Behavioral Economics and the Atheoretical Style*

Ran Spiegler†

August 6, 2018

Abstract

Behavioral Economics is widely perceived to be part of the profession's shift away from a culture that places abstract theory at its center. I present a critical discussion of the atheoretical style with which "behavioral" themes are often disseminated: a purely anecdotal style in popular expositions, simplistic cost-benefit modeling in pieces that target a wide audience of academic economists, and the practice of capturing psychological forces by distorting familiar functional forms. I argue that the subject of "psychology and economics" is intrinsically foundational, and that a heavier dose of abstract theorizing is essential for it to realize its transformative potential.

*I am grateful to Kfir Eliaz, Ariel Rubinstein and three referees for helpful discussions and comments.

†Tel-Aviv University, University College London, and Centre for Macroeconomics. URL: http://ww.tau.ac.il/~rani. E-mail: rani@post.tau.ac.il.

1 Introduction

In his scientific autobiography "*Misbehaving*", Richard Thaler suggests a link between his "anomalies" project and Thomas Kuhn's theory of scientific revolutions. Looking back to the 1980s, when his list of anomalies started to appear in print, he remarks:

> "An important aspect of Thomas Kuhn's model of scientific revolutions...is that paradigms change only once experts believe there are a large number of anomalies that are not explained by the current paradigm...As someone who had until recently still been in the "promising" stage of his career, it would be viewed as brash, unseemly, and self-destructive to talk about my own work as something that could be part of a "revolution". My goal was much more modest: just get a few more papers published and begin to establish the case that adding some psychology to economics was an activity worth pursuing. But I had certainly read Kuhn's path-breaking book *The Structure of Scientific Revolutions*, and had secretly spent idle moments wondering whether anything like a paradigm shift could ever be possible in economics." (Thaler (2015), p. 169)

Thus, in the early days of Behavioral Economics, it made sense to think (or at least daydream) about it as a movement toward a revolutionary paradigm shift, a notion that implies an overhaul of fundamental economic theory.

Times have changed. In a recent piece about teaching Behavioral Economics to undergraduates, Laibson and List (2015) define the subject as follows:

> "Behavioral economics uses variants of traditional economic assumptions (often with a psychological motivation) to explain and predict behavior, and to provide policy prescriptions."

No Kuhnian paradigm shift here. Laibson and List's definition is methodologically conservative; it emphasizes the reliance of Behavioral Economics on the existing modeling frameworks of economic theory, and does not count the search for new ones as part of its mission. Thaler himself says in an American Economic Association Presidential Address (Thaler (2016)) that "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." He goes on to describe Behavioral Economics as "simply one part of the growing importance of empirical work in economics". Despite an occasional nod to theory, his vision is quite restrictive: "Behavioral theories will be more like engineering, a set of practical enhancements that lead to better predictions about behavior."

Thus, Thaler associates the growing influence of Behavioral Economics with the profession's move away from a theory-centric culture. This conception of Behavioral Economics as an empirical antidote, rather than a catalyst to abstract theorizing is quite familiar. However, it is not self-evident. Compare it with an earlier "transformation of the culture of economics" (to borrow a phrase from Rubinstein (2006)), brought about by Game Theory. Both Behavioral Economics and Game Theory were *liberating* forces: Game Theory removed the shackles that had tied economists to competitive markets, and Behavioral Economics freed them from prior fixations on narrow self-interest and error-free decision makers. The difference is that unlike Behavioral Economics, Game Theory not only liberated economists, but also *demanded* of them to learn a new language. Ideas like Nash equilibrium and its refinements, implementation or robustness to high-order beliefs are not "variants on traditional assumptions", but a web of new concepts, modeling tools and techniques. Behavioral Economics demands relatively little in this regard, as many of its modeling ideas are reinterpretations or formerly unutilized specifications of standard frameworks. This difference is not an intrinsic feature of the two subjects, but a historical development. Had experiments been more fundamental for the early days of Game Theory, we might have seen a more empirical, less mathematically oriented subject. Likewise, the project of "psychologizing" economic theory could be carried out with a greater role for abstract, foundational theory. If anything, this project strikes me as intrinsically more foundational than the study of strategic interactions.

At the end of his Presidential Address, Richard Thaler states: "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..." This paper is about the atheoretical style in which this process is taking place. I do not focus on the development of Behavioral Economics itself as much as on how it is incorporated into the broader discourse of economics. In this context, the atheoretical style of a given piece can take various forms, depending on the piece's genre and intended audience. Popular expositions of "behavioral" themes tend to be purely anecdotal and devoid of theoretical reasoning, even by Popular Science standards. Pieces that target a general audience of academic economists make use of the most basic modeling devices in our toolkit, even when the subject matter demands (and the audience can digest) a more sophisticated approach. Incorporating "behavioral" elements into economic models in regular journal articles tends to follow an "applied" style that takes specific functional forms - rather than the modeling frameworks they belong to - as the starting point for the analysis.

Of course, there are abstract approaches to "psychology and economics" out there - sometimes by card-carrying behavioral economists, and often by theorists outside this circle. Rubinstein (1998) formulates decision processes in a clear "pure theory" style. Recent work in the wake of Gul and Pesendorfer's (2001) model of self-control preferences extends the tradition of axiomatic decision theory to new domains of choice objects, in an attempt to incorporate new psychological elements. But are these developments *part* of Behavioral Economics? I don't think that Gul, Pesendorfer or Rubinstein are viewed by anyone (themselves included) as "behavioral economists". Indeed, they have written *critiques* of the style of Behavioral Economics (Rubinstein (2006), Gul and Pesendorfer (2008)). Their approaches seldom feature in Behavioral Economics conference programs or course syllabi. Thus, when we speak of the influence of Behavioral Economics on the "psychologizing" of mainstream economics, it seems sensible to disregard these alternative approaches.

However, the key question is not whether the growing influence of Behavioral Economics has an atheoretical flavor, but whether this has any costs. I will argue that it does. Given that Behavioral Economics deals with the very building blocks of economic behavior, it has an intrinsic "foundational" character. Playing it down leads to a flatter discourse that robs "psychology and economics" of the conceptual depth and richness that the subject deserves. And at times it can stand in the way of obtaining substantive economic insights.

Full critical examination of the coevolution of Behavioral Economics and the general atheoretical trend in economics is a fascinating topic for historians and sociologists of economic thought; it would require a full-length book and lies beyond the scope of a paper like this. The best I can do is illustrate my thesis with prominent recent examples of how "behavioral" themes are absorbed in the wider discourse of economics. Given my emphasis on the dissemination of Behavioral Economics (rather than its production), I will mostly consider eminent authors who are not recognized as "full time" behavioral economists.

The structure of the paper is as follows. In Section 2, I use George Akerlof and Robert Shiller's 2015 book "*Phishing for Phools*" to discuss the anecdotal, theory-free style that is common in popular expositions of Behavioral Economics. In Section 3, I use John Campbell's 2016 Ely Lecture to demonstrate the limitations of a simplistic cost-benefit style of modeling in expositions of "behavioral" themes that target a general audience of academic economists. In Section 4, I turn to the "functional-form" style that often characterizes Behavioral Economics papers, taking as my main point of reference a recent methodological essay by Rabin (2013) that promotes this style. Thus, as the paper progresses, the genres that I examine become more technical and the targeted audiences become more narrowly professional.

In the concluding section, I come back full circle and return to Thaler's opening quote. I argue that the atheoretical style effectively denies the subject's paradigm-shifting potential that Thaler secretly dreamed of in the 1980s. It makes the subject seem more harmless than it truly is. For Behavioral Economics to fully realize this potential, it has to put a higher premium on abstract theorizing in general, and on the creation of new modeling frameworks in particular.

Given that this paper will contain a lot of talk about targeting audiences, I'd better describe my own target audience. The readers I am keen to address are economists who are interested in pure or applied theory as well as in Behavioral Economics, either as practitioners or as curious observers. My impression is that many of them, especially young ones, have grown with the "psychological realism vs. theoretical abstraction" narrative and learned to take it for granted. I hope to convince them that this separation is neither necessary nor desirable. Interest in "psychology and economics" and a taste for theoretical abstraction can and should coexist, rather than being conceived of as antagonistic alternatives.

2 The Anecdotal Style

The most extreme manifestation of the atheoretical style is expositions of "behavioral" themes that shed theoretical reasoning altogether, in favor of a loose collection of anecdotes about the economic consequences of decision biases and non-standard motivations. Naturally, this style is most likely to be seen in pieces that address a broad audience.

In this section I examine a recent example of this genre: Robert Akerlof and Robert Shiller's "*Phishing for Phools*" (Akerlof and Shiller (2015)). Their book explores the implications of consumer fallibility for the way we ought to think about the "free market". Its main thesis is that consumers' departure from rationality (their "phoolishness", to use the authors' neologism) makes the proliferation of exploitative transactions ("phishing") an inevitable feature of the market system. Akerlof and Shiller make their case with a collection of anecdotes about market exploitation of fallible consumers; their exposition is almost entirely devoid of theoretical reasoning. As one might expect from these authors, the anecdotes are illuminating and woven into an absorbing story. Nevertheless, in this section I argue that the anecdotal style has its limitations, and that incorporating *some* theorizing would have been valuable.

In the context of a popular book, I perceive the term "theorizing" in very broad terms. In particular, I do *not* identify theorizing with formal modeling, and allow for verbal abstractions that do not have a formal model in the background. Even those are very rare in *Phishing for Phools.* At any rate, the specific theoretical ideas that I will invoke in this section (and are missing from the book) are all borrowed from the existing theoretical literature on markets with "behavioral" consumers. Following the norm in academic economics, these theoretical ideas were originally presented as formal models, with varying degrees of abstraction and sophistication. Incorporating these ideas into *Phishing for Phools* would have meant popularizing these models.

And here I must get a natural objection off the table, and that is the argument that a popular book has no room for theoretical arguments that are derived from formal models. I strongly disagree. The fact that many popular books on Behavioral Economics were written by psychologists and marketing researchers accounts for their "collection of biases" style. But it does not follow that the anecdotal style must carry over to discussion of the biases' *economic* implications. By analogy, no popular exposition of Game Theory is complete without *some* description of Nash equilibrium, backward induction or signaling arguments. Of course, the expositions are verbal and entertaining, but they go beyond mere anecdotes. In an age when authors like Brian Cox and Simon Singh are writing best-selling books that contain a sketch of the derivation of $E = mc^2$ or an explanation of RSA encryption, readers of popular economics can survive a bit of non-technical theorizing.

Linking isolated anecdotes

One of the earliest stories in *Phishing for Phools* involves the famous empirical finding of DellaVigna and Malmendier (2006) that health-club customers appear to overestimate their future consumption when choosing a price plan. Many of those who select monthly subscriptions (with automatic renewal) end up paying more than if they had opted for a by–the-visit plan - they "pay not to go to the gym", as DellaVigna and Malmendier put it in the title of their paper.

Remarkably, except for two sentences at the *end* of the book, Akerlof and

Shiller remain silent about a simple theoretical argument that DellaVigna and Malmendier themselves make in a companion paper (DellaVigna and Malmendier (2004)). In their model, two firms play a simultaneous-move game in which they simultaneously offer two-part tariffs to consumers with a taste for immediate gratification. In the health-club context, this means that ex-ante, consumers would like to commit to do plenty of physical exercise in the future, but as time goes by their preferences change and they become lazier. Whether or not consumers can predict this future change in their preferences, the two-part tariffs that emerge in Nash equilibrium consist of a large lump-sum payment and a per-unit price *below* marginal cost. By comparison, if consumers had dynamically consistent preferences, firms would adopt marginal-cost pricing in Nash equilibrium.

Why is the omission of this theoretical result remarkable? Because in a later chapter, Akerlof and Shiller present yet another example of market exploitation: the pricing of credit cards (see pp. 68-69). Here, common price plans are a mirror image of the health-club case; they involve no (or effectively negative) lump sum and a *high* marginal interest rate. DellaVigna and Malmendier's model offers a simple explanation. Credit cards enable the consumer to enjoy an immediate consumption benefit and defer its cost. In contrast, attending a health club is an investment that pays off in the future. According to the DellaVigna-Malmendier model, this inversion in the temporal distribution of costs and benefits explains the direction of the equilibrium departure from marginal-cost pricing.

The logic behind this result depends on whether the consumer predicts the future change in his preferences. When he does, he seeks a commitment device to counter his taste for immediate gratification. A high marginal interest rate acts is a partial commitment device that deters excessive use of the credit card, whereas a low per-visit price acts effectively as a partial commitment device that incentivizes health-club attendance. When the consumer underestimates his future taste for immediate gratification, the equilibrium two-part tariff is an effective *bet* on the consumer's future consumption. The firm and the consumer have different prior beliefs regarding the consumer's future preferences, and therefore they have a motive to engage in speculative trade, shifting net consumer utility from the state predicted by the firm to the state predicted by the consumer.

The DellaVigna-Malmendier model thus links two, otherwise distinct examples of exploitative pricing. The model not only links them, but also explains the difference in their departures from marginal cost pricing. Luckily for authors of a popular book, this involves an undergraduate-level argument that can easily be conveyed to a broad audience. At the same time, it is pregnant with follow-up questions that feed "higher-level" theorizing: What kind of price plans would firms offer if not confined to two-part tariffs - in particular, can we explain real-life examples of complex non-linear pricing? How would firms set prices if they did not know the consumer's ability to predict future changes in his preferences? What is the effect of market competition on consumer welfare?[1]

The point is that some of the market exploitation anecdotes presented by Akerlof and Shiller cry out for a connecting thread (one that I have not mentioned, for the sake of brevity, is the add-on pricing example of Gabaix and Laibson (2006)). Such a connection requires some theorizing, however elementary. In the absence of theorizing, all we have is a loose collection of anecdotes. By refusing to theorize, Akerlof and Shiller water down their message.

Qualifying the main message

Another important role of theoretical reasoning - especially in the formal-modeling tradition - is to qualify sweeping verbal statements. Because the main thesis of "*Phishing for Phools*" is presented without any trace of formal modeling, it leaves the impression that "phoolishness" always harms consumers. But what if it could actually *mitigate* market failures that originate from other sources?

Ironically, Akerlof's celebrated "market for lemons" model provides a good illustration of this idea, since market failure in the lemons model is a consequence of uninformed buyers' *sophisticated understanding* of adverse

[1] For a few papers that address these questions and others, see Eliaz and Spiegler (2006), Grubb (2009), and Kőszegi and Heidhues (2010). For more general treatments of this class of models, see Spiegler (2011), Kőszegi (2014) and Grubb (2015).

selection. As Akerlof and Shiller point out, "phoolish" buyers have a limited understanding of the incentives behind sellers' behavior, and as a result they may form a biased estimate of the quality of the products that are traded in the market (see, for example, their discussion of mortgage-backed securities in Chapter 2). A number of authors (Eyster and Rabin (2005), Jehiel and Koessler (2008), Esponda (2008)) have proposed ways to model "markets for lemons" with such buyers. These models paint a rich picture: "phoolishness" can mitigate or exacerbate the market failure due to adverse selection, depending on the nature of consumers' limited understanding and the gains from trade. Although I believe that the argument can be (at least partly) conveyed verbally to a lay audience, in the present context it would be worthwhile to do it formally.

The above-cited papers all build on a familiar reformulation of the lemons model, following Bazerman and Samuelson (1985), where a situation in which many sellers compete for a buyer is approximated by a bilateral-trade game in which the buyer has all the bargaining power. Formally, an uninformed buyer makes a take-it-or-leave-it offer p to a seller who privately learns the value v of the object he owns, where $v \sim U[0,1]$. The buyer's valuation is $v + b$, where the constant $b \in (0,1)$ represents the gain from trade. When the buyer has rational expectations, he knows that the seller will trade if and only if $p > v$. Therefore, the buyer chooses p to maximize

$$\Pr(v < p) \cdot [E(v \mid v < p) + b - p] = p \cdot [\frac{1}{2}p + b - p].$$

The solution is $p^* = b$. Thus, although trade is efficient for all v, in equilibrium it will take only place with probability b.

Eyster and Rabin (2005) used the notion of "cursedness" to model a possible departure from rational expectations. They assumed that in equilibrium, the buyer knows the marginal distributions over v and the seller's action, but does not perceive any correlation between them. Thus, the buyer has a *coarse* perception of the seller's behavior, since he fails to account for its responsiveness to v. As a result, the buyer chooses p to maximize

$$\Pr(v < p) \cdot [E(v) + b - p] = p \cdot [\frac{1}{2} + b - p].$$

Thus, the buyer's expectations completely disregard the adverse selection consideration; his forecast of the object's value conditional on trade is given by the ex-ante distribution. The solution is $p^{ER} = \frac{1}{2}b + \frac{1}{4}$. We can see that $p^{ER} < p^*$ if and only if $b > \frac{1}{2}$ - i.e., "cursedness" exacerbates the market failure due to adverse selection only if the gain from trade is large. The intuition behind this ambiguous effect is that "cursedness" has two contradictory effects. On one hand, the buyer's expected valuation is higher than in the benchmark case because he ignores adverse selection; this raises the buyer's bid relative to the benchmark. On the other hand, the buyer does not realize that a higher bid would enhance the expected quality of the traded object; this lowers the buyer's bid relative to the benchmark. When the gains from trade are small, the former consideration outweighs the latter.

This ambiguity also implies that comparative statics with respect to the buyer's degree of "phoolishness" are not monotone. Jehiel and Koessler (2008) examined an example in which the buyer has a partially coarse perception of the seller's behavior: he partitions the set of possible realizations of v into intervals (of potentially unequal size), and he believes that the seller's strategy is measurable with respect to this partition. Using the notion of "Analogy-Based Expectations Equilibrium" (Jehiel (2005)), Jehiel and Koessler show that the equilibrium probability of trade is not monotone with respect to the fineness of this partition. In other words, greater "phoolishness" does not imply a stronger market failure.

Esponda (2008) assumed that the buyer's expectation of v conditional on trade is based on naive extrapolation from the equilibrium distribution itself. In his model, the buyer learns the traded object's value from observations of past transactions - without realizing that this sample is adversely selective, such that if the price that characterized historical observations changed, so would the observed quality distribution. The equilibrium price p^E is defined as follows:

$$\begin{aligned} p^E &\in \arg\max_p \quad \Pr(v < p) \cdot [E(v \mid v < p^E) + b - p] \\ &= \arg\max_p \quad p \cdot [\frac{1}{2}p^E + b - p] \end{aligned}$$

such that $p^E = \frac{2}{3}b$. In this case the buyer's "phoolishness" unambiguously exacerbates the market failure due to adverse selection. The reason is that of the two forces identified in our discussion of "cursed" buyers, Esponda's model shares only the force that pushes the price down.

The three models described above present different ways in which the buyer's understanding deviates from the rational-expectations ideal, and they force us to ask: "When we say that buyers don't understand the seller's incentives, what is it exactly that they don't understand?" Alternatively, they suggest that the bilateral-game reformulation of the lemons market model, which is successful in the rational-buyer case, might miss a key aspect of competition among rational sellers for a "phoolish" buyer. These question marks are a valuable corrective to a sweeping message like "phoolishness leads to bad market outcomes".

What is phishing equilibrium?

Toward the end of their book, Akerlof and Shiller give an argument that may be viewed as an explanation for their atheoretical approach:

> "This general way of thinking, with its insistence of general equilibrium, has been the central nervous system for economic thinking for almost two and a half centuries. Yet Behavioral Economics...seems oddly divorced from it. Our two examples from Behavioral Economics, of DellaVigna-Malmendier and Gabaix-Laibson, illustrate. In the style required now for a journal article, their modeling and examples are very special...In accord with the standards of economics journal articles, these articles prove that phishing for phools exists. They do so by giving models and examples, where that phishing is undeniable; but the journal's

> demand for such undeniability comes at a cost. It means that the generality of phishing for phools cannot be conveyed." (Akerlof and Shiller (2015), pp. 169-170)

As this passage demonstrates, when Akerlof and Shiller abandon the anecdotal style, it is to advocate a "think big", general-equilibrium approach to the subject of markets with "phoolish" consumers - compared with the piecemeal approach of analyzing small models that characterizes most of academic economic theory. (As an aside, I would have thought that Akerlof's lemons model proved once and for all the power of small models to convey big ideas.) They introduce the notion of "phishing equilibrium" and define it essentially as follows: *Every opportunity to exploit consumers is realized.*

Yet the meaning of this equilibrium concept is vague. An important feature of general equilibrium as we know it is linear-price taking. But as we saw in our discussion of DellaVigna and Malmendier (2004), endogenously complex price schemes are a hallmark of markets with non-rational consumers. Therefore, linear-price taking seem inappropriate. Another feature of general equilibrium is the no-arbitrage principle. Akerlof and Shiller rightly observe that firms seek every opportunity to exploit "phools". However, the no-arbitrage condition means that such activities should occur *off equilibrium*; in equilibrium, the profits from these opportunities have been competed away. Yet, game-theoretic models of competition for boundedly rational consumers often have the property that tougher competition does not dissipate profits because it strengthens firms' incentive to obfuscate and target erring consumers (Spiegler (2006), Chioveanu and Zhou (2013), Gabaix et al. (2016)). A "general equilibrium" model based on the assumption that competitive forces drive the gain from the "marginal phish" down to zero would exclude many interesting and potentially relevant situations.

Thus, while the call for a "general equilibrium" approach to the subject of market exploitation of "phools" is genuinely intriguing, it warrants a serious "pure theory" approach. In the absence of any attempt at formal modeling, it is hard to understand what "phishing equilibrium" could possibly mean or imply.

Summary

I have shown that key aspects of the "phishing for phools" argument could benefit from a modicum of theorizing, even allowing for the broad-audience factor. A more theoretical style would insightfully link the anecdotes; it would qualify sweeping claims regarding the market implications of "phoolishness"; and it would impose more discipline on conceptualizations like "phishing equilibrium". Of course, economists of Akerlof and Shiller's stature hardly need a sermon about the virtues of economic theory; as the above-quoted passage indicates, they made a deliberate choice to adopt an anecdotal style. Their choice reflects a wider sentiment that this style is appropriate to the subject matter. Yet, as I have demonstrated, this has flattened the message of their book.

It may also have diminished the book's long-run impact. A broad audience is also a variegated one: readers of a book like *Phishing for Phools* include bright undergraduate students from various disciplines. We want such students to join our ranks and move the discipline forward. Akerlof and Shiller's celebrity and absorbing anecdotal style will surely attract their attention, but a bit of abstract theorizing could better spark their imagination, by exposing them to the subject's potential depth and richness.

3 The Cost-Benefit Style

Another aspect of the atheoretical style in Behavioral Economics is the tendency to use the most elementary modeling devices in the profession's toolkit. Rather than writing down an elaborate choice model that explicitly captures a psychological mechanism, economists work out "behavioral" themes by taking a completely standard model in which choice follows a straightforward cost-benefit calculus, and then reinterpreting or relabel some of the terms as biases or errors (e.g., Bar-Gill (2012), Mullainathan et al. (2012)). In this manner, the modeler seems to have it both ways: on one hand, he can address "behavioral" phenomena and study their implications, yet on the other hand, he can conduct business as usual in terms of the modeling procedure.

A recent example of this practice is John Campbell's Ely Lecture (Campbell (2016)), which was devoted to boundedly rational decision making in the context of financial products, with possible implications for market regulation. In the lecture's theoretical part, Campbell focuses on a particular regulatory intervention: imposing a tax on complex and potentially exploitative products. To evaluate this intervention, he constructs a simple model with two products: one "simple" and the other "complex". The simple product has a fixed value, normalized to 0, which is correctly perceived by all consumers. In contrast, the complex product is characterized by heterogeneity in consumers' valuation. First, the product's subjective valuation, denoted u, varies across consumers. Second, subjective valuations may be biased. Specifically, a proportion α of consumers are sophisticated and a fraction $1 - \alpha$ are naive. Sophisticates' subjective valuations are unbiased. In contrast, when a naive consumer values the complex product at u, its *true value* for this particular consumer is $u - 1$. Thus, the valuation error committed by naive consumers is fixed at 1.

Campbell examines the consequences of imposing a fixed tax $b < 1$ on the complex product under various scenarios for the redistribution of tax revenues. For simplicity, I consider the case in which the revenues are *not* rebated. Consumers with $u \geq b$ ($u \leq 0$) choose the complex (simple) product both before and after the intervention. The only consumers whose behavior is affected by the intervention are those with $u \in (0, b)$. Turning to welfare analysis, all consumers with $u > b$ are harmed by the tax, whereas all consumers with $u < 0$ are unaffected by it. In the case of consumers with $u \in (0, b)$, we need to distinguish between sophisticates and naifs. The former are made unambiguously worse off since they switch to the simple product and earn a net payoff of 0, as compared to $u > 0$ prior to the intervention. In contrast, naive consumers with $u \in (0, b)$ are made better off since their true utility prior to the intervention is $u - 1 < u - b < 0$, as compared to 0 afterward. If there are sufficiently many consumers in the latter group, the tax improves overall consumer welfare.

From a descriptive point of view, Campbell's model is a completely standard utility-maximization model. The "behavioral" element is restricted to the welfare analysis. And while complexity of financial products is a key theme in Campbell's lecture, nothing in his model identifies the complex product as such. Consequently, it cannot tell us a story about the origin of naive consumers' errors. Campbell acknowledges that his simple model neglects various features, such as interaction between "behavioral" effects and other market failures or firms' political lobbying. However, some features surely count as more intrinsic than others: a model of complex products that does not define product complexity explicitly is the analogue of a consumption-saving model that only has one time period.

Now, the Ely Lecture is a "public" lecture that addresses a broad audience of academic economists. Although these are far more technically qualified than the lay readers of a popular book like *Phishing for Phools*, it could be argued that an elaborate behavioral model that explicitly describes product complexity would be "too much" for this forum. Perhaps a simplistic cost-benefit analysis is the best we could hope for, given the occasion. In this section I attempt to counter this claim, by presenting a simple model in the spirit of Spiegler (2006), which mimics Campbell's cost-benefit model as closely as possible while being explicit about product complexity and how it generates consumer errors. Although the model lends itself to complications that might be interesting for specialized theorists, its basic version amounts to maximization of a simple quadratic function, something that Campbell's audience should be able to digest.

The model not only defines product complexity; it also tells a story of how it comes about. In reality, the financial products that Campbell discusses are offered in *market* settings. Therefore, the most natural way to account for the origins of product complexity is to assume that it is a result of "phishing" (to use Akerlof and Shiller's terminology) by profit-maximizing firms. For simplicity, assume that the complex product is offered by a *monopolistic* firm. Think of the product as a state-contingent service contract. The state of nature is uniformly distributed over $[0, 1]$. The service is offered with two possible quality levels, 0 or 1. When the firm offers quality $q \in \{0, 1\}$ in some state, it incurs a cost of cq and the consumer earns a payoff of $q - b$, where $c \in (0, 1)$ is the cost of offering a high level of quality, and $b \in [0, 1 - c)$. The firm's strategy has two components: a price T, and a function $f : [0,1] \rightarrow \{0,1\}$ which determines the quality of service in every state. The complexity of the firm's product thus has a concrete meaning in this model: the product is a state-contingent contract with a rich state space. Let $p = \int sf(s)ds$ be the frequency of the states in which it offers high quality.

A fraction $1 - \alpha$ of consumers are naive and find it difficult to evaluate the contract. Every naive consumer follows a simplifying heuristic: he draws a state s at random, learns the value of $f(s)$, and regards it as a prediction of the level of quality he will receive if he chooses the firm's product. There is no correlation between the state the consumer draws in the course of this evaluation procedure and the state that will actually be realized. The interpretation is that the consumer, unable to fully digest the contract with its many clauses, examines a random clause and treats it as being "representative". His error lies in the fact that he exaggerates the informativeness of a very small sample - a stylized version of the phenomenon that Tversky and Kahneman (1971) called "the law of small numbers".

The remaining fraction α of the consumer population are sophisticated, in the sense that their belief regarding the level of quality they will receive is correct given their information. To mimic Campbell's assumption that the distribution over subjective valuations is the same for both the naive and sophisticated consumers, I assume that the latter are *perfectly* informed of the state of nature, and therefore know the level of quality they will receive if they choose the complex product. Thus, they also have an *informational* advantage over the naive consumers. Note that by paying attention to the procedural origins of the naive consumers' error, we get a better understanding of what might lie behind Campbell's stark assumption. Finally, the terms of the simple product are exogenous; i.e. quality 0 is offered in *all* states free of charge, and therefore, both of consumer types value it at zero. The simplicity of the simple product stems from the lack of quality variation across states.

A consumer's gross valuation of the complex product takes two possible values, 0 or 1. It follows that the firm will necessarily choose the price $T = 1-b$, such that a consumer's net subjective valuation of the complex product is either 0 (in which case he breaks the tie in favor of the complex product) or -1 (in which case he chooses the simple product). As in Campbell's model, the sophisticated consumer is always right. Unlike Campbell's model, the naive consumer's valuation is unbiased on average since it is generated by an unbiased signal. However, because the consumer will only choose the complex product when he has a high assessment of its quality, his valuation of the complex product is biased upward *conditional on choosing it.* The size of the bias is $1 - p$, since the product's true expected quality is p whereas the conditional perceived quality is 1.

The firm's problem is reduced to choosing $p \in [0, 1]$ in order to maximize

$$\alpha p(1 - b - c) + (1 - \alpha)p(1 - b - pc)$$

The first (second) term of this simple objective function represents the firm's profit from a sophisticated (naive) consumer. Every consumer chooses the firm with probability p. The firm's net profit conditional on being chosen by a sophisticated consumer is $1 - b - c$ since he chooses the complex product knowing that it will provide a high level of quality. (Our assumption that $b < 1 - c$ implies that the firm does not incur a loss on sophisticated consumers.) The firm's net expected profit conditional on being chosen by a naive consumer is $1 - b - pc$ since the actual level of quality he will obtain is independent of the level of quality in the state he sampled.

As long as α is not too large, the solution p^* to the firm's maximization problem is interior:

$$p^* = \frac{1 - b - \alpha c}{2c(1 - \alpha)}$$

By the assumption that $b < 1 - c$, $p^* > \frac{1}{2}$. This property will be instrumental in the welfare analysis presented below. Note that p^* decreases with b, i.e., *the firm responds to the tax with a lower frequency of offering a high level of quality.* Intuitively, transactions with naive consumers have an exploitative "bait and switch" flavor: with probability $p(1 - p)$, the firm attracts a consumer who sampled a high level of quality and ends up providing him with a low level, thus saving the cost. As b rises, the firm's profit margin shrinks, and its incentive to adopt the cost-saving bait-and-switch tactic becomes stronger.

Now turn to a calculation of consumer welfare as a function of b. Sophisticated consumers earn a true payoff of zero both before and after the intervention. Therefore, consumer welfare is driven by the naifs. A fraction p^* of them choose the complex product and earn a true expected payoff of $p^* - (1-b) - b = p^* - 1$, whereas their subjective payoff is 0. Thus, the valuation error of naive consumers who choose the complex product is $1 - p^*$. Unlike Campbell's model, the magnitude of a naive consumer's error *increases* with b due to the firm's *endogenous* response to the tax. When b increases, fewer naive consumers end up being exploited, but those who are get exploited to a greater degree. The latter effect is a regulatory cost that is missing from Campbell's model. Total consumer welfare is $-(1-\alpha)p^*(1-p^*)$, and because p^* is greater than $\frac{1}{2}$ and decreasing in b, consumer welfare unambiguously *decreases* with b. That is, the intervention's adverse effect due to greater exploitation of naive consumers who demand the complex product outweighs the positive effect of reducing their numbers. This equilibrium effect thus turns out to be crucial, but it is missed by the cost-benefit model.

The economic lesson is that using taxes or subsidies to make a complex product objectively less attractive may impel firms to magnify its role as a vehicle for exploiting naive consumers. Although the example was "cooked" to mimic as many features of Campbell's model as possible, the aspect it highlights would appear in competitive variations of the market model (which would be technically more intricate), as well as under different conceptualizations of product complexity. In general, when we analyze the effect of regulating "complex" products, it helps to have *some* model of what this complexity consists of and how consumers deal with it, since this may provide a clue as to the endogenous market response to the regulatory intervention.[2]

But is this merely the umpteenth demonstration that "equilibrium effects matter"? And if so, couldn't we make the same point *within* the confines of the cost-benefit style? A practitioner of that style could complicate Camp-

[2]In Spiegler (2015), I apply this methodology to regulatory interventions known as "nudges" (default architecture, disclosure).

bell's basic model by assuming that the product is offered by a firm that engages in obfuscation. He might represent obfuscation by some real-variable x and assume that the magnitude of naifs' errors is some increasing function of x. To get an interior solution, he would probably need to assume that higher values of x are more costly for the firm. The conclusion from such an extended cost-benefit model is likely to be that increasing the tax b leads to lower investment in obfuscation (because it shrinks the profit margins from this activity), and therefore smaller errors by naifs - the exact opposite of the conclusion we obtained from our procedural model.

Thus, the example is about more than the importance of analyzing equilibrium reaction to regulatory interventions - it also shows that going beyond the cost-benefit style can matter for the analysis. But even if we could somehow reproduce the concrete economic lesson with some cost-benefit-style extension of Campbell's original model, this type of "endogenization" can give us little insight into the nature of the problem, because the added functions are black boxes that tell us nothing about what product complexity is and therefore give us no guide for what assumptions to make.

A model that purports to address a "behavioral" phenomenon (such as consumer errors in the presence of product complexity) should contain an explicit account of this phenomenon. This in turn requires a style of theorizing that is conceptually more sophisticated than cost-benefit calculus. At the same time, in its simplest form, this style can be adapted to the broad audience of an Ely Lecture. The promise of "psychology and economics" lies precisely in the ability to *enrich* economic analysis in such directions, rather than in giving us permission to use the same old models while relabeling some of their components. The fact that this style of theorizing can also affect the substantive economic lessons means that there is more at stake here than one theorist's aesthetic sensibilities.

4 The Functional-Form Style

In this section I turn from formats like popular books or public lectures, which are non-technical by design, to the more narrowly targeted and technical format of the regular journal article. In this context, the atheoretical style often finds expression in the use of specific *functional forms* as a vehicle for conveying "behavioral" ideas.

We can broadly distinguish between two styles of introducing a novel psychological element into an economic model. First, the modeler can take a standard functional form that represents preferences or beliefs, and modify it so that the new behavioral element is directly seen from the modification. Second, he can target the conceptual framework to which the functional form belongs, and modify some of its fundamental assumptions or introducing new primitives. I use the terms "functional-form style" and "conceptual-framework style" to describe these approaches; the latter style is more abstract and "theoretical" than the former. Needless to say, the two styles are not mutually exclusive, and they can coexist in a given paper. The extent to which a theorist presents a new behavioral idea via modified functional forms or modified conceptual frameworks is a marker of the paper's style, and as such it influences the paper's audience and its expectations from the paper. In this section I discuss a few limitations of the functional-form style and argue that by its very nature, the topic of "psychology and economics" requires a stylistic mix that puts more weight on the "conceptual framework" style.

The case of optimal expectations

The limitations of an unadulterated functional-form style were on my mind in Spiegler (2008), where I examined the model of "optimal expectations" due to Brunnermeier and Parker (2005) (BP henceforth). The BP model is based on the idea that decision makers deliberately distort their beliefs in order to enjoy "anticipatory utility" (in addition to standard material utility). The distortion is not arbitrary and is subjected to a "no cognitive dissonance" constraint, according to which the decision maker's action maximizes his expected material utility given his chosen belief.

BP define their model in the context of an intertemporal consumption problem, since they are interested in macro/finance applications. The decision maker's objective function is

$$E_{\pi}\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\beta^{t-1}\left(\sum_{\tau=1}^{t-1}\beta^{-\tau}u(c_{t-\tau})+u(c_t)+E_{\hat{\pi}}\left(\sum_{\tau=1}^{T-\tau}\beta^{\tau}u(c_{t+\tau})\mid s_1,...,s_t\right)\right)\right]\right\}$$

where c_t is consumption in period t; u is the material utility from periodic consumption; s_t is the realization of an exogenous state variable in period t; π is the objective distribution over $(s_1, ..., s_T)$; and $\hat{\pi}$ is the decision maker's chosen belief over $(s_1, ..., s_T)$. In one of BP's applications, an investor chooses between two financial assets, one safe and the other risky; BP characterize the investor's behavior in this class of binary choice problems.

BP present their model in the functional-form style. They put the functional form front and center and get very quickly to macro/finance applications, without pausing to study the model's more foundational properties. I found (and still do) the BP model very interesting, but I felt that the style with which BP chose to present their model left a gap. A model in which decision makers *choose* what to believe is a major departure from the basic principles of rationality. Therefore, despite the model's seemingly conventional formulation as a maximization problem, it deserves some deeper digging.

The relative complexity of the above functional form makes it hard (at least for me) to gauge the model's departure from rational choice. In Spiegler (2008), I tried to get a better understanding of the BP model using a much simplified *single-period* version, where the decision maker chooses an action $a \in A$ and a belief $\hat{\pi} \in \Delta(S)$ (where S is a finite set of states of nature) in order to maximize the objective function $\alpha E_{\pi}u(a) + (1-\alpha)E_{\hat{\pi}}u(a)$ subject to the constraint that $a \in \arg\max_{a'} E_{\hat{\pi}}u(a')$, where $\alpha \in (0,1)$ is constant and $\pi \in \Delta(S)$ is assumed to have full support. I posed the following question: Can the observed choice correspondence induced by this simplified BP model be rationalized? In particular, does it satisfy the Independent-of-Irrelevant-Alternatives (IIA) axiom?

The answer turns out to be *negative*. When we take a choice set like the one that BP examined, we can generate examples that exhibit the following pattern: the decision maker selects the risky option (and optimistically distorted beliefs) in the binary-choice case, but when a third, very negatively skewed prospect is added to the choice set, he will revert to the safe option (and realistic expectations), thus violating IIA. Intuitively, the decision maker must choose a very optimistic belief if he wants to enjoy an anticipatory utility from the moderately risky action. However, in the expanded choice set, the no-cognitive-dissonance constraint requires him to react to this belief by choosing the third action, which generates lower overall utility due to its extreme downside.

This finding has several implications. In terms of economic substance, it shows that the BP predictions regarding the shift in investors' choices due to optimal expectations are not robust, since they can be overturned if we expand the choice set. At the psychological and choice-theoretic level, the violation of the IIA axiom is not arbitrary, but appears to capture an interesting and possibly general insight: people with access to more negatively skewed options are less likely to delude themselves.

At the methodological level, this exercise demonstrates the role of the conceptual-framework style in the development of decision models that exhibit "behavioral" themes. By taking the simplest possible version of the BP model and thinking about its basic choice-theoretic aspects, we obtained an interesting finding that is crucial for the interpretation of BP's results. This little exercise in choice-theoretic abstraction cannot be "outsourced" to pure-theory specialists - just as we do not expect an applied theorist who solves an optimization model with first-order conditions to "outsource" to specialists the task of checking second-order conditions. This exercise is essential to a modeler's basic understanding of his own model. And yet, it is fair to guess that because of the "applied theory" style in which BP presented their model - not to mention the fact that the authors are not choice theorists, but (highly prominent) macro/finance experts - the paper's audience did not demand a basic choice-theoretic exercise as part of the package. That is, the style in which a "behavioral" idea is presented to an audience of professional academic economists influences their expectations as to how the idea should be developed. In this case, it led to an omission of what is in my opinion a necessary ingredient.

Parametric modification of functional forms
A major strand in the functional-form style involves the distortion of familiar formulas with additional parameters. According to this approach, a researcher who wishes to explore the theoretical implications of a behavioral element takes a standard economic model, and replaces a conventional formula that represents preferences or beliefs with its parameterized distorted version. The conventionally rational case is then reduced to setting the added parameter to a particular value (zero or one, depending on whether the parametric distortion is additive or multiplicative).

Rabin (2013) presents an eloquent guide to this "parametric modification" approach. The virtues of this approach that he emphasizes - enabling empirical tests of a null hypothesis that excludes the behavioral effect in question, and quantifying the departure from the null hypothesis when it is rejected - are empirical in nature. And yet it is clear that Rabin intends this approach to be valid not only for empirical studies, but also for applied-theory investigations.

In some cases, the limitations of the parametric approach are self-evident. Suppose that we want to model the phenomenon of unawareness. Total unawareness of an event is conventionally captured by a subjective probabilistic belief that assigns zero probability to the event. However, how does one model *partial* unawareness? This is a difficult problem that requires us to probe deeper into what it means to be partially aware of something. What should be clear is that representing partial unawareness by a probabilistic belief that mixes the complete-awareness and complete-unawareness subjective distributions is a non-starter. Whatever such a representation captures, it is *not* partial unawareness.

In other cases, the point is more subtle. The example that Rabin (2013) adduces as the biggest success of the parametric approach is the (β, δ) representation of intertemporal preferences that exhibit a present bias. However, as Rabin himself points out, the (β, δ) model cannot be unambiguously implemented with standard optimization, because of the dynamic inconsistency that it implies (by design). Luckily, the theorist can invoke the multi-selves approach to analyzing behavior under dynamically inconsistent preferences. This approach was in place since Strotz (1955) and Peleg and Yaari (1973), well before the more recent surge in the popularity of the (β, δ) model. The fact that researchers were able to place the (β, δ) parametrization firmly within the multi-selves framework facilitated coherent analysis of its implications. In particular, it made it clear that a key issue in the implementation of the (β, δ) model is the solution concept one employs to analyze the resolution of the conflict among selves. Thus, while the (β, δ) model is indeed an example of a very successful implementation of the parametric approach, its power relies on our ability to relate the parametric form to a more abstract modeling framework. Because the framework was already familiar to economists, we tend to ignore its crucial role as a platform for the parametric exercise.

O'Donoghue and Rabin (2001) enriched the (β, δ) model with an additional parameterization of the notion of "partial naivete": decision makers are not oblivious to the future change in taste but underestimate it. The O'Donoghue-Rabin agent believes that in the future he will have $(\hat{\beta}, \delta)$ preferences, where $\hat{\beta} \in [\beta, 1]$. The extreme cases of $\hat{\beta} = \beta$ and $\hat{\beta} = 1$ capture perfectly sophisticated and perfectly naive decision makers, respectively; and a higher value of $\hat{\beta}$ captures greater naivete. This parameterization has become conventional. In particular, it was employed by DellaVigna and Malmendier's (2004) in their model of two-part tariffs in the presence of consumers with a taste for immediate gratification, which I discussed in Section 2.

Viewed from a slightly more abstract perspective, the $(\beta, \hat{\beta}, \delta)$ model is a special case of non-common prior beliefs in the extensive-form game played between the agent's multiple selves. A given self has an incorrect prior belief regarding the future selves' value of the present-bias parameter β. In applications with long time horizons (such that there are more than two selves), adapting the solution concept that we conventionally use under the multi-selves approach to the case of partial naivete is conceptually non-trivial, and O'Donoghue and Rabin (2001) indeed grapple with this issue. Rabin (2013) regards it as a by-product of the parametric-modification approach. To me, these conceptual considerations are not a side issue but the heart of the matter.

I would like to examine another abstract feature of the $(\beta, \hat{\beta}, \delta)$ model: any $\hat{\beta} \neq \beta$ corresponds to a belief that assigns probability *one* to the wrong state of the world. Furthermore, the predicted utility function that is induced by this prior is a convex combination of the utility functions that are induced by the two extreme values of the present-bias parameter, β and 1. Spiegler (2011, Ch. 4) shows that these two features imply that optimal contracting with such an agent (following Eliaz and Spiegler's (2006) generalization of the DellaVigna-Malmendier model) has an extreme property: the optimal contract is the same for all $\hat{\beta} > \beta$. In other words, partially naive agents are treated as if they were perfectly naive. This result does not survive alternative parameterizations of partial naivete (e.g., in Eliaz and Spiegler (2006), partial naifs assign some probability to the true future preferences).

The lesson is that treating parametric forms as if they were paradigmatic can distort our understanding of the phenomena they are meant to capture. In this case, slight "zooming out" led us to regard the $(\beta, \hat{\beta}, \delta)$ model as an instance of combining two familiar modeling frameworks: the multi-selves model and games with non-common priors. This in turn gave us a better understanding of key assumptions that drive the implications of this model in a principal-agent setting. Such an interplay between the functional-form and conceptual-framework styles is essential for the development of "behavioral" ideas.

A "conjectural variations" parable

In the Introduction, I drew an analogy between Game Theory and Behavioral Economics, and implied that the theoretical style of the former could serve as an inspiration for the latter. I would like to close the present discussion of the functional-form style with a semi-fictional example, which relates to Behavioral Economics only by way of a parable that makes use of the Game Theory analogy.

Imagine that we live in a world in which Game Theory has not been invented; moreover, the only familiar models of market structure are standard monopoly and perfect competition. Now comes along Professor X and proposes a modeling approach to oligopolistic behavior. He considers a market for a homogenous product with n firms and constant marginal cost c. The inverse demand function is $P(Q)$, where Q is the aggregate supplied quantity. Each firm chooses its production quantity q in order to maximize the following expression:

$$q \cdot [\alpha \cdot P(nq) + (1 - \alpha) \cdot p^* - c]$$

where $\alpha \in [0, 1]$ is an exogenous parameter that is allowed to vary with n, and p^* is the market equilibrium price. In equilibrium, the firms' optimal quantity q^* satisfies

$$P(nq^*) = p^*.$$

This model employs the parametric approach to capture equilibrium behavior in an oligopoly. When $\alpha = 1$, the firm plays as if it is part of a cartel that maximizes industry profits and allocates production evenly among its members. When $\alpha = 0$, the firm acts as a price taker, and the model collapses to competitive equilibrium. An interior value of α captures the intermediate case of oligopoly. Moreover, we can capture the intuition that a market with more competitors is more competitive, by assuming that α is some decreasing function of n. When we assume linear demand $P(Q) = 1 - Q$ and $c \in (0, 1)$, the equilibrium price is

$$p^* = \frac{\alpha + c}{\alpha + 1}.$$

This result is intuitively appealing: a higher value of α (which corresponds to a greater departure from perfect competition) results in a higher equilibrium price. Moreover, Professor X can make assumptions about the speed with which α decreases with n in order to derive quantitative predictions of equilibrium mark-ups and industry profits.

I said earlier that this example is partly fictional. In fact, it is very close in spirit to the actual model of "*conjectural variations*", which was a popular approach to oligopoly before the advent of Game Theory (see Tirole (1988), p. 244). That model, too, had a free parameter, which captured the firm's belief regarding the reaction of its opponent to changes in its own behavior. With the benefit of hindsight, it is clear that the game-theoretic approach to oligopoly has given us a *language* for studying many aspects of oligopolistic behavior - tacit collusion, the value of commitment, entry deterrence, etc. - that go well beyond the scope of the parametric approach (whether it takes the form of conjectural variations or the present example). The latter could continue to offer useful "reduced form" models for applied work, but its status as a Theory of Principle has clearly been diminished by the rise of the game-theoretic approach. By way of analogy, I believe that a similar diagnosis applies to Behavioral Economics. Psychological phenomena often possess intrinsic depth that calls for an enrichment of our analytical vocabulary, in ways that lie beyond the reach of the parametric approach.

Summary

The attractions of the functional-form style, and the parametric modification approach in particular, are obvious. It offers the prospect of "plug and play" applications, including tools for comparative-statics exercises and for treatment of heterogeneity (e.g., discriminating between consumers with different parameter values). The problem begins when functional forms are treated paradigmatically - rather than as interesting special cases of more abstract modeling frameworks - thus creating a false impression that the behavioral phenomena in question have been addressed with scope and generality, and depressing our appetite for a deeper understanding.

5 Conclusion

I hope that this journey has not left the impression that I am some kind of a "pure-theory fanatic". In fact, I am as uncomfortable with superfluous formalism as the next person. The range of desirable styles of theorizing varies across subjects, and not every subject requires sophisticated theorizing of the abstract, foundational variety. However, "psychology and economics" is surely one that does, since it deals with the very building blocks of economic models. The fact that Behavioral Economics has sung its music with a low-volume theory register is one of the reasons for its popularity; and it is undoubtedly a sound approach in many contexts. Nonetheless, the approach has its limitations, as I hope to have demonstrated in this paper.

Many of the examples I looked at involve the implications of Behavioral Economics for market interactions and their regulation. This reflects the centrality and topicality of this particular question, as well as my own prior preoccupation with it. But it also highlights a more general point: the need for a "high-volume theory register" is especially acute in the analysis of *interactions* with or between "behavioral" agents. The more economists try to apply "behavioral insights" to interactive systems, the greater their attention to theoretical considerations is likely to be.

In another recent Ely Lecture devoted to Behavioral Economics, Raj Chetty advocates a "pragmatic" approach to Behavioral Economics (Chetty (2015)). The following paragraph from Chetty's paper summarizes his approach well:

> "The decision about whether to incorporate behavioral features into a model should be treated like other standard modeling decisions, such as assumptions about time-separable utility or price-taking behavior by firms. In some applications, a simpler model may yield sufficiently accurate predictions; in others, it may be useful to incorporate behavioral factors, just like it may be useful to allow for time non-separable utility functions. This pragmatic, application-specific approach to Behavioral Economics may ultimately be more productive than attempting to resolve whether the assumptions of neoclassical or behavioral models are correct at a general level." (Chetty (2015), p. 3)

In some sense, I share the sentiment expressed in this passage: I am often impatient with debates about the general validity of behavioral assumptions. But the thing I find striking is the analogy between incorporating "behavioral" elements in economic analysis and the rather minor decision of whether to assume time-separable utility functions. In his attempt to make Behavioral Economics more palatable to a general audience, Chetty has also made it seem *harmless*. We have thus come back full circle: the atheoretical style of disseminating Behavioral Economics amounts to an effective denial of its revolutionary potential.

Behavioral Economics is *not* harmless. When one reads the works of Tversky and Kahneman from the 1970s or Thaler's early papers on mental accounting, one encounters insights that undermine conventional economic modeling. They attack Bayesian probabilistic sophistication as an unrealistic description of how people reason about uncertainty. And they claim that preferences are so malleable, context-specific and prone to mental accounting as to render the notion of stable preferences meaningless. Reading these impressive works, the message that I perceive is that a powerful reimagining of economic theory at the foundational level is needed, one that is comparable to and perhaps exceeds that brought about by Game Theory. Although the pragmatic approach to Behavioral Economics has its place, complementing it with a more abstractly theoretical approach is necessary in order to fully realize its transformative potential. Otherwise, behavioral economists and their followers might be committing a sort of present bias: achieving larger impact in the short run, while sacrificing its long-run influence.

References

[1] Akerlof, G. A., & Shiller, R. J. (2015). Phishing for phools: The economics of manipulation and deception. Princeton University Press.

[2] Samuelson, W. F., & Bazerman, M. H. (1985). Negotiation under the winner's curse. Research in experimental economics, 3, 105-38.

[3] Bar-Gill, O. (2012). Seduction by contract: Law, economics, and psychology in consumer markets. Oxford University Press.

[4] Campbell, J. Y. (2016). Restoring Rational Choice: The Challenge of Consumer Financial Regulation. The American Economic Review, 106(5), 1-30.

[5] Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. The American Economic Review, 105(5), 1-33.

[6] DellaVigna, S., and U. Malmendier (2004). Contract design and self-control: Theory and evidence. The Quarterly Journal of Economics (2004): 353-402.

[7] DellaVigna, S., and U. Malmendier. (2006). Paying not to go to the gym. The American Economic Review, 96(3), 694-719.

[8] Eliaz, K., & Spiegler, R. (2006). Contracting with diversely naive agents. The Review of Economic Studies, 73(3), 689-714.

[9] Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. The American Economic Review, 98(4), 1269-1291.

[10] Eyster, E., & Rabin, M. (2005). Cursed equilibrium. Econometrica, 73(5), 1623-1672.

[11] Gabaix, X., & Laibson, D. (2005). Shrouded attributes, consumer myopia, and information suppression in competitive markets. Quarterly Journal of Economics, 121 (2), 505-540.

[12] Gabaix, X., Laibson, D., Li, D., Li, H., Resnick, S., & de Vries, C. G. (2016). The impact of competition on prices with numerous firms. Journal of Economic Theory, 165, 1-24.

[13] Grubb, M. D. (2009). Selling to overconfident consumers. The American Economic Review, 99(5), 1770-1807.

[14] Grubb, M. D. (2015). Overconfident consumers in the marketplace. The Journal of Economic Perspectives, 29(4), 9-35.

[15] Gul, F., & Pesendorfer, W. (2001). Temptation and self-control. Econometrica, 69(6), 1403-1435.

[16] Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. The foundations of Positive and normative Economics: A handbook, 1.

[17] Jehiel, P. (2005). Analogy-based expectation equilibrium. Journal of Economic theory, 123(2), 81-104.

[18] Jehiel, P., & Koessler, F. (2008). Revisiting games of incomplete information with analogy-based expectations. Games and Economic Behavior, 62(2), 533-557.

[19] Kőszegi, B. (2014). Behavioral contract theory. Journal of Economic Literature, 52(4), 1075-1118.

[20] Heidhues, P., & Koszegi, B. (2010). Exploiting naivete about self-control in the credit market. The American Economic Review, 100(5), 2279-2303.

[21] Laibson, D., & List, J. A. (2015). Principles of (behavioral) economics. American Economic Review, 105(5), 385-90.

[22] Mullainathan, S., Schwartzstein, J., & Congdon, W. J. (2012). A reduced-form approach to behavioral public finance. Annual Review of Economics, 4.

[23] O'Donoghue, T. and M. Rabin (2001). Choice and procrastination. Quarterly Journal of Economics, 116(1), 121-160.

[24] Rabin, M. (2013). An approach to incorporating psychology into economics. The American Economic Review, 103(3), 617-622.

[25] Rubinstein, A. (1998). Modeling bounded rationality. MIT press.

[26] Rubinstein, A. (2006). Discussion of "Behavioral Economics". Econometric Society Monographs, 42, 246.

[27] Spiegler, R. (2006). Competition over agents with boundedly rational expectations. Theoretical Economics, 1(2), 207-231.

[28] Spiegler, R. (2008). On two points of view regarding revealed preference and Behavioral Economics. The foundations of positive and normative economics: A handbook, 95.

[29] Spiegler, R. (2011). Bounded rationality and industrial organization. Oxford University Press.

[30] Spiegler, R. (2015). On the equilibrium effects of nudging. The Journal of Legal Studies, 44(2), 389-416.

[31] Thaler, R. H. (2015). Misbehaving: The making of Behavioral Economics. WW Norton & Company.

[32] Thaler, R. H. (2016). Behavioral economics: Past, present and future. Present and Future. Mimeo, University of Chicago.

[33] Tirole, J. (1988). The theory of industrial organization. MIT press.

[34] Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. Psychological bulletin, 76(2), 105.