

Behavioral Economics: Past, Present, and Future[†]

BY RICHARD H. THALER*

May, 2016

Abstract: There has been growing interest in the field come to me known as “behavioral economics” which attempts to incorporate insights from other social sciences, especially psychology, in order to enrich the standard economic model. This interest the underlying psychology of human behavior returns economics to its earliest roots. Scholars such as Adam Smith talked about such key concepts as loss aversion, overconfidence, and self-control. Nevertheless, the modern version of behavioral economics introduced in the 1980s met with resistance by some economists, who preferred to retain the standard neo-classical model. They introduced several arguments for why psychology could safely be ignored. In this essay I show that these arguments have been rejected, both theoretically and empirically, so it is time to move on. The new approach to economics should include two different kinds of theories: normative models that characterize the optimal solution to specific problems and descriptive models that capture how humans actually behave. The latter theories will incorporate some variables I call *supposedly irrelevant factors*. By adding these factors such as framing or temptation we can improve the explanatory power of economic models. If everyone includes all the factors that do determine economic behavior, then the field of behavioral economics will no longer need to exist.

[†] This article is based on my presidential address given at the American Economic Association annual meeting in January 2016.

* University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637 (e-mail: <mailto:richard.thaler@chicagobooth.edu>).

Behavioral Economics: Past, Present, and Future[†]

BY RICHARD H. THALER*

In recent years there has been growing interest in the mixture of psychology and economics that has come to be known as “behavioral economics.” As is true with many seemingly overnight success stories, this one has been brewing for quite a while. My first paper on the subject was published in 1980, hot on the heels of Kahneman and Tversky’s (1979) blockbuster on prospect theory, and there were earlier forerunners, most notably Simon (1955, 1957) and Katona (1951, 1953).

The rise of behavioral economics is sometimes characterized as a kind of paradigm-shifting revolution within economics, but I think that is a misreading of the history of economic thought. It would be more accurate to say that the methodology of behavioral economics returns economic thinking to the way it began, with Adam Smith, and continued through the time of Irving Fisher and John Maynard Keynes in the 1930s.

In spite of this early tradition within the field, the behavioral approach to economics met with considerable resistance within the profession until relatively recently. In this essay I begin by documenting some of the historical precedents for utilizing a psychologically realistic depiction of the representative agent. I then turn to a discussion of the many arguments that have been put forward in favor of retaining the idealized model of *Homo economicus* even in the face of apparently contradictory evidence. I argue that such arguments have been refuted, both theoretically and empirically, including in the realm where we might expect rationality to abound: the financial markets. As such, it is time to move on to a more constructive approach.

On the theory side, the basic problem is that we are relying on one theory to accomplish two rather different goals, namely to characterize optimal behavior and to predict actual behavior. We should not abandon the first type of theories as they are essential building blocks for any kind of economic analysis, but we must augment them with additional descriptive theories that are derived from data rather than axioms.

As for empirical work, the behavioral approach offers the opportunity to develop better models of economic behavior by incorporating insights from other social science disciplines. To illustrate this more constructive approach, I focus on one strong prediction made by the traditional model, namely that there is a set of factors that will have no effect on economic behavior. I refer to these as supposedly irrelevant factors or SIFs. Contrary to the predictions of traditional theory, SIFs matter; in fact, in some situations the single most important determinant of behavior is a SIF. Finally, I turn to the future. Spoiler alert: I predict that behavioral economics will eventually disappear.

I. The Historical Roots of Behavioral Economics

As Simon (1987, p. 612) noted, the term “behavioral economics” is a bit odd. “The phrase

[†] This article is based on my presidential address given at the American Economic Association annual meeting in January 2016. The actual lecture can be viewed on the AEA website: <https://www.aeaweb.org/webcasts/2016/Behavioral.php> This article also draws upon my recent book *Misbehaving: The Making of Behavioral Economics*, which contains a much more extensive bibliography.

* University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637 (e-mail: <mailto:richard.thaler@chicagobooth.edu>).

'behavioral economics' appears to be a pleonasm. What 'non-behavioral' economics can we contrast with it? The answer to this question is found in the specific assumptions about human behavior that are made in neoclassical economic theory." These assumptions are familiar to all students of economic theory. (i) Agents have well-defined preferences and unbiased beliefs and expectations. (ii) They make optimal choices based on these beliefs and preferences. This in turn implies that agents have infinite cognitive abilities (or, put another way, are as smart as the smartest economist) and infinite willpower since they choose what is best, not what is momentarily tempting. (iii) Although they may act altruistically, especially toward close friends and family, their primary motivation is self-interest. It is these assumptions that define *Homo economicus*, or as I like to call them, Econs. Behavioral economics simply replaces Econs with *Homo sapiens*, otherwise known as Humans.

To many economists these assumptions, along with the concept of "equilibrium," effectively define their discipline; that is, they study Econs in an abstract economy rather than Humans in the real one. But such was not always the case. Indeed, Ashraf, Camerer, and Loewenstein (2005) convincingly document that Adam Smith, often considered the founder of economics as a discipline, was a bona fide behavioral economist. Consider just three of the most important concepts of behavioral economics: overconfidence, loss aversion, and self-control. On overconfidence Smith (1776, p. 1) commented on "the overweening conceit which the greater part of men have of their own abilities" that leads them to overestimate their chance of success. On the concept of loss aversion Smith (1759, p. 176–177) noted that "Pain...is, in almost all cases, a more pungent sensation than the opposite and correspondent pleasure." As for self-control, and what we now call "present bias," Smith (1759, p. 273) had this to say: "The pleasure which we are to enjoy ten years hence, interests us so little in comparison with that which we may enjoy today." George Stigler was fond of saying that there was nothing new in economics, it had all been said by Adam Smith. It turns out that was true for behavioral economics as well.

But Adam Smith was far from the only early economist who had good intuitions about human behavior. Many who followed Smith, shared his views about time discounting. For example, Pigou (1920, p. 21) famously wrote that "Our telescopic faculty is defective and... we therefore see future pleasures, as it were, on a diminished scale." Similarly Fisher (1930, p. 82), who offered the first truly modern economic theory of intertemporal choice, did not think it was a good description of behavior. He offered many colorful stories to support this skepticism: "This is illustrated by the story of the farmer who would never mend his leaky roof. When it rained, he could not stop the leak, and when it did not rain, there was no leak to be stopped!" Keynes (1936, p. 154) anticipated much of what is now called behavioral finance in the *General Theory*. For example, he observed that "Day-to-day fluctuations in the profits of existing investments, which are obviously of an ephemeral and non-significant character, tend to have an altogether excessive, and even absurd, influence on the market."

Many economists even thought that psychology (then still in its infancy) should play an important role in economics. Pareto (1906, p. 21) wrote that "The foundation of political economy, and, in general of every social science, is evidently psychology. A day may come when we shall be able to decide the laws of social science from the principles of psychology." John Maurice Clark (1918, p. 4), the son of John Bates Clark, went further. "The economist may attempt to ignore psychology, but it is sheer impossibility for him to ignore human nature...If the economist borrows his conception of man from the psychologist, his constructive work may have some chance of remaining purely economic in character. But if he does not, he will not thereby avoid psychology. Rather, he will force himself to make his own, and it will be bad psychology."

It has been nearly 100 years since Clark wrote those words but they still ring true, and

behavioral economists have been taking Clark's advice, which is to borrow some good psychology rather than invent bad psychology. Why did this common sense suggestion fail to gain much traction for so long?

II. Explainawaytions

In the process of making economics more mathematically rigorous after World War II, the economics profession appears to have lost its good intuition about human behavior. Defective telescopic facilities were replaced with time-consistent exponential discounting. Over-weening conceits were replaced by rational expectations. And ephemeral shifts in animal spirits were replaced by the efficient market hypothesis. Economics textbooks no longer had any Humans. How did this happen?

I believe that the most plausible explanation is that models of rational behavior became standard because they were the easiest to solve. This conjecture is not meant as a put-down. One begins learning physics by studying the behavior of objects in a vacuum; atmosphere can be added later. But physicists never denied the existence or importance of air; instead they worked harder and built more complicated models. For many years, economists reacted to questions about the realism of the basic model by doing the equivalent of either denying the existence of air, or by claiming that it just didn't matter all that much. Matthew Rabin has dubbed these defensive reactions as "explainawaytions."¹

Let's be blunt. The model of human behavior based on the premise that people optimize is and has always been highly implausible. For one thing, the model does not take into consideration the degree of difficulty of the problem that agents are assumed to be "solving." Consider two games: tic-tac-toe and chess. A reasonably bright first grader can learn to play the optimal strategy in tic-tac-toe, and so a model that assumes players choose optimally in this game will be a pretty good approximation of actual behavior for bright children and sober adults. Chess, on the other hand, is quite a different matter. Most of us play chess terribly and would have no chance of beating a free program on our smartphones, much less a grandmaster. So, it makes no sense to assume that the representative agent plays chess as well as tic-tac-toe. But that is essentially what we assume in economics.

When we assume that agents maximize utility (or profits) we do not condition that assumption on task difficulty. We assume that people are equally good at deciding how many eggs to buy for breakfast and solving for the right amount to save for retirement. That assumption is, on the face of it, preposterous. So why has it stuck? There has been a litany of explainawaytions.

A. *As If*

Grumblings within the profession about the so-called "marginalist revolution" were present in the 1940s, and this journal published several articles debating the realism of the theory that firms set output and hire workers by calculating the point at which marginal cost equals marginal revenue. One of the participants in this debate was Richard Lester of Princeton who had the temerity to ask the owners of business firms how they actually made such decisions. Whatever firms were doing did not seem to be captured by the term "equating at the margin," and Lester (1946, p. 81) ended his paper this way: "This paper raises grave doubts as to the validity of conventional marginal theory and the assumptions on which it rests." Machlup (1946) took up the defense of the traditional theory and argued

¹ Please direct all complaints about this term directly to Matthew.

that even if firm owners did not know how to calculate marginal costs and revenues, they would make decisions that would closely approximate such choices using their intuitions.

Machlup's defense was refined and polished by Friedman (1953, p. 21) in his famous essay "The Methodology of Positive Economics." Friedman brushed aside questions about the realism of assumptions and argued that instead theories should be judged based on their ability to predict behavior. He proposed what is now a well-known analogy about an expert billiard player: "excellent predictions would be yielded by the hypothesis that the billiard player made his shots *as if* he knew the complicated mathematical formulas that would give the optimum directions of travel, could estimate by eye the angles, etc., describing the location of the balls, could make lightning calculations from the formulas, and could then make the balls travel in the direction indicated by the formulas. Our confidence in this hypothesis is not based on the belief that billiard players, even expert ones, can or do go through the process described; it derives rather from the belief that, unless in some way or other they were capable of reaching essentially the same result, they would not in fact be *expert* billiard players."

Friedman had a well-deserved reputation as a brilliant communicator and debater, and those skills are on full display in this passage. Using the mere two-word phrase "as if," Friedman essentially ended the debate about the realism of assumptions in economics. But given proper scrutiny, we can see that this passage is simply a verbal sleight of hand. First of all, it is no accident that Friedman chooses to discuss an *expert* billiard player. The behavior of an expert in many activities may indeed be well captured by a model that assumes optimal behavior. But what about non-experts? Isn't economic theory supposed to be a theory about the behavior of all economic agents, not just experts? The life-cycle hypothesis is intended to be a theory of how the typical citizen saves for retirement, not just those with MBAs.

There is another problem with Friedman's defense, which is that even experts are unable to optimize when the problems are difficult. To illustrate, let's return to the game of chess. Since chess has no stochastic elements, it has long been known that if both players optimize then one of the players (either the one who goes first or second) must have a winning strategy, or neither of them do and the game will lead to a draw. However, unlike checkers, which has been "solved" (if both players optimize the game is a draw) chess matches do not yield predictable outcomes even in matches between grandmasters. Sometimes white (first player) wins, less often black wins, and there are many draws. This proves that even the best chess players in the world do not maximize. Of course one can argue that chess is a hard game, which is true. But, many economic decisions are difficult as well.

A second line of defense is to concede that we don't all do everything like experts but argue that, if our errors are randomly distributed with mean zero, then they will wash out in the aggregate, leaving the predictions of the model unbiased on average. This was often the reaction to Simon's (1955) suggestion that people "satisfice" (meaning grope for a satisfactory solution rather than solve for an optimal one). If the choices of a satisficer are not systematically different from an optimizer, then the models lead to identical average predictions (though satisficers will have more noise). This line of argument was refuted by the seminal work of Daniel Kahneman and Amos Tversky in the 1970s.

In a brilliant series of experiments on what psychologists refer to as "judgment" and what economists might call "expectations" or "beliefs," Tversky and Kahneman (1974) showed that humans make judgments that are *systematically* biased. Furthermore, these errors were predictable based on a theory of human cognition. Kahneman and Tversky's hypothesis was that people often make judgments using some kind of rule of thumb or heuristic. An example is the "availability heuristic" in which people estimate the frequency

of some event by the ease with which they can recall instances of that event. Using this heuristic is perfectly sensible since frequency and ease of recall are generally positively correlated. However, use of the heuristic will lead to predictable errors in those situations where frequency and ease of recall diverge. For example, when asked to estimate the ratio of gun deaths by homicide to gun deaths by suicide in the United States, most people think homicide gun deaths are more common, whereas there are in fact nearly twice as many gun-inflicted suicides as homicides. These are expectations that are not close to being “as if” rational—they are predictably biased.

Kahneman and Tversky’s second influential line of research was on decision making. In particular, in 1979 they published their paper on prospect theory, which was proposed as a “descriptive” (or what Milton Friedman would have called “positive”) model of decision making under uncertainty. Prospect theory was intended to be a descriptive alternative to von Neumann and Morgenstern’s (1947) expected utility theory, which is rightly considered by most economists to characterize how a rational agent should make risky choices. Kahneman and Tversky’s research documented numerous choices that violate any sensible definition of rational. This pair of problems posed to different groups of subjects offers a good illustration.

Problem 1.—Imagine that you face the following pair of concurrent decisions. First examine both decisions, and then indicate the options you prefer.

Decision (i) Choose between:

- | | |
|---|-------|
| A. A sure gain of \$240 | [84%] |
| B. 25% chance to gain \$1,000 and
75% chance to gain or lose nothing | [16%] |

Decision (ii) Choose between:

- | | |
|---|-------|
| C. A sure loss of \$750 | [13%] |
| D. A 75% chance to lose \$1,000 and a
25% chance to lose nothing | [87%] |

The numbers in brackets indicate the percentage of subjects that chose that option. We observe a pattern that was frequently displayed: subjects were risk averse in the domain of gains but risk seeking in the domain of losses. It is not immediately obvious that there is anything particularly disturbing about these choices; that is, until one studies the following problem.

Problem 2.—Choose between:

- | | |
|--|--------|
| E. 25% chance to win \$240
and 75% chance to lose \$760 | [0%] |
| F. 25% chance to win \$250
and 75% chance to lose \$750 | [100%] |

Inspection reveals that although Problem 2 is worded differently, its choices are formally identical to those in Problem 1. The difference is that some simple arithmetic has been performed for the subjects. Once these calculations are made it becomes clear to every subject that option F dominates option E, and everyone chooses accordingly. The difficulty, of course, is that option E, which no one selects, is made up of the combination of options A and D, both of which were chosen by a large majority of subjects, while option F, which everyone selects, is a combination of B and C, options that were highly unpopular in Problem 1. Thus this pair of problems illustrates two findings that are embarrassing to

rational choice adherents. First, subjects' answers depend on the way a problem is worded or "framed," behavior that is inconsistent with almost any formal model. Second, by utilizing clever framing, a majority of subjects can be induced to select a pair of options that are dominated by another pair. Once again, this behavior does not seem consistent with the idea that people are choosing *as if* they are rational.

B. Experiments, Incentives, and Learning

A second class of explanations emerged in the 1980s, in part as a reaction to the findings of Kahneman and Tversky and an early paper of mine (Thaler 1980). These retorts, usually delivered orally in workshops and conference presentations rather than in print,² were intended to be justifications for continuing business as usual. Some of the critiques were aimed at the empirical methods used in these early papers, namely hypothetical survey questions such as problems 1 and 2 above. Economists have never been very impressed by such data because the subjects have nothing on the line. Furthermore, typically these questions were just asked once, so many argued that they were not a good indication of what people would do in real-life situations in which they had an opportunity to learn from prior mistakes. So the critique was two-fold. First, if you raise the stakes people will take the questions more seriously and choose in a manner more consistent with optimization. Second, if given a chance to learn, people will get it right. Often the same person would make both of these critiques, thinking that they reinforced one another.

Of course there is no doubt that the ability to practice improves performance in most tasks. No one plays well in his first game of chess, or billiards for that matter. And most people eventually become at least competent at highly complex tasks such as riding a bike or running down a flight of stairs. Similarly, the notion that people will pay more attention when the stakes go up is intuitively appealing. Certainly we pay more attention when buying a car than when deciding what to order for lunch. But rather than these two arguments working together, they actually go in opposite directions. The reason this is so is that, as a rule, the higher the stakes, the less often we get to do something.

Consider the following list of economic activities: deciding how much milk to buy at the grocery store, choosing a sweater, buying a car, buying a home, selecting a career, choosing a spouse, saving for retirement. Most households have mastered the art of milk inventory management through trial and error. Buy too much and it spoils, buy too little and you have to make an extra trip to the convenience store. But if households do this (say) twice a week, eventually they figure it out, at least until the children move out of the house or switch to beer. Few of us buy cars often enough to get very good at it, and the really big decisions like careers, marriages, and retirement saving give very little room for learning. So critics can't have it both ways. Either the real world is mostly high stakes or it offers myriad opportunities to learn—not both.

Even in domains where there are multiple opportunities to learn, people may not make the best of those situations. Daniel Kahneman and I ran an experiment years ago that illustrates this point. (We never published the results so the details will be sketchy.) Subjects were given forms that looked something like this:

Heads: 1 2 3 4 5 ... 18 19 20

Tails: 1 2 3 4 5 ... 18 19 20

They were then shown two large manila envelopes that were labeled Heads and Tails and

² However, see the papers in Hogarth and Reder (1986, 1987) for some written versions.

were shown that each envelope contained 20 poker chips numbered from 1 to 20. The experimenter said he would first flip a coin and then, depending on the outcome, choose a poker chip from the respective envelope. Subjects were allowed to circle five numbers on their form, dividing their choices as they wished between the heads and tails rows. When the experimenter selected a chip and announced the result, for example “Heads, 17” any subject who had circled the winning coin face and number would win some money. Specifically, if the chip came from the Heads envelope winners would be paid \$2, but if the chip came from the Tails envelope they would win \$3.

Of course the optimal strategy in this game is to only circle numbers in the Tails row since those have a 50 percent higher expected payoff, but this strategy was not obvious to everyone. About half the subjects (MBA students at a top university) adopted the correct strategy of circling only Tails, but the rest used what might called an “inept mixed strategy,” dividing their choices between Heads and Tails, with the most common allocation being three Tails and two Heads, matching the ratio of the payoffs.³

The question that Kahneman and I were most interested in, however, was not these initial choices. This was an experiment about learning. So we had the subjects repeat the same task nine more times. Each time the subjects got feedback about the outcome of the coin toss and the number drawn, and the winning guessers were paid in cash immediately in plain view of the other subjects. Try to guess the results as a thought experiment.

Of the subjects that did not figure out the “all Tails” strategy immediately, how many learned to use that strategy over the course of the nine additional trials? The answer is one. One subject switched at some point to an all Tails strategy, but that subject was offset by another subject who had circled only Tails on the first trial, but then switched to the inept mixed strategy at some point during the “learning” phase.

It is instructive to consider why there was essentially no learning in this experiment. We know from psychology that learning takes place when there is useful, immediate feedback. When learning to drive we quickly see how much pressure to use on the accelerator and brake pedals in order to start and stop smoothly. In the experiment, however, the subjects were first told the outcome of the coin flip, then the number drawn. Obviously, about half the time the coin came up Heads, and those who were including Heads in their portfolio were pleased to be still in the game (if only for another few seconds). Furthermore, every time that someone won some money from a Heads outcome, there was some reinforcement for continuing to include some of that “strategy” in the portfolio.

The general point is that learning can be difficult even in a very simple environment. Those who teach an introductory course in economics know that many of the first principles that are basic to rational choice models (such as the notion of opportunity costs) are by no means intuitively obvious to the students. But our models assume they can understand much more difficult concepts such as backward induction.

As for the argument that people will do better in experimental tasks if the stakes are raised, there is little or no evidence to support this hypothesis. The first empirical test of this idea was conducted by David Grether and Charles Plott (1979) in the context of an investigation of the “preference reversal phenomenon,” discovered by psychologists Sarah Lichtenstein and Paul Slovic (1971). Lichtenstein and Slovic presented subjects with two gambles, one a near sure thing they called the p-bet (for high probability) such as a 35/36 chance to win \$10, the other more risky called the \$-bet, such as an 11/36 chance to win \$30, a higher potential payoff. Subjects were asked to value each bet by naming the lowest price

³ Likewise, when Heads and Tails aren't equally likely, people tend to engage in “probability matching” behavior instead of just picking the more likely outcome every time. See Vulkan (2000) for a survey aimed at economists.

at which they would sell it if they owned it, and also to choose which of the bets they would rather have. The term “preference reversal” emerged from the fact that of those who preferred the p-bet, a majority reported a higher selling price for the \$-bet, implying that they valued it more than the p-bet.

Grether and Plott (1979) were perplexed by this finding and set out to determine which mistake the psychologists must have made to obtain such an obviously wrong result. Since the original study was based on hypothetical questions, one of the hypotheses Grether and Plott investigated was whether the preference reversals would disappear if the bets were played for real money. (They favored this hypothesis in spite of the fact that Lichtenstein and Slovic (1973) had already replicated their findings for real money on the floor of a Las Vegas casino.) What Grether and Plott found surprised them. Raising the stakes did have the intended effect of inducing the subjects to pay more attention to their choices (so noise was reduced) but preference reversals did not thereby vanish; rather, their frequency went up! In the nearly 40 years since Grether and Plott’s seminal paper, I do not know of any findings of “cognitive errors” that were discovered and replicated with hypothetical questions but then vanished as soon as significant stakes were introduced.

C. The Invisible Handwave

There is a variation on the “if there is enough money at stake people will behave like Econs” story that is a bit more complicated. In this version markets replace the enlightening role of money. The idea is that when agents interact in a market environment, any tendencies to misbehave will be vanquished. I call this argument the “invisible handwave” because there is a vague allusion to Adam Smith embedded in there somewhere, and I claim that it is impossible to complete the argument with both hands remaining still.

Suppose, for example, that Homer falls prey to the “sunk cost fallacy” and always finishes whatever is put on his plate for dinner, since he doesn’t like to waste money. An invisible handwaver might say, fine, he can do that at home, but when Homer engages in markets, such misbehaving will be eliminated. Which raises the question: how exactly does this occur? If Homer goes to a restaurant and finishes a rich dessert “because he paid for it” all that happens to him is that he gets a bit chubbier. Competition does not solve the problem because there is no market for restaurants that whisk the food away from customers as soon as they have eaten more than X calories.

Indeed, thinking that markets will eradicate aberrant behavior shows a failure to understand how markets work. Let’s consider two possible strategies firms might adopt in the face of consumers making errors. Firms could try to teach them about the costs of their errors or could devise a strategy to exploit the error to make higher profits. The latter strategy will almost always be more profitable. *As a rule it is easier to cater to biases than to eradicate them.* DellaVigna and Malmendier (2006) provide an instructive example in their article “Paying Not to Go to the Gym.” The authors study the usage of customers of three gyms that offer members the choice of paying \$70 a month for unlimited usage, or a package of 10 entry tickets for \$100. They find that the members paying the monthly fee go to the gym an average of 4.3 times per month, implying an average cost of over \$17 per visit.

Obviously the typical monthly members have an arbitrage opportunity available. Why pay \$17 a visit when they could be paying \$10? One possible explanation for this behavior is that customers understand that they are affected by sunk costs (whether or not they realize it is a fallacy) and are strategically using the membership fee as a (rather ineffective) commitment device to try to induce more frequent gym usage. Let’s suppose that explanation is correct. What could a competing gym do to *both* make more money and

reduce or eliminate the less than fully rational behavior of their clients? It would certainly not be a great strategy to explain to customers that they could save a lot of money by switching to the 10-ticket package. Not only would the gym be losing money on a per visit basis, but they would also forego the payments from infrequent gym users who procrastinate about quitting. The average person who quits has not been to the gym in 2.3 months. So if competing gyms can't make money by turning them into Econs, who can? I suppose DellaVigna and Malmendier could have started a service convincing people to switch to paying by the visit, but I think they made a wise career choice in selecting academia over personal finance consulting.

The same analysis applies to the recent financial crisis. Many homeowners took out mortgages with initial low "teaser rates." Once the rates reset, some homeowners found they were unable to pay their mortgage payments unless home prices continued to go up *and* mortgage refinancing remained available at low interest. The mortgage lenders who initiated such mortgages and then immediately sold the loans to be securitized made lots of money while it lasted, but the subsequent financial crisis was painful to nearly everyone. Let's assume that at least some of these mortgage borrowers were fooled by fast-talking mortgage brokers.⁴ How would the market solve this problem? No one has ever gotten rich convincing people not to take out unwise mortgages.

Similarly if people fail to follow the dictates of the life-cycle hypothesis and fail to save adequately for retirement, how is the market going to help them? Yes, there are firms selling mutual funds but they are competing with other firms selling fast cars, big screen televisions, and exotic vacations. Who is going to win that battle? The bottom line is there is no magic market potion that miraculously turns Humans into Econs; in fact, the opposite pattern is more likely to occur, namely that markets will exacerbate behavioral biases by catering to those preferences.

The conclusion one should reach from this section is that the explainawaytions are not a good excuse to presume that agents will behave as if they were Econs. Instead we need to follow Milton Friedman's advice and evaluate theories based on the quality of their predictions, and, if necessary, modify some of our theories.

III. Financial Markets⁵

A good place to start in an evaluation of the potential importance of less-than-fully-rational agents is financial markets. I say this because financial markets have the features that should make it hardest to find evidence of misbehavior. These markets have low transaction costs, high stakes, lots of competition (except perhaps in some banking sectors) and crucially, the ability to sell short. It is short selling that allows for the possibility that even if most investors are fools, the activities of "smart money" arbitrageurs can assure that markets behave "as if" everyone were smart. This is the intellectual underpinning of the efficient market hypothesis (EMH).

The efficient market hypothesis really has two distinct components. The first, what I call the "no free lunch" provision, is that it is not possible to "beat the market" on a properly risk-adjusted basis. There is an enormous literature devoted to testing this hypothesis, with many arguments on each side. The difficulty in evaluating competing claims is in agreeing on the way to account for risk. For example, there is widespread agreement in the literature

⁴ Of course some borrowers might have been planning all along to default and live rent free for as long as possible before walking away. They were then acting like Econs.

⁵ This section draws on Barberis and Thaler (2003).

that a strategy of buying “value stocks,” for example those with low ratios of price to earnings or book value, earns higher returns than buying “growth stocks,” which have high price-earnings ratios. However, there is a debate about the explanation for these excess returns. Behavioralists (for example De Bondt and Thaler 1985, 1987; Lakonishok, Shleifer, and Vishny 1994) argue that the excess returns reflect mispricing of some sort. On the other side, efficient market advocates such as Fama and French (1993) argue that the high returns to value stocks occur because those stocks are risky. Although it would not be right to say that this argument has been settled to everyone’s satisfaction, I do think that no one has been able to identify a specific way in which value stocks are riskier than growth stocks. (For example, value stocks tend to have lower betas, the traditional measure of risk in the Capital Asset Pricing Model.) Still, while academics debate about the correct interpretation of these empirical results, one important fact first documented in Jensen’s (1968) PhD thesis remains true: the active mutual fund industry on average does not beat the market.

So from the point of view of an investor, this aspect of the efficient market hypothesis can safely be considered to be at least approximately true. Nevertheless, it is important not to misinterpret this finding. The lack of predictability in stock market returns does not imply that stock market prices are “correct.” This is the second aspect of the EMH, what I call the “price is right” component. The inference that unpredictability implies rational prices is what Shiller (1984, p. 459) once called “one of the most remarkable errors in the history of economic thought.” It is an error because just as the path of a toddler running around on a playground might be completely unpredictable, the path is also not likely to be the result of maximizing some well-formed objective function (other than having fun).

The price is right component of the EMH is, in my opinion, by far the more important of the two ingredients of the theory. It is important because if prices are “wrong” then capital markets are not doing an efficient job of allocating resources.⁶ The problem has been to come up with a convincing test of this part of the theory because the intrinsic value of a security is normally unknowable. If the price of Apple Inc. were too high or too low, how would we know? It turns out that there are classes of assets for which we can say something definitive, namely those for which we can use the law of one price as a test. Although we don’t know the rational price of Apple, we can say for sure that odd-numbered share certificates (if such things still exist) should sell for the same price as even-numbered shares. I have explored several such examples in work with Owen Lamont,⁷ and he recently told me about another one that I will describe here.

One type of security that has provided a fruitful source of tests of the law of one price is closed-end mutual funds. Unlike their open-ended cousins, which accept new investments that are valued at the net asset value of the securities held by the fund, and then redeem withdrawals the same way, closed-end funds are, as their name suggests, closed to new investors. Rather, when the fund starts, a certain amount of money is raised and invested, and then the shares in the fund trade on organized markets such as the New York Stock Exchange. The curious fact about closed-end funds, noted early on by Graham (1949) among others, is that the price of the shares is not always equal to the net asset value of the underlying securities. Funds typically sell at discounts of 10–15 percent, but sometimes sell at substantial premia. This is the story of one such fund.

⁶ Which, of course, is not to say that some other system would do better.

⁷ See Lamont and Thaler (2003a, 2003b).



Figure 1. Price and Net Asset Value for CUBA Fund

Note: On December 18, 2014, President Obama announced he was going to lift several restrictions against Cuba.

Source: Bloomberg

The particular fund I want to highlight here happens to have the ticker symbol CUBA. Founded in 1994, its official name is the Herzfeld Caribbean Basin Fund, which has 69 percent of its holdings in US stocks with the rest in foreign stocks, chiefly Mexican. It gave itself the ticker “CUBA” despite the fact that it owns no Cuban securities nor has it been legal for any US company to do business in Cuba since 1960 (although that may change at some point). The legal proviso, plus the fact that there are no traded securities in Cuba, means that the fund has no financial interest in the country with which it shares a name. Historically, the CUBA fund traded at a 10–15 percent discount to Net Asset Value.

Figure 1 plots both the share price and net asset value for the CUBA fund for a time period beginning in September 2014. For the first few months we can see that the share price is trading in the normal 10–15 percent discount range. Then something abruptly happens on December 18, 2014. Although the net asset value of the fund barely moves, the price of the shares jumped to a 70 percent premium. Whereas it had previously been possible to buy \$100 worth of Caribbean assets for just \$90, the next day those assets cost \$170! As readers have probably guessed, this price jump coincided with President Obama’s announcement of his intention to relax the United States’ diplomatic relations with Cuba. Although the value of the assets in the fund remained stable, the substantial premium lasted for several months, finally disappearing about a year later.

This example and others like it show that prices can diverge significantly from intrinsic value, even when intrinsic value is easily measured and reported daily. What then should we think about broader market indices? Can they also get out of whack? Certainly, the run-up of technology stocks in the late 1990s looked like a bubble at the time, with stocks selling for very high multiples of earnings (or sales for those without profits), and it was followed by a decline in prices of more than two thirds in the NASDAQ index. We experienced a similar pattern in the housing boom in the mid 2000s, especially in some cities such as Las

Vegas and Phoenix. Prices sharply diverged from their long-term trend of selling for roughly 20 times rental prices, only to fall back to the long-term trend. Because of the various forms of leverage involved, this rise and fall in prices helped create the global Great Recession.

The difference between the CUBA example and these much larger bubbles is that it is impossible to prove that prices in the latter were ever wrong. There is no clear smoking gun. But it certainly feels like asset prices can diverge significantly from fundamental value. Perhaps we should adopt the definition of market efficiency proposed by Fischer Black (1986) in his presidential address to the American Finance Association, which had the intriguing one word title “Noise.” Black (1986, p. 553) says “we might define an efficient market as one in which price is within a factor of two of value, i.e., the price is more than half of value and less than twice value. The factor of two is arbitrary, of course. Intuitively, though, it seems reasonable to me, in light of sources of uncertainty about value and the strength of the forces tending to cause price to return to value. By this definition, I think almost all markets are efficient almost all the time. ‘Almost all’ means at least 90 percent.”

One can quibble over various aspects of Black’s definition but it seems about right to me, and had Black lived to see the tech bubble of the 90s he might have revised his number up to three. I would like to make two points about this. The first is that the efficient market hypothesis has been a highly useful, indeed essential concept in the history of research on financial markets. In fact, without the EMH there would have been no benchmark with which to compare anomalous findings. The only danger created by the concept of the EMH is if people, especially policymakers, consider it to be *true*. If policymakers think that bubbles are impossible, then they may fail to take appropriate steps to dampen them. For example, I think it would have been appropriate to raise mortgage-lending requirements in cities where price to rental ratios seemed most frothy. Instead, this was a period in which lending requirements were unusually lax.

There is a broader point to make. For lots of reasons we might expect that financial markets are the most efficient of all markets. They are the only markets where it is generally possible to cheaply sell short, an essential feature if we expect prices to be “right.” Yet if financial markets can be off by a factor of two, how much confidence should we have that prices in other markets are good measures of value, where there are no realistic arbitrage opportunities?

To give just one example, consider labor markets. There has been considerable attention paid in recent years to the growing inequality in incomes and wealth around the world (Piketty 2014; Atkinson, Piketty, and Saez 2011). Although there has been much debate about the cause of this trend, most of the discussion within economics is based on the presumption that differences in income reflect differences in productivity. Is that presumption warranted? If stock prices can be off by a factor of two, might not that be true for workers, from hamburger flippers to CEOs?

There is reason for skepticism about that presumption from the bottom to the top of the income ladder. At the lower end of the wage distribution there has been a long literature begun by Slichter (1950) documenting odd inter-industry wage differentials. Simply put, some industries pay more than others, and this applies to clerical workers and janitors as well as higher paid executives. Important papers by Krueger and Summers (1988) and Dickens and Katz (1986) reignited this literature summarized in Thaler (1989). Card, Heining, and Kline (2013) have recently documented similar findings in Germany using panel data that allow for individual fixed effects. They find that when workers move from a bottom quartile paying industry to a top quartile industry their wages jump, and the opposite thing happens when workers move from a high paying industry to a low one. It seems implausible that these workers become significantly more or less productive simply

by changing industries.

At the other end of the spectrum, the ratio of CEO pay to that of the average worker has skyrocketed in the past few decades. In 1965 for large firms based in the United States this ratio was 20; by 2014 it was over 300, more than twice the ratio in any other country (Mishel and Davis 2015). Of course some economists argue that this rise simply reflects the growing productivity of the CEOs (e.g., Kaplan, Klebanov, and Sorensen 2012) but how confident should we be in this assessment? CEO pay is usually set by the compensation committees of boards of directors that rely on consultants who base their recommendations in part on the pay of other CEOs. This kind of recursive, self-fulfilling process is not one that generates high confidence that pay and performance are highly correlated. Of course there is no way to settle this argument. Rather, I just want to repeat my question. If stock prices can be off by a factor of two, why should we be confident that other markets do not diverge by that much, or more?

IV. One Theory, Two Tasks

The conclusion I reach from research in behavioral finance is that even these most efficient of markets often lead to empirical results that are inconsistent with theories based on rational investors making choices in markets with tiny transaction costs. In other words, the results we obtain are not consistent with the hypothesis that investors behave “as if” they were rational. And there should be even greater suspicion that such models will make good predictions in other markets where arbitrage is impossible. So what should happen to economic theory?

The problem is that we are asking our theories to do two different tasks. The first is to solve for optimal solutions to problems, the other is to describe how Humans actually choose. Of course in a world consisting only of Econs there would be no need for two different kinds of models. Economic agents would have the courtesy to make the optimal choices that the model determines are best (at least on average). But we are far from that world: We Humans struggle both to determine what the best choice would be and then to have enough willpower to implement that choice, especially if it requires delay of gratification. So we need descriptive economic theories.

The first and most successful such theory is Kahneman and Tversky’s (1979) prospect theory, which has had an enormous impact on both economics and social science more generally.⁸ Beyond the insights of the model itself, prospect theory provides a template for the new class of theories we need. Expected utility theory remains the gold standard for how decisions should be made in the face of risk. Prospect theory is meant to be a complement to expected utility theory, which tell us how people actually make such choices. Using one theory for both purposes makes no more sense than using a hammer both to pound nails and to apply paint.

Some economists might think that without optimization there can be no theory, but in a cogent essay Arrow (1986) rejected this idea. “Let me dismiss a point of view that is perhaps not always articulated but seems implicit in many writings. It seems to be asserted that a theory of the economy must be based on rationality, as a matter of principle. Otherwise there can be no theory.” Arrow noted that there could be many rigorous, formal theories based on behavior that economists would not be willing to call rational. He also pointed out the inconsistency of an economic theorist who toils for months to derive the optimal solution to some complex economic problem and then blithely assumes that the agents in

⁸ According to Google Scholar the paper has been cited nearly 40,000 times.

his model behave as if they are naturally capable of solving the same problem. “We have the curious situation that scientific analysis imputes scientific behavior to its subjects. This need not be a contradiction, but it does seem to lead to an infinite regress.”

This is not the place, and I am not the person, to present a detailed roadmap of what a behavioral approach to economic theory should be, but perhaps a few brief thoughts are appropriate. The first is that behavioral economic theories (or any descriptive theories) must abandon the inductive reasoning that is the core of neoclassical theories and instead adopt a deductive approach in which hypotheses and assumptions are based on observations about human behavior. In other words, behavioral economic theory must be evidence-based theory. The evidence upon which these theories can be based can come from psychology or other social sciences or it can be homemade. Some might worry about basing theories on empirical observation, but this methodology has a rich tradition in science. The Copernican revolution, which placed the sun at the center of our solar system rather than the earth, was based on data regarding the movement of the planets, not on some first principles.

A second general point is that we should not expect some new grand behavioral theory to emerge to replace the neoclassical paradigm. We already have a grand theory and it does a really good job of characterizing how optimal choices and equilibrium concepts work. Behavioral theories will be more like engineering, a set of practical enhancements that lead to better predictions about behavior. So far, most of these behavioral enhancements focus on two broad topics: preferences and beliefs.

A. Behavioral Preferences

Prospect theory is a good illustration of a model based on assumptions about preferences that differ from the ones used to derive expected utility theory. Specifically, most of prospect theory’s predictive power comes from three crucial assumptions about preferences. First, utility is derived from changes in wealth relative to some reference point, rather than levels of wealth, as is usually assumed in theories based on expected utility.⁹ Second, the “value function” which translates perceived changes in wealth into utility, has a kink at the origin, with losses weighed more heavily than gains—i.e., “loss aversion.” Third, decision weights are a function of probabilities $\Pi(p)$ where $\Pi(p) \neq p$. These aspects of the theory were inferred from studying the choices subjects made when asked to choose between gambles.

Two other research streams have been based on models of preferences. The first topic is intertemporal choice. As revealed by the quotations from Smith, Pigou, and Fisher mentioned earlier in this essay, economists have long worried that people display what we now call “present biased” preferences, meaning that the discount rate between “now” and “later” is much higher than between “later” and “even later.” Such preferences can lead to time-inconsistent behavior since we expect to be patient in choosing between a smaller reward in a year and a larger reward in a year plus a week, but when the year passes and the smaller reward is available “now,” we submit to temptation. If people realize they have such preferences, they may choose to commit themselves now to choosing the larger

⁹ It is true that von Neumann and Morgenstern do not specify what the arguments are in their utility function, and some have argued that one could simply revise expected utility to be a function of income rather than wealth to incorporate this feature. What this misses is that defining “income” depends on a theory of mental accounting in order to know over what time horizon income is being measured. If “income” is lifetime income then it is the same as wealth. But if it is daily income then one gets very different predictions. See, for example, the controversial literature on taxi cab driver labor supply (Camerer et al. 1997; Crawford and Meng 2011; Farber 2015).

delayed reward, a strategy they will later regret (at least for a week or so).

Two kinds of models have been proposed to deal with these aberrant preferences. One is based on a two-self (or “two-system”) approach that is meant to capture the inherent conflict that defines self-control problems. In the version of this type of model that Hersh Shefrin and I favor (Thaler and Shefrin 1981) individuals are assumed to have a long-sighted “planner” and myopic “doer” that interact in a model similar to agency models of organizations. Schelling (1984) and Fudenberg and Levine (2006) also proposed two-self models to characterize this behavior.

Although these two-self models provide more psychological texture, they have not been as popular among economic theorists as the simpler and more tractable “beta-delta” model originally proposed by Strotz (1955) and then refined by Laibson (1997) and O’Donoghue and Rabin (1999). In these models delta is the standard exponential discount rate and beta measures short-term impatience. The standard model is just the special case in which beta is 1.0. The beta-delta model is a good example of what Rabin (2013) calls PEEMs, which stands for “portable extensions of existing models.” The ease with which economists can incorporate such models into an otherwise standard analysis has obvious appeal.

Along with intertemporal choice, the important aspect of preferences that has received a lot of attention from behavioral economic theorists is “other-regarding preferences.” These models were all stimulated by empirical findings showing that humans are not completely selfish, even to strangers. For example, in one-shot prisoners’ dilemma games about 40–50 percent of subjects cooperate, both in laboratory experiments and even in a game show environment where the stakes are over £10,000 (van den Assem, van Dolder, and Thaler 2012). Similarly, people cooperate in public goods environments when the rational selfish strategy is to give nothing. The most prominent models in this space are by Rabin (1993) and Fehr and Schmidt (1999). The easiest way to summarize this literature is to say that Humans are nicer and more mannerly than Econs. Specifically, their first instinct is to cooperate as long as they expect others to do likewise.

B. Behavioral Beliefs

When people make choices they do so based on a set of expectations about the consequences of their choices and the many exogenous factors that can determine how the future will evolve. Traditionally, economists assume that such beliefs are unbiased. Although the rational expectations hypothesis as first formulated by Muth (1961) and elaborated upon by Lucas (1976) and many others is often considered to be a specific approach to economic modeling, especially in macroeconomics, I think it is fair to say that the essential idea is entirely mainstream. The assumption of rational expectations makes explicit an idea that is commonplace in economic theory, namely that agents act as if they understood the model (and state-of-the-art econometrics techniques as well). Whether this assumption is empirically valid is another question.

Explicit tests of rational expectations per se are uncommon because we rarely observe or elicit actual expectations data. When we do, we often find that actual expectations diverge from what would reasonably be called rational. For example, Case, Shiller, and Thompson (2012) find that homeowners during the period of rapidly rising prices from 2000–2005 expected home prices to continue to rise at double-digit rates for the next decade. While one can’t prove such expectations were irrational, they certainly seem excessively optimistic, both ex ante and ex post. Furthermore, in this domain and in many others, expectations seem to rely too much on extrapolation of recent trends. To a first approximation, people expect that what goes up will continue to go up.

We also see violations of rational expectations in the predictions of stock market returns by chief financial officers studied by Ben-David, Graham, and Harvey (2013). The CFOs were asked to predict one-year rates of return on the S&P 500 and also give 80 percent confidence limits. Perhaps unsurprisingly, the CFOs had essentially no ability to predict returns in the stock market. What is more disturbing is that they had no self-awareness of their lack of predictive skills. If the CFOs had well-calibrated forecasts the actual stock-market return would fall between their high and low estimate 80 percent of the time. Instead, their ranges included the actual outcome for just 36 percent of the forecasts recorded over a ten-year period. This is quite similar to the overconfidence observed in dozens of laboratory studies.

Overconfidence and excessive extrapolation are just two examples of biased beliefs that have been documented by psychologists studying human judgment. This literature began with the original three heuristics studied by Kahneman and Tversky—availability, representativeness, and anchoring and adjustment—but many others have been investigated and documented since then: hindsight bias, projection bias, excessive attention to whatever feature of the environment is most salient, etc. For each of these biases and many more, economists have created descriptive models to try to make the implications of the biases more specific and rigorous.

The fact that there is a long list of biases is both a blessing and a curse. The blessing is that there are a multitude of interesting ways in which human judgment diverges from rational expectations, each of which offers the possibility of providing useful insights into economic behavior. The curse is that the length of the list seems to offer theorists a dangerously large number of degrees of freedom. Although I do not dismiss this latter risk out of hand, I think good scientific practices can mitigate this degrees-of-freedom risk.

The most important thing to remember is that all these biases have empirical support, and many of the laboratory findings have subsequently been replicated in the field. Thus some discipline has already been imposed: behavioral economists can draw on a long list of potential explanatory factors, but for each there is at least some evidence that the factor is real. Compare this with the degrees of freedom available in traditional rationality-based models. For example, consider the all-purpose fudge factor: transaction costs. In the abstract such costs can explain many anomalies, but unless those costs can be measured the use of the concept is undisciplined. If we limit ourselves to variables that have an empirical basis, all of economics will become more disciplined.

Of course I do not mean to suggest that behavioral economic theory is a finished product. The field is new and growing rapidly. One goal should be to devise theories that are not just portable extensions of existing models but also *testable* extensions. I will leave it to Rabin to decide where to insert the letter T into his PEEM acronym.

V. Supposedly Irrelevant Factors

It is rare that economic theory makes predictions about magnitudes. Mostly theories make predictions about the sign of an effect. Demand curves slope down; supply curves slope up. When a clever theorist is able to extract a more precise prediction from the theory, things can get interesting. The equity premium puzzle is a case in point. The first-order prediction that stocks are riskier than bonds and so should earn a higher rate of return is resoundingly supported by the historical data. But Mehra and Prescott (1985) showed that the standard model cannot simultaneously explain the low historical risk-free rate and an equity premium in the neighborhood of 6 percent—the largest value they could justify was 0.35 percent. As a result of this calibration exercise a long and interesting literature ensued.

Although such examples of predictions about magnitude are uncommon, economic theory does make some rather precise predictions about effect sizes, namely for variables that should have no effect at all on behavior. For example the following things should not matter: the framing of a problem, the order in which options are displayed, the salience of one option over another, the presence of a prior sunk cost (or gain), whether the customer at a restaurant can see the dessert options when choosing whether to stick to the planned diet, and so forth. I call these, and a multitude of other possible variables that can and do influence choices, “supposedly irrelevant factors” or SIFs. One of the most important ways in which behavioral economics can enrich economic analyses is by pointing out the SIFs that matter most.

One domain in which the potential importance of SIFs has been best documented is retirement saving. In a standard life-cycle model Econs compute their optimal consumption path and then implement a plan of saving, investing, and eventually dis-saving that maximizes lifetime utility, fully incorporating proper actuarial probabilities of mortality rates for husband and wife as well as risks of divorce, illness, and so forth. This is a problem that makes playing world-class chess seem easy. Chess has neither uncertainty nor self-control problems to muck up the works. So it should not be surprising that many Humans have trouble dealing with retirement saving in a defined-contribution world in which they have to make all the decisions themselves. However, it has been possible to help people with this daunting task with the aid of some SIFs.

The first SIF that has been important in helping people to save for retirement is the intelligent use of the default option. In a world of Econs, especially when the stakes are as high as they are for retirement saving, it should not matter whether someone gets signed up for the plan unless he opts out or is excluded from the plan unless he opts in. The cost of ticking a box and filling out a form must be tiny compared to the benefits of receiving a company match and tax-free accumulations for decades. Nevertheless, changing the default has had an enormous impact on the utilization rates of 401(k) plans.

The first paper to document this effect was Madrian and Shea (2001) using data from a company that had adopted what is now called “automatic enrollment” in 1999. Previously, to join the 401(k) plan employees had to fill in some forms, and if they failed to do so, they were not enrolled. Madrian and Shea compared the enrollment rates for new employees in 1998 under the old “opt in” regime to those in 1999 where employees had to opt out if they did not want to join. Before automatic enrollment, only 49 percent of employees joined the plan within their first year of employment; after the switch to automatic enrollment, 86 percent of the employees were enrolled in their first year. Supposedly irrelevant indeed! By now automatic enrollment is widespread. More than half of large US employers are using the concept and the United Kingdom is in the process of rolling out a national defined contribution savings plan with this feature. Most plans, including the national UK plan, find that opt out rates are around 10 percent.

One problem with automatic enrollment is that many plans initially enroll employees at a low savings rate; in the United States it is often just 3 percent of pay. As Madrian and Shea pointed out in their initial paper, such a low initial default savings rate can have the unintended consequence of reducing the savings of those who, lacking a default, would have chosen to save more. As one solution to this problem, and more generally as a way to nudge employees to increase their savings rates, Shlomo Benartzi and I (Thaler and Benartzi 2004) introduced a plan we called “Save More Tomorrow.” Under this plan, workers are offered the option to increase their savings rate starting at some later date, ideally when they get their next raise. Once an employee enrolls in the plan, her savings rate keeps increasing until she reaches some cap or opts out.

Notice that Save More Tomorrow is just a collection of SIFs. It should not matter that the savings rate is increased in a few months rather than now, nor that the increases are linked to pay increases, nor that the default is to stay in the plan, but of course all these features help. Putting off the increase in saving to the future helps those who are present biased; linking to increases in pay mitigates loss aversion; and making staying in the plan the default puts status quo bias to good use. In the first plan Benartzi and I studied (Thaler and Benartzi 2004), savings rates more than tripled in three years. In a recent paper (Benartzi and Thaler 2013) we estimated that automatic escalation (the generic term for Save More Tomorrow, in which savings increases are not always linked to pay increases) had boosted annual savings by \$7.4 billion.

One worry about such programs has been that the increases we observe in retirement savings produced by automatic enrollment and Save More Tomorrow might be offset by reductions in savings (or increases in borrowing) in other accounts. However, there was no dataset in the United States that allowed anyone to test this hypothesis. Fortunately, such data do exist in Denmark, which, because of a history of having a wealth tax, has long kept good data on household wealth. A recent paper by Chetty et al. (2014) has made use of these data to answer this question.

The method Chetty et al. (2014) use is to see what happens to savings rates when an employee moves jobs to an employer with a more generous retirement savings plan. Using panel data with 41 million person-year observations the authors study three kinds of savings: employer contributions to tax-sheltered pensions, employee contributions to those pensions, and employee savings in taxable accounts. Their research strategy is to study those employees who have been saving a positive amount on their own and then switch to a firm whose contributions are at least 3 percentage points higher. On average these workers receive an increase in pension contributions of 5.64 percent of labor income. Do workers contribute less to compensate for this change in their employers generosity? Yes, but just by 0.56 percentage points. And saving in taxable accounts is essentially unchanged.

Chetty et al. (2014) also make use of a change in tax policy that occurred during the period for which they have data. This natural experiment allowed them to compare the effectiveness of the tax subsidy given to pension contributions in encouraging retirement savings relative to the effects of design features such as the employer contribution. The change in the law they exploit was a reduction in the subsidy given to retirement saving for roughly the top quintile of the income distribution. Even among this relatively affluent group, the vast majority did not react at all to the change in the subsidy—they were “passive savers.” About 20 percent of this segment did react and eliminated all their contributions to the tax-sheltered plans, but they did not spend that money; they just shifted it to taxable savings vehicles. This leads to a remarkable conclusion. Each \$1 of tax expenditure on retirement savings only produced a penny in increased savings. What determines savings rates is not tax policy but the design features of the employer pension plans, i.e., SIFs.

There are many other examples of the potential power of behavioral factors in policy analysis but summarizing them would be a waste of time. I cannot possibly do a better job of that than Raj Chetty (2015) did last year in his Ely lecture: “Behavioral Economics and Public Policy: A Pragmatic Perspective.” I completely endorse his view that the best way to proceed is to stop arguing about theoretical principles and just get down to work figuring out the best way of understanding the world.

VI. Conclusion

There is one central theme of this essay: it is time to fully embrace what I would call *evidence-based economics*. This should not be a hard sell. Economists use the most sophisticated statistical techniques of any social science, have access to increasingly large and rich datasets, and have embraced numerous new methods from experiments (both lab and field) to brain imaging to machine learning. Furthermore, economics has become an increasingly empirical discipline. Hamermesh (2013) finds that the percentage of “theory” papers in top economics journals has fallen from 50.7 percent in 1963 to 19.1 percent in 2011. We are undeniably an empirical discipline—so let’s embrace that.

Viewed in this context, behavioral economics is simply one part of the growing importance of empirical work in economics. There is nothing unique about incorporating psychological factors such as framing, self-control, and fairness into economics analyses. If such factors help us understand the world better and improve predictions about behavior, then why wouldn’t we use them just like we would use any other new source of data such as web searches or genetic markers?

In this sense I think it is time to stop thinking about behavioral economics as some kind of revolution. Rather, behavioral economics should be considered simply a return to the kind of open-minded, intuitively motivated discipline that was invented by Adam Smith and augmented by increasingly powerful statistical tools and datasets. This evidence-based discipline will still be theoretically grounded, but not in such a way that restricts our attention to only those factors that can be derived from our traditional normative traditions. Indeed, my sense is that we are at the beginning of a new wave of theoretical developments made possible simply by turning our attention to the study of Humans rather than Econs.

If economics does develop along these lines the term “behavioral economics” will eventually disappear from our lexicon. All economics will be as behavioral as the topic requires, and as a result, we will have an approach to economics that yields a higher R^2 .

REFERENCES

- Arrow, Kenneth J. 1986. “Rationality of Self and Others in an Economic System.” *Journal of Business* 59 (4, part 2): S385–S399.
- Ashraf, Nava, Colin F. Camerer, and George Loewenstein. 2005. “Adam Smith, Behavioral Economist.” *Journal of Economic Perspectives* 19 (3): 131–45.
- Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. “Top Incomes in the Long Run of History.” *Journal of Economic Literature* 49 (1): 3–71.
- Barberis, Nicholas, and Richard Thaler. 2003. “A Survey of Behavioral Finance.” In *Handbook of the Economics of Finance*, Vol. 1B, edited by George M. Constantinides, Milton Harris, and Rene Stulz, 1053–1123. Amsterdam: Elsevier.
- Benartzi, Shlomo, and Richard H. Thaler. 2013. “Behavioral Economics and the Retirement Savings Crisis.” *Science* 339 (6124): 1152–53.
- Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey. 2013. “Managerial Miscalibration.” *Quarterly Journal of Economics* 128 (4): 1547–84.
- Black, Fischer. 1986. “Noise.” *Journal of Finance* 41 (3): 529–43.
- Camerer, Colin. 1997. “Labor Supply of New York City Cabdrivers: One Day at a Time.” *Quarterly Journal of Economics* 112 (2): 407–41.
- Card, David, Jorg Heining, and Patrick Kline. 2013. “Workplace Heterogeneity and the Rise of West German Wage Inequality.” *Quarterly Journal of Economics* 128 (3): 967–1015.
- Case, Karl E., Robert J. Shiller, and Anne Thompson. 2012. “What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets.” National Bureau of Economic Research Working Paper 18400.
- Chetty, Raj. 2015. “Behavioral Economics and Public Policy: A Pragmatic Perspective.” *American Economic Review* 105 (5): 1–33. Video of lecture: <https://www.aeaweb.org/webcasts/2015/Ely.php>.
- Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen. 2014. “Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark.” *Quarterly*

- Journal of Economics* 129 (3): 1141–1219.
- Clark, J. M. 1918. "Economics and Modern Psychology: I." *Journal of Political Economy* 26 (1): 1–30.
- Crawford, Vincent P., and Juanjuan Meng. 2011. "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *American Economic Review* 101 (5): 1912–32.
- De Bondt, Werner F. M., and Richard Thaler. 1985. "Does the Stock Market Overreact?" *Journal of Finance* 40 (3): 793–805.
- De Bondt, Werner F. M., and Richard H. Thaler. 1987. "Further Evidence on Investor Overreaction and Stock Market Seasonality." *Journal of Finance* 42 (3): 557–81.
- DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review* 96 (3): 694–719.
- Dickens, William T., and Lawrence F. Katz. 1986. "Interindustry Wage Differences and Industry Characteristics." National Bureau of Economic Research Working Paper 2014.
- Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stock and Bonds." *Journal of Financial Economics* 33 (1): 3–56.
- Farber, Henry S. 2015. "Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers." *Quarterly Journal of Economics* 130 (4): 1975–2026.
- Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–68.
- Fisher, Irving. 1930. *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*. New York: MacMillan.
- Friedman, Milton. 1953. "The Methodology of Positive Economics." In *Essays in Positive Economics*, 3–43. Chicago: University of Chicago Press.
- Fudenberg, Drew, and David K. Levine. 2006. "A Dual Self Model of Impulse Control." Harvard Institute of Economic Research Working Paper 2112.
- Graham, Benjamin. 1973. *The Intelligent Investor, A Book of Practical Counsel*. 4th revised ed. New York: Harper & Row, (Orig. pub. 1949).
- Grether, David M., and Charles R. Plott. 1979. "Economic Theory of Choice and the Preference Reversal Phenomenon." *American Economic Review* 69 (4): 623–38.
- Hamermesh, Daniel S. 2013. "Six Decades of Top Economics Publishing: Who and How?" *Journal of Economic Literature* 51 (1): 162–72.
- Hogarth, Robin M., and Melvin W. Reder, eds. 1986. "The Behavioral Foundations of Economic Theory: Proceedings of a Conference October 13–15, 1985." *Journal of Business* 59 (4, part 2).
- Hogarth, Robin M., and Melvin W. Reder, eds. 1987. *Rational Choice: The Contrast Between Economics and Psychology*. Chicago: University of Chicago Press.
- Jensen, M. C. 1968. "The Performance of Mutual Funds in the Period 1945–1964." *Journal of Finance* 23: 389–416.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–91.
- Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen. 2012. "Which CEO Characteristics and Abilities Matter?" *Journal of Finance* 67 (3): 973–1007.
- Katona, George. 1951. *Psychological Analysis of Economic Behavior*. New York: McGraw-Hill.
- Katona, George. 1953. "Rational Behavior and Economic Behavior." *Psychological Review* 60 (5): 307–18.
- Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money*. London: Macmillan.
- Krueger, Alan B., and Lawrence H. Summers. 1988. "Efficiency Wages and the Inter-industry Wage Structure." *Econometrica* 56 (2): 259–93.
- Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112 (2): 443–77.
- Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance* 49 (5): 1541–78.
- Lamont, Owen A., and Richard H. Thaler. 2003a. "Anomalies: The Law of One Price in Financial Markets." *Journal of Economic Perspectives* 17 (4): 191–202.
- Lamont, Owen A., and Richard H. Thaler. 2003b. "Can the Market Add and Subtract? Mispricing in Tech Stock Carve-Outs." *Journal of Political Economy* 111 (2): 227–68.
- Lester, Richard A. 1946. "Shortcomings of Marginal Analysis for Wage-Employment Problems." *American Economic Review* 36 (1): 63–82.
- Lichtenstein, Sarah, and Paul Slovic. 1971. "Reversals of Preference Between Bids and Choices in Gambling Decisions." *Journal of Experimental Psychology* 89 (1): 46–55.
- Lichtenstein, Sarah, and Paul Slovic. 1973. "Response-Induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas." *Journal of Experimental Psychology* 101 (1): 16–20.
- Loewenstein, George. 1992. "The Fall and Rise of Psychological Explanations in the Economics of Intertemporal Choice." In *Choice Over Time*, edited by George Loewenstein and Jon Elster, 3–34. New York: Russell Sage Foundation.

- Lucas, Robert E., Jr. 1976. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.
- Machlup, Fritz. 1946. "Marginal Analysis and Empirical Research." *American Economic Review* 36 (4): 519–54.
- Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 1149–87.
- Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics* 15 (2): 145–61.
- Mishel, Lawrence, and Alyssa Davis. 2015. "Top CEOs Make 300 Times More than Typical Workers." Economic Policy Institute Issue Brief 399. <http://www.epi.org/publication/top-ceos-make-300-times-more-than-workers-pay-growth-surpasses-market-gains-and-the-rest-of-the-0-1-percent/>.
- Muth, J. F. 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica* 29: 315–35.
- O'Donoghue, Ted, and Matthew Rabin. 1999. "Procrastination in Preparing for Retirement." In *Behavioral Dimensions of Retirement Economics*, edited by Henry J. Aaron, 125–56. Washington, DC: Brookings Institution.
- Pareto, Vilfredo. 2014. *Manual of Political Economy: A Critical and Variorum Translation Edition*. Reprint edited by Aldo Montesano, Alberto Zanni, Luigino Bruni, John S. Chipman, and Michael McLure. Oxford, UK: Oxford University Press (Orig. pub. 1906).
- Pigou, Arthur Cecil. 1920. *The Economics of Welfare*. <http://oll.libertyfund.org/title/1410> (accessed May 23, 2016).
- Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge, MA: Harvard University Press.
- Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5): 1281–1302.
- Rabin, Matthew. 2013. "An Approach to Incorporating Psychology into Economics." *American Economic Review* 103 (3): 617–22.
- Schelling, Thomas C. 1984. "Self-Command in Practice, in Policy, and in a Theory of Rational Choice." *American Economic Review* 74 (2): 1–11.
- Shiller, Robert J. 1984. "Stock Prices and Social Dynamics." *Brookings Papers on Economic Activity* (2): 457–98.
- Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69 (1): 99–118.
- Simon, Herbert A. 1957. *Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in a Social Setting*. Oxford, UK: Wiley.
- Simon, Herbert A. 1987. "Behavioural Economics." In *The New Palgrave: A Dictionary of Economics*, edited by John Eatwell, Murray Milgate, and Peter Newman. Palgrave Macmillan.
- Slichter, Sumner H. 1950. "Notes on the Structure of Wages." *Review of Economics and Statistics* 32 (1): 80–91.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. Reprint edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Classics, 1981.
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Reprint edited by R. H. Campbell and A. S. Skinner. Indianapolis: Liberty Classics, 1981.
- Strotz, R. H. 1955. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies* 23 (3): 165–80.
- Thaler, Richard H. 1980. "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior & Organization* 1 (1): 39–60.
- Thaler, Richard H. 1989. "Interindustry Wage Differentials." *Journal of Economic Perspectives* 3 (2): 181–93.
- Thaler, Richard H. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: W. W. Norton & Company.
- Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112 (1): S164–87.
- Thaler, Richard H., and H. M. Shefrin. 1981. "An Economic Theory of Self-Control." *Journal of Political Economy* 89 (2): 392–406.
- Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–31.
- van den Assem, Martijn J., Dennie van Dolder, and Richard H. Thaler. 2012. "Split or Steal? Cooperative Behavior When the Stakes Are Large." *Management Science* 58 (1): 2–20.
- von Neumann, John, and Oskar Morgenstern. 1947. *Theory of Games and Economic Behavior*. 2nd ed. Princeton: Princeton University Press.
- Vulkan, Nir. 2000. "An Economist's Perspective on Probability Matching." *Journal of Economic Surveys* 14 (1): 101–18.