Richard Thaler and the Rise of Behavioral Economics

Nicholas Barberis

Yale University

January 2018

Abstract

The emergence of behavioral economics is one of the most significant conceptual developments in the social sciences in the past 40 years. The central figure in the field in its early years was Richard Thaler. In this article, I review and discuss his scientific contributions.

JEL classification: B2, D9, G1

Keywords: endowment effect, prospect theory, mental accounting, nudge

*I am grateful to Hunt Allcott, James Choi, Stefano DellaVigna, Keith Ericson, Owen Lamont, Ulrike Malmendier, Matthew Rabin, and Jesse Shapiro for helpful discussions about the content of this article.
I. Introduction

From the 18th century to the first half of the 20th century, the leading economists of the day – figures such as Adam Smith, John Maynard Keynes, and Irving Fisher – were known to bring aspects of human psychology into their analysis of the economy. By the middle of the 20th century, however, this practice was less common, and with the advent of the rational expectations revolution in the 1960s, economists began to focus almost exclusively on models with the same, tightly-specified assumptions about individual psychology: that people have rational beliefs, and that they make decisions according to Expected Utility.

In the early 1980s, a small group of economists began to argue that the rational expectations revolution had gone too far, and that to understand many important economic phenomena, it was critical to develop new models that made assumptions about human behavior that were psychologically more realistic, and that, in particular, allowed for less than fully rational thinking. This message was roundly dismissed at first, sometimes in scornful terms, but it gradually gained traction. Today, “behavioral economics” – the effort to improve our understanding of the economy through psychologically-realistic models – is firmly established: hundreds of papers in this tradition, many of them highly cited, have appeared in the top economics and finance journals; dozens of specialists in the area have been hired and tenured at leading universities; and conferences on the topic attract large and growing crowds. Moreover, this approach to economics has generated significant interest beyond academia, among non-academic economists, policy makers, and the public at large.

The growth of behavioral economics over the decades is the result of a collective effort by many researchers. But if there is one person who was central to the rise of behavioral eco-
nomics, particularly in its early years, it would be Richard Thaler, the 2017 Nobel Laureate in economics. To appreciate how central he was, consider the four factors that have driven the growth of the field. First, researchers documented numerous “anomalies” – empirical facts that are hard to square with the traditional, rational model of economic behavior. Second, they developed a new generation of models based on more psychologically realistic assumptions. Third, they found ways of helping people to make better economic decisions. And fourth, the behavioral economics movement attracted talented young researchers who accelerated the development of the field.

Thaler played an important role in all four of these factors. First, he documented a number of anomalies, and emphasized these and other anomalies relentlessly, most famously in a series of columns in the *Journal of Economic Perspectives*. Second, he began the effort to develop new, more psychologically realistic frameworks, and several of the intuitions he described decades ago remain at the core of the latest models. Third, he was at the forefront of finding ways to improve people’s economic decisions, most notably through his efforts to raise saving rates in U.S. retirement plans. And fourth, he was the primary mentor to the generation of behavioral economists that followed him. In the course of making these broad contributions, Thaler came up with a number of specific concepts and ideas that have endured, including the endowment effect, mental accounting, and the Save More Tomorrow plan for increasing saving. But perhaps his single most influential insight was his recognition, after encountering the work of Daniel Kahneman and Amos Tversky, that their research on the psychology of judgment and decision-making offered a foundation for a new generation of more psychologically realistic economic models.

In this essay, I review Thaler’s contributions in more detail. The article is structured
around the four factors I listed above. Section II discusses the anomalies that Thaler studied. Section III describes the many ways in which he influenced the development of behavioral economics models. Section IV reviews his research on ways to improve economic decision-making. And in Section V, I discuss his efforts to train and mentor a new generation of behavioral economists.¹

II. Anomalies

In the early 1970s, when Thaler was a graduate student at the University of Rochester, the rational expectations revolution had begun in earnest. Not surprisingly, then, his dissertation, in which he estimated the economic value of a human life, took a traditional, rational approach, one based on comparing wages across professions with different rates of accidental death. One day, it occurred to him that he might learn something by conducting some surveys. Specifically, he asked survey participants how much they would be willing to pay to reduce their probability of dying over the next year by 0.001, but also how much they would need to be paid in order to accept an 0.001 increase in this probability. When reviewing people’s answers, Thaler noticed something curious: the amount people were willing to pay to reduce their probability of dying was much lower than the amount they required in order to accept an increase in this probability, even though traditional economic theory predicted that the two quantities would be roughly equal. This was Thaler’s first encounter with the “endowment effect,” the most famous of the anomalies he studied: the finding that the amount people are willing to pay for an object of economic value is much lower than the

¹For a review of Thaler’s work, it is hard to beat his own book, Misbehaving (Thaler, 2015); some of the historical background in this article is drawn from there. I try to add value by organizing the material differently: while Misbehaving proceeds chronologically, this essay is structured by conceptual theme.
amount they are willing to accept in order to give the object up.

Thaler first described the endowment effect in a 1980 paper (Thaler, 1980). He used the above example, but also others he had come across: the colleague who wouldn’t sell a bottle of wine in his collection for $100, but who also wouldn’t spend more than $35 for a new one; or the acquaintance who wouldn’t pay $8 for someone to mow his lawn, but who also wouldn’t mow his neighbor’s lawn for $20. In the decade that followed, there were a number of studies of the effect, but it was an experiment that Thaler described in a 1990 paper with Kahneman and Jack Knetsch that has become the focus of attention (Kahneman et al., 1990). In this experiment, half of the participants are given a mug and are asked, for each price on a list of prices, whether they would be willing to give up the mug in exchange for that amount of money. The remaining participants are not given a mug; instead, they are asked, for each price on a list of prices, whether they would be willing to pay that amount to receive a mug. The key finding is that the average amount participants require in order to give up their mug is roughly double the average amount participants without mugs are willing to pay to obtain one.

While the endowment effect is at odds with the traditional framework, the question remains: How should it be interpreted? When Thaler first encountered the effect in the 1970s, he was not sure of the answer, but within a few years he had one, one that would prove very influential. I return to this, and to current thinking about the endowment effect, in Section III.

In the 1980 paper where he first described the endowment effect, Thaler also emphasized some other anomalies culled from his observations of human behavior. One is the sunk cost effect, illustrated by the memorable example of a family who, having previously paid $40 for
tickets to a basketball game, decide to drive through a snowstorm to attend the game, even though, had they received the tickets for free, they would have stayed home. Another is a violation of traditional models of search: a shopper is willing to drive across town to save $5 on a $25 item, but not on a $500 item. And another is limited self-control, exemplified by the guests who, unable to stop eating the cashews brought out as a pre-dinner snack, feel relieved when the host removes the nuts and hides them in the pantry. All of these anomalies continue to be discussed today, some more than others. Self-control problems, in particular, are a major topic of study that I return to in Section III.

In the early 1980s, Thaler became increasingly interested in financial economics. Researchers in that field were working with high-quality data on asset prices – data that Thaler thought might allow him to identify anomalies through systematic empirical analysis rather than “just” through surveys, experiments, or observing others’ behavior. Moreover, any anomaly he uncovered in financial markets would be very striking: at the time, most academic economists believed that irrational thinking had little to no impact on the prices of financial assets.

A fundamental question in finance asks why some assets have higher average returns than others. The benchmark rational model for thinking about this is the Capital Asset Pricing Model, or CAPM. According to this model, an asset’s average return is determined by a measure of the asset’s risk known as “beta” – the (scaled) covariance of the asset’s return with the return on the overall market portfolio – and by beta alone. Tests of this prediction in the early 1970s found some support for it, leading economists at the time to feel that the variation in average returns across U.S. stocks was well understood.

In the late 1970s and early 1980s, some problems with this view began to appear. Re-
searchers found that variables other than beta have predictive power for stock returns: even after controlling for beta, stocks with low market capitalizations have higher average returns than stocks with high market capitalizations, and stocks with low price-to-earnings ratios have higher average returns than stocks with high price-to-earnings ratios (Banz, 1981; Basu, 1983). The papers that reported these findings adopted a moderate tone: the authors noted that their results might reflect mispricing, but, perhaps because of the dominance of the rational paradigm at the time, they did so only briefly and with little fanfare.

Given this state of affairs, Thaler’s first finance paper, De Bondt and Thaler (1985), caused a stir. The paper showed that, on average, stocks with big losses over the previous three years subsequently outperform stocks with big gains over the previous three years in a way that is not captured by beta. Moreover, the authors gave this finding an unabashedly behavioral interpretation: that stocks with big losses over the previous three years are stocks that investors have become irrationally pessimistic about, and that the high subsequent returns of these stocks represent a correction back up to fair value.

Today, De Bondt and Thaler’s (1985) finding continues to be seen as a robust fact about the returns of stocks, both in the U.S. market and in most international markets. However, after years of debate, there is still no consensus on how best to interpret it. While many economists subscribe to De Bondt and Thaler’s interpretation, others have sought to explain the finding within the rational framework, arguing that past loser stocks have higher average returns because they are riskier, although the sense in which they are riskier has not yet been clearly demonstrated.

The importance of De Bondt and Thaler’s paper goes beyond the finding of return predictability. In 1981, a paper by Robert Shiller – one in which he showed that stock market
prices fluctuate more than can be explained by rationally-varying forecasts of firms’ future cash flows – shook economists’ confidence in the prevailing rational framework for thinking about financial markets. Coming on the heels of Shiller’s (1981) work, De Bondt and Thaler’s paper further undermined the dominance of the rational paradigm. As such, these studies helped to pave the way for behavioral finance, which took off in earnest in the 1990s.

Two other anomalies that Thaler studied are also drawn from the field of finance. The first is about “closed-end funds” – funds that, at inception, raise capital from investors and then allocate this money to stocks or other assets. After inception, a closed-end fund’s shares are traded on an exchange; investors wanting to buy or sell fund shares do so there at the prevailing market price. Over the years, finance practitioners, as well as a few academics, had pointed out something odd about the prices of closed-end funds: these prices often diverge from the value of the funds’ assets. For example, if a fund holds assets with a market value of $100 and there are 50 traded shares in the fund, the price of a share is typically not $2, but might be as low as $1.8 or as high as $2.2.

Thaler heard about this anomaly and joined forces with Charles Lee and Andrei Shleifer to study it more deeply (Lee et al., 1991). Their paper accomplishes three things. First, it documents several puzzling facts about the prices of closed-end funds, most notably that, on average, the price of a fund is lower than the value of the assets it holds. Second, it tests some rational explanations of this “discount,” and finds that these are unable to fully explain the facts. Third, it proposes a behavioral framework for understanding closed-end fund prices, one in which these prices fluctuate due to the changing sentiment of individual investors. Individual investors constitute a larger fraction of the ownership base of closed-end funds than of the assets held by these funds. As a result, if these investors become excessively
exuberant (pessimistic) about the future prospects of the assets they own, they will push the prices of closed-end funds up (down) relative to the prices of the assets held by the funds. The authors test this hypothesis in the data, with supportive results.

This paper might have received less attention than it did were it not for the fact that Merton Miller, a staunch defender of the rational paradigm at the University of Chicago, took umbrage at Lee, Shleifer, and Thaler’s message. After trying but failing to stop the *Journal of Finance* from publishing their paper, Miller co-authored a paper of his own attacking their claims. This led to a series of sharp exchanges in the pages of the journal, which naturally drew a lot of attention to Lee, Shleifer, and Thaler’s results. While closed-end funds are not a major focus of study in financial economics, many researchers have been persuaded that the facts documented by Lee et al. (1991) are hard to understand in a rational model. As such, this study has further weakened the rational framework’s grip on academic thinking.

One other financial anomaly that Thaler studied relates to the U.S. technology-stock boom of the late 1990s. During this time, the prices of technology stocks rose dramatically, leading many observers to argue that these stocks were severely overvalued. While this was a plausible hypothesis, it was hard to prove. However, some specific events that occurred during this time allowed researchers to make a stronger case for overvaluation. Some of these events involved “equity carve-outs,” and Thaler studied them in a paper with Owen Lamont (Lamont and Thaler, 2003).

Consider a parent company P that has a wholly-owned subsidiary S. In an equity carve-out, P sells part of S – 10% of S, say – into the public markets through an Initial Public Offering (IPO); after the IPO, then, shareholders in P continue to own 90% of S. This structure allows for an interesting calculation. Suppose that the publicly-traded shares
corresponding to 10% of S have a market value of $1bn. The remaining 90% of S owned by P is therefore implicitly valued at $9bn. As a result, the market value of the shares of P must be at least $9bn: simply put, a parent company must be worth at least as much as its subsidiary holdings.\textsuperscript{2} Lamont and Thaler (2003) document a number of instances during the technology-stock boom in which this was not the case: parent companies were sometimes worth less than the implied market value of their subsidiary holdings, a sign of severe mispricing.\textsuperscript{3}

The impact of Lamont and Thaler’s (2003) paper goes beyond documenting episodes of mispricing. In the examples the authors study, the relatively overvalued subsidiary is typically a technology firm. Their findings therefore add credence to the view that technology stocks as a whole were overvalued in the late 1990s, and that their dramatic rise and fall in price constituted a major financial bubble.

Aside from documenting and studying anomalies, Thaler played a major role in bringing them to the attention of economists. In the mid-1980s, he and Hal Varian, an advisory editor at the newly-founded Journal of Economic Perspectives, came up with the idea of having the journal feature a regular “Anomalies” column in which Thaler and collaborators would discuss some evidence that challenged the rational paradigm. The journal’s editor, Joseph Stiglitz, readily accepted the idea. Over the next few years, Thaler, with many co-authors, wrote 20 columns for the journal. These attracted a much wider readership than most economics articles: in a survey that the American Economic Association conducted of its...
members, more than half of those who responded said that they read the column regularly.

III. Building a New Framework

One of the reasons for the growth of behavioral economics over the past few decades is the accumulation of empirical facts that are hard to square with the rational paradigm. As discussed in Section II, Thaler played a key role in documenting such facts and in bringing them to the attention of the economics profession. However, if behavioral economics was to move forward, it had to come up with an alternative paradigm – with a new generation of models that allowed for departures from full rationality, captured the puzzling facts in a parsimonious way, and made new predictions that could be brought to the data. The problem was that it was unclear how to begin: Of the many ways in which an individual could be irrational, which were the most promising to explore?

It is here that Thaler had his single most influential insight. In the 1970s, after discovering that, unbeknownst to economists, psychologists – most notably Daniel Kahneman and Amos Tversky – had been cataloguing the ways in which people depart from full rationality, he recognized that this research was the key to progress in behavioral economics. It is hard to overstate the impact of this insight. Over the next few decades, one of the principal approaches behavioral economists took to developing new models was precisely to incorporate ideas from the area of psychology known as “judgment and decision-making,” and particularly from the work of Kahneman and Tversky, into otherwise traditional economic frameworks. This process continues unabated today. For example, some recent work formalizes and applies Kahneman and Tversky’s notion of representativeness (Bordalo et al., 2017a), while other papers explore aspects of prospect theory (see Barberis, 2013).
Given the impact of Thaler’s insight on the development of behavioral economics, it is worth recounting the events of the time. In the summer of 1976, Thaler, then an assistant professor at the University of Rochester, accompanied his dissertation advisor Sherwin Rosen to a conference at Stanford. There, he met a psychologist, Baruch Fischhoff, who worked in a field called judgment and decision-making whose existence Thaler was not previously aware of. Hearing about Thaler’s interests, Fischhoff pointed him to the work of Kahneman and Tversky, whom Thaler had also not heard of. On his return to Rochester, Thaler got hold of Tversky and Kahneman’s (1974) article in *Science* magazine in which they summarize their work on judgment heuristics. In *Misbehaving*, Thaler describes this moment as a life-changing experience; as he read the article and recognized its significance for economics, his hands shook and heart raced. In a phone call a few days later, Fischhoff encouraged Thaler to seek out a recent working paper by Kahneman and Tversky; after getting hold of the paper – an early draft of their classic 1979 paper on prospect theory – Thaler was again stunned: here was a theory that appeared to offer a foundation for new models of economic decision-making.

While Thaler’s biggest influence on the development of behavioral economics models came through his vision that they should be grounded in research on judgment and decision-making, he did not stop at this broad insight. Rather, he himself began to think about the form that these new models might take. While the frameworks he described were typically not highly structured models, it is remarkable that, in many cases, the economic intuitions he laid out decades ago remain at the core of the more sophisticated models being developed today. It is also striking how broad Thaler’s reach is: almost every class of behavioral economics model has his fingerprints on it.
To see Thaler’s influence, it is helpful to put behavioral economics models into three categories: models that make psychology-based assumptions about individual preferences; models that make psychology-based assumptions about individual beliefs; and models that make psychology-based assumptions about the process by which an individual makes decisions. The work on preferences can itself be divided into four sub-categories: research on preferences over riskless outcomes; preferences over risky outcomes; time preferences; and social preferences. If we count the four preference-related categories separately, this makes for a total of six streams of research. Remarkably, Thaler made early contributions to all six. To demonstrate this, I take each one in turn.

Preferences: Riskless Choice

Kahneman and Tversky’s (1979) prospect theory was originally developed as a model of risky choice. However, upon reading it, Thaler realized that it could also be used to build new models of riskless choice. This insight has been influential: a prominent behavioral model of riskless choice in use today, one developed by Tversky and Kahneman (1991) and Koszegi and Rabin (2006) among others, is indeed based on prospect theory.

Thaler planted the seeds of this model in his discussion of the endowment effect in Thaler (1980). The two most important elements of prospect theory are “reference dependence” – the idea that people derive utility not only from levels of wealth or consumption, but also from gains and losses measured relative to a reference point often taken to be the status quo – and “loss aversion,” the idea that people are much more sensitive to potential losses than to potential gains. As soon as he read about loss aversion, Thaler realized that it could be the key to understanding the endowment effect. To see this, think of the endowment effect
as Kahneman et al.’s (1990) finding that the average amount participants are willing to pay to obtain a mug is much lower than the average amount they require in order to give up a mug. Thaler’s reasoning is that people view obtaining a mug as a “gain” and giving up a mug as a “loss”. If they are much more sensitive to losses than to gains, they will demand more money to give up the mug than they are willing to pay to get one.

This reasoning, in more general form, lies at the heart of cutting-edge behavioral models of riskless choice (Koszegi and Rabin, 2006). In these models, the individual derives utility not only from consumption levels, as in the traditional framework, but also from gains and losses in the various components of consumption. Moreover, he is more sensitive to losses than to gains.

Some 25 years after Thaler first wrote about the endowment effect, some authors began to question its robustness. Plott and Zeiler (2005) showed that changes in experimental conditions could reduce the magnitude of the effect, while List (2003, 2004) found that the effect was less pronounced for experienced traders in the field. These results are striking. However, it turns out that we can make sense of them within the loss-aversion framework by leveraging advances in our understanding of reference points. Koszegi and Rabin (2006) argue that reference points are often based on expectations: an outcome is a “gain” if it is better than the expected outcome, and a “loss” otherwise. This leads to a re-interpretation of the endowment effect: people demand a lot of money to give up a mug because, having been given the mug at the start of the experiment, they expect to leave with it at the end; parting with the mug is therefore coded as a loss. This, in turn, offers a way of understanding Plott and Zeiler (2005) and List’s (2003, 2004) results: the participants in these studies may have exhibited smaller endowment effects because, as a consequence of the specific experimental
conditions or market context, they did not expect to keep the objects in their possession. As such, giving the objects up was not seen as a loss.

Today, most economists who have studied the topic believe that the endowment effect is a real phenomenon. Moreover, loss aversion remains the most prominent explanation of the effect, although other explanations have also been proposed (see Ericson and Fuster, 2014). However, Thaler’s work on the endowment effect is significant not so much because the effect itself is important, but because his research on it pointed the way to a new, psychology-based model of riskless choice.

In later work, Thaler and his co-authors Linda Babcock, Colin Camerer, and George Loewenstein applied the concepts of reference dependence and loss aversion in the context of labor supply. The traditional model predicts that an increase in wages will lead workers to exert more effort. However, using data on the labor supply of New York City cab drivers, Camerer et al. (1997) find evidence of a negative wage elasticity: on days when a driver’s hourly wage is higher, the driver works fewer hours. To interpret this finding, the authors propose a framework in which cab drivers derive utility from daily earnings measured relative to a target level of earnings, and are loss averse, in other words, more sensitive to falling short of the target than to exceeding it. On days when the hourly wage is higher, a cab driver can reach his target in fewer hours; since, due to loss aversion, the marginal benefit of working an additional hour once he has reached the target is lower, he often stops working once the target is attained.

Camerer et al.’s (1997) interpretation has been controversial; Farber (2005, 2008) uses a different methodology to argue that there is little evidence of reference dependence. However, more recent work by Crawford and Meng (2011) suggests that Thaler and his co-authors were
on the right track after all. Following Koszegi and Rabin’s (2006) proposal that reference points are based on expectations, Crawford and Meng (2011) define a driver’s target level of earnings to be his expected earnings, and measure this as his average earnings over the previous days. In addition, they modify the driver’s utility function so that it incorporates not only utility from earnings relative to expected earnings, but also (dis)utility from hours worked relative to the expected number of hours of work. With this specification, Crawford and Meng (2011) find significant evidence of reference dependence: a driver is more likely to stop once he reaches his target earnings or target number of hours of work.\footnote{Research on this topic continues; see Thakral and To (2017).}

Preferences: Risky Choice

Thaler has also influenced behavioral models of risky choice – for example, models of how people make financial decisions. Here too, his idea is that prospect theory is the key to building a new generation of models. This idea has been impactful: a vibrant stream of research in behavioral finance explores the predictions of prospect theory for financial decisions and asset prices (see Barberis, 2013). Models based on prospect theory are also being used to think about household insurance choices (Barseghyan et al., 2013).

Benartzi and Thaler (1995), the first paper to use prospect theory to think about asset prices, captures Thaler’s vision for what behavioral models of risky choice might look like. The paper addresses a long-standing puzzle in financial economics, the “equity premium puzzle.” Over the past two centuries, the average return of the U.S. stock market has greatly exceeded the average return of Treasury Bills. Of course, since the stock market is riskier than T-Bills, it is not surprising that its average return is higher. However, traditional
models of asset prices have had trouble explaining the sheer size of the difference in the average returns of the two asset classes.

Benartzi and Thaler (1995) argue that we can make sense of the high historical average return of the stock market by incorporating elements of prospect theory – specifically, reference dependence and loss aversion – into our description of investors’ decision-making. Suppose that, in line with reference dependence, investors derive utility from annual changes in the value of their financial wealth – annual because, in this framework, investors are assumed to track their financial wealth on an annual basis, perhaps because they receive their most comprehensive brokerage statements once a year or because they file their taxes once a year. Suppose also that investors are loss averse – in other words, much more sensitive to a potential drop in financial wealth than to a potential increase. Benartzi and Thaler (1995) perform some calculations to show that, if investors take this source of utility into account when allocating their wealth, this can explain the stock market’s high historical average return. To these investors, the stock market is unappealing: if they allocate a large fraction of wealth to it, they face the possibility of a painful loss over the next year. As a result, they require a high average return to hold the available supply of stocks.

Benartzi and Thaler’s (1995) work has helped to inspire a sizable literature on applications of prospect theory in finance. Some papers further explore the concepts of reference dependence and loss aversion that were the focus of Benartzi and Thaler (1995) (Barberis et al., 2001; Pagel, 2016), while other papers investigate the “diminishing sensitivity” component of prospect theory (Barberis and Xiong, 2009; Li and Yang, 2013) or its “probability weighting” feature (Barberis and Huang, 2008). By and large, this research has confirmed Thaler’s early intuition that prospect theory would be helpful for thinking about empirical
facts in finance.

*Time Preferences*

The traditional model of intertemporal choice posits that people maximize the expected value of a time-additive utility function, where future utility flows are discounted in an exponential fashion. In mathematical terms, an individual solves

$$\max_{\{c_t\}} E \sum_{t=0}^{\infty} \delta^t u(c_t),$$

subject to a budget constraint, where $c_t$ is time $t$ consumption and $\delta$ is the discount factor.

Early in his career, Thaler emphasized that this model fails to capture some basic aspects of behavior, often referred to as problems of self-control (Thaler, 1980). First, people are often time-inconsistent, taking actions that differ from those they planned to take. For example, they eat fattening desserts at dinner even after telling themselves earlier in the day that they wouldn’t, or fail to go to the gym on a regular basis even after resolving to do so (DellaVigna and Malmendier, 2009). Second, and relatedly, people sometimes use commitment devices which restrict the options available to them. For example, they put money into saving accounts that penalize early withdrawals, even when these accounts offer interest rates no higher than those on more liquid investments (Ashraf et al., 2006). The model in (1) captures neither of these observations: it predicts time-consistent behavior, and does not explain why someone would willingly restrict his choice set.

Thaler did not simply emphasize the shortcomings of the traditional model. Together with Hersh Shefrin, he pioneered an alternative framework that captures self-control prob-
lems (Thaler and Shefrin, 1981). In this framework, self-control problems are due to people having “multiple selves,” each of which is trying to control behavior. The idea of such competing selves is an old one, going back to Adam Smith and even to Plato, but Thaler was particularly inspired by the social scientist Donald McIntosh, who argued that self-control problems arise because “the psyche contains more than one energy system” (McIntosh, 1969).

In Thaler and Shefrin’s (1981) model, there are multiple periods of time. In each period, there is a “doer” who maximizes the utility of consumption in that period alone. Meanwhile, there is also a “planner” who cares about the welfare of all the doers taken together. This sets up a conflict: the doer in the first period would like to consume as much as possible right then, but the planner would like him to consume much less so as to leave something for future doers to consume.

The framework that Thaler and Shefrin (1981) pioneered – a model of multiple selves with overlapping periods of control – remains one of the leading approaches to analyzing self-control problems; the others are models of multiple selves with non-overlapping periods of control – the so-called $\beta$-$\delta$ framework studied by Laibson (1997) and O’Donoghue and Rabin (1999) – and models of temptation utility (Gul and Pesendorfer, 2001). While the $\beta$-$\delta$ framework is the most widely used of the three, new models of the Thaler and Shefrin (1981) variety continue to be developed; see, for example, Fudenberg and Levine (2006).

Thaler made a second important contribution to research on time preferences. To understand how people discount payoffs, he conducted an experiment in which he asked participants questions of the form: “What amount $X$ would you need in one year in order to be indifferent between that and $250$ today?” The literature is now replete with papers that run such experiments, often known as Money-Early-or-Later (MEL) experiments, but
Thaler (1981) appears to be the first study in an economics journal to conduct one. In this paper, Thaler presents three findings that, to this day, are seen as some of the most robust facts about payoff discounting; moreover, his interpretations of these findings remain the leading interpretations today. Specifically, he shows that people’s implied discount rate \( r \) is lower, the longer the delay in the later payoff: his average participant is indifferent between $15 today and $30 in three months (\( r = 277\% \)), but also indifferent between $15 today and $100 in three years (\( r = 63\% \)), something that he takes as evidence that discounting has a “hyperbolic” form – one that, relative to the exponential form in (1), puts greater weight on near-term payoffs.\(^5\) Second, discount rates are lower for larger monetary amounts: the average participant is indifferent between $3000 today and $3500 in three months (\( r = 62\% \)), and also between $3000 today and $6000 in three years (\( r = 23\% \)). Thaler’s explanation is that the psychological effort a person exerts in order to be patient has a fixed component, one that arises in any instance of delayed gratification. Finally, discount rates are lower for losses than for gains: Thaler’s average participant is indifferent between -$15 today and -$16 in three months (\( r = 26\% \)), something he attributes to “anticipation utility”; simply put, people want to get the loss over with.

While there are now many papers on MEL experiments, the past few years have seen a growing debate as to how much we can learn from them. What we are really trying to understand is how people discount the future utility flows that come from consumption – but since monetary payoffs are not necessarily consumed at the moment they are received, it is not clear how much MEL experiments can tell us about this. Nonetheless, the discount rates

\(^5\)Here, \( r \) is the annualized discount rate: if an individual is indifferent between \( $X \) at time \( t_1 \) and \( $Y \) at time \( t_2 > t_1 \), where \( t_1 \) and \( t_2 \) are measured in years, \( r \) is given by \( \log(Y/X)/(t_2 - t_1) \). The exponential discounting framework, in conjunction with some additional assumptions, predicts a constant \( r \) whatever the values of \( t_1 \) and \( t_2 \) (Cohen et al., 2016).
extracted from MEL experiments do correlate with certain behaviors, suggesting that there is at least some information in them. The debate is in full force. As Cohen et al. (2016) put it, the role of MEL experiments in the temporal discounting literature is currently “up for grabs.”

Social Preferences

There is now a large body of work on “social” preferences which studies phenomena such as charitable giving using models in which people are not purely self-interested, but also care about the payoffs of others. Thaler made an early contribution to this literature, too, through his work on fairness. In this research, he and his collaborators Kahneman and Knetsch were trying to understand what consumers perceive as “fair” behavior on the part of firms, and whether firms take these perceptions into account.

By surveying consumers, Thaler and his co-authors obtained a number of findings (Kahneman et al., 1986a, 1986b). First, if a firm raises prices in response to a short-term increase in demand that is not accompanied by an increase in its costs – for example, if a store increases the price of snow shovels after a snowstorm – this is deemed unfair. If firms take this perception into account, they will limit price increases in such situations, which will result in temporary shortages. Kahneman et al. (1986a) use this prediction to interpret some historical episodes.

Second, consider a firm whose business is stable, but which is located in an area with unemployment where there are people who would be willing to work at the firm for lower wages than it currently pays. Kahneman et al. (1986a) find that, in this situation, people deem it unfair for the firm to lower wages for its existing employees. If firms take this
perception into account, this may explain the “downward wage rigidity” that is thought to contribute to unemployment during recessions. There are two interesting caveats to this finding. In an illustration of nominal illusion, it is reductions in nominal wages, not real wages, that are deemed unfair: if inflation is 12%, it is seen as fair for the firm to raise wages for its employees by only 5%. Also, if the firm itself is struggling, then people consider it fair for it to lower wages.

Third, Kahneman et al. (1986a) find that it is seen as fair for a firm to pass a cost increase on in full to consumers: a furniture maker whose costs rise by $40 can fairly increase its price by $40. However, it is also seen as fair for a firm to keep part of the surplus from a cost decrease: if the firm’s costs drop by $40, it is deemed fair for the firm to reduce its price by only $20.

One reason why firms may take perceptions of fairness into account is because they fear retaliation from consumers. Kahneman et al. (1986b) provide evidence of such retaliation using experimental games such as the “ultimatum game” and the “dictator game” which ask people to divide a sum of money between themselves and another person. Specifically, they find that people are willing to pay a personal cost in order to punish someone who they believe has behaved unfairly.

Thaler’s intuition that fairness considerations are important in economic transactions is regularly validated, both by new academic research but also by events in the business world. For example, fairness perceptions can help explain the backlash against “surge pricing” by the ride-sharing company Uber – the sometimes dramatic rise in the price of an Uber fare
during times of peak demand.

Beliefs

Thaler’s influence on behavioral economics models has been felt most keenly in the preference-based approaches discussed above. However, from the earliest days of his career, Thaler has also strongly advocated, and worked on, frameworks in which people are not fully rational in their beliefs. For example, in his papers on financial anomalies – on mean-reversion in stock returns, on closed-end funds, and on equity carve-outs – he singles out irrational beliefs as the likely culprit: stocks with prior losses have high average returns because they are stocks that investors have excessively pessimistic beliefs about; the prices of closed-end funds move around because of irrational fluctuations in individual investor beliefs; and a subsidiary can be overpriced relative to its parent company because people are making overly exuberant forecasts about the subsidiary’s line of business.

In his papers about beliefs, Thaler typically does not take a strong stand on the psychological source of the mistaken belief. By contrast, more recent papers in behavioral economics tend to explore the implications of specific biases: of overconfidence, say, or of representativeness or a belief in a law of small numbers (Rabin, 2002; Malmendier and Tate, 2015; Bordalo et al., 2017a). By focusing on a particular bias, we can derive more precise predictions. However, even Thaler’s more “agnostic” approach can lead to testable predictions. For example, in their work on closed-end funds, Lee et al. (1991) are able to predict, even without positing a particular psychological bias, that changes in the average discount on closed-end funds should covary with the returns on stocks with low market capitalizations, something that they test and confirm in the data.
More recently, Thaler has himself focused on more specific biases. For example, Massey and Thaler (2013) present evidence that overconfidence affects even professionals facing high financial stakes. The authors predict that teams in the National Football League will overvalue high draft picks because they are too confident in their ability to pick out the players who will contribute the most in the future. The empirical facts are consistent with this prediction.

The Decision-making Process

Behavioral economics models aim to make realistic assumptions not only about people’s beliefs and preferences, but also about the process by which people make decisions. It is here that Thaler had one of his most famous ideas, the concept of “mental accounting” (Thaler, 1985; Thaler, 1999).

Mental accounting is a term Thaler coined to describe the process or system that households use to track their finances – their income, wealth, and expenditures. There is a nice analogy to financial accounting: if financial accounting is the set of rules that firms use to monitor their financial position, so mental accounting is the set of rules that households use to track their financial situation. The problem is that, while the rules of financial accounting are known and codified, the rules of mental accounting are not; instead, we have to figure them out, a task that Thaler pioneered.

One idea – a hypothesis that Thaler formed in part based on interviews with lower-income families living near Cornell University, where he was on the faculty at the time – is that people have temporal and category budgets: total expenditures over some period of time must not exceed $X; expenditures on clothes in particular must not exceed $X_c;
expenditures on entertainment must not exceed $X_e$; and so on. One consequence of such a system is that money is not fungible: money in the clothes account cannot be used to pay for entertainment, and vice-versa. Mental accounting of this type has a clear cost: if a person needs new clothes but has exhausted his clothing budget, he will have to do without, even if there is unused money in the entertainment account – that money is off-limits for clothes. The reason why people would nonetheless use this system is because its benefits outweigh the costs: it forces people to keep track of their expenditures and prevents over-spending.

The idea that households use systems of mental accounting is widely seen as plausible, but has proven hard to study: we lack a fully-developed model of mental accounting, and the data that would be needed to test such a model are not easy to come by. Nonetheless, every so often, new evidence emerges that is very consistent with mental accounting. Hastings and Shapiro (2017) study the purchases of people who participate in the Supplemental Nutrition Assistance Program (SNAP), the successor to the Food Stamp Program. SNAP benefits can only be spent on eligible food items. The traditional model predicts that an individual who joins the SNAP program will only *slightly* increase his purchases of these food items: he will use the benefits to buy roughly the same amount of food as before, and will then use the freed-up cash to purchase *other* things that he has higher marginal utility for. Hastings and Shapiro (2017) show that, in fact, people who join the SNAP program substantially increase their purchases of the eligible food items. This is consistent with mental accounting: the arrival of SNAP benefits significantly enlarges the individual’s perceived “food budget” – a
budget that, psychologically, feels like it needs to be spent on food.  

IV. Guiding People to Better Decisions

Another reason for the growing interest in behavioral economics over the past few decades is that researchers have found ways of helping people to make better decisions. Here, too, Thaler has played a central role.

In the 1990s, Thaler was one of a number of economists to voice the concern that U.S. workers are not saving enough for retirement. Workers themselves share this view. In a survey of the employees of a large U.S. corporation, two-thirds of respondents felt that their saving rate was “too low” (Choi et al., 2002). Limited self-control appears to be a factor in this under-saving. One-third of those who felt they were saving too little expressed an intention to increase their saving rate in the near future, but very few followed through on this plan.

One well-known approach to increasing saving rates is “auto-enrollment.” For many years, it was the case that, when a new employee joined a firm with a defined-contribution retirement plan, she was not automatically enrolled in the plan; in order to enroll, she had to fill out a form. Under auto-enrollment, when a new employee joins a firm, she is automatically enrolled in the retirement plan. She can leave the plan if she chooses, but needs to fill out a form to do so. In one case, then, the “default” is that she is enrolled; in the other, the default is that she is not.

In a traditional rational framework, whether or not a firm uses auto-enrollment should not have much impact on the fraction of employees who participate in its retirement plan:

---

6See Heath and Soll (1996) and Hastings and Shapiro (2013) for related findings.
whatever the default, each employee will weigh up the costs and benefits of participating and will choose her preferred course of action. However, influential studies by Madrian and Shea (2001) and Choi et al. (2004) show that, in fact, the fraction of employees who, a few months after joining a firm, are enrolled in the firm’s retirement plan is significantly higher in firms that use auto-enrollment than in those that do not. These findings have persuaded many firms to take up auto-enrollment; by some estimates, more than 50% of retirement plans now use it.

Thaler was supportive of auto-enrollment – he advocated it in Thaler (1994) – but he was not instrumental in its widespread adoption; in part, he was concerned that its effectiveness might be limited by the relatively low default saving rate that it often uses. Together with Shlomo Benartzi, he continued to look for other ways of encouraging saving.

Benartzi and Thaler reasoned that, to increase saving rates, they had to find a way to overcome two psychological impediments to saving. The first is loss aversion: since saving more means a reduction, or “loss,” in take-home pay, it is unappealing. The second is present bias, a type of time discounting in which extra weight is given to current payoffs: since an increase in saving means a reduction in take-home pay today, it is all the more aversive. Benartzi and Thaler came up with a plan called Save More Tomorrow which neutralizes these biases. When an employee enrolls in the plan, she agrees to increase her saving rate in the future by a fixed amount every time she receives a pay increase. Notice two things. First, since, at the time of enrollment, the saving increase is in the future, present bias does not stop the employee from enrolling. Second, since the saving increases are tied to pay increases, the employee never experiences a drop in (nominal) take-home pay; loss aversion therefore has no bite.
Thaler first presented this idea at a conference for investment professionals. Initially, it received little attention. Then, one day, Benartzi and Thaler found out that a financial consultant had implemented their plan at a mid-size manufacturing firm. The results were striking: the employees who enrolled in the program – people who had typically not been saving much before and who had resisted previous entreaties to save more – increased their saving rate from 3.5% to 13.6% after their fourth pay raise (Thaler and Benartzi, 2004).

The findings in Thaler and Benartzi (2004), together with support for the approach in the 2006 Pension Protection Act, have led a version of Save More Tomorrow known as “auto-escalation” to become widely adopted in U.S. retirement plans. Auto-escalation is a simplified form of Save More Tomorrow: it involves increasing the employee’s saving rate in fixed steps up to a pre-specified maximum; however, the saving increases are not tied to increases in pay, but simply occur at regular intervals.

In plans like auto-enrollment and Save More Tomorrow, Thaler saw something broader than “just” a way of helping people to save more – he saw a new approach to tackling policy problems in general, one that he called “libertarian paternalism.” In short, this is an approach that nudges people toward a sensible action (the paternalistic part), but without taking any options away from them (the libertarian part). Sensing a potentially wide range of applications for this approach, Thaler joined with Cass Sunstein to write some papers on the topic, and eventually a best-selling book called *Nudge* (Thaler and Sunstein, 2008).

*Nudge* caught the attention of governments around the world, most notably the Conservative Party in the U.K. When the party came to power in 2010, it established a government department called the Behavioral Insights Team which, in the spirit of libertarian paternalism, tried to use ideas from behavioral science to find solutions to policy problems. For
example, it was able to increase timely payment of taxes using a social-comparison technique pioneered by the psychologist Robert Cialdini, who showed, in a number of studies, that an individual is more likely to comply with a rule when told that a majority of her peers comply with the rule. To reduce the late payment of taxes, then, the government sent people letters informing them, truthfully, that a majority of their neighbors paid their taxes on time.

Today, there is a large literature exploring the effectiveness of various kinds of “nudges.” Several have been found to work in useful ways. For example, the social-comparison technique has also been used to reduce energy consumption and to motivate participation in elections (Allcott, 2011; Gerber and Rogers, 2009), and something as simple as sending parents a text message about their children’s missed assignments can lead to improved performance at school (Bergman, 2017). Nonetheless, in Misbehaving, Thaler sounds a note of caution. For one, the “nudge” approach is hard to scale: there is no single technique that can address many policy problems; rather, each problem has its own unique features, and hence often requires a tailor-made solution. Even after years of effort, the most successful type of libertarian paternalism remains the use of default options, the best example of which is still auto-enrollment, an idea that is now decades old.

V. Mentoring the Next Generation

In the 1980s, Thaler was one of just a handful of people working on behavioral economics; Colin Camerer and George Loewenstein were almost the only other economists pursuing a similar agenda. If the field was to grow, it would have to attract a new generation of researchers – researchers who would need training and mentoring. Here, too, Thaler played an important role. Many of those who started working in behavioral economics in the 1990s
viewed him as a mentor, if not their primary mentor.

In the mid-1980s, Thaler met someone who would help him train the next generation. Eric Wanner was the President of the Russell Sage Foundation, and was interested in supporting research at the interface of psychology and economics. He and Thaler decided to work together. The foundation supported a number of initiatives, but the most influential was the Summer Institute in Behavioral Economics, an intensive two-week course for PhD students taught by leading researchers in the field. The first event was held in 1994 and was run by Thaler, Daniel Kahneman, and Colin Camerer. The institute has convened every two years since then, and is now led by David Laibson and Matthew Rabin. Its exact impact is hard to measure, but it seems likely that it has played a major role in training and encouraging the new generation of behavioral economists.\(^7\)

Thaler was also involved in other initiatives to support research in behavioral economics. In 1991, he and Robert Shiller founded the National Bureau of Economic Research Working Group in Behavioral Finance; the two of them ran it for the next 24 years. The Working Group continues to be an important locus for the presentation of new research in behavioral finance and is as well-attended as ever.

VI. Conclusion

The rise of behavioral economics is one of the biggest conceptual developments in the social sciences in the past 40 years. Several factors have contributed to the growth of the field: the discovery of anomalies which challenge the traditional paradigm; the development of new, psychology-based models of economic behavior; advances in helping people to make

\(^7\)The Institute inspired a summer school in behavioral finance, held every two years at Yale University, and may also have spurred the recent proliferation of summer schools in economics.
better decisions; and an influx of talented researchers into the field. As I have outlined in this essay, Thaler played a central role in all four factors.

Today, behavioral economics is thriving: new anomalies are being discovered; new models are being built; new nudges are being implemented; and new talent is joining the research effort. The greater availability of large-scale datasets is likely to have a positive impact on the field, as in many other areas of economics. And there is a clear trend toward more “structural” behavioral economics – taking fully-specified models to the data in a way that allows for the estimation of key model parameters (DellaVigna, 2018).

It is less obvious how behavioral economics will evolve conceptually in the years ahead. Thaler’s single biggest insight was his recognition that the work of Kahneman and Tversky offered a foundation for a new generation of models, and a lot of the research in behavioral economics over the past 30 years has embraced that vision. However, while there is still much left to understand about the role of judgment and decision-making biases in economics, many of Kahneman and Tversky’s key ideas have now been explored. Where do we go from here? One possibility is that economists will begin to connect more fully with other areas of psychology beyond the narrow field of judgment and decision-making – with work on memory or attention, for example – but also with research in decision neuroscience. Some papers are already taking this approach (Bordalo et al., 2017b). It remains to be seen whether this will lead to a fundamentally new set of economic models or instead to a deeper understanding of the models we already have.

In his talks and writings, Thaler has often noted his wish for “the end” of behavioral economics. His hope is that economics will reach a point where there is no need for separate courses or conferences in behavioral economics. Rather, the ideas of behavioral economics
will be fully integrated into existing courses on financial economics, labor economics, macroeconomics, and so on; moreover, all research economists will be familiar with these ideas and will apply them as appropriate in their work. While there is some way to go before this goal is reached, the underlying vision is starting to be realized. More and more often, researchers who do not identify particularly as behavioral economists are nonetheless incorporating behavioral ideas into their analyses. The end of behavioral economics is in sight, and Richard Thaler is surely heartened to know it.

References


