"But Can't we Get the Same Thing with a Standard Model?" Rationalizing Bounded-Rationality Models

Spiegler, Ran

University College London, Tel Aviv University

15 March 2010

"But Can't we Get the Same Thing with a Standard Model?" Rationalizing Bounded-Rationality Models*

Ran Spiegler†

March 15, 2010

Abstract

This paper discusses a common criticism of economic models that depart from the standard rational-choice paradigm - namely, that the phenomena addressed by such models can be "rationalized" by some standard model. I criticize this criterion for evaluating bounded-rationality models. Using a market model with boundedly rational consumers due to Spiegler (2006a) as a test case, I show that even when it initially appears that a bounded-rationality model can be rationalized by a standard model, the rationalizing models tend to come with unwarranted "extra baggage". I conclude that we should impose a greater burden of proof on rationalizations that are offered in refutation of such models.

*I thank Ayala Arad, Yves Guéron, Barton Lipman, Ariel Rubinstein and participants at the Summer School in Economics and Philosophy, San Sebastian 2009, for their comments. Financial support from the European Research Council, Grant no. 230251, as well as the ESRC (UK), is gratefully acknowledged.

†Tel Aviv University and University College London. E-mail: r.spiegler@ucl.ac.uk. URL: http://www.homepages.ucl.ac.uk/~uctprsp

- I am not the Messiah, would you please listen, I am not the Messiah, do you understand? Honestly!

- Only the true Messiah denies his divinity.

- What? Well, what sort of chance does that give me? All right! I am the Messiah!

- He is! He is the Messiah!

(Monty Python's *Life of Brian*)

1 Introduction

One side effect of the growing popularity and influence of behavioral economics has been the profusion of discussions of the methodology and rhetoric of economic theory (see Caplin and Schotter (2008) for a representative collection of essays, as well as Rabin (2002), Rubinstein (2006), Bernheim (2009), Dekel and Lipman (2009), Binmore and Shaked (2010)). This paper aims to add a dimension to this debate.

To motivate the discussion, think of the following familiar situation. You are sitting in the audience of an economic theory seminar. The speaker is presenting a model of a certain economic phenomenon, in which some of the economic agents are boundedly rational in some way. As the speaker is going through her model, you are beginning to sense that although this may be an interesting exercise, the model could be entirely recast in terms of a standard model with rational agents, possibly with an added conventional source of friction such as imperfect information or search costs. You are preparing to raise your hand...

This paper is about what happens next - or, to be more precise, about what *should* happen next. How should we conduct our debates about explanations of economic behavior that are based on non-standard behavioral assumptions? In particular, how should we compare these explanations with more conventional explanations based on rational choice? Should we devalue a bounded-rationality model (BRM henceforth) when it the economic phenomenon it addresses seem to be explicable by a rational-choice model?

As the above not-so-imaginary scenario suggests, these questions are not motivated by abstract philosophizing, but on my own direct experience on both sides of the fence, both as a member of seminar audiences and as a theorist who has been preoccupied with economic models in which at least some agents are boundedly rational. From

this experience (as well as others' - see Rabin (2002)), a major share of the comments contributed by referees and seminar audiences in response to a BRM can be read as attempts to "rationalize" the model. As an inventor of bounded-rationality models, I often feel an internal need to compare such models with more conventional ones based on rational choice. And when I do not feel this need myself, I can always count on seminar audiences, referees and coffee-machine conversation partners to fill the gap.

The audience's basic criticism can be summarized as follows:

> Although BRMs may shed some light on economic phenomena, in many cases one could think of a rational-choice model that could account for these phenomena. And if we can "get the same thing" with a standard model, why should we depart from the rational-choice paradigm? Moreover, since rational-choice models and BRMs tend to have dramatically different welfare implications, a switch from rational-choice to BRMs is not only problematic methodologically, but also carries a significant cost in terms of its implied policy prescriptions.

This paper is an attempt to come to terms with this "*can't we get the same thing with a standard model?*" critique. The methodological problem at hand is fundamentally *theory selection*: how should we choose from a number of competing models that provide different explanations for a given phenomenon? The need to choose is only magnified by the models' diverging welfare implications. The normatively scientific way of making this choice is to tease out cases in which the models generate different predictions, and subject these predictions to an empirical test. However, economics being the dismal science that it is, such empirical tests are difficult and rare. Indeed, the whole point of the "can't we get the same thing with a standard model" critique is that since the two types of explanations are empirically hard to distinguish, the conventional rational-choice explanation should be given priority.

Therefore, for the purpose of our discussion here, I will set aside the question of empirical tests and take it for granted that the account a BRM in question provides for certain economic phenomena is sound: the "story" it tells "rings true"; its behavioral assumptions seem to fit generally known psychological principles and the market situation in question; and its predictions are broadly consistent with known (stylized) facts. What the "can't we get the same thing with a standard model" critique maintains is that for the purpose of making these predictions, one does not have to abandon conventional behavioral assumptions, and therefore one ought not to; even if there is

some truth in the bounded-rationality story, the same truth could be captured equally well by a rational-choice model. I refer to such a rational-choice model that is offered in refutation of a BRM as a "*rationalization*", or as a "*rationalizing model*".

My discussion so far may have given the impression that every BRM faces a single, well-defined rationalization. This is obviously not the case. The rational-choice paradigm is famously flexible, and there is a variety of conventional models that can be offered in refutation of any given bounded-rationality model. Rationalizing models tend to come in one of the following three forms:

Rationalization via modified information. The rationalizing model modifies the bounded-rationality model by replacing the boundedly rational agents with conventionally rational agents who happen to have different information.

Rationalization via modified preferences. The rationalizing model modifies the bounded-rationality model by replacing the boundedly rational agents with conventionally rational agents who happen to have different preferences.

Rationalization via endogenization. The rationalizing model refuses to take the behavioral rule assumed by the bounded-rationality model as truly exogenous,and instead *derives* it as a rational equilibrium response in a larger model that introduces frictions which are not explicitly included in the original model.

I could illustrate these forms with any number of BRMs from the literature. For expositional effectiveness, however, I adopt a case study approach and restrict attention to a single model, due to Spiegler (2006a), of price competition in markets for credence goods when profit-maximizing firms face consumers who use naive anecdotal reasoning to evaluate stochastic variables. This model was proposed to highlight aspects of industries such as alternative medicine, consulting and mutual funds. There are several reasons for this expositional strategy, apart from my obvious familiarity with the model in question. First, a case study approach is useful because fine details turn out to matter. Second, the model in question is extremely simple to begin with. This facilitates the formulation of explicit rationalizing models and enables an intelligible comparison with the original model. Were the BRM itself more complex to begin with, detailed comparison with its rationalizations would quickly become intractable.

The rationalizing models that are offered in refutation of Spiegler (2006a) are themselves based on ideas made by flesh-and-blood seminar audiences, referees and colleagues in corridor conversations. I hope that I have done full justice to these suggestions. The sparseness of the original model makes it a particularly easy target for

the "can't we get the same thing with a Standard Model?" critique, because it is easy to think of conventional frictions that are excluded from the original model. Nevertheless, I show that even in this ideal case, the "can't we get the same thing with a standard model" criticism is plagued with several difficulties, which can be summarized as follows:

- The rationalizing model changes not only assumptions regarding individual behavior, but also assumptions regarding the *external environment* that individuals face.
- The rationalizing model introduces new parameters into the model. Replication of the original BRM's predictions hinges on proper selection of parameter values.
- The rationalizing model may give rise to multiple equilibria, whereas the original BRM has a unique equilibrium. Replication of the BRM's predictions hinges on proper selection of equilibrium.
- Changed assumptions about individual behavior mandated by the rationalizing model may be implausible in the context of the applications of the original BRM.
- Natural extensions of the original BRM can be meaningless under the rationalizing model.

These are essentially problems of how to assign *burden of proof* in a debate: which desiderata should the rationalizing model satisfy in order to count as a successful refutation of the BRM? For instance, when the rationalization introduces new unobservable parameters and replicates the original BRM's predictions under a suitable selection of parameter values, does this diminish its power as a "devastating criticism" of the BRM? In the concluding section, I offer my own opinion about how we should regard "rational explanations" that are offered as a criticism of BRMs. But my main objective in this paper is simply to expose the burden-of-proof problems themselves, because I believe that heightened awareness of these problems could improve the quality of our debates over BRMs.

I should to make three caveats:

- We need to distinguish the program of rationalizing *models* of economic behavior that depart from the rational-choice paradigm from the time-honored tradition of rationalizing *observed* economic behavior that superficially contradicts rational choice. This paper is entirely about the former.

- My discomfort with the "can't we get the same thing with a Standard Model?" critique of BRMs does not imply a rejection of other criticisms of BRM: their scope tends to be limited in comparison with the impressive generality of the basic rational-choice models; they are perceived as arbitrary and post-hoc relative to their rational-choice counterparts; and they violate the cherished revealed preference principle. These concerns are often justified but have been debated elsewhere, whereas the "can't we get the same thing with a Standard Model?" critique has not been subjected to careful methodological scrutiny.

- I refrain from considering rationalizations that restore the rationality of an agent's behavior by describing it as an optimal response to an *incorrect* subjective model. The reason is simple: I do not view such rationalizations as being standard at all. Conventional economic models invariably assume a common understanding of the model, such that any asymmetry in the agents' beliefs is fully embedded in the model itself. The only subjectivity that standard models admit is non-common priors. (However, violating the common-priors assumption would probably be deemed "non-standard" twenty years ago.)

The remainder of the paper is structured as follows. Section 2 presents the simplest version of the model due to Spiegler (2006a). Sections 3-5 subject this model to the three forms of the "can't we get the same thing with a Standard Model?" critique. I discuss my lessons from these rationalization exercises in Section 6.

2 A Market Model with Boundedly Rational Consumers

Imagine a market that consists of a continuum of identical consumers and n identical firms. Consumers enter the market with some problem. The value of fixing it is 1 for all consumers. Each firm i sells at zero cost a product that fixes the problem with independent probability $\alpha \in (0,1)$. Consumers also have an outside option ("doing nothing"), labeled $i = 0$, which fixes their problem with the same probability α. Firms are standard profit maximizers. They compete by choosing prices simultaneously. Let $p_i \in [0,1]$ denote the price chosen by firm i. Assume that $p_0 = 0$ - that is, "doing nothing" costs nothing.

I will refer to the firms as "*quacks*", as they display no skills relative to "doing nothing". There are several real-life situations that seem to fit this specification. Ac-

tively managed mutual funds are a case in point. According to the *Efficient Market Hypothesis*, prices in financial markets fully reveal private information. Consequently, an actively managed mutual fund cannot generate (risk-adjusted) returns in excess of the market portfolio. Thus, under the Efficient Market Hypothesis, the market for actively managed mutual funds is a "market for quacks". And of course, as the term "quacks" indicates, practitioners of non-scientific medicine often fall into this category.

If consumers chose rationally with respect to a correct understanding of the above market model, the market for quacks would be inactive, as all consumers would choose the outside option.[1] I refer to this outcome as the *rational-consumer benchmark*. Instead, let us assume that consumers choose according to the following procedure. Each consumer independently draws one sample point from each alternative (including the outside option). For every $i = 0, 1, ..., n$, let x_i denote the outcome of the consumer's sampling of alternative i: $x_i = 1$ (the problem is fixed) with probability α and $x_i = 0$ (the problem is not fixed) with probability $1-\alpha$. Given a sample, the consumer chooses an alternative i that maximizes $x_i - p_i$. (Let us ignore the case of ties.) The outcome of the consumer's choice i is a new, independent draw; therefore, his expected payoff is $\alpha - p_i$.

The consumers' choice procedure induces a complete-information, simultaneous-move game played by the firms. To illustrate the firms' payoff function, suppose that $p_n > p_{n-1} > \cdots > p_1 > p_0 = 0$. Then, firm k's expected payoff is

$$p \cdot \alpha \cdot (1-\alpha)^k$$

The reason is that the firm's clientele consists of all consumers who obtained a good sample point about the firm's product and a bad sample point about all of cheaper alternative. Thus, if all firms play a mixed strategy given by a continuous *cdf* G, then the expected payoff that a price p in the support of G generates for an individual firm is

$$p \cdot \alpha \cdot (1-\alpha) \cdot [\alpha(1-G(p)) + (1-\alpha)]^{n-1}$$

The max-min payoff in this game is $\alpha(1-\alpha)^n$.

This game has a unique Nash equilibrium. The equilibrium is symmetric: each firm plays a mixed strategy given by the following *cdf*:

$$G(p) = \frac{1}{\alpha} - \frac{1-\alpha}{\alpha} \cdot p^{-1/(n-1)}$$

[1] There are also Nash equilibria in which a non-empty subset of firms set $p = 0$ and consumers choose these firms.

defined over the support $[(1-\alpha)^{n-1}, 1]$. Each firm earns its max-min payoff in equilibrium, hence equilibrium industry profits are

$$n\alpha(1-\alpha)^n$$

The equilibrium has several noteworthy features:

Consumer behavior. The market for quacks is active and consumers pay positive prices for what is ultimately a useless product. Moreover, there is a positive clientele for each firm, including the most expensive one. Given a realization of the strategy profile, the size of the clientele of the firm that charges the k^{th}-lowest price is $\alpha \cdot (1-\alpha)^k$.
Comparative statics: prices. Expected price goes up as α goes down, and converges to the monopoly level $p = 1$ as α tends to zero. The reason is that as α gets closer to zero, the probability of multiple successes in the consumer's sample goes down, and therefore each firm is effectively unlikely to face competition.
Comparative statics: industry profits. Industry profits are hump-shaped with respect to the number of competitors n. The intuition for this effect is straightforward. On one hand, a greater number of firms increases the incentive to cut prices. This is a standard "competitive" effect. On the other hand, a greater number of market alternatives increases demand for the industry as a whole, because there is a higher chance of hearing a good anecdote about some product. This is an "exploitative" effect. Fixing α, the exploitative effect outweighs the competitive effect when n is relatively small (the critical value of n for which this is overturned increases as α decreases). Note that industry profits are a pure transfer from consumers to firms, given our assumption that the probability that the consumer's problem gets fixed is independent of his decision. Thus, all the statements regarding industry profits are at the same time statements about consumer welfare.

Having presented the basic BRM, let us turn to its rationalizations.

3 Rationalization via Modified Information

Replacing imperfect rationality with imperfect information is perhaps the most immediate and common of traditional responses to models of bounded rationality. The idea is to replace what seems like a decision error resulting from bounded rationality with a rational response to limited information. For instance, choosing a low-quality product over an identically priced, high-quality product can be interpreted as evidence

of imperfect information regarding product characteristics.

In the case of the market-for-quacks model, this rationalization is very naturally suggested by the sampling-based procedure itself. Instead of viewing the samples as part of the choice procedure in a complete-information model, we can re-interpret the samples as information sets in a model in which consumers are imperfectly informed. The rationalization turns the model into an incomplete-information extensive-form game: firms move first (making simultaneous pricing decisions) and consumers move second, after receiving partially informative signals of the firms' success rates. This model's predictions are given by applying the solution concept of sequential equilibrium to the incomplete-information game. In the BRM, consumers confront their market environment with a decision procedure that generates systematic inference errors. In contrast, the imperfect-information rationalizing model rules out systematic inference errors because the solution concept of sequential equilibrium embodies "rational expectations".

This rationalization sounds highly plausible. As we shall see, it gives rise to symmetric Nash equilibria in which firms play mixed pricing strategies and a positive fraction of consumers choose firms over the outside option. Nevertheless, the rationalization suffers from a number of difficulties.

Changed assumptions about the external environment

In order for consumers' imperfect information to have any relevance, we must assume that firms have multiple payoff-relevant *types* - e.g., a high-quality firm and a low-quality firm. Thus, in order to apply the rationalization, we also need to modify our assumptions about the *external* market environment.

This is not an innocuous modification. For example, consider the market for homeopaths. A homeopathic medicine is based on a solution so diluted that it is, to an excellent approximation, pure water. To claim that there are high-quality and low-quality homeopaths is to claim that different types of water have different therapeutic properties.[2]

In another context, if we consider the money-management application of the model, the assumption that there are high-quality and low-quality money managers means that some managers can systematically beat the market. This is an important substantive assumption, which is not taken lightly by financial economists. One should continue not to take it lightly when using it to rationalize models of money management markets

[2] Alternatively, a higher-quality homeopath could be regarded as one who is better at generating a Placebo effect. However, admitting Placebo effects reintroduces irrationality through the back door, and therefore I ignore this possibility.

with boundedly rational investors.

New unobservable parameters

The incomplete-information game designed to rationalize the market-for-quacks model requires us to introduce new parameters that describe the distribution of firm types and the consumers' signal structure. The following specification is minimalistic in this regard. Each market alternative (the firms as well as the outside option) has a type $t \in \{L, H\}$. The prior probability of L is λ, independently across market alternatives. When a consumer chooses an alternative of type t, his need is satisfied with probability α_t, where $1 \geq \alpha_H > \alpha_L > 0$. Thus, each alternative's ex-ante success rate is $\alpha = \lambda\alpha_L + (1-\lambda)\alpha_H$. Each consumer observes a signal $s_i \in \{L, H\}$ about each alternative. The signals are distributed independently across market alternatives and across consumers. Let q_{ts} denote the probability that the consumer observes the signal s conditional on the alternative's type being t. Assume that $q_{LL} > q_{LH}$ and $q_{HH} > q_{HL}$ - i.e., signals have some informational content.

Notice how many new parameters have been introduced, even in such a minimalistic two-type, two-signal rationalization: $\alpha_L, \alpha_H, q_{LL}, q_{LH}, q_{HL}, q_{HH}$ (λ is not an independent parameter, as it is determined by α_L, α_H and α.) In contrast, the market-for-quacks model essentially had a single parameter, namely the firms' ex-ante success rate α. Also, note that if $\alpha_H = \alpha_L$, consumers know with certainty that they are in a market for quacks, and the model collapses to what I referred to as the rational-consumer benchmark. Therefore, in what follows we must insist that $\alpha_H > \alpha_L$.

Does the rationalizing model replicate the original model's key predictions?

Let us explore sequential equilibria in this incomplete-information game, and compare them to the unique Nash equilibrium in the market-for-quacks model. Formal near-equivalence between the original model and its rationalization is attained for the following parameter values: $\alpha_H = 1$, $\alpha_L = 0$, and $q_{HH} = q_{LL} = 1$. That is, high-quality (low-quality) alternatives satisfy the consumer's need with probability one (zero), and the consumer is perfectly informed of the alternative's type. The equilibrium strategy for high type firms is the same as in the basic model. Low types are always recognized as such and are never chosen, and so their pricing strategy is indeterminate as well as irrelevant for the market outcome.

The reason I refer to this as "near-equivalence" is two-fold. First, we should have compared the firms' ex-ante pricing strategy in the rationalizing model to the equilibrium strategy in the original market-for-quacks model. Instead, we compared the latter to the equilibrium behavior of *high-quality* firms in the rationalizing model. Second,

and more importantly, the two models generate different consumer behavior. In the rationalizing model, all consumers make the same decision in equilibrium. They are all informed of the firms' types, and since firms play continuous pricing strategies, price ties occur with probability zero. Therefore, all consumers make the same choice: they select the cheapest high-quality alternative (or the outside option, if no high-quality alternative is available). In contrast, recall that a salient feature of the market-for-quacks model was that each firm - including the most expensive one - had a positive clientele. Here, the most expensive firm ends up with either no clients or with all clients. In light of this discrepancy in the predictions that these two models make regarding consumer behavior, should we view this rationalization as successful?

The parameter values in this specification of the rationalizing model are also problematic in terms of interpretation. They imply that consumers are perfectly informed of the types of all market alternatives, while firms receive no signal about their opponents' types. It is not easy to think of market situations for which this would be a plausible assumption. And in any case, recall that our motivation was to replace imperfect rationality with imperfect information about firms' types, yet consumers turn out to be perfectly informed under these parameter values.

When we turn to more plausible parameter specifications, the rationalizing model fails to reproduce salient features of firm behavior in the original model. Recall that in the market-for-quacks model, the firms' equilibrium pricing strategy is a continuously increasing *cdf* over the interval $[(1-\alpha)^{n-1}, 1]$. That is, firms charge prices that range all the way up to consumers' willingness to pay for guaranteed satisfaction of their need, and these prices generate a positive clientele. Can sequential equilibrium in the rationalizing model reproduce this effect?

When consumers are imperfectly informed of the firms' types, the firms' pricing strategy in equilibrium is independent of their type, because there is nothing in the incentive structure in the model that enables firms to signal their type. In other words, equilibrium must be *pooling*. (Note, however, that in other cases, rationalization via modified information does introduce signalling issues that give rise to multiple equilibria. In this case, the rationalizing model's ability to replicate the predictions of the target bounded-rationality model relies on suitable equilibrium selection, thus raising a burden-of-proof problem similar to the parameter selection problem discussed here.)

By Bayes' rule:

$$\Pr(t = H \mid s = H) = \frac{\lambda q_{HH}}{\lambda q_{HH} + (1-\lambda) q_{LH}}$$

Therefore, when a consumer observes a good signal about an alternative, the alternative's posterior success rate is

$$\bar{\alpha} \mid H = \frac{\lambda q_{HH}\alpha_H + (1-\lambda)q_{LH}\alpha_L}{\lambda q_{HH} + (1-\lambda)q_{LH}}$$

Similarly, when a consumer observes a bad signal about an alternative, the alternative's posterior success rate is

$$\bar{\alpha} \mid L = \frac{\lambda q_{HL}\alpha_H + (1-\lambda)q_{LL}\alpha_L}{\lambda q_{HL} + (1-\lambda)q_{LL}}$$

In order for a consumer to be willing to pay a positive price for a firm, it must be the case that he received a bad signal about the outside option. This is just as in the market-for-quacks model. The reason is that all market alternatives are symmetric in terms of ex-ante quality, but the outside option comes free whereas firms charge positive prices. Therefore, the maximal price that consumers are willing to pay to firms is $(\bar{\alpha} \mid H) - (\bar{\alpha} \mid L)$. This has to be the maximal price in the support of the marginal equilibrium pricing strategy. It is easy to see that $(\bar{\alpha} \mid H) - (\bar{\alpha} \mid L) \leq 1$. This inequality is strict unless $q_{HH} = \alpha_H = 1$ and $q_{HL} = \alpha_L = 0$, which is the case we already covered above. Thus, the price range cannot be replicated under reasonable assumptions on the signal structure.

It could be argued that the range of equilibrium prices is not a key prediction of the original model, because of the difficulty of observing the consumers' underlying willingness to pay. But the rationalizing model also fails to reproduce the market-for-quacks model's comparative statics. As α_L and α_H go down, it is easy to see that since $q_{LL} > q_{LH}$ and $q_{HH} > q_{LH}$, $(\bar{\alpha} \mid H) - (\bar{\alpha} \mid L)$ decreases as well. Therefore, the maximal price that consumers are willing to pay in the rationalizing model goes down. In the limit, as α_L and α_H tend to zero, the maximal price converges to zero. This is in marked contrast to the effect of lowering ex-ante success rates on equilibrium prices in the original market-for-quacks model.[3]

To summarize, in the zoo of new parameters that are needed to specify the rationalizing model, there is a configuration of parameter values that roughly replicates the firms' equilibrium behavior in the market-for-quacks model. However, this configuration is inconsistent with the motivation of imperfect informed consumers. Indeed, it has a difficult-to-interpret property that consumers are fully informed of firms' quality,

[3] There are other ways to manipulate the rationalizing model's parameters in a way that lowers the ex-ante success rate. For instance, we can reduce λ without changing α_L or α_H. This would have a similar effect.

while firms do not receive any signal about their opponents' quality. Furthermore, consumer behavior in equilibrium differs from the market-for-quacks model. For all other configurations of parameter values, the rationalizing model fails to replicate the original model's range of equilibrium prices, and the comparative statics of expected prices with respect to the ex-ante success rate are diametrically opposed to what the original model predicts.

Summary

Our analysis has raised several questions regarding the burden of proof we may wish to impose on the rationalizing model. How should we evaluate a rationalization when it requires us to modify assumptions about the external environment, particularly when these are essential to the "moral" of the original story? Should we discount the rationalizing model if it forces us to introduce a number of new parameters? Is it enough to replicate the firms' behavior, or do we need to reproduce consumer behavior as well, in order for the rationalization to count as a success? Is it enough to find particular parameter values that replicate certain aspects of the original model's equilibrium? Or should the replication hold for a large range of parameter values? What is the interpretational burden on the parameter values that are used for replicating the original model's predictions?

4 Rationalization via Modified Preferences

When a certain choice pattern appears like a decision error that results from bounded rationality, we should entertain the possibility that what seems like an error is in fact a perfectly rational decision, and the only reason it seems erroneous is that we, as outside observers, attribute the wrong preferences to the agent. For instance, a manager's failure to choose a profit-maximizing project can be interpreted as evidence of a career concern. Replacing a behavioral model based on boundedly rational reasoning with a rational-choice model in which the consumer's preferences are re-specified is another common form of rationalizing BRMs.

Unlike the rationalization via modified information, rationalization via modified preferences turns out to be extremely effective in the case of the market-for-quacks model. Drop the assumption that consumers are interested in the firms' products only because they expect it to fix their problem. Instead, assume that there is an idiosyncratic value for each consumer-firm match. Specifically, a consumer's evaluation of each alternative $i \in \{0, 1, ..., n\}$ is $u_i - p_i$, where u_i gets the values 1 or 0, with

probabilities α and $1-\alpha$, independently across alternatives and consumers.

Rational consumers with this specification of independent private values behave *exactly* like boundedly rational consumers who evaluate alternatives according to the sampling procedure. Therefore, the rationalizing model is formally - and therefore behaviorally - equivalent to the market-for-quacks model. This is an extreme case of the methodological dilemma which motivated this paper. The formal equivalence between the two models means that every prediction about market outcomes that we make in one model is *perfectly* mimicked by the other. However, the welfare implications are radically different. The BRM implies that in a "market for quacks" (where the success rate of any product traded in the market is no different from the outside option of doing nothing), industry profits are a pure welfare loss for consumers. This loss can grow with the number of competitors. In contrast, in the rationalization, the market serves genuine consumer needs. It is welfare enhancing, and a greater number of firms is unambiguously better for consumers because it gives consumers access to a greater set of alternatives to choose from, while lowering their prices.

How should we compare these two accounts?

Prior plausibility of behavioral assumptions

I do not see any escape from the need to judge the prior plausibility of the behavioral assumptions that underlie two, formally equivalent models, in the context of their intended application. For instance, when the market in question is for forecasting services, then assuming independent private values makes little sense: every rational consumer should prefer a forecaster who makes more accurate predictions. In contrast, when the market in question is for self-help guides, both explanations are plausible. On one hand, independent private values make sense because different self-help guides may contain different pieces of advice that fit different people. On the other hand, casual observation suggests that people extrapolate naively from anecdotal evidence to evaluate self-help guides.

The only reason that I mention this obvious point is that there is a strong tradition in economic methodology (following Friedman (1953)) that is opposed to a priori judgments of behavioral assumptions and preaches that we evaluate models exclusively by their predictive success. However, when we need to choose between two formally equivalent models having different welfare implications, Friedman's positivistic criterion is too weak, and it seems clear to me that we should favor the model that is based on behavioral assumptions that make better sense in the relevant context.

Extended models

Even when two different models appear equivalent, they may differ when we move outside the original environment for which they were formulated. That is because different models tend to suggest different extensions. For example, Spiegler (2006b) extends the model of market competition with consumers who follow the sampling procedure, by allowing firms to randomize over prices. The extension presupposes that the same element of bounded rationality that made consumers extrapolate naively from small samples when evaluating the random performance of credence goods is going to make them extrapolate naively from small samples to evaluate random pricing strategies. Spiegler (2006b) shows that this behavioral model implies a strict incentive for firms to randomize over prices. Moreover, a greater number of competitors results in a mean-preserving spread in the equilibrium price distribution that firms adopt. In contrast, it is hard to think of an organic extension of the differentiated-taste rationalization of the market-for-quacks model that will generate these effects. Should the observation that the BRM and its rationalization become behaviorally distinct in an enlarged domain affect our judgment of the rationalizing model in the original domain?

Although the idea that extensions can break formal equivalence between two models is familiar, it has certain subtleties. Consider another extension of the market-for-quacks model, discussed in Spiegler (2006a), in which firms choose not only prices but also (simultaneously) whether to disclose their success rates to consumers. Disclosure is meaningless under the differentiated-taste rationalization, because its premise is that consumers are better informed than firms, and not the other way around. One could argue that this by itself provides a meaningful distinction between the sampling-based model and its differentiated-taste rationalization. However, it turns out that the equilibrium prediction of the extended model is that firms choose not to disclose their success rates. Therefore, as far as equilibrium behavior is concerned, the two models are equivalent after all. In the sampling-based model, disclosure is meaningful but fails to occur in equilibrium, while in the differentiated-taste rationalization, disclosure does not occur because it is meaningless in the first place. Can we legitimately say that the differentiated-taste rationalization replicates the sampling-based model's predictions in the extended domain that includes disclosure?

Summary

In this section we examined a rationalization that looks perfect at first glance, as it is formally equivalent to the original BRM. However, we identified two burden-of-proof issues. First, the behavioral assumptions underlying the rationalizing model may be implausible in the context of the original model's intended domain applications. Second, determining the relevant domain itself is not straightforward, because the

original model has natural extensions that are either nonsensical from the point of view of the rationalizing model, or generate predictions which are distinct from those of an analogous similar extension of the rationalizing model. How should we evaluate the rationalizing model in light of these observations?

5 Rationalization via Endogenization

Another way of rationalizing a BRM is to argue that the behavior it generates appears to be non-rational only because the model leaves out certain costs associated with the decision process. Once these are explicitly incorporated into the model, rationality is restored. In the extended model, the consumer chooses how much mental resources to spend on the decision problem, on the basis of "rational expectations" of the benefits of information processing. Note that it is not so much the friction as its formal treatment that is conventional. Decision costs are rarely incorporated into standard economic models. However, the type of extended model described above is conventional in that it treats decision costs as if they were search costs, or costs of acquiring information.

The following example has become almost canonical in methodological discussions of BRMs. (see Caplin and Schotter (2008)). An American tourist visits London (the tourist is invariably American in tellings of this story). Before crossing a street, he looks left, sees that the road is clear, starts walking, and gets hit by a car coming from his right. The bounded-rationality interpretation of the tourist's behavior is that he does not deliberate over his decision problem (when to cross the street). Instead, he follows an automatic rule that may be optimal in his home environment. A rationalization that incorporates information-processing costs would proceed as follows. The tourist realizes that he is on foreign soil and that he needs some time to remember which side the cars are coming from. However, spending time on this mental task is costly. The tourist rationally trades off this cost against the benefit of safe crossing.

As this example is only meant to illustrate a methodological dilemma, I will not get into a detailed discussion of the plausibility of the rationalization of the tourist's behavior. I will only comment that the two explanations of the tourist's behavior are in principle distinguishable. For instance, one could put up a sign for pedestrians saying "don't cross the street without thinking first". This intervention would have an impact only under the bounded-rationality interpretation.

Let us turn back to the market-for-quacks model. The consumer's procedure of sampling each market alternative and selecting the best alternative in the sample can be viewed as an optimal strategy in a larger model, in which we introduce search costs.

In such an extended model, there is an arbitrarily large number of firms. The consumer optimally designs a sample, taking into account the cost of obtaining information about the firms' quality and pricing decision. In this way we endogenize n as the size of the consumers' sample, given their correct expectation of the firms' equilibrium pricing strategy.

This rationalization shares all the problems of the rationalization method discussed in Section 3. In particular, the equilibrium pricing strategy can be replicated only if we assume that when a consumer samples a firm, he obtains perfect information about its type (whereas the firm's opponents are uninformed). We have already commented on the implausibility of this informational assumption. At any rate, in this case, it is easy to come up with a cost of obtaining a single sample point, for which the number of firms that the consumer samples will be optimal ex-ante, given their pricing strategy. However, the assumption that the consumer commits ex-ante to the size of his sample is problematic. Suppose that the consumer has sampled n firms, and all of them - as well as the outside option - turned out to be of low quality. As the cost of obtaining these sample points is sunk, what prevents the consumer from obtaining new sample points?

Rationalization via endogenization is an interesting modelling exercise. However, as a criticism of a BRM, it suffers from several methodological difficulties.

When should we endogenize informational constraints?
Any informational constraint in any economic model could be endogenized, by enabling agents to invest resources that help them relax their constraint. Economists address this endogeneity only when they wish to focus on the information acquisition process, and in most applications they are happy to treat the informational limitations as exogenous, because this is a good modeling strategy. The same standard should hold for rationality constraints. If the modeler has good reasons for certain restrictions on the process by which the consumer receives word-of-mouth information in the form of anecdotes, and a much fuzzier notion of the restrictions that could be imposed on the *costs* of actively looking for such anecdotes, then it is a good modelling strategy to take the sampling procedure as primitive.

At which scale should we endogenize informational constraints?
Even if the sampling procedure is a result of some optimizing process that takes into account information-processing costs, the optimization often takes place not on a case-by-case basis, but at a "general equilibrium" level, or on an "evolutionary" time scale. Consumers devise heuristics and calculational short-cuts that are meant to work well

on average across a large number of market (as well as non-market) situations. The sampling procedure is such a calculational short-cut. For the kind of partial equilibrium analysis that economists apply in Industrial organization, taking the consumer's decision rule as given is a good approximation.

A "Lucas critique"

In order for rationalization via endogenization to be operational, the consumer's optimization must be made on the basis of correct knowledge of the market equilibrium. Otherwise, we are not dealing with a rationalization, but merely shifting the element of bounded rationality to another level. The rational-expectations restriction often restricts the class of behavioral parameters of the original model which can be derived. For example, in the case of the market for quacks, we could imagine that certain pairs of parameter values (α, n) are simply inconsistent with equilibrium under the rationalizing model. In contrast, the market-for-quacks model does not impose any constraint on the range of values that they can get.

One point of view is that such a failure to rationalize certain specifications of the BRM should be regarded as a *criticism* of the BRM, akin to the famous *Lucas critique* of traditional Keynesian macroeconomic models (Lucas (1976)). According to this interpretation, the fact that certain specifications cannot be justified as equilibrium responses in a larger model that incorporates explicit information-processing costs means that these specifications are illegitimate. Note the rhetorical cunning at play here. So far, we have viewed success at rationalizing bounded-rationality models as a vindication of the "can't we get the same thing with a standard model" critique. Now, the "Lucas critique" turns the rationalization program on its head, and sees its failure as a reason to detract the bounded-rationality model. Nothing better illustrates the trickiness of debates over bounded-rationality models.

6 Discussion

This paper has been concerned with the following dilemma: what is the burden of proof that should be imposed on a rational-choice model (referred to as RM in the sequel) that is offered in refutation of a given BRM? The following scenarios abstract from the details of the specific case study we examined. Each scenario raises a difficulty in the evaluation of RM as a successful rationalization/refutation of BRM.

- RM mimics key predictions of BRM, but RM differs from BRM not only in behavioral assumptions, but also in assumptions about the external environment.

- RM introduces new parameters that were not included in BRM. Whether RM mimics key predictions of BRM depends on a suitable selection of parameter values.

- BRM and RM are observationally equivalent in a certain domain, but BR is based on behavioral assumptions that appear more plausible in this particular domain.

- BRM and RM are observationally equivalent in a certain domain of market situations, but become distinct when we extend (in a "natural" way) the behavioral models underlying BRM and RM to a broader domain.

- BRM and RM are observationally distinct out of equilibrium, but they imply identical equilibrium behavior. An extreme case is where a certain action is meaningful in BRM yet not taken in equilibrium, while the same action is meaningless and inconceivable a priori in RM.

- RM has multiple equilibria whereas BRM has a unique equilibrium. Whether RM mimics the predictions of BRM depends on a suitable equilibrium selection.

These difficulties should make us wary of the "can't we get the same thing with a standard model" critique. In my opinion, we should subject "rational explanations" are offered in refutation of BRMs to stricter burden-of-proof requirements. At the very least, the criteria that we use to evaluate rationalizing models should not differ from those that we use to evaluate two rational-choice explanations. If a certain prediction of a given model is qualified by a particular choice of parameter values or a particular assumption about the external environment, we should not discard these qualifications simply because the model is offered in refutation of a BRM.

We should also be reluctant to tamper with assumptions of a bounded-rationality model that concern the *domain* of market situations that agents face. Whether a model is static or dynamic, whether its informational constraints are endogenous or exogenous, whether agents are assumed to be homogenous or heterogeneous - these are all modeling choices that the theorist makes to *define* the limits of the theoretical exercise he wishes to pursue. We are not forced to respect these assumptions when discussing the model's merits and drawbacks, but we should try to accept them as given when advancing alternative "rational explanations".

If the "can't we get the same thing with a standard model" critique is so problematic, why is it so popular? I believe that the reason is simple: are we *used* to looking for "rational explanations". It is what we do for a living. Rationalizing behavior is

an important part of what defines economic theory, and it has had great successes. Rationalizations of superficially non-rational behavior continue to the subject of very interesting works (e.g., Samuelson (2001), Compte and Postelwaite (2005), Kamenica (2008), Baliga and Ely (2009)). The "can't we get the same thing with a standard model" critique simply extends this type of reasoning. We are so effortless in our search for rational explanations that we tend to overlook the rough edges, especially when we have a prior inclination to reject explanations of economic behavior that depart from the rational-choice paradigm. But I hope that this paper has shown that a modeling style can be very useful for understanding phenomena, and yet quite weak as a way of criticizing models that follow alternative modeling styles.

References

[1] Baliga, S. and J. Ely (2009), "Mnemonics: The Sunk Cost Fallacy as a Memory Kludge", Northwestern University, mimeo.

[2] Bernheim, D. (2009), "On the Potential of Neuroeconomics: A Sober (but Hopeful) Appraisal", American Economic Journal: Microeconomics 1 (2), 1-41.

[3] Binmore, K. and A. Shaked (2010), "Experimental Economics: Where Next?" Journal of Economic Behavior & Organization 3 (1), 87-100.

[4] Caplin A. and A. Schotter, eds. (2008), Perspectives on the Future of Economics: Positive and Normative Foundations, Handbook of Economic Methodology, Vol. 1, Oxford University Press.

[5] Compte, O. and A. Postelwaite (2005), "Confidence-Enhanced Performance", American Economic Review 94, 1536-1557.

[6] Dekel E. and B. Lipman (2009), "How (Not) to Do Decision Theory", Annual Review of Economics, forthcoming.

[7] Kamenica, E. (2008), "Contextual Inference in Markets: On the Informational Content of Product Lines", American Economic Review 98, 2127–2149.

[8] Lucas, R. (1976), "Econometric Policy Evaluation: A Critique", in Carnegie-Rochester conference series on public policy, Volume 1, pp. 19–46, Elsevier.

[9] Rabin, M. (2002), "A Perspective on Psychology and Economics", European Economic Review 46, 657–685.

[10] Rubinstein, A. (2006), "Dilemmas of An Economic Theorist", Econometrica 74, 865-883.

[11] Sameulson, L. (2001), "Analogies, Adaptations, and Anomalies," Journal of Economic Theory 97, 320-367.

[12] Spiegler, R. (2006a), "The Market for Quacks", Review of Economic Studies 73, 1113–1131.

[13] Spiegler, R. (2006b), "Competition Over Agents with Boundedly Rational Expectations", Theoretical Economics 1, 207–231.