers is a snake or a fish? The unbeliever recoils at what God offers; for the unbeliever, God offers a snake. For the believer, God offers a fish. Whether we want what God offers depends on our underlying preferences. What is a snake to the unbeliever is a fish to the believer. But Pascal's wager doesn't work if the stakes change only if one accepts the wager. One would not take the gamble for an eternity of snake.

It is at this point as well that there is a telling difference between the popular take on Pascal's wager, and the way Pascal developed it in *Pensees*. Pascal did not conclude his discussion of the wager with the conclusion suggested at the end of many popular invocations of the wager, something like, "So, therefore one ought to believe." Pascal's own conclusion is to consider the wager as one whether to be open to conversion, open to a dramatic change in preferences.

"I confess it, I admit it. But, still, is there no means of seeing the faces of the cards?"—
Yes, Scripture and the rest, etc. "Yes, but I have my hands tied and my mouth closed; I am forced to wager, and am not free. I am not released, and am so made that I cannot believe. What, then, would you have me do?"

True. But at least learn your inability to believe, since reason brings you to this, and yet you cannot believe. Endeavour then to convince yourself, not by increase of proofs of God, but by the abatement of your passions. You would like to attain faith, and do not know the way; you would like to cure yourself of unbelief, and ask the remedy for it. Learn of those who have been bound like you, and who now stake all their possessions. These are people who know the way which you would follow, and who are cured of an ill of which you would be cured. Follow the way by which they began; by acting as if they believed, taking the holy water, having masses said, etc. Even this will naturally make you believe, and deaden your acuteness.—"But this is what I am afraid of."—And why? What have you to lose?

But to show you that this leads you there, it is this which will lessen the passions, which

are your stumbling-blocks.

The end of this discourse.—Now, what harm will befall you in taking this side? You will be faithful, honest, humble, grateful, generous, a sincere friend, truthful. Certainly you will not have those poisonous pleasures, glory and luxury; but will you not have others? I will tell you that you will thereby gain in this life, and that, at each step you take on this road, you will see so great certainty of gain, so much nothingness in what you risk, that you will at last recognize that you have wagered for something certain and infinite, for which you have given nothing.

"Ah! This discourse transports me, charms me," etc. that it is made by a man who has knelt, both before and after it, in prayer to that Being, infinite and without parts, before whom he lays all he has, for you also to lay before Him all you have for your own good and for His glory, that so strength may be given to lowliness.

In Pascal's argument, one's preferences could be, would be, shaped by participating in the life and liturgy of the faithful. This is not an argument for faith, but an argument to be open to faith. Or more precisely, an argument to give its practices a shot. To live as if transformed, and to see if it works from without to within as neatly as it works from within to without. Sort of a "taste and see that the LORD is good" (Ps 34.8).

Here, too, Pascal's rational choice argument only gets us so far. But it's all the more winsome in its more-modest form than in its grander version. It recognizes that it can take its interlocutor only so far. This is not a reason to reject it, but to appreciate the distance it can move a person and yet appreciate the limitations of the argument.

This, too, with social science models. To understand something one does not need to understand everything. One does not need a grand theory of person or society to say something interesting about human domains we inhabit. Practical wisdom can flourish with a more-modest, local epistemology. A recognition that we're simply attempting to figure out how people usually behave in common areas of life. To be sure, there is a lot about life beyond these domains that is interesting. But I do insist that human behavior is intrinsically interesting, and that we can learn interesting things about it with theoretical, and empirical, models that assume away a lot of reality "out there" in order better to understand our shared life together. In this way, social science (as well as the physical sciences) represents nothing more than a refinement of common ways in which to think about the human world.

To be sure, many of us pine for grander theories that make sense of it all. There's nothing wrong with that. These theories are too diverse to generalize. It is nonetheless true that some are little more than secular faiths, replete with their own ethics, sins, and eschatologies. But I think it's a mistake for those inclined to non-secular faiths to think that these sorts of grand theories are the sum of social science, or the equivalent. Most social scientists, at least the good ones, inhabit a more-modest world in which we delight over puzzles of human social interaction, and in which we attempt to simplify away from details extraneous to those puzzles, in order to understand and explain some of how we live together, and how we fail to live together. In these very practical tasks the faithful are as welcomed as those who aren't.