



VOLUME 6 | NUMBER 1 | SPRING 2010

Competition Policy International

The Impact of Behavioral Economics on Consumer and Competition Policies

Eliana Garces

The Role of Behavioral Economics in Competition Law: A Judicial Perspective

Vivien Rose

The Failed Resurrection of the Single Monopoly Profit Theory

Einer Elhauge

What Does Behavioral Economics Mean for Competition Policy?

Matthew Bennett

The Future of Behavioral Economics in Antitrust Jurisprudence

Douglas H. Ginsburg

A Behavioral Model of Rational Choice

Cecilia Chaing



Edited by David S. Evans

From the Editor

David S. Evans

University of Chicago &
University College London



From the Editor

CPI's Spring 2010 issue takes a comprehensive look at behavioral economics and its implications, if any, for the practice of competition policy.

Since the late 19th century economists have generally analyzed markets under the view that people and businesses have specific values they care about—utility for consumers and profits for businesses—and operate in systematic ways to maximize those values over time. The short-hand term for the various assumptions that go into this framework is "people and businesses behave rationally." This approach has resulted in a vast and influential body of work that underlies all economic fields, as well as related subjects including finance, marketing, and antitrust.

Over the years, however, many economists have questioned this assumption of rationality. The field that is now known as "behavioral economics" grew rapidly following a 1980 paper by Richard Thaler. He argued that people systematically deviate from rational utility-maximizing behavior and this deviation should be taken into account in predicting actual consumer behavior. Much subsequent work has relied on laboratory experiments—games played with undergraduates for example—both to identify systematic patterns of behavior and to document departures from the rational actor model. Behavioral economics writings have flourished and now form the basis for an increasing number of policy proposals: from "sin taxes" to discourage people from making bad choices they will come to regret to "opt-out" policies that encourage people to make choices that are better for them.

The traditional view that economic actors are rational has influenced the economics of competition policy. The analysis of the effect of business behavior on consumer welfare, for example, assumes that people are worse off when they have to pay higher prices or get less output. Given that premise, it is then assumed that people make the "right" decisions when they purchase goods and services. But what if, as some research in behavioral economics suggests, people make systematic mistakes when making purchase decisions and consume too much of some goods or too little of others. A merger of fast-food restaurants could make consumers better off by increasing prices and discouraging them from engaging in bad behavior. To take another example, the assumption that businesses maximize profits is behind every economic

premise from critical loss analysis to price-cost tests for predation. But what if, again as some research suggests, firms do not focus on marginal costs in determining profit-maximizing prices?

Behavioral economics is still controversial and its implications for public policy are just now being fully explored. The articles in this issue make a substantial contribution in assessing where, if anywhere, behavioral economics is relevant to antitrust and the increasingly related field of consumer protection. The first three articles provide perspectives by several academic economists. Mark Armstrong & Steffen Huck lead the discussion with an initial review of what behavioral economics literature tells us about the behavior of firms and their management, and then look at the implications for antitrust. Roman Inderst & Marco Ottaviani focus on behavioral biases by consumers, how businesses can exploit those biases, and how to design consumer protection policies to prevent potentially harmful behavior. Michael Salinger then turns to an overview of behavioral economics and argues that, although it is relevant to consumer protection, it offers few insights for the practice of antitrust.

The next group of articles evolved from presentations made at the 2009 Jevons Colloquium. D.C. Circuit Court of Appeals Judge Douglas Ginsburg & Derek Moore examine what impact behavioral economics has had on court decisions, and whether that impact is likely to grow. Vivien Rose, a chairman of the U.K. Competition Appeals Tribunal, utilizes actual cases to argue that the courts have followed approaches that are closer to behavioral than traditional economics. Then we turn to the perspectives of several officials at competition authorities: Matthew Bennett, John Fingleton, Amelia Fletcher, Liz Hurley, & David Ruck from the Office of Fair Trading, which handles both antitrust and consumer protection issues in the United Kingdom; Alison Oldale from the U.K. Competition Commission; and Eliana Garcés who is in the cabinet of the Commissioner for Competition Policy of the European Commission.

In the Autumn 2009 issue of CPI, Paul Seabright, Harry First, Daniel Crane & Joshua Wright, and Barry Nalebuff discussed a recent Einer Elhauge article claiming that the Chicago single-monopoly profit theorem has been repudiated, and that many forms of tying and bundling should be considered suspect practices. Each of these authors disagreed with Elhauge in various ways. Seabright and Crane & Wright, in particular, disputed Elhauge's conclusions on the single monopoly profit theorem. The spirited debate continues in this issue with a rejoinder by Elhauge who, in particular, insists that the Chicago single monopoly profit theorem really is dead.

We continue our practice of presenting recent cases of special interest with the analysis by Antonio Bavasso & Mark Friend of Barclays' successful appeal of the Competition Commission's judgment regarding personal protection insurance.

Bavasso & Friend examine the implication of the Competition Appeals Tribunal's repudiation of the remedies adopted by the Competition Commission, relating this decision and related judgments to the larger topic of judicial review of competition authority decisions.

We conclude this issue with one of the early pieces in behavioral economics: Herbert Simon's 1955 article on bounded rationality. As Lindsay McSweeney notes in her introduction, Simon imagines how actors that have limited information and computational abilities might behave. For many years, this article provided one of the leading alternatives to the traditional model of rational profit-maximizing behavior, bringing to the forefront of academic attention several of the key issues on which behavioral economics would eventually focus.

On behalf of CPI's readers and its editorial team, I am delighted to extend my thanks to all the contributors to this issue.

David S. Evans
University of Chicago and University College London

Behavioral Economics as Applied to Firms: A Primer

Mark Armstrong & Steffen Huck
University College London

Behavioral Economics as Applied to Firms: A Primer

*Mark Armstrong and Steffen Huck**

We discuss the literatures on behavioral economics, bounded rationality, and experimental economics as they apply to firm behavior in markets. Topics discussed include the impact of imitative and satisficing behavior by firms, outcomes when managers care about their position relative to peers, the benefits of employing managers whose objective diverges from profit-maximization (including managers who are overconfident or base pricing decisions on sunk costs), the impact of social preferences on the ability to collude, and the incentive for profit-maximizing firms to mimic irrational behavior.

*Department of Economics, University College London. We have benefited from the non-selfish behavior of many friends and colleagues, and we are grateful for advice received from Matthew Bennett, Emanuele Giovannetti, Antonio Guarino, Joe Harrington, Bruce Lyons, Adam Rosen, David Ruck, David Sappington, Burkhard Schipper, John Vickers, Eric Van den Steen, Maurice Stucke, Mark Williams, Jidong Zhou, Liva Zorgenfrei, and the editorial team at this journal. We also want to thank the Office of Fair Trading and the Economic and Social Research Council (UK) for funding assistance.

I. Introduction

In recent years there has been a good deal of research investigating how poor or non-standard decision making by consumers might affect market outcomes. In much of this work, the assumption is that firms are fully rational and aim to maximize their profits (and sometimes they do this by exploiting the behavioral biases of consumers). Some of this work points to situations where there is a role for policy which protects consumers from their own failings and from exploitative firms.¹

In this article we focus instead on non-standard approaches to firm behavior. Consumers are kept in the background, and are present merely to generate in some fashion a demand curve for the firms' products. We present evidence—both real world and experimental—that firms (or experimental subjects playing the role

WE PRESENT EVIDENCE—
BOTH REAL WORLD AND
EXPERIMENTAL—THAT FIRMS
(OR EXPERIMENTAL SUBJECTS
PLAYING THE ROLE OF FIRMS)
SOMETIMES DEPART FROM THE
PROFIT-MAXIMIZING PARADIGM.

of firms) sometimes depart from the profit-maximizing paradigm. For instance: firms may be content to achieve “satisfactory” rather than optimal profits; firms might rely on simple rules of thumb—such as imitating the strategies of well-performing rivals, or changing strategies only when profit falls below some acceptable threshold—rather than on explicit calculation of complex optimal strategies; firms may base pricing decisions on sunk costs as well as avoidable costs; CEOs may be overly optimistic about the profitability of mergers or other actions they undertake; managers might face incentives which induce them to care about relative rather than absolute profits; firms might punish rivals who behave “unfairly” towards them; and so on. We believe that many mainstream industrial economists and policy-makers are not yet fully aware of the substantial literature on these topics, and our aim in this survey is to bring some of the insights of behavioral economics as applied to firms to wider attention.²

There are, of course, a number of reasons why one might expect firms to be better decision makers than consumers, and this helps to explain the recent focus on consumer failings. First, there are economies of scale in making good decisions. A consumer may have to decide whether to buy a given product just a few times and it may not be worthwhile to invest much effort in making the right decision, while a firm selling to millions of customers has more at stake in getting it right. Relatedly, since firms often do the same things repeatedly, they may quickly learn how to do it right, while a consumer buying a rarely-purchased product may not have that opportunity. Second, firms compete with one another while consumers usually do not, and firms that are better at generating profits may succeed and prosper at the expense of firms which make worse decisions. In modern society, consumers rarely “exit” when they make poor market decisions.

However, there are also a number of considerations which go the other way:

- Firms typically operate within a highly complex and uncertain environment, and may need to resort to decision-making short-cuts and rules-of-thumb.
- Firms often face the added complexity stemming from strategic interaction with rivals, which consumers rarely do when buying products.³ For instance, in even the simplest situations, theories of tacit collusion or reputation-building require firms to follow highly intricate strategies.
- In practice, under-performing firms may take a long time to exit, and there is often a long period of decline before an established firm actually leaves the market. More generally, the complexity of the environment may mean that crucial decisions are taken with significant delay.⁴
- Group decision making (as practiced more often by firms rather than by consumers) could introduce extra biases. For instance, the separation of ownership from control could leave managers free to pursue

THAT IS TO SAY, ALTHOUGH
IT APPEARS PARADOXICAL,
ACTUAL PROFITS MIGHT BE
ENHANCED WHEN THE FIRM'S
OBJECTIVE DEPARTS FROM
PROFIT-MAXIMIZATION.

their own objectives that may differ from maximizing shareholder value. In fact, this last point can be turned on its head: In imperfectly competitive markets, shareholder returns might be enhanced by (deliberately or not) hiring managers whose objectives differ from maximization of profit. That is to say, although it appears paradoxical, actual profits might be enhanced when the firm's objective departs from profit-maximization. For instance, hiring

an aggressive or over-optimistic CEO, rewarding a CEO based on her performance relative to peers, or employing a CEO who bases pricing decisions on sunk costs, might all have strategic benefits in terms of affecting rival responses.

- Relatedly, the people who succeed in the tough career competition to manage firms might have these kinds of personality traits more frequently than the general population. Sometimes it seems that managers are motivated in part by personal animosity—or respect—towards a rival.
- Different managers have differing skills, and in particular some managers may be better at thinking strategically than others.⁵ More generally, it seems clear that an individual manager's "style" can be important—for good or ill—for a firm's performance.⁶
- A manager may on occasion have a personal interest in the firm's activities (say, a sports team, newspaper, or fine wine⁷) beyond the profit generated, and this may cause a divergence from profit-maximization.⁸ (Captain Ahab, in Melville's *Moby Dick*, is an extreme, if fictional, instance.)

- Rather than affecting rival responses as mentioned above, explicitly non-profit aims might act to boost consumer demand, as appears often to be the case with “fair-trade” or “green” products, for instance.
- Illegal cartels need to find ways to resolve disagreements about market shares, whether cheating has occurred, and so on, without recourse to legally binding agreements. As such, issues such as building trust and *esprit de corps* among conspirators are important.
- Finally, the potential presence of a “behavioral” type of firm in a market could induce a profit-maximizing firm to mimic irrational behavior. For instance, a firm might wish to gain a reputation to fight entry come what may, in order to deter future entry.

Milton Friedman is perhaps the most famous, and often quoted, exponent of the pervasive view that competition acts to discipline firms to maximize their profits:

“Let the apparent determinant of business behavior be anything at all—habitual reaction, random chance or what not. Whenever this determinant happens to lead to behavior consistent with rational and informed maximization of returns, the business will prosper and acquire resources with which to expand; whenever it does not, the business will tend to lose resources and can be kept in existence only by the addition of resources from outside. The process of ‘natural selection’ helps to validate the hypothesis [of ‘rational and informed maximization of returns’]—or, rather, given natural selection, acceptance of the hypothesis can be based largely on the judgment that it summarizes appropriately the conditions for survival.”⁹

As we discuss in our conclusions, this view has much merit when it comes to many kinds of decision errors and behavioral biases, and some decision problems exhibited by consumers (such as procrastination, for instance) are less plausible in firms. Nevertheless, at various points in this survey we will see examples of situations where firms prosper when their objective diverges from maximizing profits, when “behavioral” managers are put in place, or when firms are content to achieve merely a satisfactory profit level.

In the remainder of this article, we discuss several of these points in greater detail. Section II surveys the experimental evidence on the ability of firms to collude, and how public policy could affect this ability. Section III discusses the related issues of imitative behavior by firms and concerns for relative (not absolute) profit, both of which often make an oligopolistic market more competitive than the orthodox theory suggests. Other forms of social preferences—a

desire for vengeance if a rival “cheats” on a collusive agreement, say—are presented in section IV, which we argue may help sustain collusive agreements. Various kinds of satisficing behavior are discussed in section V, where we show how imperfectly optimizing firms may actually end up with greater profits than their profit-maximizing counterparts. Section VI collects together discussion of the impact of over-optimism by entrepreneurs and managers, the imitation of irrational behavior by profit-maximizing firms, and the possible benefits of including fixed costs in a firm’s pricing decisions. Concluding comments are contained in section VII.

We do not attempt in this article to survey the entire terrain of behavioral economics as applied to firms. In particular, our focus is on firm behavior in markets rather than on the internal organization of firms. Similarly, we do not emphasize the potential policy implications of non-standard behavior by firms, although a few remarks about this are made throughout the paper and in our conclusions.

Before continuing, however, it is worthwhile discussing the role of laboratory experiments—which supply much of the empirical evidence we present in the following discussion—as an aid to understanding firm behavior. Laboratory experiments usually involve students playing the role of firms, who make decisions in minutes and for relatively low stakes, whereas real firms employ man-

agers who are highly remunerated, experienced, and carefully selected. Why should these experiments tell us anything about the performance of actual markets? There are a number of reasons why we cautiously believe that data from experiments are indeed useful.¹⁰

First, experiments have been run with business people as subjects instead of students, and the latter do not appear to perform less (or more)

rationally than the former.¹¹ In addition, experimenters are careful to ensure that their subjects have a good deal of experience in playing the chosen game before they start giving weight to the generated data. Indeed, it is possible that subjects have more experience playing the highly stylized games in the laboratory than real managers have in their own naturally occurring markets. Second, as Plott¹² writes:

WHY SHOULD THESE EXPERIMENTS
TELL US ANYTHING ABOUT THE
PERFORMANCE OF ACTUAL
MARKETS? THERE ARE A NUMBER
OF REASONS WHY WE CAUTIOUSLY
BELIEVE THAT DATA FROM
EXPERIMENTS ARE INDEED USEFUL.

“The relevance of experimental methods rests on the proposition that laboratory markets are ‘real’ markets in the sense that principles of economics apply there as well as elsewhere. Real people pursue real profits within the context of real rules. The simplicity of laboratory markets in comparison with naturally occurring markets must not be confused with questions about their reality as markets.”

Thus, theories and models that work generally should also work in the particular cases of laboratory markets.

Third, many theories in industrial organization are very subtle, and depend on detailed assumptions about what firms observe about their rivals, and what firms can say to their rivals. Information flows can be carefully controlled in the laboratory, while the environment of naturally occurring markets is harder to pin down, and it is often impractical to test complex theories about oligopoly behavior using non-experimental data. For instance, experiments are able to examine tacit collusion, in which no communication is possible between firms, while it is hard to be sure there is no communication going on in a real market. It is also hard to measure such important parameters as marginal costs in natural markets, while these can be accurately generated in the laboratory.¹³ Nevertheless, worries about the relevance of laboratory experiments should not be dismissed too casually: The concern that student behavior does not always closely match CEO behavior is probably more serious for oligopoly experiments than for other experiments (such as those concerning shopping or bargaining), and it is hard convincingly to replicate the detailed institutional structure of firms—such as hierarchies and group dynamics—in the laboratory.

II. Ability to Sustain Collusion

Collusion among oligopolists occurs when firms can sustain relatively high prices by credibly threatening a price war should one firm undercut the prevailing price (or boost its output). Firms need to interact repeatedly for high prices to be sustained, since a firm must trade off high profits now (if it deviates and undercuts its rivals) with low profits in the future (after a price war is triggered). A typical “trigger” strategy to sustain high prices takes this form: Each firm sets the collusive price as long as all firms continue to do so, but if one or more firms sets a price below the collusive price, all firms subsequently set prices according to a one-shot equilibrium strategy (e.g., set price equal to marginal cost if the firms supply homogenous products). Although even this simplest setting involves complex strategies, including the need somehow to coordinate on the particular collusive price, matters become vastly more complex if: firms are asymmetric; demand varies over time; firms cannot observe each others’ price (and so cannot tell if low demand now is due to an adverse demand shock or due to one firm offering a low price); and so on.¹⁴

Fully rational firms (who are sure their rivals are also rational) cannot sustain collusion if there is a known end-point to their interaction, no matter how far in the future. (If the market ends after 100 periods, in the 100th period firms know there is no future to punish them, and so set low prices to undercut their rivals, and the one-shot equilibrium is played. In the 99th period, firms know what will happen next period, so again have no incentive to cooperate then. And the whole repeated interaction unravels, with the result that no collusion

is ever sustained.) When there is open-ended interaction, collusion is harder to sustain when there are more firms in the market. With many firms in the market, the short-run benefits of a price cut are relatively large compared to getting a small share of the on-going collusive profits. This means that collusion among many firms requires a higher discount factor (i.e., more weight placed on future

MANY EXPERIMENTS FIND THAT COLLUSION CAN STILL BE OBSERVED EVEN WHEN THERE IS A KNOWN FIXED NUMBER OF PERIODS; A FINDING THAT GOES AGAINST THE PREDICTION OF RATIONAL PLAY.

collusive profits) than does collusion among few firms. Nevertheless, theory suggests that in plausible environments, collusion is achievable with relatively large numbers of rational firms.¹⁵ Collusion is also likely to be harder to sustain if firms cannot observe each others' actions.¹⁶ For instance, if a firm cannot tell if low demand is due to an adverse market shock or due to an undercutting rival, then it may be

reluctant to punish the rival harshly (since punishment hurts the punisher as well), and this could make collusion hard to achieve. But if a firm can observe a rival's actual price, it can punish harshly when that price is low, and so make collusion more effective. For this reason, frequent information exchange is a vital part of many cartels.¹⁷

A. COLLUSIVE BEHAVIOR IN THE LABORATORY

How do firms behave in repeated interactions in experiments? A typical experiment to analyze this question is conducted as follows:

- The same subjects repeatedly play a stylized oligopoly game (say, Cournot quantity competition or Bertrand price competition) for a number of periods. The number of periods is either fixed (and announced in advance) or random (with, say, a 90 percent chance of another interaction after each period's play). In the latter case, the likelihood of having another interaction plays the role of the discount factor, representing how important the future is.
- In each period, the actions available to each firm might be very restricted (e.g., with just two possibilities: "collude" or "compete" as in the Prisoner's Dilemma) or more numerous (e.g., choosing any integer output from 0 to 100).
- The payoff structure of each period's game might be described in full to each player if the actions are not too numerous (e.g., if playing a Cournot game a payoff table shows how a player's profit depends on her own chosen output along with the aggregate output of the rivals). If actions are numerous, then subjects may be given a "profit calculator" which gives the subject's profit as a function of some specified values for her own action and those of her rivals. Alternatively, the payoff structure might not be revealed at all, and after each period's play a subject sees her own realized profit, as well as (possibly) the actions and realized profits of her rivals.

- Particular market features could be introduced. For instance, players could be permitted to make (non-binding) announcements in advance about what actions they will follow, or players could decide whether to form a cartel.¹⁸

A focus of the literature has been to observe the extent of collusion in the laboratory market (say, measured as the ratio of actual profits to the theoretical one-shot Nash equilibrium profits), and how this depends on: i) the discount factor (or number of periods in the fixed-period case); ii) the number of oligopolists; iii) the kind of information revealed to participants; or iv) whether they can make pre-play announcements about their actions. Note that it is only in a sterile laboratory setting, where the information and communication flows can be precisely controlled, that one could investigate some of these questions.

This is not the place for a detailed summary of the large literature on experimental repeated oligopoly. But some highlights include the following observations.¹⁹ First, the importance of the future does affect the extent of collusion. Feinberg & Husted²⁰ performed Cournot experiments with two different probabilities of continuing interaction, and observed that collusion was more prevalent with the higher probability. This is qualitatively in line with predictions of the behavior of fully rational firms in indefinitely repeated games.²¹ However, many experiments find that collusion can still be observed even when there is a known fixed number of periods; a finding that goes against the prediction of rational play.²² Of course, even with a finite horizon it is perfectly rational to cooperate for a time if one thinks that there is a chance one may be playing against a naïve opponent who is playing something like a trigger strategy (see section VI below for further discussion). But it is common to observe that collusion breaks down as the known endpoint nears.

In experimental markets, it appears that the number of oligopolists is crucial in terms of the ability to collude tacitly, i.e., where firms cannot communicate directly with each other. Huck, Normann, & Oechssler²³ document how there is very little tacit collusion in Cournot markets with three or more firms, which conflicts with the theoretical predictions discussed earlier. Indeed, if anything, it is more common to observe that markets with more than two firms are more competitive than the static Nash prediction. On the other hand, there is a considerable amount of collusion in duopolies.²⁴ One obvious difference between duopolies and oligopolies is that when there is a deviation from the collusive strategy, it is obvious to the two suppliers which of them deviated, presuming a deviation is known to have occurred. Moreover, it is possible to inflict punishment on that firm without hurting innocent rivals who were colluding according to plan.

IN EXPERIMENTAL MARKETS, IT APPEARS THAT THE NUMBER OF OLIGOPOLISTS IS CRUCIAL IN TERMS OF THE ABILITY TO COLLUDE TACITLY, I.E., WHERE FIRMS CANNOT COMMUNICATE DIRECTLY WITH EACH OTHER.

In Bertrand markets with homogeneous products, prices above competitive levels are routinely observed in experiments, even with more than two firms or with a one-shot market interaction.²⁵ However, these markets are very special, as in the one-shot Bertrand equilibrium firms make zero profits (when unit costs are constant), and so the only chance firms have to make positive profits is to price above cost (and they have little to lose if they fail to offer the lowest price). These markets are discussed further in section V, where we argue that satisficing behavior by firms is a plausible explanation for observed prices.

Huck, Normann, & Oechssler²⁶ and Offerman, Potters, & Sonnemans²⁷ examine the impact of the experimenter revealing information about each rival's

HOWEVER, ALL THREE STUDIES
FIND THAT REVEALING
INFORMATION ABOUT RIVAL
OUTPUTS AND PROFITS
ACTUALLY RENDERS MARKETS
MORE COMPETITIVE.

actions and payoffs to firms. (Huck et al. investigate four-firm Cournot markets, while Offerman et al. look at three-firm Cournot markets.) In some models of tacit collusion (though not the ones studied in these laboratory experiments), when firms can directly observe each other's past actions, this enables deviations to the collusive agreement to be detected, and so helps collusion be achieved.²⁸ However, all

three studies find that revealing information about rival outputs and profits actually renders markets more competitive.²⁹ Compared to the setting in which no firm-specific information is revealed, they find significantly lower prices when firms can observe each other's actions and profits.³⁰ While standard economic theory cannot easily explain this empirical finding—that competition in the presence of information about others is often tougher than in the static Nash equilibrium, sometimes approaching the fully competitive outcome where price equals marginal cost—there is a class of models in the evolutionary and learning literature that does predict such patterns, which we discuss in section III.

A related issue concerns the impact of communication between firms, and whether the ability to engage in some form of communication before market interaction aids collusion in the laboratory. (Of course, any agreements made in the communication stage are non-binding and purely “cheap talk.”) This literature is surveyed in Potters,³¹ who concludes that given the opportunity to communicate, firms do use this opportunity to conspire to fix prices, and this ability often has the effect of raising prices in the market (even with more than two firms). It appears that face-to-face talk has more of a collusive impact than computer-mediated communication, somewhat consistent with the camaraderie among conspirators we discuss later in section IV. Andersson & Wengstrom³² report results from an experiment in which it is costly to send a message from one firm to another. Intriguingly, they find that as the cost of communication rises (which could be interpreted as the outcome of a more vigilant antitrust policy) fewer messages are sent, as is intuitive, but when messages are sent collusion is more effective. One interpretation is that firms feel more committed to a collu-

sive agreement if costs are needed to reach that agreement. The net impact of making it hard to communicate is that collusion is substantially *more* prevalent, so that certain forms of competition policy might perversely turn out to aid cartel formation and stability.³³

B. LENIENCY POLICIES

Here we continue the discussion of cartel formation in the presence of a competition authority, but in the context of leniency policies. Both U.S. and European antitrust authorities currently make use of leniency policies for whistle blowers, and in recent years the number of detected cartels has increased considerably. The idea is that providing an incentive to whistle blowers should render cartels less stable, as well as providing antitrust authorities with evidence which would otherwise be hard to obtain. While these benefits appear intuitive, theoretical work by Spagnolo³⁴ and Motta & Polo³⁵ have cast doubt on this intuition. The logic of their argument is that firms can use whistle blowing as a way to punish those members of a cartel which undercut agreed prices. Ironically, the ability to blow the whistle may then help to sustain collusion.

IRONICALLY, THE ABILITY TO
BLOW THE WHISTLE MAY THEN
HELP TO SUSTAIN COLLUSION.

Antitrust authorities have a number of possible policies to fight cartels, including (i) fining the participants when misconduct is proved, (ii) granting leniency to cartel members who bring evidence of the cartel to the authority, or (iii) rewarding a cartel member if it brings forward evidence of the cartel. Apesteguia, Dufwenberg, & Selten³⁶ experimentally investigated the impact of various leniency policies. They studied three-firm Bertrand markets under four experimental treatments: Standard, Leniency, Bonus, and Ideal. In Standard, Leniency, and Bonus, firms have the opportunity to form a cartel before they interact in the market. (A cartel is formed only if all three firms agree, and firms interact just once in the market to avoid issues of tacit collusion.) If a cartel is formed, the three firms can communicate in an unstructured way for 10 minutes, presumably to discuss their collusive strategy. (For instance, one possible strategy is to agree on a collusive price and to suggest that if one firm undercuts the price its rivals report the cartel.) In Standard there is no leniency clause and every firm (including the firm that blows the whistle) is fined a fraction of their turnover if a cartel has been formed and reported to the authorities. In Leniency the whistle blower gets a discount on the fine. (The discount is 100 percent if there is only one whistle blower, 50 percent if there are two simultaneous whistle blowers, and 33.3 percent if all three do so.) In Bonus the whistle blowers share the fines paid by the cartel members who kept quiet. In the fourth treatment, Ideal, the design is such that firms simply do not have the opportunity to form a cartel.

The parameters chosen in the experiments imply that with rational players the Standard and Leniency treatments can sustain collusion while Bonus and Ideal cannot. However, the data draw a different picture, and the most effective treat-

ment is not Bonus but Leniency despite its potential theoretical weaknesses. In fact, Leniency generates prices that are statistically indistinguishable from prices in Ideal where cartel formation is not possible. But theory not only fails to predict the performance of Leniency, it also gets the effect of Bonus wrong. Predicted to be the most effective remedy against collusion, Bonus, in fact, leads to prices which are as high as those seen in the Standard treatment. In addition, the Bonus treatment leads to the greatest incidence of cartel formation, while Leniency has the least. The authors conjecture that subjects in the Bonus treat-

THUS, BOTH OF THESE STUDIES
SHOW THAT LENIENCY PROGRAMS,
DESPITE THEIR POTENTIAL
WEAKNESSES IN THEORY, ARE
SUCCESSFUL IN FIGHTING CARTELS.

ment are tempted to form a cartel with an agreement to set high prices, and then also being rewarded for reporting the cartel (although this strategy is not in fact profitable when many firms form a cartel with this duplicitous intent). Apesteguia, et al. conclude: "Our findings in this paper provide no reason for Gary Spratling and Mario Monti to feel disappointed with the leniency clauses that have recently been incorporated into the anti-trust legislation in most member states of the OECD."

While Apesteguia, et al. study a one-shot interaction, Hinloopen & Soetevent³⁷ examine a setting where three firms repeatedly interact in the same market for 20 periods. (With a repeated interaction, firms may be more reluctant to report a cartel, since they then forgo the future benefits of collusion. In addition, even if cartel formation is impossible, there is the possibility of tacit collusion.) They report that leniency programs not only reduce the frequency with which cartels are formed but also reduce the stability of cartels that do get formed. Moreover, cartels that do get formed in the presence of the leniency programs charge relatively lower prices. Thus, both of these studies show that leniency programs, despite their potential weaknesses in theory, are successful in fighting cartels.

III. Imitative Behavior and Concerns for Relative profit

Rather than each firm laboriously calculating its own optimal strategy, even if that were feasible, it is plausible that firms may sometimes choose to imitate the strategies of their more successful peers. As Alchian put it:

"[W]henver successful enterprises are observed, the elements common to these observable successes will also be associated with success and copied by others in their pursuit of profits or success. [...] What would otherwise appear to be customary 'orthodox,' non-rational rules of behavior turns out to be

codified imitations of observed success, e.g., ‘conventional’ markup, [...] ‘orthodox’ accounting and operating ratios, ‘proper’ advertising policy, etc.”³⁸

Alchian suggests that imitation enables firms to make use of other firms’ private information and optimizing behavior, and to enjoy the benefits of conformity (as imitating firms are likely to do as well as the average of their peers). To discuss these and other points further, it is useful to distinguish between imitation by non-rivalrous firms and imitation by firms competing in the same market.

A. IMITATION BY NON-RIVALROUS FIRMS

For now suppose that the relevant firms do not interact strategically in a market, e.g., they may be local monopolists, or they may be price-taking firms facing some given exogenous price (such as farmers deciding on an agricultural technique). It is plausible that “optimizing” and “imitating” firms might co-exist, and each firm chooses which of the two broad kinds of strategy to follow. If firms incur a cost for calculating the optimal action, then when most firms are optimizers and firms operate in a similar environment, it may benefit a firm simply to copy these optimizing firms’ actions and save itself the calculation cost. Of course, if too many firms free-ride and copy average observed behavior, there may re-emerge a benefit to investing in optimizing. Since optimizers provide a free service to imitators, in general there will be too few optimizers in equilibrium.³⁹

In other situations, firms may be unsure about the optimal action even after significant deliberation, and instead they may have some kind of noisy signal about what is the best action. (For instance, a firm might read a trade report or attend a conference about a new technology in the market.) If firms choose their actions sequentially, and can observe the actions chosen by earlier firms, then a (rational) firm should take into account what other firms did before choosing its own action. In such cases, there is a danger that firms will become locked into taking the wrong action.⁴⁰ Such a framework might be used to explain inefficient technology adoption or merger waves by profit-maximizing firms. However, the predicted equilibrium depends on very sophisticated reasoning by firms. (For instance, the third firm needs to work out what the second firm’s action implies in terms of that firm’s private signal, and the required “depth of reasoning” gets progressively more onerous for firms which are further back.) In experiments, it appears that firms do not play this equilibrium, and instead put more weight on their private signal than they should.⁴¹ This could be interpreted as a form of over-confidence, a topic to which we return in section VI.

Career concerns of managers might give a reason why a manager mimics the action of a peer, even if there is no extra cost to calculating the optimal action from scratch. Scharfstein & Stein⁴² suggest a model with two managers (who do not compete in the same product market), each of whom may have private infor-

mation about the desirability of choosing action X or action Y. One manager chooses her action first, and the other then chooses after taking account of the first decision. Managers are either “smart” (i.e., they have genuine information about which strategy is better) or “dumb” (i.e., they know nothing), and managers do not know which type they are when they decide on their action. Since their private information is correlated, a pair of smart managers tends to choose the same action. Therefore, if a manager’s subsequent career depends on whether or not she is perceived to be smart, the second manager has an incentive to imitate to the first so as to be considered smart, even if that means going against her

CAREER CONCERNS OF MANAGERS
MIGHT GIVE A REASON WHY A
MANAGER MIMICS THE ACTION OF
A PEER, EVEN IF THERE IS NO
EXTRA COST TO CALCULATING THE
OPTIMAL ACTION FROM SCRATCH.

own private information. As the authors put it: “an unprofitable decision is not as bad for reputation when others make the same mistake—they can share the blame if there are systematically unpredictable shocks.”⁴³ This managerial herding is often socially inefficient.⁴⁴

However, imitation can also be socially useful. A model which several researchers have investigated supposes that each firm usually imitates its best-performing peer, but with some (perhaps small) probability it “experiments” and chooses a random strategy. A firm—if it has access to the necessary information—can observe the actions and profits of the other firms, and when it next has the opportunity to change its own action it can choose the action of the most profitable firm. In a stable and symmetric environment, this process of imitation will likely lead over time to approximately optimal behavior.⁴⁵

B. IMITATION BY COMPETING FIRMS

Additional effects come into play when firms are rivals operating in the same market, so that firms compete as well as observe.⁴⁶ In practice, oligopolists may have little idea of the consumer demand function, or how closely substitutable their rivals’ products are with their own. Nevertheless, they may observe their rivals’ actions and realized profits. When oligopolists imitate the most profitable actions observed in the market, it is possible that the market moves over time to a highly competitive outcome (more competitive than the one-shot Nash equilibrium).

To see this, consider a Cournot market where several firms with identical constant marginal costs compete to supply a homogenous product. Suppose, whenever they have the opportunity to change their output, that firms imitate the output decision of the most profitable firm of the previous period. Then, when the market price is above cost the most profitable firm will be the one with the largest output. Hence, firms with low output will increase their output, imitating the profitable firm, which pushes the price down. (If price is below cost, the most profitable firm is the one with the smallest output, and so imitation will then drive prices up.) Thus, imitation pushes prices towards cost and the market

evolves towards the perfectly competitive outcome where price equals marginal cost. In sum, when firms myopically imitate the most profitable strategy, the industry as a whole moves to an unprofitable, highly competitive outcome. See Vega-Redondo⁴⁷ for more details of this model, as well as Schaffer⁴⁸ for a related model.⁴⁹

Many markets are better modeled as firms choosing prices rather than quantities, and where firms offer differentiated products. In these markets, imitation can also induce firms to compete aggressively, although not to the extent of perfect competition as seen in the (homogeneous product) Cournot case. To illustrate, consider a duopoly with firms labeled 1 and 2, where if the two prices are p_1 and p_2 then firm 1's demand is

$$q_1 = 1 + p_2/2 - p_1 \quad (1)$$

(and similarly for firm 2). For simplicity, suppose costs are normalized to zero. Then profit-maximizing firms are predicted to choose the Bertrand equilibrium prices $p_1 = p_2 = 2/3$, and each firm sells quantity $2/3$. One can check that when both prices are above $1/2$ then it is the lower-price firm which makes greater profit. (When both prices are below $1/2$ then the higher-price firm has more profit.) Therefore, when the less profitable firm chooses the price of the more profitable firm (and firms experiment every so often as discussed in footnote 45), prices will settle over time at $p_1 = p_2 = 1/2$, so that the mark-up on cost falls by one-third with imitative behavior relative to profit-maximizing behavior.⁵⁰

An interesting corollary of imitative behavior is that with symmetric firms the familiar distinction between price-setting and quantity-setting behavior vanishes, and the long-run performance of markets with imitative firms does not depend on whether firms choose to compete in prices or in quantities. With profit-maximizing behavior, when firms compete in quantities the outcome differs—typically it is less competitive—from when they compete in prices. (For instance, in the linear demand example of the previous paragraph, when firms compete in quantities one can show that the equilibrium involves each firm setting the output $3/5$, which induces each firm to set the price $4/5$, which is higher than when the firms compete in prices.) However, when firms choose quantities the process of imitation converges to the situation where each firm chooses output $3/4$, which induces each firm to set the price $1/2$, which is exactly the same as when firms chose prices.⁵¹

AN INTERESTING COROLLARY OF IMITATIVE BEHAVIOR IS THAT WITH SYMMETRIC FIRMS THE FAMILIAR DISTINCTION BETWEEN PRICE-SETTING AND QUANTITY-SETTING BEHAVIOR VANISHES.

As discussed earlier in the context of non-rivalrous firms, it is natural to consider situations where profit-maximizing firms compete against naïve imitators to see which type of behavior performs better. Schipper⁵² investigates a Cournot model where “imitators” compete against “optimizers,” where the latter firms are

not really forward-looking profit-maximizers but rather myopic, and choose their output in one period as the most profitable response to the previous period's outputs. He shows that imitators make greater long-run profits than optimizers. A sterner test of the benefits of naïve imitation is when such a firm competes against firms who are forward-looking strategic players. It turns out that in a wide class of games, firms which imitate the best past performers do "essentially" as well as profit-maximizing rivals. For instance, consider a repeated Prisoner's Dilemma in which one firm plays the most profitable actions used in the previous period (starting initially by playing "collude", say). Then, unless there is very severe discounting of the future, the profit-maximizing way to play against this imitator is to collude in all periods except the last (if there is a last period), when it is optimal to compete. The profit-maximizing firm therefore obtains identical profits to the imitator, except just for one period (if there is a last period) when it makes higher profits.⁵³

IT TURNS OUT THAT IN A WIDE CLASS OF GAMES, FIRMS WHICH IMITATE THE BEST PAST PERFORMERS DO "ESSENTIALLY" AS WELL AS PROFIT-MAXIMIZING RIVALS.

Of course, firms can only imitate the most successful firm when they are provided with the necessary market information (namely, the profits of each firm, together with individual outputs in the Cournot case). But as soon as this information is present, a process as described by Vega-Redondo can make markets very competitive. As discussed in section II, Vega-Redondo's paper has inspired several experimental tests,⁵⁴ all of which found some support for the basic qualitative prediction of the imitation model. Even when firms have enough information to calculate the Cournot equilibrium, say, if they are then given the extra firm-specific information which allows them to imitate the best, this makes their behavior more competitive. However, the perfectly competitive outcome was not always achieved, indicating that not all subjects were following the "imitate the best" strategy. (An exception is, Huck et al.,⁵⁵ who in their treatment denoted "IMIT+" where firms did not know much about the functional forms of market demand or costs but did observe individual rival quantities and profits, found that the outcome was statistically indistinguishable from the perfectly competitive outcome.) Offerman, et al.,⁵⁶ suppose that a fraction of firms imitate, while the remainder are assumed to choose the most profitable output given the observed outputs of the previous period. In their data, the best fit for this mixed model is that two thirds of firms are imitators.

C. CONCERNS FOR RELATIVE PROFITS

There is a close connection between situations in which firms imitate the most profitable action played by their rivals and situations in which firms (or their managers) have as their objective the maximization of relative profits. The reason for this close connection is as follows. Consider the differentiated product duopoly just discussed, and suppose one firm reduces its price to undercut its

rival. If the lower-price firm makes greater profit than its rival, this means that the price cut harms it less than its rival, i.e., the difference between its profit and its rival's rises. In the specific linear demand example in expression (1) above, if firm 1 wishes to maximize its relative profit, it chooses price p_1 to maximize the profit difference

$$p_1(1 + p_2/2 - p_1) - p_2(1 + p_1/2 - p_2),$$

and it is therefore optimal for it to set the price $p_1 = 1/2$, regardless of the rival's price p_2 . If both firms care purely about their relative profit, they will both set the price $1/2$, just as eventually occurred when firms imitate. Thus, the long-run outcome in markets in which firms imitate coincides with the (one-shot) outcome in markets in which firms are fully rational except that their objective is relative rather than absolute profit.

The importance of this result is that, while purely imitative behavior on the part of sophisticated firms perhaps seems unlikely in many markets, the assumption that firms care about their relative position in a market has strong intuitive appeal.⁵⁷ First, there is by now no doubt that many individuals are strongly driven by relative pay. This is evident from the behavioral and experimental economics literature, as well as from the emerging happiness literature.⁵⁸ Indeed, CEOs as a group may have a greater proportion of "rivalrous" people than the population as whole, and such people may put particular weight on their relative standing. Secondly, managers often have placed on them (either explicitly or implicitly) incentives which induce them to care at least in part about relative as well as absolute performance. (One reason for this might be to insure managers against common shocks impacting the market as a whole.) For instance, Gibbons & Murphy⁵⁹ document empirically how a CEO's pay rise and likelihood of retention depend positively on the firm's performance and negatively on the overall industry performance. Vickers⁶⁰ shows that a firm in a Cournot market can improve its equilibrium profits (both in absolute and relative terms) when competing against profit-maximizing firms by inducing its managers to care about relative rather than absolute profit.⁶¹ For the same reason, if potential managers differ in their intrinsic preferences for absolute or relative profit performance, Miller & Pazgal⁶² argue that a firm may wish to hire a manager with a known behavioral bias towards relative profit so as to gain strategic advantage. By contrast, in Bertrand markets rather than Cournot markets, a firm's profits typically fall if its manager cares about relative instead of absolute profits; a manager who cares about relative profits will set a low price, which in turn will induce low prices from its rivals, and each firms' profits decline.⁶³

THE IMPORTANCE OF THIS RESULT IS THAT, WHILE PURELY IMITATIVE BEHAVIOR ON THE PART OF SOPHISTICATED FIRMS PERHAPS SEEMS UNLIKELY IN MANY MARKETS, THE ASSUMPTION THAT FIRMS CARE ABOUT THEIR RELATIVE POSITION IN A MARKET HAS STRONG INTUITIVE APPEAL.

So far, we have argued that imitation or concern for relative profits makes the static interaction among oligopolists more competitive, but what does this behavior imply for the sustainability of collusion? Suppose that each oligopolist cares about its profits relative to its rivals. Compared to the setting with absolute profit maximization, there are two effects:

- the static Nash equilibrium—which is used to punish firms if they deviate from the collusive agreement—is more competitive, but
- the short-run benefits to deviating are larger, since by undercutting its rivals a firm not only boosts its own profits but also harms its rivals.

SO FAR, WE HAVE ARGUED THAT IMITATION OR CONCERN FOR RELATIVE PROFITS MAKES THE STATIC INTERACTION AMONG OLIGOPOLISTS MORE COMPETITIVE, BUT WHAT DOES THIS BEHAVIOR IMPLY FOR THE SUSTAINABILITY OF COLLUSION?

The first effect makes collusion easier, while the second makes it harder, and *a priori* it is unclear which dominates. However, in a Cournot setting, Lundgren⁶⁴ shows that when firms care purely about relative profit, collusion is impossible to sustain.

We can summarize this discussion as follows. In markets where firms offer substitute products (i.e., in which firms compete), the market is made more competitive relative to the textbook situation of firms maximizing profits if firms either (i) imitate the previous most profitable strategies or (ii) aim to maximize their relative profits. The (eventual) outcomes under scenarios (i) and (ii) coincide, and in both cases the outcome is the same whether firms compete in prices or quantities. In mixed settings, where imitators (or relative profit maximizers) compete against profit-maximizing firms, the former will often perform in absolute terms at least as well as the latter.

IV. Vengeful Behavior and *Esprit de Corps*

The previous section discussed situations in which firms and their managers care about relative rather than absolute performance. Another kind of social preferences is present when firms care when their rivals obtain an “unfair” share of industry profits, for instance by cheating on a collusive agreement.

It is clear that many people are willing to incur costs in order to harm others who are perceived to have behaved unfairly towards them, a phenomenon which could be called vengeance or spite. Among the most famous and robust experiments in economics are those that study the “ultimatum game.”⁶⁵ Here, two players must share some specified prize, and one player (the “proposer”) suggests a way to share the prize between the two players. If the second player (the “responder”) agrees with the proposed shares, the prize is divided accordingly, while if the responder does not agree then neither player gets anything. If players are interested in obtaining as much of the prize as possible, the predicted outcome

(from a one-shot interaction) is that the proposer offers only a tiny share to the responder, who accepts since she prefers a small positive payoff to nothing. However, it is commonly observed that the responder will reject offers she finds too small, e.g., less than 25 percent of the prize. Because of this, proposers learn or anticipate that it is too risky to make such unfair offers, and outcomes where the proposer offers the responder 40 or 50 percent are often the norm.⁶⁶ These experiments have been performed (in poorer countries) when the prize is very significant in terms of monthly salary, and the results are similar.

The same effect is observed in markets. Huck, Muller, & Normann⁶⁷ show that a firm with a theoretical commitment advantage (a Stackelberg leader in a Cournot market) finds it hard to exploit that advantage in experimental markets. The reason is that the theoretically disadvantaged firm (the Stackelberg follower) acts more aggressively than predicted by the subgame perfect equilibrium of these market games. In fact, followers appear to punish the leader when the leader supplies a quantity above the symmetric Cournot quantity. (They punish the leader by themselves supplying a higher quantity than their most profitable response to the leader's quantity.) This behavior is in line with the vengeful behavior seen in the ultimatum game when the proposer tries to exploit his first-mover advantage.

Similar behavioral effects are observed in experiments that study strategic delegation. As already mentioned, Vickers⁶⁸ showed theoretically that firm owners might want to employ managers simply for strategic reasons. By writing an appropriate incentive contract for the manager, the owner can in effect commit his firm to Stackelberg-like aggressive behavior. However, if all firms do this a highly competitive outcome results: All firms are worse off than when managers care solely about profit. Huck, Muller, & Normann⁶⁹ tested this theory in an experiment, and do not confirm its predictions. If an owner does offer his manager the aggressive contract (and other owners do not), then managers in the weaker position are not content with simply accepting their "equilibrium fate" which would give them a lower salary than their opponent. Rather they behave more aggressively, in line with the observed behavior of the Stackelberg followers mentioned in the previous paragraph. This could mean that attempts to hire "behavioral" types of managers, as in Miller & Pazgal,⁷⁰ may also backfire.

A NATURAL QUESTION IS
WHETHER VENGEANCE MIGHT
HELP TO SUSTAIN COLLUSION

A natural question is whether vengeance might help to sustain collusion. For instance, if the collusive arrangement (a high price and equal market shares) is considered to represent the "fair" outcome, then if one firm reneges on the agreement and undercuts the price, its rivals may be offended and hence punish the deviator especially aggressively (even at extra cost to themselves). In the previous section we argued that concern for relative profits did not help sustain collusion, since it induced an extra benefit to deviating. But vengeance is subtly different from a concern for relative profit, since it only comes into effect when

someone is treated unfairly. (There is no extra benefit from harming your rival, unless he has first harmed you.) Thus, because vengeance induces more aggressive punishments for deviating, it could help sustain collusion relative to situations in which firms care only about their own profits.⁷¹

While the previous section showed how concerns for relative performance could increase competitiveness in a market (thus benefiting consumers), these examples show how vengeful behavior can reduce competitiveness (to the detriment of consumers). In settings where firms are symmetrically positioned, concerns for relative profits increase competition. On the other hand, in environments where markets are predicted to be particularly competitive due to strategic asymmetries (as in the Stackelberg model), vengeance against unfairly advantaged rivals tends to mitigate these advantages, rendering market outcomes more symmetric and less consumer-friendly.

The impact of vengeful behavior is less likely to be present when agents are very asymmetric or if some are perceived to be entitled to a bigger share of the surplus.⁷² Thus, the potential for spiteful behavior may affect conduct between rival firms, but

WHILE VENGEFUL BEHAVIOR MAY
SOMETIMES BE A WAY TO SUSTAIN
COLLUSION, AN ALTERNATIVE
METHOD IS TO FOSTER A SENSE OF
LOYALTY, OF *ESPRIT DE CORPS*,
AMONGST THE CONSPIRATORS.

not so much between a firm and its consumers, for instance, or between a large supermarket and a small supplier. Consider modifying the ultimatum game so that there are many responders. (More precisely, the single proposer offers a share, and the various responders simultaneously decide whether to accept. If some respon-

ders accept, one is picked at random and given her offered share; if none accepts then no one gets anything.) Then orthodox theory predicts—and experiments confirm—that the single proposer will be able to offer the responders very little and still find a willing responder. In a sense, competition forces the responders to act as if they were purely self-interested, and so “vengeance” is a rent which competition dissipates.⁷³

While vengeful behavior may sometimes be a way to sustain collusion, an alternative method is to foster a sense of loyalty, of *esprit de corps*, amongst the conspirators. Although information sharing is an important ingredient for collusive schemes to work, a useful by-product of regular meetings is that loyalty and friendship may be inculcated, which may make it socially costly to cheat on agreements. Moreover, since illegal cartels cannot enforce agreements with legally binding contracts, trust plays a central role in their operation.⁷⁴ (Recall from section II that face-to-face communication seemed to foster collusion in the laboratory more effectively than computer-mediated communication.) An important role for antitrust, via leniency programs and the like, is indeed to foster distrust among conspirators.

A clear example of an attempt to run a cartel (or cartel-like operation) by means of encouraging a sense of camaraderie is the U.S. steel industry during the years 1907-1911, in which cooperation was fostered through a series of social events and meetings which have become known as the “Gary Dinners” after the chairman of U.S. Steel, Judge Elbert Gary. It is worth quoting one of Gary’s speeches at length:⁷⁵

“[W]e have something better to guide and control us in our business methods than a contract which depends upon written or verbal promises with a penalty attached. We as men, as gentlemen, as friends, as neighbors, having been in close communication and contact during the last few years, have reached a point where we entertain for one another respect and affectionate regard. We have reached a position so high in our lines of activity that we are bound to protect one another; and when a man reaches a position where his honor is at stake, where even more than life itself is concerned, where he cannot act or fail to act except with a distinct and clear understanding that his honor is involved, then he has reached a position that is more binding on him than any written or verbal contract.”

In essence, if conspirators can find a way to increase the social or psychological cost of cheating, collusion will be more readily achieved.

Further historic evidence of the impact of social preferences on anticompetitive behavior comes from the UK shipping cartels in the period 1879 to 1929. Podolny & Scott Morton⁷⁶ document how established cartels behaved significantly more aggressively toward entrants who had low social status (or who were foreign), relative to the situation where an entrant had a knighthood, say. (The study controls for the correlation between social status and “deep pockets,” which may make predation less profitable.) The authors argue that the likely reason is not so much pure snobbishness on the part of the cartel (who typically comprised high status individuals), but rather that social status was used as a signal of the entrant’s likelihood of cooperating within the cartel, and the expected “transactions costs” of having the entrant as a member of the cartel, if admitted.

V. Satisfactory, Not Maximum, Profits

Rather than denying the importance of (absolute) profits, one might instead question whether firms really *maximize* profits. For reasons of complexity, ignorance, or the “easy life,” firms might instead engage in satisficing behavior to secure a target level (or “aspiration level”) of profit. In its starkest form, the util-

ity of a firm then has just two values: good enough, and not good enough. This idea goes back at least to Rothschild, Gordon, and Simon.⁷⁷ Just as a consumer might not change her bank, say, until the perceived level of service quality she receives falls below some threshold, a firm might not revise its strategy while it continues to obtain satisfactory profits. Only if the realized profit falls below the target level will a firm resort to experimenting with an alternative strategy. Such behavior could give rise to a degree of price rigidity for instance, even if underlying cost or demand conditions vary.⁷⁸ Experimentation could be purely random in extreme cases, or firms may be able to target their attention towards new strategies which will likely boost profits. (For instance, an idea in the early literature on satisficing was that firms often had a degree of “managerial slack,” and when times were hard managers could focus their energies on cutting costs or expanding their markets.) What level of profits a firm considers to be “satisfactory” is likely to depend on its historical returns as well as the performance of its peers and the economy as a whole.

Cyert & March⁷⁹ discuss some implications of satisficing behavior and test these ideas empirically analyzing data from (a few) manufacturers of farm implements. They conclude that a firm is more likely to resort to aggressive strategies (seeking cost reductions or sales expansions) if it is operating close to its

RATHER THAN DENYING THE IMPORTANCE OF (ABSOLUTE) PROFITS, ONE MIGHT INSTEAD QUESTION WHETHER FIRMS REALLY MAXIMIZE PROFITS.

breakeven point or if its costs are higher than the industry average; behavior which accords with the idea that firms are most likely to change strategy after they realize low profits.

Huck, Konrad, Muller, & Normann⁸⁰ document related effects in an experiment concerning mergers in Cournot markets. As shown by

Salant, Switzer, & Reynolds,⁸¹ bilateral mergers in linear Cournot markets with more than two firms are predicted to reduce the joint profits of merging firms when firms are profit-maximizers. Since this prediction appears counterintuitive at first sight, the result has been dubbed the “merger paradox.” However, the intuition for the result is simple. In Cournot markets firms’ outputs exert negative externalities on their rivals. Hence, after a merger the two insiders have an incentive to internalize some of these negative externalities; that is, they have an incentive to reduce their combined output. (Obviously, they would keep their pre-merger profits if they simply kept their output constant, but this is not the most profitable response to the un-merged firms’ outputs.) Further, as Cournot interactions are games with strategic substitutes, the un-merged firms’ optimal responses to this contraction in output are to increase their own outputs (in response to which the merged firm will reduce its output even further, and so on). This leads into a new equilibrium where the market price is higher, the outsiders are better off, and the two merging insiders are worse off. The reduced market share of the merging firms (from $2/n$ to $1/(n-1)$ if there are n firms initially) outweighs the benefits from increased market concentration.⁸²

Huck, et al. test this prediction in laboratory markets with initially three or four firms. Mergers are implemented exogenously after 25 periods in which firms learn to play the Cournot equilibrium. Firms are randomly selected to merge and then play for a further 25 periods. The merged firm is managed by one of the two subjects playing the role of the two merging firms. Profits are shared equally between the manager-owner and the passive owner. After the merger, profit-maximizing firms are predicted to behave as if one of the firms had vanished, and in particular, a merged firm is predicted to choose the same output as the non-merged firms. However, a different behavior emerges in the laboratory; the merged firm systematically supplies more than the outsiders. (The outsiders are seen to respond approximately optimally to the behavior of the merged firm.) This implies that the losses from the merger are smaller than predicted. In fact, when there are four firms initially present the merged firm does not lose at all. Rather, it experiences a significant increase in profits for a few periods and then, when the outsiders are fully adjusted, its per-period profits fall slightly compared to their pre-merger levels.

Huck, et al. discuss a number of reasons for this effect, and by conducting suitable control experiments they suggest something akin to satisficing behavior explains observed behavior.⁸³ After two firms merge, the firms perceive the danger that their joint profits fall, and pursue aggressive strategies to avoid this (as the firms in Cyert & March did). The result is that the merged firm's output is significantly higher than its rivals' outputs, even though post-merger a merged firm and an unmerged firm are in a symmetric position. (However, the merged firm's output is lower than their combined outputs in the pre-merger phase, and so their strategies are not entirely inert.)

Dixon and Oechssler discuss an interesting application of satisficing behavior.⁸⁴ They consider a number of oligopolies (rather than just one), and suppose that firms follow the following rule of thumb: "If my profits are no lower than the average profits observed across all markets, I keep my strategy unchanged; if my profits are lower than average profits across all markets, I experiment and choose a random strategy." Thus, firms use the average payoff of all firms as their aspiration level. Such a policy requires information about average profitability, but no information about individual firm actions or profits or about consumer demand. It turns out that when firms behave in this manner (with a small amount of noise added) then behavior in each market eventually becomes collusive. Consequently, practices that look innocent from the orthodox point of view—myopic adjustments towards better strategies—may lead to undesirable outcomes. There can be collusive effect without any collusive intent.⁸⁵

To understand this surprising result, consider for simplicity a pair of identical duopolies (i.e., there are four firms in all), which interact over time in a Prisoner's Dilemma manner using the two possible actions of "compete" or "collude." Suppose that all firms are initially colluding, in which case all firms are satisfied with their profits. Next, imagine one firm "trembles" and changes its

action to “compete.” A duopoly in which both firms collude has higher profits than a duopoly where one firm colludes and the other competes, and so firms in the colluding duopoly remain satisfied with their profits and do not experiment. The only firm dissatisfied with its profit is the colluding firm playing against the competing firm, which will experiment and end up playing “compete.” We then have two firms in a market playing “compete,” both of whom will be dissatisfied with their profits, and will experiment and end up both playing “collude.” Thus all firms colluding is the stable steady state of this process (unlike all firms competing, where one “tremble” will eventually induce all firms to collude). In contrast to Adam Smith’s invisible hand, which guides profit-maximizing firms

towards outcomes which benefit consumers, with this form of satisficing behavior, a second invisible hand guides firms toward outcomes which exploit consumers.

IN CONTRAST TO ADAM SMITH’S
INVISIBLE HAND, WHICH GUIDES
PROFIT-MAXIMIZING FIRMS
TOWARDS OUTCOMES WHICH
BENEFIT CONSUMERS, WITH THIS
FORM OF SATISFICING BEHAVIOR, A
SECOND INVISIBLE HAND GUIDES
FIRMS TOWARD OUTCOMES
WHICH EXPLOIT CONSUMERS.

A concept related to satisficing behavior is “approximately” optimal behavior. Firms may not find it worthwhile to calculate their optimal strategy precisely, but instead to cease their search for a good strategy when they get to within some tolerance of the optimal strategy. In technical terms, the outcome of approximately optimal behavior is termed an “ ϵ -equi-

librium,” where $\epsilon > 0$ is the (perhaps small) tolerance which firms have for short-fall in optimal profits. An important insight is that even a small tolerance by firms for sub-optimal strategies—that is, a small departure from fully optimizing behavior—can result in significant departures from the outcomes corresponding to optimizing behavior. (In the neighborhood of a firm’s optimal strategy its profits are roughly flat, and so a moderate change in its strategy might have little impact on its own profit, and yet could have a significant impact on its rivals’ profits and strategies.)

To illustrate, consider two symmetric firms supplying a homogeneous product engaged in quantity competition. To be concrete, suppose that the consumer demand curve is $q(p) = 1 - p$ and that production is costless. Then the symmetric collusive outcome in this market involves each firm supplying quantity $1/4$ inducing the monopoly price $p = 1/2$. How small can a firm’s tolerance for sub-optimal behavior be for this monopoly outcome to be an ϵ -equilibrium? The most profitable response to the rival’s monopoly output of $1/4$ is to supply output $3/8$ rather than $1/4$. However, its gain in profit from pursuing the optimal strategy rather than the satisficing strategy of supplying output $1/4$ is only $1/64$. Thus, the collusive outcome (where each firm supplies quantity $1/4$) is an ϵ -equilibrium provided that $\epsilon > 1/64$. Note that $1/64$ is about 6 percent of a firm’s share of the monopoly profit, and so if firms are prepared to optimize to within 6 percent of their exact optimal profits, the monopoly outcome can be sustained without

any collusive intent.⁸⁶ Of course, though, a problem with the notion of ε -equilibrium is its lack of predictability: firms could also accidentally play more competitively than the Cournot duopoly equilibrium.

Baye & Morgan⁸⁷ present an interesting account of how almost-optimal behavior can lead to significant departures from the predictions of fully-optimal models. They analyze a static homogenous product Bertrand market, so that profit-maximizing firms are predicted to set prices at marginal cost. They obtain an ε -equilibrium involving mixed strategies for choosing prices that yields relatively high profits even for small ε . (With duopoly, when ε is just 1 percent of monopoly profit, their chosen ε -equilibrium yields expected profit which is more than 25 percent of the monopoly profit.)

Baye & Morgan also discuss an alternative model of bounded rationality, which is that firms play a so-called quantal-response equilibrium.⁸⁸ This model supposes that a firm is more likely to choose a price which yields higher profit, but is not sure to choose the most profitable price. (This model nests the fully rational model and a model with purely random behavior as special cases.) While the details of the equilibria in the two models of bounded rationality differ, the broad conclusions—that prices are above cost, profits are positive, and profits fall with more rivals—coincide. Baye & Morgan run laboratory experiments to generate data, and estimate which model of firm behavior best fits the data. Especially for duopoly, they find that both the ε -equilibrium and the quantal-response equilibrium model fit the data better than either of the extreme fully-rational or fully-random models.⁸⁹

VI. Other Topics

A. OVER-OPTIMISM

As is well-documented in the psychology literature, over-optimism (or over-confidence) about one's own ability or about the probability of favorable outcomes is apparently common in the population.⁹⁰ Adam Smith⁹¹ wrote: "the chance of gain is by every man more or less over-valued, and the chance of loss is by most men under-valued." There are good reasons for thinking over-optimism is still more common among entrepreneurs than the population as whole. There is a "winner's curse" aspect to launching a new business: It is likely that others have already thought about launching a similar product (be it a new restaurant in a particular locale, or something more ambitious), and the entrepreneur who actually decides to start the new business is likely to be more optimistic than others. As a result, even if beliefs about the likely return from the investment are unbiased on average, the entrepreneur will typically be overly optimistic. Unless entrepreneurs rationally take full account of the fact that others have decided not to enter this market, we expect to see: (i) high failure rates for new businesses, and (ii) credit rationing, or loans being offered only with collateral.⁹²

Internal promotion procedures may also have a tendency to favor the over-optimistic, so that CEOs as well as entrepreneurs may be disproportionately over-optimistic. Consider a situation where management promotions are driven through rank-order tournaments, as in Lazear & Rosen.⁹³ To be promoted to a higher level, managers tend to require both skill and luck. In Lazear & Rosen's original model luck was exogenous noise, but the model can be extended so that managers can affect the riskiness of projects they undertake. The consequence is that the best performing manager in a tournament is likely to have high skill and to have chosen a risky set of projects. Over-optimism in a manager might mean that that manager downplays the true riskiness of projects. In this setting, managers who rise to the top of a firm are likely to be the highly-skilled optimists who were lucky.⁹⁴

Using a more abstract selection mechanism, Heifetz, Shannon, & Spiegel⁹⁵ also argue that optimists will systematically outperform realists in competitive environments, and so will predominate in the pool of successful agents. The key intuition is that optimism can serve as a commitment device if observed by rivals (much as making a manager maximize relative profits does in Vickers).⁹⁶ Translated into a market context, Heifetz et al.'s analysis suggests that particular forms of managerial over-optimism may sometimes act to soften competition.

For instance, consider a Bertrand oligopoly with differentiated products where there is some uncertainty about demand. The more optimistic a manager is—say, about the scale of demand, or the extent of product differentiation among suppliers—the higher the price she will charge. If a manager's optimism is observable or if competitors at least have some informative signal about her degree of

optimism, they will rationally anticipate these higher prices and optimally adjust their own prices upwards as well. Hence, one optimist is enough to increase the prices charged by all firms. And crucially the optimist will earn more money for her firm than the realist would have done (although her rivals will earn still more).⁹⁷

IN WHAT IS PERHAPS THE BEST-
KNOWN PAPER ABOUT MANAGERIAL
IRRATIONALITY, ROLL SUGGESTS
THAT EXCESSIVE MERGER
ACTIVITY MAY BE GENERATED BY
MANAGERIAL OVER-CONFIDENCE.

In what is perhaps the best-known paper about managerial irrationality, Roll⁹⁸ suggests that excessive merger activity may be generated by managerial over-confidence. In his theory, a CEO is too confident in the accuracy of his information about the potential profitability of a takeover, and so will be too quick to launch a takeover bid. Predictions of this theory are that the combined gain to bidder and target will be close to zero, and that the bidder's value will fall on announcement of the bid. Roll suggests that the available empirical evidence does not reject these predictions. Over-confidence by CEOs in their ability to choose investment projects and merger opportunities has been detected by Malmendier & Tate.⁹⁹ CEOs will differ in the extent of their overconfidence, and Malmendier & Tate identify as over-confident those CEOs who hold on to their stock options

until they expire. They interpret this as evidence of the CEO's over-estimation of their firm's future performance. They show that this group of CEOs is 55 per cent more likely to undertake a merger.¹⁰⁰ Managerial over-confidence could explain why companies that undertake mergers seem to under-perform.

Another form of potential over-confidence concerns a manager's faith that he will not be caught if he pursues an illegal activity such as price-fixing.¹⁰¹ If many cartelists downplay the risk of detection, then policy measures to increase the penalties for illegal cartel activity could be less effective than would be predicted by an approach based on rational profit-maximizing behavior.

ANOTHER FORM OF POTENTIAL
OVER-CONFIDENCE CONCERNS
A MANAGER'S FAITH THAT HE
WILL NOT BE CAUGHT IF
HE PURSUES AN ILLEGAL
ACTIVITY SUCH AS PRICE-FIXING.

There is also a rich literature in finance about whether overconfident securities traders can survive in the long term or whether, *à la* Friedman, they are driven out by more rational traders. If an overconfident trader underestimates the riskiness of an asset, she will buy more of the asset than rational traders. If the risky asset also has higher expected return, then the over-confident traders can become wealthier than more rational investors (although their expected utility is lower).¹⁰² If an over-confident trader believes that her estimate of the expected return of a risky asset is more precise than it really is, such a trader will trade more aggressively than an unbiased trader, and since she trades based on useful information, her expected wealth is higher than unbiased traders (although her expected utility is lower). As a result, over-confident traders can persist in, and even dominate, the market.¹⁰³ In such cases, Darwinian selection actually selects the biased traders. Kyle & Wang¹⁰⁴ present a model with just two traders, one of whom is over-confident and known to be so by the rational trader. They show, in a very similar manner to our previous discussion about Heifetz, et al., that the over-confident trader gains strategic advantage by being known to trade aggressively, which induces the rational trader to scale back her own trades. The result is that the overconfident trader performs better than the rational trader.

B. ACCOUNTING ANOMALIES

It is a fundamental tenet of profit-maximizing behavior that fixed and sunk costs, while they are important for entry and exit decisions, should not play a role in the determination of prices to customers. For instance, competition authorities do not always put much weight on claimed synergies between merging firms which act to reduce fixed costs, since such synergies are not expected to feed through into lower prices. However, this policy seems to be widely flouted by managers. Long ago, Hall & Hitch¹⁰⁵ interviewed 38 business executives about their methods for setting prices. Instead of equating marginal revenue to marginal cost, the authors concluded that: "... there is a strong tendency among business men to fix prices directly at a level which they regard as their 'full cost'." Al-

Najjar, Baliga, & Besanko¹⁰⁶ describe a number of more recent surveys that report similar attitudes among managers, the majority of whom claim to take fixed and sunk costs into account when setting prices. (They also report a managerial accounting textbook which argues against basing prices on marginal costs.)

Experimental work confirms that supposedly irrelevant sunk costs can have an impact on how prices are actually chosen. Offerman & Potters¹⁰⁷ conduct an experiment to investigate pricing in a Bertrand-type duopoly market with product differentiation. In one treatment, there are no sunk costs and pricing behavior is observed to converge to the Bertrand equilibrium. In a second treatment, participants must pay a sunk entry fee to join the market. (This is designed to model auctioning a license for the right to enter a market.) In this second treatment, the average mark-up of prices over marginal cost is substantially higher than in the first treatment. Interestingly, when they perform the same pair of experiments but with monopoly instead of oligopoly, there is no impact of sunk costs on the chosen monopoly price (which is observed to be close to the profit-maximizing level). This is somewhat reminiscent of our discussion of communi-

cation costs in section II, where we saw that firms were more inclined to stick to a non-binding collusive agreement if they incurred costs to reach that agreement.

EXPERIMENTAL WORK CONFIRMS
THAT SUPPOSEDLY IRRELEVANT
SUNK COSTS CAN HAVE
AN IMPACT ON HOW PRICES
ARE ACTUALLY CHOSEN.

This observation is consistent with Al-Najjar, et al., who present a theoretical model to show how the use of full-cost pricing policies might persist in the long-run in oligopoly markets. The main result is by now a familiar one in this survey: By introducing a behavioral bias in managerial decision-making, a firm can gain strategic advantage. In more detail, Al-Najjar et al. suppose that firms compete in a Bertrand market with product differentiation. In such a market, if a manager somehow commits to set a high price, its rivals will also set a high price, and all firms will make higher profits.¹⁰⁸ (The effect is akin to hiring an over-optimistic manager or, as previously discussed, placing a suitable incentive scheme on the manager.) But a manager who bases prices in part on fixed and sunk costs effectively commits to set a high price, and so hiring a manager who practices this “naïve” pricing policy (or instilling a corporate culture where this form of pricing is used) boosts the firm’s profits.¹⁰⁹ The effect is akin to the strategic tax policy analyzed by Eaton & Grossman,¹¹⁰ where a country has an incentive to tax the output of a home firm in order to relax competition with a foreign rival. If, for whatever reason, firms take fixed costs into account when setting their prices, this has potentially important implications for merger policy when the merger is expected to generate only synergies in fixed costs. Such mergers would in fact lead to lower prices than would be predicted in a profit-maximizing world.

Other alleged accounting anomalies might be explained by similar myopic learning or evolutionary pressures, or reasons of strategic delegation. For

instance, a firm might be organized into separate profit centers, each of which is given the task of maximizing its own profit (despite the competitive or contracting externalities which might exist between these profit centers). Thus, a manufacturer might choose to supply its products through an independent retailer rather than sell directly to consumers via an integrated retailer. To see one reason why this might be so, despite the apparent dangers of double marginalization, suppose that two manufacturers are competing to sell to consumers. If they sell directly to consumers, then in the absence of collusion we expect to see the Bertrand equilibrium emerge. However, if each firm delegates its retailing operation to a separate division (and each retailing division sees the terms at which the rival retailer sources its product), then by setting a wholesale price above its production cost a firm can induce its retailing division to price high, which softens competition and boosts profits.¹¹¹

C. UNCERTAINTY ABOUT THE RATIONALITY OF RIVALS

Even if a firm is fully rational, its behavior will be affected when it believes that its rivals may not be rational. Because of this, a rational firm may have an incentive to mimic non-rational behavior so as to induce its rival to think it may be a non-rational type of firm. Consequently, even the potential for behavioral biases can have a substantial effect on market outcomes.

CONSEQUENTLY, EVEN THE
POTENTIAL FOR BEHAVIORAL
BIASES CAN HAVE A SUBSTANTIAL
EFFECT ON MARKET OUTCOMES.

As we have discussed already, if two rational firms (who know for sure that each other is rational) play a Prisoner's Dilemma game for a known finite number of rounds, with the two actions "collude" and "compete," there is no way to sustain collusion even at the start. But if one firm is unsure about the rationality of its rival, it may be unwise to "compete" in each round. For instance, suppose the rival is for some reason believed to be using the strategy of "tit-for-tat," i.e., this rival starts off colluding, and then, in each subsequent round, imitates the previous action of the rival. (Note that the "tit-for-tat" strategy cannot be a rational strategy for a profit-maximizing firm, since in the final round it is a dominant strategy to "compete" regardless of the rival's penultimate action.) If one firm thinks it is playing against a "tit-for-tat" firm, it will typically be optimal to collude for initial periods so as to induce the rival to continue with collusion. Only toward the endpoint will the rational firm start to compete. A tempting strategy is for one or both players to mimic the behavior of "tit-for-tat," so as to induce the rival to believe she is more likely to be this irrational type. This will induce collusive behavior for some rounds.¹¹² Thus, the (perhaps small) possibility that one or both firms are irrational can induce play that is better for both firms (and worse for consumers).¹¹³

Another way in which a rational firm might wish to mimic an irrational firm is in the context of predatory pricing, as analyzed in the model of Kreps & Wilson.¹¹⁴ If a monopoly incumbent faces a sequence of potential entrants to its

market, it may wish to establish a reputation for fighting entry whenever it occurs, even when fighting is actually more costly than accommodating entry. The predictions of the Kreps-Wilson model of predatory pricing are broadly (i) there is little entry in the early period, (ii) when there is entry in early periods, the incumbent “fights” even if it is a rational firm, and (iii) towards the endpoint there is more entry and less willingness to fight on the part of the rational incumbent. This model, which involves highly sophisticated reasoning on the part of firms, is tested in an experiment by Jung et al.¹¹⁵ The experiment found widespread predatory pricing—defined to be either no entry in the early stages, or a rational incumbent fighting entry if it does occur in the early stages—although the more detailed predictions of the Kreps-Wilson model did not fit the data. (For instance, experimental subjects did not enter more frequently if there had been no previous entry as compared to the situation where there had been fought entry, whereas the Kreps-Wilson model predicts there should be more entry in the former case.)

Rather than go to the trouble of mimicking the behavior of behavioral type of firm, which in any case may only deter entry for some of the time, a more straightforward method to deter entry could be simply to hire an “aggressive” manager, who is known by potential entrants intrinsically to wish to fight entry whenever it occurs. (For instance, an aggressive manager might be someone who is over-optimistic about the ease by which entry can be successfully fought, and so is willing to fight more often than an unbiased manager would.) The impact of this policy, assuming it is credible to keep this aggressive manager in place after entry occurs, is akin to hiring a “conservative” central banker who is less

THUS, IT CAN BE RATIONAL
TO BEHAVE IRRATIONALLY.

likely to be tempted to cause inflation when unemployment rises.¹¹⁶ As Schelling¹¹⁷ wrote: “the conspicuous delegation of authority to a military commander of known motivation, exemplifies a common means of making credible a response pattern that the original source of decision might have been thought to shrink from or to find profitless.” Thus, it can be rational to behave irrationally.¹¹⁸

VII. Discussion

This paper has discussed a number of reasons why firms might not pursue maximum profits:

1. In some complex, uncertain environments, the optimization problem is simply too hard, and firms must resort instead to satisficing and the use of rules of thumb. Decision shortcuts included imitating the actions of well-performing peers, satisficing, or being content to achieve profits to within “ ϵ ” of the maximum. These rules of thumb appear attractive since in many situations that require strategic sophistication, such as herding, reputation-building, or collusion among

many firms, the complex strategies needed to maximize profits are rarely observed in even in the most simplified laboratory settings and even if participants are highly experienced.

2. Alternatively, optimization might still occur, but with alternative aims or under mistaken beliefs. Thus, a manager might maximize her profits relative to those of her peers, or a manager could be over-optimistic about the profitability of some action. The reason why managers have aims different from maximizing profits could be due to selection effects (e.g., only “competitive” or over-optimistic people rise to become CEOs, or because firms which aim to maximize their relative standing actually obtain greater absolute profits than their profit-maximizing rivals), or because profit-maximizing principals choose to give their managers distorted incentives to gain strategic advantage.
3. Social preferences (other than caring about relative profit as above) may play a role, and a firm might punish a rival if that rival obtains an “unfair” share of profits. Alternatively, face-to-face communication between firms, or having a similar social background among firms, may generate solidarity among conspirators that makes it socially hard to cheat on collusive agreements.

IN SOME SITUATIONS, MARKETS
ARE MORE COMPETITIVE
WHEN FIRMS DO NOT AIM TO
MAXIMIZE THEIR PROFITS.

In some situations, markets are more competitive when firms do not aim to maximize their profits. For example, if firms in a market myopically imitate the action of the most profitable rival, then the market may paradoxically move towards a highly competitive outcome. Alternatively, we saw at least in laboratory settings that firms were often unable to achieve tacit collusion, despite this being an equilibrium option for profit-maximizing firms. Some of the more complex strategies that foster collusion in theory are perhaps too subtle to matter empirically.

In other situations, when firms cannot maximize profits, their realized profits are actually increased. For these firms, following Voltaire’s dictum, the best is the enemy of the good. This can be clearly seen in the case of static Bertrand competition with homogenous products. Here, the only equilibrium involves firms setting prices equal to marginal costs, leaving them with no profits. But if satisficing firms are content to choose actions which are only approximately optimal, they may all be able to enjoy substantial profits. Likewise, when satisficing firms alter their actions only when they under-perform relative to average performance, the result may be as if firms were successfully colluding.

If a firm adopts a non-standard objective, it may gain strategic advantage in the market since competitors’ behavior will change in a desirable way. As such, it could be rational to behave irrationally. For instance, in Cournot markets a firm which aims to maximize its relative profits may do better in equilibrium

than a profit-maximizing rival. Alternatively, a firm which chooses to base its price on “full costs” rather than marginal costs may do better than if it followed textbook profit-maximizing precepts. Or a firm led by an aggressive manager may deter entry more often than a profit-maximizing manager. These non-standard objectives could be put in place by far-sighted profit-maximizing shareholders (as emphasized in the strategic delegation literature) or they could arise myopically due to evolutionary selection of better performing managers and/or firms.¹¹⁹ Regardless of the reason why firms have non-profit objectives, the presence of non-profit objectives is likely to affect competitive interaction. For instance, a manager operating under an incentive scheme which rewards relative performance is likely to behave more competitively than orthodox analysis would indicate, and this is a factor which a competition authority should take into account.

FOR INSTANCE, A MANAGER
OPERATING UNDER AN INCENTIVE
SCHEME WHICH REWARDS RELATIVE
PERFORMANCE IS LIKELY TO BEHAVE
MORE COMPETITIVELY THAN
ORTHODOX ANALYSIS WOULD
INDICATE, AND THIS IS A FACTOR
WHICH A COMPETITION AUTHORITY
SHOULD TAKE INTO ACCOUNT.

(The effect is the reverse of the situation when there is cross-ownership in an oligopoly, which competition authorities already recognize leads to a blunting of competition.)

Thus we see there are several situations in which Friedman’s Darwinian critique of non-profit maximizing behavior appears to fail. Nevertheless, there are a number of situations in which market competition and market experience do seem to diminish those behavioral biases which do not confer evolutionary advantage.

For instance, competitive versions of the ultimatum game appear in experiments to conform to more orthodox models of selfish behavior. In addition, as shown in the work of List,¹²⁰ market experience can dampen the bias known as the endowment effect, where the valuation of a good increases when it is owned. Such a bias cannot easily improve own performance in markets with many traders (rather it hinders agents from making otherwise beneficial trades). Likewise, we are not aware of many situations in which procrastination, impulsive behavior, and other self-control problems play a major role in firm behavior (e.g., price fixing is not a “crime of passion”), although these behaviors are prominent in discussions of behavioral economics as applied to consumers.

We have seen a number of reasons why firms may not maximize their profits, and this potentially has implications for empirical studies of markets, including the use of merger simulation for competition policy. Empirical market studies typically assume profit-maximizing behavior on the part of firms to produce their estimates of, say, marginal costs. In such cases, the analysis may lead to biased estimates of the welfare impact of a merger if, in fact, the merging firms were not profit-maximizers. To illustrate, consider the merger situation in Huck et al.¹²¹ which we discussed in section V. If one took the data from this experiment pre- and post-merger and, in line with the structural approach, assumed these data were generated by profit-maximizing firms, one would conclude that the merger

must have induced substantial cost synergies. Only with reduced costs would it be possible for the merged firm to have higher outputs than the outsiders. This would affect estimates of the merger's welfare consequences. While the true consequences are unambiguously negative (after all, in the experiment the merged firm does operate with the same costs), sufficiently big synergies could offset the loss in consumer welfare. Hence, if there are systematic deviations from orthodox profit-maximizing assumptions, structural approaches that assume profit-maximization might detect increases in welfare when, in fact, welfare is reduced. Likewise, a merger simulation exercise may be inaccurate if the assumption is that firms maximize profits.

More fundamentally, much of competition policy is founded on an assumption of profit-maximizing behavior by firms. Some go so far as to say that the "entire antitrust enterprise is dedicated to the proposition that business firms behave rationally."¹²² Courts and regulators in some jurisdictions may not consider seriously conduct (such as predatory pricing, for instance) which does not appear to make "business sense" according to their judgment. Leslie¹²³ reports that "if a plaintiff's complaint describes a conspiracy that the judge concludes is irrational, then the court rules that the conspiracy must not have happened as a matter of law, regardless of the evidence presented by the plaintiff to support its claim." In the light of the theories and evidence reported in this article, we suggest that a dogmatic attitude towards the pervasiveness of business rationality may lead to instances where harmful behavior goes unpunished. Behavioral economics may sometimes pose at least as many questions as it answers, and complicate antitrust debate. But it also sheds light on important market phenomena, and if competition policy is to reflect market realities, behavioral economics cannot be ignored.

WE SUGGEST THAT A DOGMATIC
ATTITUDE TOWARDS THE
PERVASIVENESS OF BUSINESS
RATIONALITY MAY LEAD TO
INSTANCES WHERE HARMFUL
BEHAVIOR GOES UNPUNISHED.

Finally, while this article has surveyed behavioral economics as it applies to firms, in future it would be interesting to investigate how it applies also to policy-makers.¹²⁴ Competition authorities, like firms, operate within a complex and strategic environment, and may need to resort to rules-of-thumb and satisficing behavior. Instances of this could include an authority's use of *per se* rules, or a reliance on relatively rigid market definitions and market share thresholds. It may also be advantageous to induce competition authorities to have an objective or institutional focus which differs from social welfare, in order to alter the response from the firms subject to regulation.¹²⁵ Imitative strategies may sometimes be employed, and safety in numbers may be enjoyed, as policy-makers look around the world for current "best practice." (Indeed, the recent emphasis on behavioral economics may be an instance of this.) Friedman's point about competitive pressures may have less force in terms of constraining good decision-making by public officials, and the result may be that behavioral biases are more prevalent among policy-makers than in the firms they oversee. ▼

- 1 See Armstrong for a review of some of this literature and its implications for consumer protection policy, Mark Armstrong, *Interactions Between Competition and Consumer Policy*, COMPETITION POL'Y INT'L 4, 97-148 (2008). See Spiegler for an account of economic models which assume profit-maximizing behavior by firms but bounded rationality on the part of consumers, RAN SPIEGLER, BOUNDED RATIONALITY AND INDUSTRIAL ORGANIZATION (in press). As observed by Ellison, fifty years ago the focus was instead very much on non-optimizing behavior by firms rather than consumers, Glenn Ellison, *Bounded Rationality in Industrial Organization*, ADVANCES IN ECONOMICS AND ECONOMETRICS: THEORY AND APPLICATIONS, 145, (R. Blundell, W. Newey, & T. Persson, eds.) (2006).
- 2 The recent financial crisis may well have stemmed in part from a variety of behavioral biases of some of the banks' executives. For instance, some managers may have been over-optimistic about the risks they were taking in their lending strategies, and there may have been a herd mentality among some managers, who imitated apparently successful lending strategies and who may have felt there was "safety in numbers."
- 3 Hayek argued that a central merit of competitive markets, specifically price-taking behavior by consumers and firms, is that agents' strategies are then relatively simple. (Agents need to know only their endowments, preferences, and the market prices to optimize.) Frederick Hayek in *The Use of Knowledge in Society*, AMER. ECON. R. 35, 519-530 (1945), at §VI. Gale & Sabourian argue that in oligopolistic markets, where optimal strategies may be extremely complex, if agents incur "complexity costs" when they pursue complex strategies then the outcome might be more competitive than the standard theory suggests. Douglas Gale & Hamid Sabourian, *Complexity and Competition*, ECONOMETRICA 73, 739-769 (2005).
- 4 The take-up of technological improvements in agriculture provides useful evidence. Ellison & Fudenberg quote a historian of the English agricultural revolution as writing "land tilled in very ancient ways lay next to fields in which crop rotations were followed." They report that the agricultural practices known as the "new husbandry" diffused through England and France at the rate of just one mile per year. Glenn Ellison & Drew Fudenberg, *Rules of Thumb for Social Learning*, J. POL. ECON. 101, 612-643 (1993).
- 5 A recent pair of papers by Goldfarb & Yang and Goldfarb & Xiao illustrate this. For instance, it appears that older and better-educated managers tended to enter markets with fewer competitors after the U.S. telecommunications market was deregulated in 1996. Avi Goldfarb & Botao Yang, *Are All Managers Created Equal?*, J. MARKETING RES. 46, 612-622 (2009); Avi Goldfarb & Mo Xiao, *Who Thinks About the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets*, working paper University of Toronto (2010).
- 6 See Bertrand & Schoar for an empirical analysis of the importance of "manager fixed effects." For instance, they find that managers with an MBA tend to follow strategies that are more aggressive. They find that managers differ in their attitude to mergers, dividend policy, and cost-cutting policy. Marianne Bertrand & Antoinette Schoar, *Managing with Style: The Effect of Managers on Firm Policies*, Q. J. ECON. 118, 1169-1208 (2003).
- 7 Apparently, a proportion of wine producers in California do not care purely about the profit they generate, and instead enjoy producing high-quality, high-price wine. (Profit-maximizing wineries tend to offer lower quality wine.) See Fiona Scott Morton & Joel Podolny, *Love or Money? The Effects of Owner Motivation in the California Wine Industry*, J. INDUS. ECON. 50, 431-456 (2002).
- 8 An interesting example of this was seen in the Genzyme-Novazyme merger-to-monopoly which was approved by the FTC in 2004. This was a merger of two firms both engaged in R&D for treating a rare disease, where the prime danger from the merger was whether the discovery of a successful treatment would be delayed relative to the duopoly outcome. One factor in the decision was that the proposed CEO of the merged entity had two children with the disease, who may therefore not have wished to delay discovery. See the statement by the then FTC chairman Timothy Muris, available at www.ftc.gov/os/2004/01/murisgenzymestmt.pdf.

- 9 MILTON FRIEDMAN, *ESSAYS IN POSITIVE ECONOMICS*, 22 (1953).
- 10 For further details, see Charles Holt, *Industrial Organization: A Survey of Laboratory Research*, THE HANDBOOK OF EXPERIMENTAL ECONOMICS, §3 (John Kagel & Alvin Roth eds.) (1995). For a collection of papers discussing the use of experiments for competition policy, see EXPERIMENTS AND COMPETITION POLICY, (Jeroen Hinloopen & Hans-Theo Normann eds.) (2009).
- 11 For instance, see Douglas DeJong, Robert Forsythe, & Wilfred Uecker, *A Note on the Use of Businessmen as Subjects in Sealed Offer Markets*, J. ECON. BEHAVIOR AND ORG. 9, 87-100 (1988). More recently, for an illustration of how competition agency practitioners behaved on average in a similar way to students in a particular regulatory decision problem, see Bruce Lyons, Gordon Menzies, & Daniel Zizzo, *Professional Interpretation of the Standard of Proof: An Experimental Test on Merger Regulation*, Centre for Competition Policy Working Paper 10-2 (2010).
- 12 Charles Plott, *An Updated Review of Industrial Organization Applications of Experimental Methods*, THE HANDBOOK OF INDUSTRIAL ORGANIZATION, Vol. II, at 1165, (Richard Schmalensee & Robert Willig eds.) (1989).
- 13 In much empirical work on naturally occurring markets, marginal costs are inferred from observed data such as prices by *assuming* firms maximize their profit. At the end of this paper we discuss the dangers of this method when firms might potentially not be optimizers. The fact that marginal costs are rarely directly observable makes it hard to perform tests of the profit-maximization hypothesis. Steven Levitt in *Bagels and Donuts for Sale: A Case Study in Profit Maximization*, mimeo (2008), is a rare example of such a test. He obtained data from a relatively "simple" firm which supplies bagels and donuts to businesses, and where marginal costs are known. He finds that the firm is extremely good at predicting demand for given prices, but apparently prices too low given the estimated demand elasticity. (Interestingly, the decision-maker for this firm is a well-trained economist, who has published in the *Journal of Political Economy*.)
- 14 For an account of the theory of collusion, assuming rational firms, see Marc Ivaldi, Bruno Jullien, Patrick Rey, Paul Seabright, & Jean Tirole, *The Economics of Tacit Collusion*, Report for DG Competition, European Commission, (2003).
- 15 For instance, in the case of Bertrand price competition and homogenous products, the fully collusive outcome can be sustained in an infinitely repeated interaction with n symmetric suppliers if the discount factor δ satisfies $\delta > 1 - 1/n$. With reasonable choices for the discount factor and say, monthly price adjustment, collusion should be possible in oligopolies consisting of a hundred firms.
- 16 George Stigler, *A Theory of Oligopoly*, J. POL. ECON. 72, 44-61 (1964).
- 17 For instance, see Margaret Levenstein & Valerie Suslow, *Cartel Bargaining and Monitoring: The Role of Information Sharing*, THE PROS AND CONS OF INFORMATION SHARING (Mats Bergman ed.), (2006).
- 18 Some experiments randomly match subjects in each period, so that firms play against different rivals in each period. However, real markets do not frequently operate like this, and so we mainly focus on experiments where subjects interact repeatedly in the same groups.
- 19 For a brief survey, see Steffen Huck, Hans-Theo Normann, & Jorg Oechssler, *Two are Few and Four are Many: Number Effects in Experimental Oligopolies*, J. ECON. BEHAVIOR AND ORG. 53, 435-446 §§ 2 and 3 (2004).
- 20 Robert Feinberg & Thomas Husted, *An Experimental Test of Discount Rate Effects on Collusive Behavior in Duopoly Markets*, J. INDUS. ECON. 41, 153-160 (1993).

- 21 Feinberg argues that social welfare may sometimes be higher when firms' managers are relatively myopic or short-termist, perhaps because of high managerial turnover or the kinds of incentive schemes they are offered, since collusion thereby becomes harder to sustain. For the same reason, it seems plausible that shareholders who wish to achieve collusion would not wish to put in place a manager who was myopic or had hyperbolic time preferences. Robert Feinberg, *In Defense of Corporate Myopia*, *MANAGERIAL AND DECISION ECON.* 16, 205-210 (1995).
- 22 It might seem that this apparent falsification of rational play is not necessarily important in practice, as most real markets do not have a known "endpoint." Nevertheless, the unravelling argument applies even if the endpoint is uncertain, but it is known for sure that the interaction will have ceased by some date (e.g., if subjects in the laboratory do not believe that the experiment could possibly go on for more than a day, or if it is common knowledge that the world will have ended in 10 billion years).
- 23 *Supra* note 19.
- 24 As an aside, it is interesting in this regard to recall Robert Bork's assertion that any merger that left at least three rivals should be presumptively lawful. However, tacit collusion is not the only problem for concentrated markets; even the one-shot interaction could be insufficiently competitive when there are few firms. See ROBERT BORK, *THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF* (1978).
- 25 For instance, see Martin Dufwenberg & Uri Gneezy, *Price Competition and Market Concentration: An Experimental Study*, *INT'L J. INDUS. ORG.* 18, 7-22 (2000); Michael Baye & John Morgan, *Price Dispersion in the Lab and on the Internet: Theory and Evidence*, *RAND* 35, 449-466 (2004); and Jeroen Hinloopen & Adriaan Soetevent, *Laboratory Evidence on the Effectiveness of Corporate Leniency Programs*, *RAND* 39, 607-616 (2008).
- 26 Steffen Huck, Hans-Theo Normann, & Jorg Oechssler, *Learning in Cournot Oligopoly: An Experiment*, *ECON. J.* 109, C80-C95 (1999) and Steffen Huck, Hans-Theo Normann, & Jorg Oechssler, *Does Information About Competitors' Actions Increase or Decrease Competition in Experimental Oligopoly Markets?*, *INT'L J. INDUS. ORG.* 18, 39-57 (2000).
- 27 Theo Offerman, Jan Potters, & Joep Sonnemans, *Imitation and Belief Learning in an Oligopoly Experiment*, *REV. ECON. STUDIES* 69, 973-997 (2002).
- 28 An interesting real-world experiment is reported in Svend Albaek, Peter Mollgaard, & Per Overgaard, *Government-Assisted Oligopoly Coordination? A Concrete Case*, *J. INDUS. ECON.* 45, 429-443 (1997). The Danish antitrust authority changed policy so that transaction prices in the concrete market were published, and subsequently the average prices rose significantly. The authors argue that the most plausible reason is that this enabled the firms to coordinate their prices at a high level.
- 29 Offerman et al., *supra* note 27, also investigated the case where rival outputs but not rival profits were reported, and found that the market was less competitive than when no firm-specific information was reported. It would be interesting in future work to see what happens when rival firms' profits (or even just the average profits achieved by firms) but not their actions were revealed. It is possible that firms with higher profits than average might keep their strategy unchanged, but firms who do poorly revise their strategies, in accordance with a model of satisficing behavior discussed in section V. Bigoni allows firms to choose the kinds of information they see (e.g., aggregate output of rivals, individual rival outputs, individual rival profits) when operating under a time constraint, see Maria Bigoni, *Information and Learning in Oligopoly: An Experiment*, mimeo, (2008).
- 30 Huck et al., (2000), *supra* note 26, find that when firms set prices rather than quantities, the revelation of firm-specific data reduces prices only slightly.
- 31 Jan Potters, *Transparency About Past, Present and Future Conduct*, *EXPERIMENTS AND COMPETITION POLICY* (Jeroen Hinloopen & Hans-Theo Normann eds.) (2009).

- 32 Ola Andersson & Erik Wengstrom, *Do Antitrust Laws Facilitate Collusion? Experimental Evidence of Costly Communication in Duopolies*, SCANDINAVIAN J. ECON. 109, 321-339 (2007).
- 33 Explicit collusion in naturally occurring markets appears to be feasible with large numbers of participants. For instance, in their study of 41 cartels in Europe, Levenstein & Suslow, *supra* note 17, find that 18 involved more than five firms.
- 34 Giancarlo Spagnolo, *Divide et Impera: Optimal Leniency Programs*, CEPR Discussion Paper 4840 (2004).
- 35 Massimo Motta & Michele Polo, *Leniency Programs and Cartel Prosecution*, INT'L J. INDUS. ORG. 21, 347-379 (2003).
- 36 Jose Apesteguia, Martin Dufwenberg, & Reinhard Selten, *Blowing the Whistle*, ECON. THEORY 31, 143-166 (2007).
- 37 Hinloopen & Soetevent, *supra* note 25.
- 38 Armen Alchian, *Uncertainty, Evolution, and Economic Theory*, J. POL. ECON. 57, 211-22, at 218, (1950).
- 39 For a model along these lines, see John Conlisk, *Costly Optimizers Versus Cheap Imitators*, J. ECON. BEHAVIOR AND ORG. 1, 275-293 (1980). The situation is somewhat related to Grossman & Stiglitz's analysis of the incentives for investors to become better informed about the return of an uncertain asset: If all investors choose to be better informed, the asset's price reflects the information, and there is no need for any individual investor to become informed if there is a cost to doing so, Sanford Grossman & Joseph Stiglitz, *Information and Competitive Price Systems*, AMER. ECON. REV. 66, 246-253 (1976). It is also related to Burdett & Judd's analysis of a consumer's incentive to search for a low price: If all consumers choose to search then the market is highly competitive, and there is no need for any individual consumer to search if there is a cost to doing so, Kenneth Burdett & Kenneth Judd, *Equilibrium Price Dispersion*, ECONOMETRICA 51, 955-969 (1983).
- 40 Banerjee & Bikhchandani, et al., consider a situation in which similar agents need to decide between two options (say, whether to adopt technology X or technology Y), and each agent has a private signal as to which option is better and can also observe the previous choices (but not the payoffs) made by other agents. (The order in which agents have to make their choice is pre-determined in these models.) Even if agents are completely rational, it is possible that they become locked into the wrong choice. For instance, if technology X is in fact superior, but by chance the first few agents have private signals which induce them to choose Y, then subsequent agents will infer that the superior action is likely to be Y despite their own private signals to the contrary. (If instead, agents could observe the private signals of the earlier adopters, this inefficient herding could not occur.) Abhijit Banerjee, *A Simple Model of Herd Behavior*, Q. J. ECON 107, 797-817 (1992), and Sushil Bikhchandani, David Hirshleifer, & Ivo Welch, *A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades*, J. POL. ECON. 100, 992-1026 (1992).
- 41 See Georg Weizsacker, *Do We Follow Others When We Should? A Simple Test of Rational Expectations*, AMER. ECON. REV. (forthcoming), for an analysis of data from several experiments on the Bikhchandani et al. model, *Id.*
- 42 David Scharfstein & Jeremy Stein, *Herd Behavior and Investment*, AMER. ECON. REV. 80, 465-479 (1990).
- 43 *Id.* at 466.
- 44 The authors (*Id.*) quote Keynes as writing: "Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally." Tse & Tucker empirically investigate the timing of

earnings warnings, and find that a manager is more likely to issue an earnings warning if a peer has done so in the previous days. They conclude that the data are better explained by managers attempting to maintain their reputations than by the impact of a common shock. Senyo Tse & Jennifer Wu Tucker, *Within-Industry Timing of Earnings Warnings: Do Managers Herd?* REV. OF ACCT. STUDIES (forthcoming).

- 45 It is important that firms experiment occasionally, otherwise the process will grind to a halt after one period when all firms imitate the most profitable firm in the first period (which is unlikely to have chosen the optimal price immediately), and then all prices are unchanged thereafter. An alternative framework is presented in Ellison & Fudenberg, *supra* note 4, §II where firms choose between one of two technologies, and the relative payoff from using one technology rather than the other is uncertain. If in any period firms choose the technology which performed best in the previous period, then the chosen technology will flip over time depending on which one happened to work best one period earlier. Thus, instead of converging to the consistent use of the superior technology, the outcome is merely that the better technology is used more frequently. The authors go on to investigate less naïve rules of thumb—where a firm conditions its choice on how many firms use that technology—which have superior efficiency properties. (The reason why market shares matter for firms is that they reveal information about the relative performance of the two technologies for more than just the single previous period.)
- 46 David Ridley analyzes a model in which a second firm sometimes decides to enter a market only if its rival has first entered, in order to save on the costs of acquiring its own market information. He provides some anecdotes about how competitors of McDonald's often locate near a new McDonald's franchise, and he quotes a manager of a coffee shop chain as saying: "The reason we want to open across the street from every Starbucks is they do a great job at finding good locations." David Ridley, *Herding Versus Hotelling: Market Entry with Costly Information*, J. ECON. AND MGMT. STRATEGY 17, 607-631 (2008).
- 47 Fernando Vega-Redondo, *The Evolution of Walrasian Behavior*, ECONOMETRICA 65, 375-384 (1997).
- 48 Mark Schaffer, *Are Profit-Maximizers the Best Survivors?*, J. ECON. BEHAVIOR AND ORG. 12, 29-45 (1989).
- 49 Karl Schlag proposes an alternative model of imitation which yields a contrasting long-run prediction. In Schlag's model, there are many distinct oligopolies, and a firm in one market imitates the best-performing firm in another oligopoly, not the best-performing rival in the same market (as in Vega-Redondo's model). The result of this alternative specification is that firms move towards the Cournot-Nash equilibrium rather than the perfectly competitive outcome. Karl Schlag, *Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-Armed Bandits*, J. ECON. THEORY 78, 130-156 (1998), and Karl Schlag, *Which One Should I Imitate?*, J. MATH. ECON. 31, 493-522 (1999). See Jose Apesteguia, Steffen Huck, & Jorg Oechssler, *Imitation: Theory and Experimental Evidence*, J. ECON. THEORY 136, 217-235 (2007) for further discussion of the two approaches. In addition, Vega-Redondo's result that imitation leads to competitive outcomes does not necessarily hold when firms have longer memory. If, say, firms choose the most profitable action over the previous two periods, the Cournot equilibrium is then a stable outcome (although the perfectly competitive outcome remains stable as well). See Carlos Alos-Ferrer, *Cournot Versus Walras in Dynamic Oligopolies with Memory*, INT'L J. INDUS. ORG. 22, 193-217 (2004).
- 50 See Paul Rhode & Mark Stegeman, *Non-Nash Equilibria of Darwinian Dynamics with Application to Duopoly*, INT'L J. INDUS. ORG. 19, 415-453 (2001) for more details.
- 51 See Yasuhito Tanaka, *Stochastically Stable States in an Oligopoly with Differentiated Products: Equivalence of Price and Quantity Strategies*, J. MATH. ECON. 34, 235-253 (2000) and Rhode & Stegeman, *Id.* for a general argument. If firms are not symmetrically placed, then imitation may not tend towards competitive outcomes, nor is imitation equivalent to the maximization of relative profits. See, Jose Apesteguia, Steffen Huck, Jorg Oechssler, & Simon Weidenholzer, *Imitation and the Evolution of Walrasian Behavior: Theoretically Fragile but Behaviorally Robust*, J. ECON. THEORY (forthcoming).

- 52 Burkhard Schipper, *Imitators and Optimizers in Cournot Oligopoly*, J. ECON. DYNAMICS AND CONTROL 33, 1981-1990 (2009).
- 53 See Peter Duersch, Jorg Oechssler, & Burkhard Schipper, *Unbeatable Imitation*, mimeo, (2009) for further details about how imitators do almost as well against even the smartest opponents in a wide class of games. The main kind of game where imitators do poorly is something like "rock-paper-scissors," where a smart player can systematically trick the naïve imitator into always playing the wrong action. (These kinds of games do not seem common in market situations, however.)
- 54 Huck, Normann, & Oechssler (*supra* note 26); Offerman, Potters, Sonnemans (*supra* note 27); and Apesteguia, Huck, and Oechssler (*supra* note 49).
- 55 Huck et al., *supra* note 26.
- 56 Offerman et al., *supra* note 27 at 989.
- 57 Goel & Thakor propose a model in which merger waves can be caused by concerns for relative compensation by CEOs. If CEO compensation is based in part on firm size, then if one merger occurs which boosts that CEO's pay, other CEO feel envious and set about finding their own take-over targets. Anand Goel & Anjan Thakor, *Do Envious CEOs Cause Merger Waves?*, REV. FIN. STUDIES 23, 487-517 (2010).
- 58 For instance, see Andrew Clark, Paul Frijters, & Michael Shields, *Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles*, J. ECON. LIT. 46, 95-144 (2008).
- 59 Robert Gibbons & Kevin Murphy, *Relative Performance Evaluation for Chief Executive Officers*, INDUS. AND LABOR REL. REV. 43, 30-51 (1990).
- 60 John Vickers, *Delegation and the Theory of the Firm*, ECON. J. 95, 138-147, at 143, (1985).
- 61 The impact of making the firm wish to maximize relative profits is that the firm then behaves as the Stackelberg leader, even though both firms in fact choose quantities simultaneously. It is important that rivals observe the incentive scheme so that they know the firm's objective and can react to it accordingly.
- 62 Nolan Miller & Amit Pazgal, *Relative Performance as a Strategic Commitment Mechanism*, MANAGERIAL AND DECISION ECON. 23, 51-68 (2002).
- 63 Vicente Salas Fumas, *Relative Performance Evaluation of Management*, INT'L J. OF INDUS. ORG. 10, 473-489 (1992).
- 64 Carl Lundgren, *Using Relative Profit Incentives to Prevent Collusion*, REV. INDUS. ORG. 11, 533-550 (1996).
- 65 See Güth et al.'s path-breaking experiment and the many papers which followed, Werner Güth, Rolf Schmittberger, & Bernd Schwarze, *An Experimental Analysis of Ultimatum Bargaining*, J. ECON. BEHAVIOR AND ORG. 3, 367-388 (1982). See Ernst Fehr & Klaus Schmidt, *A Theory of Fairness, Competition, and Cooperation*, Q. J. ECON. 114, 817-868, §IIIA (1999) for a review of this literature.
- 66 The proposer might offer a significant proportion of the prize to the responder because he is purely self-interested and is afraid of lower offers being rejected by spiteful responders, or because he cares directly for fair allocations and is willing to sacrifice some of the prize to achieve a more equitable outcome. A variant of the ultimatum game—the dictator game—can discriminate between these two hypotheses. The dictator game does not allow the responder to reject the offer. Experimental comparisons of the two games reveal that offers are considerably less generous (and often zero) in the dictator game than in the ultimatum game, suggesting that generosity on the part of many proposers is

purely strategic. For more details, see Robert Forsythe, Joel Horowitz, N. E. Slavin, & Martin Sefton, *Fairness in Simple Bargaining Experiments*, GAMES AND ECON. BEHAVIOR 6, 347-369 (1994).

67 Steffen Huck, Wieland Muller, & Hans-Theo Normann, *Stackelberg Beats Cournot: On Collusion and Efficiency in Experimental Markets*, ECON. J. 111, 749-765 (2001).

68 *Supra* note 60.

69 Steffen Huck, Wieland Muller, & Hans-Theo Normann, *Strategic Delegation in Experimental Markets*, INT'L J. INDUS. ORG. 22, 561-574 (2004).

70 *Supra* note 62.

71 See Doruk Iris & Luis Santos-Pinto, *Tacit Collusion Under Fairness and Reciprocity*, mimeo, (2009), for a model along these lines. However, we are aware of no experimental test of this hypothesis.

72 For instance, if the ultimatum game is modified so that the proposer has first to run around a running track, or if the proposer is chosen by his performance in a general knowledge quiz, then the responder may feel that the proposer "owns" the initial stake, and so be more willing to accept a lower share of the prize.

73 See §§ IIIB and IIIC of Fehr & Schmidt (*supra* note 65) for further details and references.

74 See Levenstein & Suslow, *supra* note 17, and, especially, Christopher Leslie, *Trust, Distrust and Antitrust*, TEXAS L. REV. 82, 515-680 (2004), for further discussion of these points and illustrations of the role of trust and distrust in cartel stability.

75 William Page, *The Gary Dinners and the Meaning of Concerted Action*, SMU L. REV. 62, 597-619, at 608, (2009).

76 Joel Podolny & Fiona Scott Morton, *Social Status, Entry and Predation: The Case of British Shipping Cartels 1879-1929*, J. INDUS. ECON. 47, 41-67 (1999).

77 For instance, Gordon writes:

The fear of bankruptcy and the even more widespread fear of temporary financial embarrassment are probably more powerful drives than the desire for the absolute maximum in profits. [...] Given the fog of uncertainty within which [the businessman] must operate, the limited number of variables his mind can juggle at one time, and his desire to play safe, it would not be at all surprising if he adopted a set of yardsticks that promised reasonably satisfactory profits, R. A. Gordon, *Short-Period Price Determination in Theory and Practice*, AMER. ECON. REV. 3, 265-280 at 271 (1948)

See also K.W. Rothschild, *Price Theory and Oligopoly*, ECON. J. 57, 299-320 (1947) and Herbert Simon, *A Behavioral Model of Rational Choice*, Q. J. ECON. 69, 99-118 (1955).

78 Simon makes a joke about those economists who believe that departures from the predictions of rational behavior do not matter if the predictions of the standard models are good enough: "economists who are zealous in insisting that economic agents maximize turn around and become satisficers when the evaluation of their own theories is concerned." Herbert Simon, *Rational Decision Making in Business Organizations*, AMER. ECON. REV. 69, 493-513, at 495, (1979).

79 Richard Cyert & James March, *Organizational Factors in the Theory of Oligopoly*, Q. J. ECON. 70, 44-64 (1956).

- 80 Steffen Huck, Kai Konrad, Wieland Muller, & Hans-Theo Normann, *The Merger Paradox and Why Aspiration Levels Let it Fail in the Laboratory*, *ECON. J.* 117, 1073-1095 (2007).
- 81 Stephen Salant, Sheldon Switzer, & Robert Reynolds, *Losses from Horizontal Merger: The Effects of an Exogenous Change in Market Structure on Cournot-Nash Equilibrium*, *Q. J. ECON.* 98, 185-199 (1983).
- 82 The prediction is typically reversed when firms have differentiated products and compete in prices. See Raymond Deneckere & Carl Davidson, *Incentives to Form Coalitions with Bertrand Competition*, *RAND* 16, 474-486 (1985).
- 83 Levin presents a related theoretical analysis of mergers in Cournot markets, in which he allows that a merged firm changes its behavior from a Cournot-Nash player to a Stackelberg player (among other possible behavioral changes), in which case a merger becomes profitable. Daniel Levin, *Horizontal Mergers: The 50-Percent Benchmark*, *AMER. ECON. REV.* 80, 1238-1245 (1990).
- 84 Huw Dixon, *Keeping Up with the Jones's: Competition and the Evolution of Collusion*, *J. ECON. BEHAVIOR AND ORG.* 43, 223-238 (2000), and Jorg Oechssler, *Cooperation as a Result of Learning with Aspiration Levels*, *J. ECON. BEHAVIOR AND ORG.* 49, 405-409 (2002).
- 85 An alternative adjustment mechanism which does not require any information about rival actions or profits, or indeed about the demand and cost functions, but which may nevertheless lead again to collusive outcomes, is discussed by Huck, Normann, & Oechssler. They consider the rule of thumb: "If your last increase in output [or price] increased your profit, increase your output [or price] again; if your last increase in output [or price] decreased your profit, now decrease your output [or price]." They show that if firms are constrained to change their strategy at a fixed rate over time, then the market moves towards the collusive outcome. Steffen Huck, Hans-Theo Normann, & Jorg Oechssler, *Through Trial & Error to Collusion*, *INT'L ECON. REV.* 45, 205-224 (2004).
- 86 See Ellison (*supra* note 1 at 170-171). Note that if there are n firms instead of just two, the monopoly outcome can be implemented as an ε -equilibrium provided that $\varepsilon > (n-1)^2/(16n^2)$, and so with more firms it does become harder to sustain the monopoly outcome with satisficing behavior. Relatedly, approximately optimal behavior by firms can sustain tacit collusion when they interact only finitely often, at least for the early periods of the interaction. (In simple models, fully rational firms cannot sustain any collusion when they meet a known, finite number of times.) For further details, see Roy Radnor, *Collusive Behavior in Non-Cooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives*, *J. ECON. THEORY* 22, 136-154 (1980).
- 87 Baye & Norman, *supra* note 25.
- 88 For further details of this form of approximately optimal behaviour, see Richard McKelvey & Thomas Palfrey, *Quantal Response Equilibria for Normal Form Games*, *GAMES AND ECON. BEHAVIOR* 10, 6-38 (1995).
- 89 Renou & Schlag propose an alternative model for the Bertrand market, which is that firms are unsure about the rationality of their opponents, and a firm aims to minimize the maximum "regret" it experiences when competing with rivals. (A firm feels regret if it sets a price far below the minimum price of its rivals, since it could have made more profit with a higher price, and it feels regret if it sets a price above the minimum of its rivals' prices.) Using this model, they predict that firms set random prices, and expected profits are positive but decreasing with the number of competitors. They argue that the data reported in Baye & Morgan, *supra* note 25, conform well with their predictions. See, Ludovic Renou & Karl Schlag, *Minimax Regret and Strategic Uncertainty*, *J. ECON. THEORY* 145, 264-286 (2010).
- 90 For example, Svenson finds that 93 percent of respondents report that they are above the median in terms of driving ability. Ola Svenson, *Are We All Less risky and More Skilful Than our Fellow Drivers?*, *ACTA PSYCHOLOGICA* 47, 143-148 (1981).

- 91 For a persuasive account of how Adam Smith anticipated many of the central ideas of behavioral economics, see Nava Ashraf, Colin Camerer & George Loewenstein, *Adam Smith, Behavioral Economist*, J. ECON. PERSP. 19, 131-145 (2005).
- 92 See De Meza & Southey for empirical evidence and a model which predicts that entrepreneurs are more optimistic than the general population, David De Meza & Clive Southey, *The Borrower's Curse: Optimism, Finance and Entrepreneurship*, ECON. J. 106, 375-386 (1996). Camerer & Lovallo document in an experimental study that excessive entry can be driven by over-confidence in own ability, Colin Camerer & Dan Lovallo, *Overconfidence and Excess Entry: An Experimental Approach*, AMER. ECON. REV. 89, 306-318 (1999). See Tor for an extended discussion of behavioral economics and entry decisions, and the implications for competition policy, Avishalom Tor, *The Fable of Entry: Bounded Rationality, Market Discipline, and Legal Policy*, MICH. L. REV. 101, 482-568 (2002).
- 93 Edward Lazear & Sherwin Rosen, *Rank-Order Tournaments as Optimal Labor Contracts*, J. POL. ECON. 89, 841-864 (1981).
- 94 For a model along these lines, see Anand Goel & Anjan Thakor, *Rationality, Overconfidence and Leadership*, mimeo, (2000).
- 95 Aviad Heifetz, Chris Shannon, & Yossi Spiegel, *The Dynamic Evolution of Preferences*, ECON. THEORY 32, 251-286 (2007).
- 96 Vickers, *supra* note 60.
- 97 See Florian Englmaier, *A Strategic Rationale of Having Overoptimistic Managers*, mimeo, (2007) for such a model. In a Cournot market, hiring a manager who is overoptimistic (say, about the scale of market demand) again confers strategic advantage to the firm, but if all firms do this they will all be worse off relative to the situation with unbiased managers.
- 98 Richard Roll, *The Hubris Hypothesis of Corporate Takeovers*, J. BUS. 59, 197-216 (1986).
- 99 Ulrike Malmendier & Geoffrey Tate, *CEO Overconfidence and Corporate Investment*, J. FIN. 60, 2661-2700 (2005) and Ulrike Malmendier & Geoffrey Tate, *Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction*, J. FIN. ECON. 89, 20-43 (2008).
- 100 This correlation between holding onto stock options and corporate behavior could also be due to the CEO's inside information. However, Malmendier & Tate argue that this is unlikely to be the explanation since the CEOs who hold onto their stock options do not gain money from doing so, *Id.*
- 101 For further discuss of this point, see Maurice Stucke, *Am I a Price-Fixer? A Behavioral Economics Analysis of Cartels*, mimeo, University of Tennessee College of Law (2010).
- 102 For more details, see Bradford De Long, Andrei Shleifer, Laurence Summers, & Robert Waldmann, *The Survival of Noise Traders in Financial Markets*, J. POL. ECON. 64, 703-738 (1991).
- 103 See Luo Hirshleifer & Guo Ying Luo, *On the Survival of Overconfident Traders in a Competitive Securities Market*, J. FIN. MARKETS 4, 73-84 (2001).
- 104 Albert Kyle & Albert Wang, *Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?*, J. FIN. 52, 2073-2090 (1997).
- 105 Robert Hall & Charles Hitch, *Price Theory and Business Behavior*, OXFORD ECON. PAPERS 2, 12-45 (1939).
- 106 Nabil Al-Najjar, Sandeep Baliga, & David Besanko, *Market Forces Meet Behavioral Biases: Cost Misallocation and Irrational Pricing*, RAND 39, 214-237 (2008).

- 107 Theo Offerman & Jan Potters, *Does Auctioning of Entry Licences Induce Collusion? An Experimental Study*, R. ECON. STUDIES 73, 769-792 (2006).
- 108 A numerical example may help to fix ideas. Suppose there is a duopoly with differentiated products where the two firms choose prices and where a firm's demand is given by expression (1) as in section III. But instead of profit-maximizing behavior, suppose that firm $i = 1, 2$ behaves as if its marginal cost is c_i rather than the true cost (which is zero). Suppose that far-sighted profit-maximizing share-holders of firm i (or some form of evolutionary pressures) choose a corporate culture such that their manager behaves as if cost was c_i , and both groups of share-holders do this before price competition takes place. Then one can show that the equilibrium chosen "costs" are $c_1 = c_2 = 1/11$, and costs are artificially boosted so as to raise equilibrium prices. In this example, prices rise by approximately 10 percent relative to the situation in which managers base their prices on the true marginal costs.
- 109 Al-Najjar et al., *supra* note 106, suppose that the oligopolists reach equilibrium prices via a myopic adjustment process, and this means that rivals do not need to observe the "bias" of the manager. In addition, they assume that each firm observes only its own profits and not the profits of its rivals, and are more likely to choose cost methodologies that are performed well for it in the past.
- 110 Jonathan Eaton & Gene Grossman, *Optimal Trade and Industrial Policy Under Oligopoly*, Q. J. ECON. 101, 383-406 (1986).
- 111 See Giacomo Bonanno & John Vickers, *Vertical Separation*, J. INDUS. ECON. 36, 257-265 (1988) for this model.
- 112 See David Kreps, Paul Milgrom, John Roberts, & Robert Wilson, *Rational Cooperation in the Finitely Repeated Prisoners' Dilemma*, J. ECON. THEORY 27, 245-252 (1982).
- 113 However, if the potential irrationality took a different form then collusion might not be sustained. For instance, if the irrational player has a strategy of always colluding, then a rational player's best response to this is always to defect (just as if she were playing against a rational agent).
- 114 See David Kreps & Robert Wilson, *Reputation and Imperfect Information*, J. ECON THEORY 27, 253-279 (1982).
- 115 Yun Joo Jung, John Kagel, & Dan Levin, *On the Existence of Predatory pricing: An Experimental Study of Reputation and Entry Deterrence in the Chain-Store Game*, RAND 25, 72-93 (1994).
- 116 Kenneth Rogoff, *The Optimal Degree of Commitment to an Intermediate Monetary Target*, Q. J. ECON. 100, 1169-1189 (1985).
- 117 THOMAS SCHELLING, *THE STRATEGY OF CONFLICT*, 142 (1960).
- 118 For more analysis of this and related points, see Christopher Leslie, *Rationality Analysis in Antitrust*, U. PA. L. REV. 158, 262-353 (2010).
- 119 However, as discussed in section III, the effect may be limited if rivals react punitively when one firm puts in place an aggressive manager or an incentive scheme which induces aggressive behavior by its manager.
- 120 John List, *Does Market Experience Eliminate Market Anomalies?* Q. J. ECON. 118, 41-71 (2003) and John List, *Neoclassical Theory Versus Prospect Theory: Evidence From the Marketplace*, ECONOMETRICA 72, 615-625 (2004).
- 121 Huck et al., *supra* note 80.

122 HERBERT HOVENKAMP, THE ANTITRUST ENTERPRISE: PRINCIPLE AND EXECUTION, 134 (2005).

123 Leslie, *supra* note 118 at 269.

124 See also Leslie, *supra* note 118 at § III.F.

125 For instance, Armstrong & Vickers show how it can be optimal for a competition authority to permit only those mergers which do not harm consumers, even if society places equal weight on profit and consumer surplus. The reason is that a "consumer standard" affects the merger opportunities considered by firms, in a way which enhances total welfare. Mark Armstrong & John Vickers, *A Model of Delegated Project Choice*, *ECONOMETRICA* 78, 213-244 (2010).

Consumer Protection in Markets with Advice

Roman Inderst &

Marco Ottaviani

University of Frankfurt &

Northwestern University

Consumer Protection in Markets with Advice

*Roman Inderst & Marco Ottaviani**

Economists have long been interested in the performance of markets with imperfect information—and in the role of intermediation services in bridging the information gap between product providers and customers. Still, the classic information-economics framework for studying markets may fail to account for another role through which advice can affect market efficiency. Customers may suffer from “behavioral biases” in how they process information and make decisions. Thus, it is natural to ask whether advisors help households make better decisions or whether they, instead, exploit the biases and naïveté of their customers.

In this article we present some of the reasons why markets with advice may malfunction, and explore the potential rationales for some of the policy proposals that are on the table. We focus on the role of mandatory disclosure policies, the regulation of cancellation terms for service contracts (and refund policies for products), the imposition of liability standards for product providers and intermediaries, and the outright regulation of the size and structure of commissions.

*Roman Inderst, University of Frankfurt and Imperial College London & Marco Ottaviani, Kellogg School of Management, Northwestern University

I. Introduction

The recent subprime mortgage debacle has generated active debate on the role of the advice households receive from brokers and other information intermediaries when purchasing mortgages and other financial services such as consumer credit, life insurance, and investment products. In the current legislative proposal to reform investor protection and establish a Consumer Financial Protection Agency (“CFPA”) in the United States, advice features prominently:

“Impartial advice represents one of the most important financial services consumers can receive . . . Mortgage brokers often advertise their trustworthiness as advisors on difficult mortgage decisions. When these intermediaries accept side payments from product providers, they can compromise their ability to be impartial. Consumers, however, may retain faith that the intermediary is working for them and placing their interests above his or her own, even if the conflict of interest is disclosed. Accordingly, in some cases consumers may reasonably but mistakenly rely on advice from conflicted intermediaries. It is unfair for intermediaries to take advantage of that trust.

To address this problem, we propose granting the CFPA authority to impose carefully crafted duties of care on financial intermediaries. For example, the CFPA could impose a duty of care to counteract an intermediary’s patent conflict of interest, or to align an intermediary’s conduct with consumers’ reasonable expectations as demonstrated by empirical evidence. The CFPA could also consider imposing on originators a requirement to disclose material information such as the consumer’s likely ability to qualify for a lower interest rate based on her risk profile. In that regard, the CFPA could impose on mortgage brokers a duty of best execution with respect to available mortgage loans and a duty to determine affordability for borrowers.”¹

The United Kingdom, Germany, and many other jurisdictions are also considering the introduction of new regulation and policies in the face of potentially widespread “misselling” of retail financial services.²

The problem of unsuitable advice is clearly not confined to the financial industry—although the common use of commissions in that industry as well as the lack of customer sophistication (“financial capability”) certainly aggravate the problem. Outside the financial industry, for instance, payments and gifts made by pharmaceutical companies to physicians are also attracting closer scrutiny around the world, driven by concerns about both consumer protection and bulging health budgets.³ In fact, many of the current policy proposals in the area of retail finance, such as the disclosure of commissions and other inducements,

are also being debated or have already been enacted for the provision of medical services.⁴

Economists have long been interested in the performance of markets with imperfect information—and in the role of intermediation services in bridging the information gap between product providers and customers. Still, the classic information-economics framework for studying markets may fail to account for another role through which advice can affect market efficiency. Customers may suffer from “behavioral biases” in how they process information and make decisions.

THUS, IT IS NATURAL TO ASK
WHETHER ADVISORS HELP
HOUSEHOLDS MAKING BETTER
DECISIONS OR WHETHER THEY,
INSTEAD, EXPLOIT THE BIASES AND
NAÏVETÉ OF THEIR CUSTOMERS.

Thus, it is natural to ask whether advisors help households make better decisions or whether they, instead, exploit the biases and naïveté of their customers.⁵

In this article we present some of the reasons why markets with advice may malfunction, and explore the potential rationales for some of the policy proposals that are on the table.⁶ We focus

on the role of mandatory disclosure policies, the regulation of cancellation terms for service contracts (and refund policies for products), the imposition of liability standards for product providers and intermediaries, and the outright regulation of the size and structure of commissions. For examples of such policies, consider the following:

- As part of their occupational licensing procedures, various U.S. states require mortgage brokers to maintain a minimum net worth or to post a “surety bond.”⁷ The risk of losing this surety bond—or the imposition of penalties or close scrutiny by supervisory authorities and professional associations—should have a disciplining role on the market for advice.⁸
- There is presently a lively debate about regulating the structure of payments to those who sell financial products to households.⁹ To better align the interests of customers with those of their advisors, regulation could impose limits on the fraction of an advisor’s commission that is paid up-front instead of over the duration of a contract (“trail commission”). Intervention could also be directed at banning commissions altogether and steering the industry toward a more direct form of compensation for advice (for example, through hourly fees).
- Policies mandating the disclosure of payments between product providers and intermediary agents have been commonly adopted. In November 2008 the U.S. Department of Housing and Urban Development strengthened the requirement imposed on third-party brokers to disclose the payments they received for intermediated mortgage agreements to homeowners.¹⁰ Also, in 2008, the Federal Trade Commission proposed rules that would require brokers to join with customers in an initial agreement that:

“must state that the consumer will pay the entire compensation even if all or part is paid directly by the lender, and that a lender’s payment to a broker can influence the broker to offer the consumer loan terms or products that are not in the consumer’s interest or are not the most favorable the consumer could obtain.”¹¹

The article proceeds as follows. Section II discusses in detail the role of advice in markets for retail financial services, our leading example. Providing reliable advice may be essential to ensure that retail financial services provide benefits to a wide range of consumers. Section III discusses various policy interventions. While the discussion of specific policies is dictated by the current debate, these policies generally fall into three groups:

1. Policies that aim at reducing the need for advice, either by improving the quality of information or sophistication of consumers or by reducing the (perceived) complexity of products and services;
2. Policies that ensure that the quality of advice improves, e.g., by requiring that advisors meet higher standards of qualification or that they are given appropriate incentives to gather information and provide unbiased advice; and
3. Policies that target the way consumers deal with advice, for instance, through the provision of information about advisors’ incentives.

Section IV offers concluding remarks.

II. Financial Advice to Households

Currently, in the United States, there is widespread concern about the quality of advice regarding consumer credit products, most notably subprime mortgages.¹² In Europe there seems to be an equal—if not greater—concern for the (mal)functioning of the market for retail investment services.¹³ Although the role of advice may be equally important in other industries, here we’re focusing on the example of retail financial services, describing some of the key issues.

A. PERVASIVENESS OF ADVICE

Advice is ubiquitous in the retail finance industry. In the United States, mutual funds and equities (apart from employer-sponsored plans) are overwhelmingly purchased after receiving financial advice.¹⁴ According to a broad survey of retail investors in Germany, more than 80 percent of investors consult a financial advi-

sor.¹⁵ Further, a large cross-country survey in Europe showed that close to 90 percent of respondents in several countries expected advice to be provided by financial institutions—and the vast majority of customers said they trust the advice they received.¹⁶ While these observations relate only to investment services, advice obviously also plays a key role in the market for credit products.¹⁷ However, there is little evidence regarding how customers process financial advice, whether from banks or independent advisors.¹⁸

Until recently, despite its pervasiveness, the role of advice has been largely overlooked by much of the empirical literature that deals with the analysis of households' borrowing, saving, and investment decisions. The standard "household finance" paradigm features active investors who make individual decisions, possibly after acquiring costly information. That paradigm may describe some investors, most notably those trading frequently through online brokers,¹⁹ but it fails to capture a key determinant of the behavior of other, less sophisticated investors who heavily rely on financial advice.

B. FINANCIAL CAPABILITY AND CUSTOMER BIASES

Many people seem to lack even a basic level of financial literacy. Policy studies have been conducted in several countries to map out possible ways to overcome this deficiency.²⁰ At the same time, it is not clear that even a substantial fraction of people are capable of learning and retaining the necessary concepts and tools to make informed financial decisions, in particular in the area of investments.

MANY PEOPLE SEEM TO LACK
EVEN A BASIC LEVEL OF
FINANCIAL LITERACY.

In addition, financial capability may suffer less from a lack of knowledge than from decision-makers' "behavioral biases."²¹ In fact, when making financial decisions, people suffer from the same mistakes they commonly make in

other areas. For instance, they may be influenced by irrelevant aspects of the decision problem, such as the way in which the decision is presented (e.g., the format and material used to present the products).²² Or, in order to save on decision-making costs, people apply (ill-suited) heuristics—which may work well in some situations but lead to serious mistakes in others. For instance, they may simply stick to the status quo or choose whatever option is provided as the default.²³

Research in behavioral economics has recently pointed to various decision-making biases that may be particularly applicable to financial decisions. Some people procrastinate, delaying an action even though they are (or should be) aware that acting promptly would be better. A common explanation in cognitive psychology is that, for these people, immediate costs and benefits are unduly salient in comparison to future costs and benefits. People who are not aware of their tendency to procrastinate are liable to suffer significant welfare losses, e.g., as they incur high debt.²⁴

A second bias is demonstrated by people who are myopically loss averse. They may see risks in isolation rather than considering the “whole picture,” disregarding that a particular investment only accounts for a small fraction of their total wealth. Such a tendency may explain why some people shy away altogether from higher-risk investments such as stocks.²⁵

Finally, behavioral biases in the way people process information may have important implications for the financial services industry. Some people may refrain altogether from investing when there are too many choices available, and the large array of different financial assets that retail investors have access to may generate just such “choice overload.”²⁶

Households’ financial decisions may remain poor when they lack relevant knowledge or when they suffer from behavioral biases that affect how they process information. In the case of inadequate knowledge, the role of advice is immediate—households lacking financial sophistication should be willing to seek and follow advice. In the case of behavioral biases, advice could, in principle, be equally effective—in particular when the improvement of customers’ decisions does not negatively impact the advisor’s revenues.

C. PAYING FOR ADVICE AND CUSTOMER NAIVETÉ

It is common practice in the retail finance industry not to charge customers directly for advice, but for customers to pay indirectly for that advice through distribution fees, commissions, and other inducements that flow from product providers to brokers and (supposedly) independent financial advisors. These inducements often take the form of “kickbacks” that are not directly observed by the customers.²⁷ When advice represents, at least to some extent, a “credence good,” then the value of that advice is potentially compromised by advisors’ private interest in eliciting purchases.²⁸

There is much anecdotal evidence that the fee structure of investment products, rather than their suitability, drives customer sales.²⁹ In the United States, evidence suggests that mutual funds sold through broker/agent networks underperform, and that funds with higher fees (“loads”) are sold harder because of higher commissions, thus negatively affecting fund return.³⁰ Financial advisors may also have an interest in increasing the turnover in their clients’ portfolio (“churning”) when they earn additional fees or commissions with every new purchase.³¹

THERE IS MUCH ANECDOTAL EVIDENCE THAT THE FEE STRUCTURE OF INVESTMENT PRODUCTS, RATHER THAN THEIR SUITABILITY, DRIVES CUSTOMER SALES.

The impact of commissions on the quality of advice depends not only on whether commission fees are made transparent to customers but also on customers’ wariness. Do customers rationally evaluate the impact that such payments may have on a possi-

ble conflict of interest with their advisor and, thus, on the resulting quality of advice? Casual evidence indicates that not all customers are equally wary. In the United States, the Federal Trade Commission's staff report (2008) on disclosure rules for mortgage brokers suggests that "many consumers purportedly view mortgage brokers as trusted advisors who shop for the best loan for the consumer."³² Findings in the academic literature support the view that some people are naïve about how the quality of advice is impacted by conflicts of interest. For example, studies of investors' reactions to analysts' recommendations suggest that at least

FINDINGS IN THE ACADEMIC
LITERATURE SUPPORT THE VIEW
THAT SOME PEOPLE ARE NAÏVE
ABOUT HOW THE QUALITY
OF ADVICE IS IMPACTED BY
CONFLICTS OF INTEREST.

some investors are naïve about analysts' incentives.³³ There is also some experimental evidence to the effect that many subjects are willing to blindly follow advice. Interestingly, even when subjects are informed about the divergence of interests between them and their advisors, this knowledge does not seem to always make them sufficiently wary.³⁴

III. Policies to Make Advice Work

Firms and intermediaries may have information that can lead customers to make better decisions. In this context, it is worth distinguishing between two broad scenarios. In the first scenario, the customer may be in a position to understand and, when necessary, validate the information obtained. The role of advisor is then essentially that of facilitator, who provides information to the customer in the most convenient way and subsequently assists with the transaction. Our discussion, however, mostly concerns a second scenario, in which the customer needs help in either fully processing the available product information or in overcoming behavioral biases in information-processing and decision-making. This customer is in a weaker position and must rely on the advisor's recommendation.³⁵

A. HOW TO PAY FOR ADVICE

As we have noted, customers commonly pay "indirectly" for advice through higher product prices, some fraction of which is passed on to advisors in the form of commissions or other contingent payments. When a customer must rely on an advisor's recommendation at least to some extent, and when reputational or liability concerns are not sufficiently strong, this practice can lead to biased advice where the advisor biases his recommendations toward making a sale while favoring products that pay higher commissions.

Recent research seeks to analyze why this remuneration structure is dominant and persists, despite the inefficiency created by the bias. There is some evidence that some customers are not sufficiently wary of the conflict of interest that affects advice. In this case, in the equilibrium market outcome, customers are not charged a direct fee for advice but, rather, end up paying higher product

prices. In turn, advisors are induced to provide biased advice through high commissions.³⁶ Intuitively, while customers fully take into account any additional amount that they have to pay up-front, they underestimate the likelihood of their ultimately purchasing the product and paying the corresponding price. Thus, providers have an incentive to charge for advice through a higher product price (paid contingent on purchase) rather than through an up-front fee (paid regardless the purchase). The business practice of not charging customers directly for advice then persists, whether or not firms enjoy market power and whether or not competition prevails, as long as customers remain naïve about the conflict of interest generated by the commission compensation. In this scenario, there is a positive role for policy intervention that makes customers sufficiently wary through a “warning” (see Section B below) or imposes restrictions on how advice is paid for.

In contrast, if customers are wary of the seller’s strategic incentives, any policy that interferes with business practice is bound to reduce efficiency and ultimately reduce consumer surplus if competition is sufficiently intense. Moreover, there may be an efficiency rationale for the practice of paying advisors a higher margin when their advice results in a sale: A sales commission may induce advisors to acquire information by reading detailed material about particular products, keeping themselves informed about market developments, and acquiring customer-specific information so as tailor the advice toward the specific needs of their customer. When expecting to earn a commission only if the customer subsequently makes a purchase, the advisor may be motivated to work hard and may be able to credibly convince the customer with a superior recommendation. An advisor who, instead, is paid only a fixed fee (or is paid by the hour) and who has little at stake in business and reputation has a limited incentive to work hard.³⁷

The prospect of policy interference by (for instance) mandating caps on commissions to force advisors to charge customers directly for their service should, thus, depend heavily on the perception of whether customers who buy the particular product through the particular sales channel are sufficiently wary of how advisors are compensated and how this compensation potentially compromises the value of the advice they’re receiving.

B. DISCLOSING COMMISSIONS

Wary customers should discount advisors’ recommendations to a greater degree when they must pay a high price for the respective product, given that they should rationally infer that the product provider pays a high commission to the advisor. In this case, product providers could benefit from not giving advisors steep incentives, thereby enhancing the quality of advice and increasing customers’ willingness to pay for their products. However, out of a desire to push sales further, firms could provide advisors with additional, secret kickbacks. In such a situation the enforcement of a mandatory disclosure of commissions could benefit both firms and customers.³⁸

Firms' and customers' incentives with respect to disclosing commissions are, however, no longer aligned when, without such disclosure, customers remain naïve about the conflict of interest. As we have noted, firms may then be able to maximally exploit customers by reducing any direct fees for advice and increasing both product prices and commissions. Naïve customers may, however, be made aware of the conflict of interest when firms are forced to disclose commissions or to provide a general "warning" that such commissions are paid and may compromise the quality of advice.³⁹

SOME RECENT EXPERIMENTAL AND
THEORETICAL WORK SHOWS THAT
MANDATORY DISCLOSURE OF
COMMISSIONS MAY HAVE
UNINTENDED CONSEQUENCES.

Some recent experimental and theoretical work shows that mandatory disclosure of commissions may have unintended consequences. Customers who are unfamiliar with such disclosure may fail to make appropriate use of the disclosed information. In fact, this information may distract their attention from attractive product characteristics and induce them to choose inferior products associated with lower commissions.⁴⁰ Furthermore, disclosing commissions could undermine the trust relationship between advisors and clients.

Advisors who experience mistrust from their customers may then feel "morally licensed" to maximize only their own profits.⁴¹

In practice, such a possible "information overload" of customers or a change in the "framing" of the advice relationship may, however, have only temporary implications until customers and advisors adjust. More research is needed in this area. Research is also lacking on how advice works differently in face-to-face situations where it may be combined with high-pressure sales techniques.⁴² Sales people who also provide advice may apply psychological tactics (or even make false claims) to build an image of expertise and use their influence to exert undue sales pressure on customers.⁴³

Finally, it is important to bear in mind that besides steering advisors' recommendations toward particular products, commissions have a much wider range of purposes, such as rewarding intermediaries who prospect for new purchasers. Dampening commissions could also inefficiently interfere with any efficiency-enhancing functions that commissions serve, leading to negative welfare implications, e.g., by slowing the roll-out of products or preventing reaching certain segments of the market that require more sales effort.⁴⁴

C. REGULATING CONTRACTS: COMMISSION AND PRODUCTS

In the wake of the financial crisis there is much debate on how to regulate compensation to better align the interests of top managers in financial institutions with those of society. The most common proposals are to limit the steepness of incentives and to make incentives longer term. Interestingly, these suggestions mirror some of the policies that have already been implemented for retail finan-

cial and insurance services. In fact, policymakers in a number of countries are pushing toward a higher share of trail commissions, and some countries have already taken steps in this direction.⁴⁵

To our knowledge, economics research has paid little attention to the issue of the structure of commissions and its implications for consumer welfare. However, some of the more general insights regarding the (mandatory) disclosure of commissions, which we discussed above, apply here as well. Take the matter of upfront commissions versus trail commissions. Generally, when contracts are long-term and may be cancelled by consumers, or when the product provider can expect feedback from consumers over the long term, then the firm could better align the interests of customers with those of the agent by postponing part of an intermediary agent's commissions. Such an action may, however, be quite costly if the agent has a strong (time or liquidity) preference for being paid immediately. If customers are sufficiently wary and can observe the structure of commissions, they may correctly infer that more long-term and less short-term commissions should increase the quality of advice. Policies mandating the postponement of a given fraction of commissions may then serve for firms as commitment devices, much like the mandatory disclosure of commissions, and may provide disincentives for churning.

In markets where the quality of advice is the linchpin to deliver customer value, policies could be directed both toward increasing the quality of advice, e.g., by regulating the size and structure of commissions, and toward making advice less necessary. This latter goal could be achieved by either making the customer better informed and more sophisticated or by reducing the need for complex decisions. With respect to financial products, for example, a government's decision to privatize much of its pension system could substantially complicate households' decision problems but the increased complexity could be reined in through a pre-selection of providers and products that would enjoy preferential tax treatment. Regulation of products may also ensure that firms and their intermediary agents have incentives to provide better advice. We explore this next.

IN MARKETS WHERE THE QUALITY OF ADVICE IS THE LINCHPIN TO DELIVER CUSTOMER VALUE, POLICIES COULD BE DIRECTED TOWARD INCREASING THE QUALITY OF ADVICE.

Take the case of termination and cancellation clauses in contracts.⁴⁶ When customers need time before becoming sure a particular product or service is indeed suitable for their needs, they benefit from the right to return the product or to cancel the contract. When the initial decision to purchase a product or to enter into a contract is made under advice, the provision of generous rights of refund or cancellation can help the seller commit to high quality and fair advice. Intuitively, the margin lost from early cancellations (or returns) disciplines the seller to advise the customer to purchase only following the observation of a sufficiently favorable signal about the product's suitability. But this commitment

mechanism again requires that customers are wary. If credulous customers put too much faith in an advisor's inflated statements, they may wrongly presume that it is very likely they will be unable to terminate a contract. Indeed, firms are aware of this possibility and can maximally exploit customers' misperceptions by granting them very unfavorable terms of refund and cancellation, which in turn leads to a low quality of advice.

Both consumer surplus and welfare can be increased through a policy that mandates a minimum level for consumers' right to cancel.⁴⁷ In fact, unconditional refund periods are commonly imposed for the sale of life insurance policies and annuity contracts (typically sold following advice) and are often combined with suitability rules.⁴⁸ Regulations of cancellation terms and "free-look periods" tend to apply to retail channels populated by more vulnerable buyers such as senior citizens who can easily fall prey to aggressive marketing techniques.⁴⁹

D. OVERSIGHT AND LIABILITY

Conceptually, the problem of unsuitable advice has much in common with that of the provision of faulty or inferior products, ("lemons"). In addition to learning and reputational mechanisms, the market ameliorates the lemons problem through contractual provisions such as warranties that allow customers to return faulty products or have them repaired at the seller's expense. However, outright unsuitability or inferiority of advice is likely to be much more difficult to establish. Purely contractual solutions may then simply become insufficient. If customers are wary of this deficiency, both customers and firms can benefit if firms

CONCEPTUALLY, THE PROBLEM
OF UNSUITABLE ADVICE HAS
MUCH IN COMMON WITH THAT OF
THE PROVISION OF FAULTY OR
INFERIOR PRODUCTS, ("LEMONS").

are incentivized to work toward higher standards; for example, through self-regulation or the oversight and interference of agencies that protect consumers.⁵⁰

Recent research has analyzed the benefits of imposing higher liability standards and tighter oversight. A major, though often overlooked, determinant of the need for a more interventionist policy is the severity of the "internal" agency problem between a product provider and the agent who is responsible for offering advice. This agency problem becomes more severe when the same agent is responsible both for providing advice and for eliciting new sales, for example by prospecting for new customers. In economics terminology, the agency relationship then features "multiple-tasks," with possibly conflicting implications for the firm's incentive structure: High rewards for sales are needed to generate new prospective customers, but they lead to biased advice.

Hence, the appropriate standard of liability and oversight that agencies would want to impose should depend on how products are sold and advice is given. Also affecting the agency problem and the standard of advice that would prevail in equilibrium without policy interference is whether firms have access to early cus-

customer feedback, which in turn depends on customers' sophistication and the nature of the products.⁵¹

Further, when selling and advice of complex products are undertaken by independent agents, such as mortgage brokers or independent financial advisors, the question arises which party should be ultimately liable following an unsuitability claim. Clearly, some form of "vicarious liability" is called for when intermediary agents are unlikely to pay compensation to customers, given that their size and chosen organizational form make them essentially "judgment proof."⁵² Policy makers must then decide which party should be responsible to implement their chosen standard of responsible selling.⁵³

However, when tightening liability for particular products or channels, care must be taken to avoid unintended consequences. Product providers may simply cease to develop or roll out the products they deem to have high legal risk, irrespective of the ultimate benefits to customers. With respect to advice, agents may shift toward selling products without advice, even though customers may fail to understand the distinction.⁵⁴ Also, the special treatment of independent agents may distort the market by imposing a penalty on vertical separation and open-architecture sales. Such penalties can clearly distort competition and lead to less consumer welfare. And, in the long term, increasingly active policy intervention may act as a disincentive for customers to take sufficient care themselves!

AND, IN THE LONG TERM,
INCREASINGLY ACTIVE
POLICY INTERVENTION MAY
ACT AS A DISINCENTIVE FOR
CUSTOMERS TO TAKE
SUFFICIENT CARE THEMSELVES!

IV. Conclusion

This article is a progress report on our current research that seeks to explain widely-observed compensation methods for advice and to analyze the effects of common policy measures that are meant to enhance the quality of advice. A key challenge is the intrinsic difficulty in evaluating the quality of advice tailored to the particular needs of customers. Quality level is important to understand not only for consumers but also for product providers, who may wish to implement a high standard of advice so as to be able to charge high product prices, as well as for policymakers and regulators.

In this respect, technological improvements in the way products are sold (such as through electronic platforms) may provide new opportunities for the industry as well as for policy makers. For example, the advent of smart agents might allow product providers to enforce rigid rules when giving personalized advice and to closely monitor internal compliance to chosen rules. At the same time, advances in information technology may also assist policy makers in developing more objective ways to evaluate and enforce the implementation of suitable rules.

However, particularly in the case of retail financial services, policy intervention may also be called for to establish reliable comparisons on how well customers fare when relying on advice from different sources.

Overall, more research is certainly needed to evaluate the costs and benefits of different policy interventions. To build policies on solid foundations, further theoretical, empirical, and experimental work must be done to better understand the role of advice in retail markets. Looking ahead, a particularly promising area is the integration of new advances from economics with marketing research. Case studies tailored to particular industries, products, and customer segments could also prove useful. ▼

-
- 1 FINANCIAL REGULATORY REFORM. A NEW FOUNDATION: REBUILDING FINANCIAL SUPERVISION AND REGULATION, U.S. DEPARTMENT OF TREASURY, 68, (June 2009).
 - 2 In the EU, the regulation of retail financial sales and advice is framed by the Markets in Financial Instruments Directive (MiFID) and the Insurance Mediation Directive (IMD).
 - 3 See M. Millenson, *Getting Doctors to Say Yes to Drugs: The Cost and Quality Impact of Drug Company Marketing to Physicians*, BlueCross, Blue Shield Association, (2003) for an overview of the practice of detailing drugs and physician preference items.
 - 4 For instance, in the United States, the proposed Physician Payment Sunshine Act would require certain manufacturers of drugs and medical devices to report inducements given to physicians through consultant fees, educational grants, or travel gifts. According to this proposal, disclosure would go to the Secretary of Health and Human Services rather than to customers.
 - 5 For example, financial advisors who are paid commissions may find it easier to increase revenues through "churning" when customers already have a bias toward excessive trading.
 - 6 Broadly speaking, these policies fall into the area of consumer protection. Vickers defines consumer policy in terms of the underlying problems that it seeks to remedy, namely: (1) duress and undue sales pressure; (2) pre-purchase information problems; and (3) undue surprises after purchase. The topic of (unsuitable) advice covers all these problems, see John Vickers, *Economics for Consumer Policy*, Proceedings of the British Academy, 125, 287-310, (2004). However, in this article we will not address how consumer protection, sector specific regulation, and competition policy can jointly address these problems. We refer to Armstrong for a discussion of the interaction of competition and consumer policy, see Mark Armstrong, *Interactions Between Competition and Consumer Policy*, 4(1) COMPETITION POL'Y INT'L, 97-145 (2008).
 - 7 More precisely, as documented by Pahl, surety bonds are typically posted through third parties that initially check whether the broker has sufficient net wealth. While these third parties are the first to be liable, they are then compelled by regulation to seek redress from the broker, see C. Pahl, *A Compilation of State Mortgage Broker Laws and Regulations*, 1996-2006, Federal Reserve Bank of Minnesota, Community Affairs Report No. 2007-2.
 - 8 The compensation that may have to be paid to customers once government agencies wade in and aggressively prosecute alleged misbehavior could be substantial. For instance, after a full review of private pension sales in 1994, financial institutions in the United Kingdom reportedly paid out a total compensation of £12 billion. In the more recent case of endowment mortgages, which bundle mortgages with risky (stock market) investment, U.K. consumers have already obtained in excess of £2 billion in redress. For additional information, see the website of the U.K.'s Financial Services Authority (FSA) at www.fsa.gov.uk.

- 9 For instance, consider the following: "The CFPA should also be authorized to ban often invisible side payments to mortgage originators... that are tied to the borrower receiving worse terms than she qualifies for, if the CFPA finds that disclosure is not an adequate remedy. These payments incentivize originators to steer consumers to higher-priced or inappropriate mortgages. In addition, the CFPA could consider requiring that originators receive a portion of their compensation over time, contingent on loan performance, rather than in a lump sum at origination." *Financial Regulatory Reform. A New Foundation: Rebuilding Financial Supervision and Regulation*, U.S. Department of Treasury, 68, (June 2009).
- 10 For details, see www.hud.gov.
- 11 In the European Union, the Markets in Financial Instruments Directive (MiFID) has required since January 2008 the disclosure of commissions on retail financial products. In the United Kingdom, similar provisions were imposed earlier by the Financial Services Authority.
- 12 "Many borrowers whose credit scores might have qualified them for more conventional loans say they were pushed into risky subprime loans. . . . The subprime sales pitch sometimes was fueled with faxes and emails from lenders to brokers touting easier qualification for borrowers and attractive payouts for mortgage brokers who brought in business. One of the biggest weapons: a compensation structure that rewarded brokers for persuading borrowers to take a loan with an interest rate higher than the borrower might have qualified for." *Subprime Debacle Traps Even Very Credit-Worthy As Housing Boomed, Industry Pushed Loans To a Broader Market*, WALL STREET J., (December 3, 2007).
- 13 For updates on-going work by the EU Commission on the market for Packaged Retail Investment Products (PRIPs) see http://ec.europa.eu/internal_market/finservices-retail/investment_products_en.htm.
- 14 See D. Bergstresser, J. Chalmers, & P. Tufano, *Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry*, REV. FINANCIAL STUDIES, forthcoming, and *Equity Ownership in America 2005* at http://www.ici.org/pdf/rpt_05_equity_owners.pdf. In a survey conducted by the Investment Company Institute (ICI 2007) over 80 percent of respondents stated that they obtained financial advice from professional advisors or other sources.
- 15 See DABANK, FASZINATION WERTPAPIER: FAKTEN UND HINTERGRÜNDE ZUM ANLEGERVERHALTEN IN DEUTSCHLAND, München, (2004). Two thirds respond that they obtain financial advice from their main bank. For a comparison, only one fifth (also) obtains advice from an independent financial advisor.
- 16 See Eurobarometer 60.2, Nov-Dec 2003. For instance, regarding households, 95 percent in Germany, 90 percent in Denmark, 95 percent in Austria, 91 percent in the Netherlands, and 80 percent in Finland expect to receive advice from financial institutions. (However, only 40 percent of Greek households expect to receive advice.) Furthermore, 65 percent of German respondents trust advice, which compares with 76 percent in Denmark, 75 percent in Austria, 60 percent in the Netherlands, and 79 percent in Finland, but only 22 percent in Greece.
- 17 For instance, in the United Kingdom, 91 percent of intermediary mortgage sales are "with advice" (see Financial Service Authority, *Mortgage Market Review*, Discussion Paper 09/3, (2009).
- 18 In a recent study, Hackethal et al. used trading as well as survey data from a sample of customers of a large German bank and found that over half of the surveyed customers stated that they consistently rely on the advice of their personal advisor. These customers are, predictably, less informed about financial products and do not perceive there to be a large conflict of interest—and they end up trading substantially more, thereby generating higher revenues for the bank, see A. Hackethal, R. Inderst & S. Meyer, *Trading on Financial Advice*, Mimeo, (2009). Georgarakos & Inderst find, using a pan-European survey, that trust in financial advice has a significant impact on the decision of less-educated households to buy stock or other risky and more information-sensitive "collective investment" products. Instead, for more educated households or those who do not perceive financial decisions to

be particularly complex, trust in financial advice does not significantly impact these decisions, see D. Georgarakos & R. Inderst, *Trust in Financial Advice*, Mimeo, (2010).

- 19 Incidentally, much empirical research that has access to detailed, micro-level portfolio and trading data comes from such online brokers, e.g., T. Odean, *Do Investors Trade Too Much?* AMER. ECON. REV. 89, 1279-298, (1999).
- 20 For instance, for the United Kingdom the Thoresen Review in 2008 proposes to increase the level of "generic financial advice," so as to establish a minimum level of knowledge for all households that make financial decisions, see THORESEN REVIEW OF GENERIC FINANCIAL ADVICE, (2008).
- 21 Indeed, some studies have found that the provision of financial literacy education has only very limited and short-term effects, if at all (see the reviews in L. MANDELL, FINANCIAL LITERACY: ARE WE IMPROVING? (2004). The authors of a FSA report on financial capability thus conclude: "Making people better informed is hard and expensive and is of minimal value if it has no effect on behavior. This would be the case if low financial capability is more to do with psychological factors than lack of knowledge." Financial Service Authority, *A Review of Retail Distribution*, Discussion Paper 07/1, (2007).
- 22 For instance, Choi *et al.* show that such mistakes in the choice between index-tracking funds are common even among MBA students, see J. Choi, D. Laibson, & B. Madrian, *Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds*, NBER, (2006).
- 23 For pioneering work in economics on the tendency to favor the status quo, see W. Samuelson & R. Zeckhauser, *Status Quo Bias in Decision Making*. J. RISK AND UNCERTAINTY 1, 7-59, (1988).
- 24 To illustrate the consequences of procrastination, take a consumer who decides not to return a rented video today as the immediate disutility from walking to the shop exceeds the small charge for an additional day's rent. This decision may be fully rational for a consumer who expects the opportunity cost of returning the video to be higher tomorrow. But if the consumer naively underestimates the possibility of procrastinating again tomorrow, long and costly delay may be sustained. In economics, procrastination preferences go back at least to Strotz, and may account for low savings rates and the reluctance to participate in government or company sponsored savings plans. See R. Strotz, *Myopia and Inconsistency in Dynamic Utility Maximization*, REV. ECON. STUDIES 23, 165-80, (1956), and also T. O'Donoghue & M. Rabin, *Doing It Now or Later*, AMER. ECON. REV. 89, 103-24, (1999). Despite the foundations of procrastination in neuroscience (e.g. S. McClure, D. Laibson, G. Loewenstein, & J. Cohen, *Separate Neural Systems Value Immediate and Delayed Monetary Rewards*, SCIENCE 306, 503-07, (2004)), it is remarkable that countries such as Germany neither have a low savings rate nor high (revolving) consumer debt, which are empirical regularities that are commonly cited in support of procrastination.
- 25 For instance, in an interesting experiment, Fellner & Suttner show that subjects are willing to take on more risk when they receive less feedback on their investment choices and have to take a more long-term decision, (G. Fellner & M. Suttner, *Causes, Consequences and Cures of Myopic Loss Aversion—An Experimental Investigation*, ECON. J. 119, 900-16, (2009)). Similarly, such a tendency to avoid risky choices has also been linked to regret aversion, which makes people avoid situations where they appear to have made the wrong decision even when the decision was a priori correct (e.g., G. Loomes & R. Sugden, *Regret Theory: An Alternative Theory of Rational Choice under Uncertainty*, ECON. J. 92, 805-24, (1982)). Generally, "mental accounting" refers to the cognitive method of treating different decisions in isolation, e.g., the decision to save for retirement and the decision to borrow for short-term consumption (see, R. Thaler, *Mental Accounting and Consumer Choice*, MARKETING SCIENCE 4, 199-214, (1990)).
- 26 A classic experiment in this vein was performed by Iyengar & Lepper, who showed that when the number of tasting booths for jam in a shop was increased from six to twenty-four flavors, then the fraction of customers who bought after tasting dropped from 20 percent to 3 percent, see S. Iyengar & M. Lepper, *When Choice is Demotivating: Can One Desire Too Much of a Good Thing?* J. PERSONALITY AND SOCIAL PSYCHOLOGY 76, 995-1006, (2000).

- 27 At least in some countries, when a customer pays directly for advice, the advisor is legally bound to pass on to the customer these benefits, implying that for the customer there is an immediate tradeoff. Also, the payments made to intermediaries may be funded by fees that are directly collected from the respective investment vehicles or that are funded from the additional interest ("yield spread") that a customer pays (see the discussion in H.E. Jackson & L. Burlingame, *Kickbacks and Compensation: The Case of Yield Spread Premiums*, STANFORD J. L., BUS. & FIN. 12, 289-361, (2007) and E. Keith, D. Bocian, & W. Li, *Steered Wrong: Brokers, Borrowers, and Subprime Loans*, Center for Responsible Lending, (2008).
- 28 Bolton *et al.* and Inderst & Ottaviani show this in a model of "cheap talk" applied to the financial industry, see P. Bolton, X. Freixas, & J. Shapiro, *Conflicts of Interest, Information Provision, and Competition in Banking*, J. FIN. ECON. 85, 297-330 (2007) and R. Inderst & M. Ottaviani, *Mis-selling through Agents*, AMER. ECON. REV. 99, 883-908, (2009). What mitigates this conflict of interest are reputational concerns as well as the threat of legal prosecution. In the analysis of intermediated investment management of Stoughton *et al.* a fund advisor charges an advisory fee based on the end-of-year value of the client's portfolio. According to a prevalent practice in the industry, investors are charged indirectly for advice through high loads that, in turn, give intermediary agents steep incentives to sell., see N. Stoughton, Y. Wu, & J. Zechner, *Intermediated Investment Management*, Mimeo, (2008). To this effect, it is indicative to note, in the United States, the low membership (of around one thousand professionals) in the National Association of Personal Financial Advisors (NAPFA), which admits only agents working on a fee-only compensation (see <http://www.napfa.org>). There are also legal obstacles, given that when receiving payment for advice, agents are subject to a stricter fiduciary duty. The U.K.'s financial services regulator has proposed plans to steer independent financial advisors fully toward direct charges for advice (Financial Services Authority, *supra* note 17).
- 29 See also the survey among EU members of the CFA Institute in which 64 percent of respondents agreed that the prevailing fee structure serves the purpose of steering sales rather than serving customers' needs, CFA Institute, *European Union Member Poll on Retail Investment Products: Summary Report*, (2009).
- 30 See Bergstresser *et al.* (*supra* note 14), R. Edelen, R. Evans & G. Kadelec, *What Do Soft Dollars Buy? Performance, Expense Shifting, Agency Costs*, Mimeo (2008), and J. CHEN, H. HONG & J. KUBIK, *OUTSOURCING MUTUAL FUND MANAGEMENT: FIRM BOUNDARIES, INCENTIVES, AND PERFORMANCE*, Mimeo, (2006).
- 31 As noted previously, payments to brokers have reportedly led to distortions also in the U.S. mortgage market. Generally, such distortions are likely when commissions vary between different products and product groups. For instance, in the United Kingdom, the Financial Services Authority suggests that unsuitable advice in the prime mortgage market may be a lesser concern, because their fees are typically flat between different products.
- 32 As noted above, the evidence in Hackethal *et al.*, *supra* note 18, also suggests that customers differ in their perception.
- 33 See U. Malmendier & D. Shanthikumar, *Are Small Investors Naive About Incentives?* J. FIN. ECON. 85, 457-489, (2007) and H. Hong, J. Scheinkman, & W. Xiong, *Advisors and Asset Prices: A Model of the Origins of Bubbles*, J. FIN. ECON. 89, 268-287, (2008). Various theoretical attempts have been made to model the underlying bounded strategic rationality (e.g., N. Kartik, M. Ottaviani & F. Squintani, *Credulity, Lies, and Costly Talk*, J. ECON. THEORY 134, 93-116, (2007)). In general, such boundedly rational individuals may not understand (even if they had the necessary information) what the actions of the various players imply for the resulting payoffs.
- 34 In Cain *et al.* subjects are paid for the precision of the estimates of the number of coins in a jar. They can rely on the additional judgment of an advisor, who can closely inspect the jar. In a first treatment, advisors are paid for the accuracy of the subjects' guesses of the number of coins; in a second treatment they are paid more when the guess is high. The estimate of the subjects is 28 percent higher in the second treatment, see D. Cain, G. Loewenstein, & D. Moore, *The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest*, J. LEGAL STUDIES 34, 1-25, (2005). See also U. Gneezy,

Deception: The Role of Consequences, AMER. ECON. REV. 95, 384-394, (2005) on subjects' trust in advice.

- 35 This distinction between the provision of information, on the one hand, and making a recommendation, on the other, also underlies much of the applicable regulation (such as MiFID in Europe).
- 36 Inderst & Ottaviani distinguish between two types of unsuitable advice: advice on whether to purchase a particular product or not, and advice on which product to choose. Even when customers do not pay a direct fee and when commissions are high, there may be little bias in the choice between different products if the product providers compete for the agent's recommendation by offering counteracting commissions, see R. Inderst & M. Ottaviani, *Intermediary Commissions and Kickbacks*, Mimeo, (2008) and *supra* note 28.
- 37 This trade-off between a potential bias in the advisor's recommendation and the quality of the information gathered is analyzed in R. Inderst & M. Ottaviani, *How (Not) to Pay for Advice*, Mimeo, (2009).
- 38 Inderst & Ottaviani, *supra* note 28, explore this commitment feature of disclosed commissions.
- 39 See Inderst & Ottaviani, *supra* note 37.
- 40 For a discussion of disclosure with mortgage contracts, see J. Lacko & J. Pappalardo, *Improving Consumer Mortgage Disclosures: An Empirical Assessment of Current and Prototype Disclosure Forms*, Federal Trade Commission Bureau of Economics Staff Report, (2007).
- 41 In an experimental study, Cain *et al.*, *supra* note 34, find that advisors seem to be more prone to provide worse advice when the conflict of interest is disclosed.
- 42 In fact, the face-to-face situation may often be the factor that distinguishes advice from the provision of information at a distance, e.g., on the phone or through internet services. Furthermore, when advice is not customer-specific but regards general features of a product or service, such as quality or costs, issues may be similar to those arising with respect to deceptive marketing, e.g., through making false claims, denigrating rivals' products, or trying to pass off the product as another firm's.
- 43 OFT discusses various such techniques, building on Cialdini's "principles of influence." See Office of Fair Trading, *Psychology of Buying and Selling in the Home*, DOORSTEP SELLING REPORT, Annexe F, (2004) and R. B. CIALDINI, *INFLUENCE: SCIENCE AND PRACTICE*, (2001).
- 44 For details see Inderst & Ottaviani, *supra* notes 36 and 28.
- 45 As reported by the Financial Services Authority, *supra* note 21, this is the case in the Netherlands, Denmark, Finland, Sweden, Israel, and Australia, albeit to different degrees and not always through the imposition of formal requirements. Denmark and Finland have moved, according to this report, toward a full ban of initial commissions on life and pension sales through independent brokers, while the Netherlands has limited the initial commission on life-and-protection insurance contracts to 50 percent of the total compensation.
- 46 This discussion is based on Inderst & Ottaviani's theoretical analysis, see R. Inderst & M. Ottaviani, *Sales Talk, Cancellation Terms, and the Role of Consumer Protection*, Mimeo, (2008).
- 47 "Cooling-off rules" are used to target purchases that require an active marketing effort by sellers and for which buyers learn their utility only after purchase, as in the case of doorstep sales. These rules protect customers from purchasing under inflated perceptions. Interestingly, they typically apply less to situations in which there is no advice or face-to-face contact involved (with the exception of internet commerce). For instance, in the United State, the Federal Trade Commission requires sellers concluding transactions away from their premises to give buyers three days to cancel purchases of \$25 or

more, with the exception of some goods (such as arts or crafts) or services that are subject to other regulation (such as insurance). In the EU, the “Doorstep Selling” Directive 85/577/EEC protects consumers who purchase goods or services during an unsolicited visit by a seller at their doorstep (or otherwise away from the seller’s business premises). This regulation provides a cooling-off period of seven days, enabling the buyer to cancel the contract within that period and making the contract unenforceable if the buyer is not informed in writing of this right. Similar regulations are in place in most industrialized countries; see *supra* note 43, Annex E.

- 48 For instance, New York Insurance Department’s Regulation 60 on “Replacement of Life Insurance Policies and Annuity Contracts” grants buyers an unconditional cancellation right for sixty days. Insurance Commissioners in many U.S. states have adopted a model regulation issued by the National Association of Insurance Commissioners that mandates an unconditional refund period of, typically, thirty days for life insurance and annuity replacements.
- 49 For instance, New York State Bill A8965 extends the duration of the mandatory “free-look” period (during which the insured may pull out of a purchased insurance contract and obtain a refund) from thirty to ninety days for individual accident and health insurance policies or contracts that cover an insured who is 65 years of age or older on the effective date of coverage. Similarly, the Omnibus Budget Reconciliation Act of 1990 mandates a thirty day free-look period to allow beneficiaries time to decide whether the Medigap plan they selected is appropriate.
- 50 As an example of self regulation, the Financial Industry Regulatory Authority (FINRA), the major self-regulatory organization for securities firms operating in the United States, mandates that broker-dealers make a reasonable effort to obtain information about the individual characteristics of their (non-institutional) customers and to ensure that their recommendations are “suitable” to customers’ financial situation and needs. FINRA was formed in 2007 through a consolidation of the enforcement arm of the New York Stock Exchange, NYSE Regulation, Inc., and the National Association of Security Dealers (NASD). NASD Conduct Rule 2310(a) “Recommendation to Customers (Suitability),” originally adopted in 1939, prescribes: “In recommending to a customer the purchase, sale or exchange of any security, a member shall have reasonable grounds for believing that the recommendation is suitable for such customer upon the basis of the facts, if any, disclosed by such customer as to his other security holdings and as to his financial situation and needs.” Added in 1991, Rule 2310(b) “Broker’s Duty of Inquiry,” further requires: “Prior to the execution of a transaction recommended to a non-institutional customer, other than transactions with customers where investments are limited to money market mutual funds, a member shall make reasonable efforts to obtain information concerning: (1) the customer’s financial status; (2) the customer’s tax status; (3) the customer’s investment objectives; and (4) such other information used or considered to be reasonable by such member or registered representative in making recommendations to the customer.” In addition, NASD Rule 3010 imposes a duty of supervision on the firm employing the broker-dealer. See R.H. Mundheim, *Professional Suitability of Broker-Dealers: The Suitability Doctrine*, DUKE L. REV. 3, 445-480 (1965) for an early account of the suitability doctrine and for a discussion of the evolution of NASD Rule 2310. For a more recent overview of the main legal issues relating to the enforcement of suitability regulation, see L. Lowenfels & A.R. Bromberg, *Suitability in Securities Transactions*, BUSINESS LAWYER 54, 1557-1597, (1999).
- 51 See Inderst & Ottaviani, *supra* note 28 for a detailed discussion.
- 52 For instance, especially outside the prime mortgage market, lenders rather than intermediaries may be held responsible for establishing an “affordability standard,” even when they do not have direct contact with customers (see Financial Services Authority, *supra* note 17).
- 53 Carlin & Gervais present a formal analysis of the “team production” problem when both a product provider and an intermediary contribute towards making the supply of financial products suitable for consumers, see B. Carlin & S. Gervais, *Legal Protection in Retail Financial Markets*, Mimeo, (2008).
- 54 Selling without advice typically results in a loosening of regulatory requirements and of the liability risk for both intermediaries and product providers. For instance, in the United Kingdom only the buyers of advised mortgages have special rights of access to an Ombudsman.

Behavioral Economics, Consumer Protection, and Antitrust

Michael A. Salinger
Boston University

Behavioral Economics, Consumer Protection, and Antitrust

*Michael A. Salinger**

In both consumer protection and antitrust, the use of standard economic analysis has generally been to limit the scope of government intervention. The interest in behavioral economics (and some of the resistance to it) stems from the belief that it justifies intervention that conventional economic analysis suggests is unwarranted. Proponents see behavioral economics as the antidote to the Chicago School poison. Opponents see it as a mutated bacterium, resistant to the economic medicine that has led to improved policy. In this article, I will provide some background on behavioral economics and assess what insights it provides for consumer protection and antitrust policy.

*Michael Salinger is Professor of Economics and Everett W. Lord Distinguished Faculty Scholar at Boston University's School of Management. He is also Managing Director at LECG. From 2005-2007, he was Director of the Bureau of Economics at the FTC.

I. Introduction

“Behavioral economics” refers to economic analysis based on a richer model of individual behavior than the rational actor model underlying mainstream economic analysis.¹ The field has attracted widespread attention for its possible relevance particularly to consumer protection regulation but also to antitrust.

Intense interest in developments in the academic economics literature can have more to do with ideology than intellectual curiosity. The interest among policy makers in behavioral economics may be a case in point. In both consumer protection and antitrust, the use of standard economic analysis has generally been to limit the scope of government intervention. The interest in behavioral economics (and some of the resistance to it) stems from the belief that it justifies intervention that conventional economic analysis suggests is unwarranted. Proponents see behavioral economics as the antidote to the Chicago School poison. Opponents see it as a mutated bacterium, resistant to the economic medicine that has led to improved policy. In this article, I will provide some background on behavioral economics and assess what insights it provides for consumer protection and antitrust policy.

A discussion of behavioral economics must start with some background on economics as a discipline and the role of the assumption of rational behavior in it. Section II provides this background. Section III then describes the deviations from rational behavior documented in the behavioral literature. Section IV discusses analyses of markets in which some consumers are rational and others are not, focusing on the extent to which the presence of informed, rational consumers protects those that are poorly informed and/or irrational. Section V, which is divided into three sections, discusses public policy implications of the behavioral economics literature. One of the concerns about formulating policy based on behavioral approaches concerns how to articulate and impose limiting principles. The first subsection describes two proposals in the literature. The next two subsections then turn specifically to consumer protection and antitrust policy. Section VI briefly concludes by arguing that the behavioral economics literature is not likely to give current policy makers much insight. Consumer protection policy is arguably far ahead of the literature in recognizing how individuals are sometimes irrational and in considering the trade-offs in government intervention that takes irrationality into account. With antitrust, the behavioral economics literature may provide insight into the extent to which vigorous antitrust enforcement is sufficient to protect consumers (thus making more direct regulation unnecessary), but it does not provide much guidance on what antitrust interventions are appropriate.

THE INTEREST IN BEHAVIORAL ECONOMICS (AND SOME OF THE RESISTANCE TO IT) STEMS FROM THE BELIEF THAT IT JUSTIFIES INTERVENTION THAT CONVENTIONAL ECONOMIC ANALYSIS SUGGESTS IS UNWARRANTED.

II. Economics and the Role of Rationality

Common perception notwithstanding, economists do not believe that real human beings are rational optimizers. Most economic analysis concerns market phenomena, not individual behavior. Because market phenomena are inherently too complex to understand in every detail, economic analysis necessarily relies on simplifying assumptions that sacrifice realism for tractability. The rationality assumption plays so prominently in the literature because it is tractable (for people sufficiently proficient in mathematics) and yields some quite accurate predictions.² How accurate they are is subject to debate. Still, for predicting, say, how a mandate to increase the use of ethanol in gasoline will affect the price of corn, standard economics based on the elasticities of supply and demand is likely to be the best approach available; and richer assumptions about how people really behave are not likely to add much. This principle of making simplifying assumptions is not peculiar to economics. Cartographers do not believe that the land they are mapping is flat.

This justification for the role of rationality in economics is susceptible to three broad objections. The first concerns how well models based on the assumption of rational behavior in fact predict economic phenomena. To the extent that the justification for an unrealistic assumption lies in predictive accuracy, economists should be open to alternative assumptions that yield more accurate predictions.³

ECONOMISTS DO NOT MERELY
PREDICT AND EXPLAIN ECONOMIC
PHENOMENA; THEY ALSO ASSESS
WHETHER SOME ECONOMIC
OUTCOMES ARE MORE
DESIRABLE THAN OTHERS.

The second possible objection is that, unlike natural sciences, economic analysis has a normative as well as a positive dimension. Economists do not merely predict and explain economic phenomena; they also assess whether some economic outcomes are more desirable

than others. Arguably, this feature of economics is central to its influence on policy; and it is also the source of controversy. Even if the “predictive power” argument is persuasive with respect to the rationality assumption for positive economic analysis, the normative conclusions about market outcomes do not follow as a matter of pure logic.

A third objection is that the rationality assumption has surprisingly little empirical content in the sense that it is hard to refute. A vivid example that illustrates the point is that pedestrian accidents involving foreigners are a bigger problem in London than in most cities because visitors from countries where cars drive on the right side of the road often look the wrong way when crossing the street. This behavior is an example of what Herbert Simon referred to as “bounded rationality.”⁴ It reflects a purposeful pursuit of self-interest. (The pedestrians looking to the left are trying to avoid being hit by a car.) But it is based on “heuristics” or “rules of thumb,” which are mental short cuts people use to make decisions that they do not have either the time or the mental capacity to think

through completely rationally. Remarkably, though, by introducing a personal preference for looking to the left, it is hard to reject the hypothesis that the pedestrian behavior is the rational pursuit of self-interest.⁵

Without a willingness to reject a “revealed preference” as being irrational, the only way to reject rational behavior is to observe actual inconsistencies. Even this is hard to do with actual behavior, as apparently inconsistent choices made at different times might reflect a change in preferences. However, it is possible to demonstrate inconsistencies in choices in laboratory settings. As will be discussed in the next section, the focus of much of the behavioral literature has been to document inconsistencies.

The combination of the three objections to the role of the rationality assumption in economics is one of the explanations for why the debates about the relative value of “Chicago-school” and behavioral approaches can be ideological. Much of the objection to economics concerns the normative conclusions about market outcomes that economists draw. But the difficulty of testing the assumptions underlying normative conclusions makes it hard to resolve scientifically the dispute among people with different predispositions about the efficiency of market outcomes and the prospects for government intervention to improve upon them.

There is another (and related) difficulty associated with resolving rationally the debate about what model of individual behavior should underlie economic analysis. The debate over the rationality assumption is likely a proxy for a more nuanced issue. If individuals behave rationally, voluntary market exchange makes both parties to a transaction better off. If so, then government intervention is unwarranted in the absence of externalities. While this argument preserves some role for the government,⁶ it carves out a substantial fraction of economic activity that the government should leave alone. A proper role for government intervention becomes the relatively rare exception, not the norm.

But a dogmatic belief in the rationality of consumer decisions may not be the main basis for beliefs in limits to government intervention.

Consider policy toward smoking, behavior that is arguably not much different from the behavior of pedestrians from foreign countries in London. In contrast to a preference for looking to the left, however, some people do get pleasure from smoking; it is therefore harder to conclude as a matter of economic science that the decision to enjoy the short-term benefit and accept the long-term risk is irrational. Yet, in the United States and no doubt elsewhere, there is likely a consensus that starting to smoke is irrational and that it would be desirable to prevent anyone else from starting to smoke (and to help all current smokers break

BUT A DOGMATIC BELIEF IN THE RATIONALITY OF CONSUMER DECISIONS MAY NOT BE THE MAIN BASIS FOR BELIEFS IN LIMITS TO GOVERNMENT INTERVENTION.

the habit) if doing so were practical. Yet, there is no consensus among either the public or among professional economists—including behavioral economists—that smoking should be banned. Indeed, there may well be a consensus against a ban on tobacco products as being impractical. As with alcohol under Prohibition, a ban on smoking would likely result in a black market for cigarettes. Supply in that market would be provided by criminal organizations. Vigorous enforcement of the laws would result in prison terms for people who would not otherwise engage in criminal activity. Without vigorous enforcement, the laws would be entirely ineffective (which would reduce general respect for the law). The argument against a ban on smoking is more about the limits of government effectiveness than about consumer rationality.

While there may be a consensus that governments can legitimately discourage smoking by taxing it and disseminating information, there is likely substantial unease about how far the government should go in helping people make better decisions. What should government policy be toward decisions to drink alcohol or eat at fast food restaurants, activities that most people are able to enjoy in moderation without suffering substantial long-term harm? Yet, there is evidence that some people end up making choices they ultimately regret with respect to both. Is there a governmental role in preventing alcoholism or overeating and what should that role be? If it does have such a role, how extensive should it be? And if the government has a role in influencing people's smoking and diet decisions, what else might it involve itself in? Government policies to alter individual choices would generally affect commercial interests and therefore become the source of lobbying and political contributions. Much of the controversy surrounding behavioral economics is not about

whether individuals are rational but over the competence of government to protect people against their own irrationality.

THE QUESTION THAT NEEDS TO BE
ASKED ABOUT THE BEHAVIORAL
ECONOMICS LITERATURE IS
WHETHER IT CLARIFIES THE
LIMITING PRINCIPLES THAT DIVIDE
WHERE GOVERNMENT
INTERVENTION YIELDS BENEFITS
THAT EXCEED THE COSTS.

In evaluating the literature on behavioral economics, the issue of whether it proves that people in fact behave irrationally is a straw man. It is not just that we already know that people behave irrationally. It is that policy has long recognized that people sometimes behave irrationally and that government has some role

in modifying individual behavior. The question that needs to be asked about the behavioral economics literature is whether it clarifies the limiting principles that divide where government intervention yields benefits that exceed the costs.

III. Deviations from the Rationality Assumption

If, by behavioral economics, one means any economic analysis in which individuals are not assumed to be both perfectly informed and rational, the field is vast

and a complete review is far beyond the scope of this article. The literature on deviations from rationality is somewhat more limited (although not as new as some might suspect), but imperfect consumer information is at the heart of a great deal of consumer protection policy; and some of the modern behavioral economics literature concerns the effect of imperfect information as opposed to irrational responses to the information people have.

A. IMPERFECTLY INFORMED CONSUMERS

A critic of the existing economics literature of his day once wrote, “One should hardly need to tell academicians that information is a valuable resource: knowledge is power. And yet it occupies a slum dwelling in the town of economics. Mostly it is ignored: the best technology is assumed to be known; the relationship of commodities to consumer preferences is datum.” One might suspect that this quote comes from a noted Chicago school critic. In fact, it is the opening sentence to George Stigler’s *The Economics of Information*.⁷

Once one recognizes that information is itself a scarce good, several questions naturally follow. First, to what extent will individuals devote the right amount of effort to becoming informed? Second, do firms have the right incentives to provide consumers with information that will help them make rational choices?⁸ Third, do firms have an incentive to provide false information to consumers? Fourth, what role should the government play in the information consumers receive? Should it regulate private efforts and/or should it be an independent source of information? Fifth, is it ever more efficient for the government to make decisions for consumers (rather than rely on informed consumer decisions) and, if so, what are the principles underlying those cases?

IS IT EVER MORE EFFICIENT
FOR THE GOVERNMENT TO MAKE
DECISIONS FOR CONSUMERS
(RATHER THAN RELY ON
INFORMED CONSUMER DECISIONS)
AND, IF SO, WHAT ARE
THE PRINCIPLES UNDERLYING
THOSE CASES?

B. DEVIATIONS FROM RATIONALITY

When James Tobin won the Nobel Prize in economics in part for his theory of portfolio choice, he was asked at a news conference to explain his work in layman’s terms. After repeated attempts to get him to further simplify his explanation, he said that his Nobel Prize was for explaining the principle that investors should not put all their eggs in one basket. That prompted a cartoon of a spokesman for the Nobel committee announcing a subsequent prize for demonstrating that “An apple a day keeps the doctor away.” When stated as simply as possible, academic advances particularly in the social sciences can sound trivial.

Those who have heard of behavioral economics and suspect it must have great value but have not yet read the literature might have a similar reaction to the ways in which the literature has demonstrated deviations from rational behavior.

In the behavioral economics literature, there are three major ways in which individual behavior deviates from rationality: bounded rationality, incomplete self-interest, and incomplete self-control.⁹

1. Bounded Rationality

As described above, bounded rationality means that individuals (or firms) act purposefully, but not necessarily as if they are both fully informed and perfectly rational. In the bounded rationality literature, two deviations from rational behavior play prominently. The first is that individuals exhibit systematic biases

when making decisions under uncertainty. The second is that the decisions people make depend on how they are “framed.”¹⁰

IN THE BEHAVIORAL ECONOMICS LITERATURE, THERE ARE THREE MAJOR WAYS IN WHICH INDIVIDUAL BEHAVIOR DEVIATES FROM RATIONALITY: BOUNDED RATIONALITY, INCOMPLETE SELF-INTEREST, AND INCOMPLETE SELF-CONTROL.

Neither of these results is new. Both date back at least to the work of Allais, who developed pairs of choices that demonstrated the inconsistency of choices people make. In what is known as “Allais’ Paradox,”¹¹ people are presented with two sets of choices. The first choice is between a certain outcome of \$1 million and a random outcome of an 89 percent chance of

\$1 million, a 10 percent chance of \$5 million, and a 1 percent chance of \$0. Most respondents choose the sure option of \$1 million.

The second choice is between an 11 percent chance to get \$1 million and a 10 percent chance to get \$5 million (and \$0 the rest of the time). Most people say they prefer the 10% percent chance at \$5 million.

The paradox is that the choices are inconsistent with each other. In both cases, the choice the individual makes is irrelevant 89 percent of the time. (In the first case, he gets \$1 million whether he takes the safe or the risky outcome. In the second case, he gets \$0 regardless of his choice). Thus, the choice only matters the other 11 percent of the time. In both cases, the first choice gives \$1 million over the entire 11 percent probability that the choice matters. In the second choice, the 11 percent chance of the choice mattering is divided between a 10 percent chance of \$5 million and a 1 percent chance of \$0. Neither choice is inherently irrational, but the choice between them cannot rationally depend on what the individual gets the 89 percent of the time that the choice does not matter.¹² Allais’ paradox is an example of both themes. The irrational decision concerns risk and it appears that the decision is affected by how it is framed.¹³

The work of Daniel Kahneman & Amos Tversky, for which Kahneman received the Nobel Prize in economics in 2002,¹⁴ was another important antecedent to the modern behavioral economics literature. In a seminal article published in *Science* entitled *Judgment under Uncertainty: Heuristics and Biases*,

they explored ways in which individual decisions under uncertainty deviate systematically from rationality and categorized the deviations as arising from three sources of bias: “representativeness,” “availability,” and “adjustment and anchoring.”¹⁵ “Representativeness” concerns how people assess the relative probabilities of possible explanations for information they are given. They tend to treat all possibilities as being equally likely *ex ante*, and then judge the relative *ex post* probabilities based just on how representative the facts were of the candidate explanations.¹⁶ “Availability” is the phenomenon that people assess the relative likelihood of events based on their ability to think of examples. “Anchoring” refers to the phenomenon that people sometimes solve problems by starting with some reference point and then making an adjustment to it.

Thaler & Sunstein illustrate this anchoring phenomenon with the example of asking people to estimate the size of Milwaukee, Wisconsin, which is the biggest city in Wisconsin. Residents of Chicago on average give a higher estimate than residents of Green Bay (a small city in Wisconsin). Thaler & Sunstein’s explanation is that residents of Chicago start with the population of Chicago and adjust downward. Residents of Green Bay start with the population of Green Bay and adjust upward. Both, however, make adjustments that are too small. This phenomenon of anchoring and adjustments suggest how the framing of a question can affect responses, since the question can be asked so as to provide an anchor.¹⁷

Modern work on bounded rationality has emphasized several behavioral biases. One is “status quo bias,” which is a bias toward inaction.¹⁸ An implication of status quo bias is that the choices people make are affected by defaults. For example, voluntary contributions to savings plans depend on the default option. If people were fully rational, their contributions would not depend on whether they have to sign up for the plan in order to participate or opt out if they wish not to participate.¹⁹

Two points about the bounded rationality detected in the literature are worth noting. First, as described in the previous section, proving irrationality without making value judgments about what constitutes a rational preference is difficult. Many of the proofs of irrationality have been proofs of inconsistency. (The other way is to show that people fail to account for objective laws of probability.) These are certainly forms of irrationality, but they may not be the most important ways in which people behave irrationally. Second, the definition of irrationality focuses on systematic biases. For example, people may systematically overestimate the risk of some types of rare events. An exclusive focus on how the average response differs from the correct response necessarily misses cases in which people are right on average but some individuals make large errors in both directions.

PROVING IRRATIONALITY
WITHOUT MAKING VALUE
JUDGMENTS ABOUT WHAT
CONSTITUTES A RATIONAL
PREFERENCE IS DIFFICULT.

Consider whether there is need for government regulation of skydiving. Obviously, skydiving entails some risk, but some people may enjoy it enough to justify taking the risk. Suppose some people substantially overestimate the risk of skydiving while others underestimate it with the average assessment being approximately correct. Is government intervention warranted? Those who substantially overestimate the risk will presumably choose not to jump from an airplane in flight. They might end up foregoing something they would enjoy, but few would suggest that the government should educate people that skydiving may well be safer than some of them realize. Those who underestimate the risk might, however, make a catastrophic decision. Whether or not people misestimate the risk on average is less important than whether some people substantially underestimate it.

If the skydiving example seems far-fetched, similar issues arise with respect to loans with severe default penalties. One of the systematic biases documented in the literature is that people are on average overly optimistic about their own prospects. Poll a class of students at the start of the semester about what grade they expect in the class, and the resulting distribution will entail more high grades than the actual distribution. Ask a group of people with a particular type of loan what they think their probability of default is; the average answer may well be lower than what the lender knows to be the statistical probability of default. Whatever role the government might have in protecting people from taking out loans for which they do not understand the full consequences, the need for such a policy does not require bias. It just requires that some people make large errors.

2. Imperfect Self-Control

The second major deviation from rationality documented in the behavioral economics literature is imperfect self-control. As with bounded rationality, the demonstration of imperfect self-control is from an inconsistency between choices people actually make in the short run and the decisions they say they would make. For example, suppose someone is told that they can either have \$100 in one year or \$101 in one year plus a day. Most people prefer the latter. Then, suppose after a year, they are told they can have the \$100 immediately or \$101 the next day. Many people opt for the \$100 immediately. The inconsistency is that both choices entail choosing whether their one-day discount rate is greater or less than 1 percent.

Moving outside experimental settings, the phenomenon of imperfect self-control can explain why people decide to limit the options available to them. For example, some banks used to have Christmas Clubs, which were savings accounts that allowed customers to retrieve the funds only near Christmas. Since they did not yield higher returns than regular savings accounts, it is not clear why anyone with complete self-control would place limits on when they could retrieve the funds.²⁰ Models of lack of self-control assume that people operate in

two different states—a hot state and a cool state. When in a cool state, they recognize that they will make choices in a hot state that they would regret in a cool state. As a result, they take actions to restrict the choices that might be available to them in the hot state.

3. Incomplete Self-Interest

Lack of self-interest means that individual decision-making often reflects an interest in the well-being of others, perhaps due to some underlying sense of fairness. A well-known example concerns experimental results in the “ultimatum game.”²¹ Two people, John and Mary, are given \$10 to divide between them. The rules of the game are that John gets to propose a split to Mary. Mary can either accept or reject it. If she accepts, each player gets the split proposed by John. If she rejects, they both get \$0. In the fully rational solution to the game, John should offer to give Mary \$0.01, taking \$9.99 for himself. Mary would rationally accept since \$0.01 is better than \$0.00. When the game is played experimentally, however, the most common outcome is that the player who gets to make the ultimatum offers a more even split of the money. Whether John offers Mary more than \$0.01 because of his own sense of fairness or whether he is merely protecting himself from the possibility that Mary would “irrationally” reject the \$0.01 (perhaps based on her own sense of fairness), the model of fully rational behavior is at substantial odds with extensive experimental evidence.

4. Reprisal

Reduced to very simple terms, the three deviations from rational behavior documented in the behavioral economics literature come down to this: the average IQ is only 100, people sometimes make short-run decisions they regret,²² and people sometimes act selflessly and with a sense of fairness.

REDUCED TO VERY SIMPLE TERMS,
THE THREE DEVIATIONS FROM
RATIONAL BEHAVIOR
DOCUMENTED IN THE
BEHAVIORAL ECONOMICS
LITERATURE COME DOWN TO
THIS: THE AVERAGE IQ IS ONLY
100, PEOPLE SOMETIMES MAKE
SHORT-RUN DECISIONS THEY
REGRET, AND PEOPLE SOMETIMES
ACT SELFLESSLY AND WITH
A SENSE OF FAIRNESS.

C. ACADEMIC CONTRIBUTIONS OF BEHAVIORAL ECONOMICS

To be sure, the previous paragraph over-simplifies behavioral economics and its contributions. The quote from James Tobin provides a suitable analogy. Tobin did not discover the principle that people should not “put their eggs in one basket” or, to put the matter less colloquially, that people should diversify their portfolios. His contribution was to develop a tractable mathematical approach to modeling such behavior.

A similar point can be made about the behavioral literature that suggests that some people make short-run decisions with respect to current consumption and savings that they ultimately regret. The literature on “hyperbolic discounting”

concerns a mathematical representation of such behavior. The mathematical representation is itself a contribution as it may lay an important foundation for models that predict economic variables (like savings and interest rates) more accurately and some of that analysis might ultimately have policy implications.

THE AREAS OF ECONOMICS WHERE
BEHAVIORAL APPROACHES HAVE
HAD THE BIGGEST EFFECT HAVE
BEEN FIELDS WITH DOCUMENTED
ANOMALIES TO EXPLAIN. THE
LEADING EXAMPLE IS FINANCE.

The areas of economics where behavioral approaches have had the biggest effect have been fields with documented anomalies to explain. The leading example is finance. Because it is the field of economics with the most abundant data, it has a relatively large

number of anomalies (or, at least, alleged anomalies). While some of these anomalies, such as the equity premium puzzle²³ and the excessive volatility of stock prices,²⁴ are controversial, enough scholars believe them to be true that they have explored whether behavioral phenomena could explain them. Another field in which behavioral approaches have been influential is macroeconomics. One of the puzzles in macroeconomics is the low rate of saving in the United States. The literature on hyperbolic discounting is an obvious explanation to consider once rational explanations have been exhausted.

In 2005, a high-profile conference in Helsinki was devoted to behavioral approaches to areas of economics besides finance and macroeconomics. The organizers of the conference did not include a paper on industrial economics because, as conference organizer Peter Diamond explained, “There was no behavioral industrial organization covered at the conference because there is not yet an audience there.”²⁵ We will turn to this observation below in the discussion of the implications of the behavioral economics literature for antitrust.

IV. Do Markets Protect Irrational Consumers?

You find yourself in the unfamiliar city Nowheresville for the evening and it is time for dinner. You come across a restaurant called “Joe’s Diner.” You have no other information about it other than the sign, “Proudly Serving Nowheresville for over 15 years.” How confident can you be that you will not regret the decision to eat there? How persuasive is the argument that Joe’s only could have survived in the market for over 15 years if its regular customers who live in Nowheresville like the food? And if the argument is persuasive with respect to Joe’s, how general is the argument that the market protects uninformed customers because only businesses that provide service that informed customers like can survive?

Analysis of markets with some informed and some uninformed customers falls under the label “behavioral economics.” For example, in a financial market with some rational and some irrational investors, do market prices reflect the informa-

tion of informed investors? At one point, it had been argued informally that informed investors would be able to take advantage of any mispricing and that they would ultimately drive uninformed investors from the market. Recent theoretical analysis in finance has shown that this is not necessarily the case.²⁶

With respect to non-financial markets, Salop & Stiglitz demonstrated that in a market with some informed and some uninformed buyers, businesses offering bad deals can survive.²⁷ No informed customers would shop there, but some uninformed customers would; and the prospect of selling to some uninformed buyers could make it possible for a business offering relatively bad deals to survive.

The more recent behavioral economics literature that falls into this category is the work on “shrouding,” or hidden charges.²⁸ An example of a “shrouded” charge is the price rental car companies charge for refueling cars returned without a full tank. Informed customers can avoid the charge by refueling just before returning the car. But “uninformed”²⁹ customers end up buying gasoline at far above the market price. The prospect that some customers will pay the high refueling charges gives the rental companies an incentive to offer lower daily rates than they otherwise could. Informed customers get a good deal on the daily rate. But those who pay for refueling might end up paying much more than they expected.

THE INSIGHT FROM THE SHROUDING LITERATURE IS THAT THE COMBINATION OF INFORMED CUSTOMERS AND COMPETITION DO NOT NECESSARILY PROTECT THE UNINFORMED CUSTOMERS.

The insight from the shrouding literature is that the combination of informed customers and competition do not necessarily protect the uninformed customers.³⁰ The shrouded charges are a way of giving informed customers a good deal and uninformed customers a bad deal. Thaler & Sunstein illustrate the point with extended service contracts for electronic equipment, which are generally offered on unattractive terms. Informed customers refuse the offer, uninformed customers do not. The prospect of selling extended service contracts to uninformed customers induces electronics retailers to offer base prices below what they otherwise would. Informed customers get a good deal, so the inability to attract informed customers does not drive the retailers out of business. As Thaler & Sunstein explain, it is hard to make money by explaining to people that they do not need something that they would like to purchase. Similar issues can arise with virtually any form of credit (credit cards, mortgages, auto loans), with attractive base terms but high penalties for late payments, cell phones (with charges for extra minutes being shrouded), hotels (with charges for parking, internet access, and telephone usage being shrouded), video rentals (with late charges being shrouded), and so on.

V. Policy Implications

A. LIMITING PRINCIPLES

As discussed above, the question of whether real people are fully rational is a straw man. Skepticism about the policy implications of behavioral approaches stems less from a faith in human rationality than from a recognition that sources of irrationality are abundant. Ideally, policy implications would be based on a careful weighing of the costs from consumer irrationality and government imperfections. Since such an analysis is typically impractical, behavioral approaches to policy must admit some more practical limiting principle.

This point is not merely theoretical. In the 1970's, the Federal Trade Commission ("FTC") entered into a series of rule makings that at least implicitly embodied a behavioral economics approach. A particularly notorious example was the so-called "kid-vid" proceeding, in which the Commission sought to ban

all advertising to children as an unfair trade practice. The Commission was subjected to substantial public and political criticism for these efforts. The Washington Post ran an editorial accusing the FTC of being a "nanny."³¹ For many years, Congress failed to pass the FTC authorization. There was serious discussion of eliminating the agency. One manifestation of the self-evaluation that emerged from that period was a major change in how the Agency defined an unfair trade practice.³²

IDEALLY, POLICY IMPLICATIONS
WOULD BE BASED ON A CAREFUL
WEIGHING OF THE COSTS FROM
CONSUMER IRRATIONALITY AND
GOVERNMENT IMPERFECTIONS.
SINCE SUCH AN ANALYSIS IS
TYPICALLY IMPRACTICAL,
BEHAVIORAL APPROACHES TO
POLICY MUST ADMIT SOME MORE
PRACTICAL LIMITING PRINCIPLE.

One approach to limiting principles put forward by Camerer, Issacharoff, Loewenstein, O'Donoghue, & Rabin (henceforth, Camerer et al.) is what they refer to as "asymmetric paternalism."³³ Paternalism necessarily entails policies that prevent people from taking actions they would like to take, based on the rationale that some people are either sufficiently badly informed or irrational enough to need protection from themselves. A fundamental risk with paternalistic policies is that in trying to prevent people from taking actions they regret, they prevent people from taking actions that do in fact reflect their rational self-interest. The "asymmetry" in the term "asymmetric paternalism" refers to the difference in how policies affect the informed and the uninformed. The ideal asymmetrically paternalistic policy allows rational, informed individuals to make the choice they want and prevents irrational or uninformed individuals from making mistakes.³⁴

Based on this ideal, Camerer et al. categorize interventions into four categories reflecting different degrees of intrusiveness. These are defaults, framing and information disclosure, cooling off periods, and restriction of consumer choice.

An example of a default might concern magazine or newspaper subscriptions. When a subscriber buys a one-year subscription with a credit card, the magazine might give the consumer the option of automatically renewing the subscription, requiring the subscriber to go to the effort to cancel the subscription. Since some consumers might prefer automatic renewal to avoid the effort of re-subscribing, banning automatic renewal is arguably overly intrusive. However, one might argue that public policy should force magazines to make non-automatic renewal the default, allowing those who want automatic renewal to opt in. Changing the default in principle does not prevent anyone with an informed (or, for that matter, uninformed) preference from choosing his preferred option.

The second least intrusive intervention according to Camerer et al. is information provision or framing. An example is the requirement in the Truth in Lending Act that credit providers must state loan terms as an annual percentage rate, or “APR.” The rationale is that many consumers might misinterpret loan terms (by, for example, thinking that a 1 1/2 percent monthly credit card interest rate is low by failing to distinguish between a monthly and annual rate of interest). The disclosure does not hurt in any significant way those who understand loan terms but might help those who do not.³⁵

The third in the Camerer et al. hierarchy of intrusiveness is cooling-off periods, which are mandated delays. An example would be a delay of a few days between when a couple applies for a marriage license and when the state issues it. The common-sense motivation for such rules is to prevent people from rushing into such a long-range commitment in the heat of passion. The underlying foundation in the psychological literature is the distinction between decisions made in “hot” and “cool” states. To the extent that delays are short, they would appear not to impose significant costs.

The fourth and most intrusive level in the Camerer et al. hierarchy is actual (and permanent) limitations on the choices (by making certain transactions illegal). Examples would include bans on products (or employment conditions) considered to be too unsafe.

Related to the Camerer et al. definition of “asymmetric paternalism” is Thaler & Sunstein’s advocacy of “libertarian paternalism,” a term that many will consider to be an oxymoron. (Thaler is an *eminence grise* of the behavioral economics literature. Sunstein, in addition to being a noted scholar in the area, is currently Director of the Office of Information and Regulatory Affairs (“OIRA”) in the Office of Management and Budget.) They discuss the concept extensively in their recent book, *Nudge*.³⁶ The emphasis of the book is cases where framing seems to have a significant effect on individual behavior. Examples include changing the placement of milk relative to soda in school cafeteria lines and changing the default provisions on contributions to retirement plans by requiring people who do not want to participate to opt out rather than requiring those who do want to participate to opt in.

DESPITE THE SIMILARITY OF THE
TERMS THEY USE, THE LIMITATION
INHERENT IN “LIBERTARIAN
PATERNALISM” IS MUCH
DIFFERENT FROM THOSE IN
“ASYMMETRIC PATERNALISM.”

Despite the similarity of the terms they use, the limitation inherent in “libertarian paternalism” is much different from those in “asymmetric paternalism.” The limitation in “libertarian paternalism” entails the form of the intervention. By focusing on framing and defaults, the intervention does not prevent anyone who is determined on a particular course of action from doing so. There is nothing

inherent in this conceptualization that limits the situations in which the government might try to intervene. The concept seems innocuous when applied to inducing students to eat healthier food because there is likely a strong consensus (at least among adults not in the business of selling fat- and sugar-laden foods) that healthier diets for students would be desirable. Without strong principles to limit

when it is appropriate for the government to try to (gently) engineer individual behavior, however, the term “libertarian paternalism” is unlikely to assuage concerns that it is simply a way of framing “big brother” to make it seem innocuous.

B. INSIGHTS INTO CONSUMER PROTECTION POLICY

Consumer protection policy existed before the renewed interest in behavioral economics. A question to ask about the literature is what insight it provides into that policy.³⁷

Some of the existing policy fits into Camerer et al.’s fourth category of intervention. An example is FDA approval of drugs and the requirement that people obtain a prescription for some drugs. The presumed rationale for FDA regulation is that drug safety and efficacy are difficult for individuals to evaluate and there is relative homogeneity of preferences to avoid drugs that are unsafe and/or ineffective. In broad terms, FDA regulation is consistent with the new literature. In this regard, the literature might provide a useful framework for understanding why regulation is what it is, but it does not necessarily yield insights into how to improve it.

One might suspect that the new literature might yield more guidance for enforcement of the more amorphous consumer protection mission of the FTC, which is allowed to bring actions against “unfair or deceptive acts or practices.” Again, though, the literature might do more to explain why FTC enforcement has developed the way it has than to provide insights into how it might improve policy. Consumer protection enforcement at the FTC has always (or at least for decades) been premised on the notion that consumers sometimes behave irrationally. If they did not, the Commission would not need to devote scarce enforcement resources to actions against sellers of products that promise weight loss without either diet modification or exercise (but with payment). In its regulation of advertising, it has long been aware of what advertisers also understand—that framing affects the choices people make. The challenge for the FTC

is to identify when the framing companies do cross the line of being either deceptive or unfair. It is not clear that the behavioral economics literature has advanced to the point where it can help the FTC identify the boundary more accurately.

Formal economics plays less of a role in consumer protection enforcement at the FTC than in antitrust enforcement. Of the several reasons why this is the case,³⁸ one is that the formal behavioral economics underlying the consumer protection mission is relatively underdeveloped. While it might seem obvious that further development of the behavioral economics literature should help improve policy, there are risks as well. In a discussion of the implications of behavioral economics to consumer protection regulation, Jolls argues that price ceilings such as bans on usury or price gouging reflect social norms against letting markets work.³⁹ Her argument may well be persuasive as a matter of positive law and economics. That is, it might explain why such laws have been passed. Most economists would view this behavioral view of the law as an argument for conventional economics as a tool for avoiding harmful economic policy. Jolls argues for an alternative use of behavioral economics in law and economics. She sees a role for legal standards to “debias” consumer decisions. In other words, she sees legal standards as protecting consumers from their own bad decisions.⁴⁰

FORMAL ECONOMICS PLAYS LESS
OF A ROLE IN CONSUMER
PROTECTION ENFORCEMENT
AT THE FTC THAN IN
ANTITRUST ENFORCEMENT.

The recent financial crisis provides an interesting case to evaluate whether the behavioral economics literature yields insights into how to improve policies. The FTC shares responsibility with the Federal Reserve Board for enforcing Section 5 of the FTC Act with respect to financial products. In that capacity, it has attacked predatory lending practices as well as deceptive practices with respect to a wide variety of financial products from home and car loans to credit cards and payday loans. With respect to home loans, it is widely accepted that one of the factors that precipitated the financial crisis was that home loans were given to people who had high probabilities of default, particularly if housing prices fell. This raises the question of whether the crisis could have been avoided with stronger consumer protection policies that would have prevented people from taking out loans that they ultimately regretted.

The issue is not whether it would have been better if those loans had not been granted. Rather, it is whether the way to prevent such loans in the future is through consumer protection regulation and whether the failure to impose the appropriate regulations to prevent the crisis stemmed from a failure to recognize that individuals are not always rational profit-maximizers. In evaluating those issues, one needs to consider the behavior of the lenders as well as the borrowers. It should come as a surprise to no one that individuals are willing to take out loans that they have a poor chance of repaying. Whether such behavior is ration-

al or irrational, this is why banks have loan officers and why borrowers have to fill out financial disclosure forms in order to get a loan.

Why lenders issued the loans is a bigger puzzle than why borrowers were willing to take out loans for houses they could only afford if housing prices continued to rise. Some pieces of the puzzle are understood. Mortgage brokers were willing to arrange for bad loans because they received origination fees and did not bear the default risk. Why investors were willing to buy the securitized loans is more of a puzzle. At least part of the answer is that they relied on the major ratings agencies. Why the ratings agencies miscalculated the risk so badly is also subject to debate. They might have faced perverse incentives to understate risk. They might have made honest errors in the difficult science of predicting the probability of rare events. The deviations from rationality documented in behavioral economics literature do not make the short list of likely explanations for what went wrong at the ratings agencies.

THIS RAISES THE QUESTION OF
WHETHER THE CRISIS COULD HAVE
BEEN AVOIDED WITH STRONGER
CONSUMER PROTECTION POLICIES
THAT WOULD HAVE PREVENTED
PEOPLE FROM TAKING OUT
LOANS THAT THEY
ULTIMATELY REGRETTED.

Indeed, looking at the crisis as a whole, a common explanation is that a failure in regulation and corporate governance created situations in which individuals could make large “heads-I-win-tails-you-lose” bets. The fact that, confronted with such incentives, people in financial institutions took those bets is exactly what a model based on the rational pursuit of self-interest would predict.

C. IMPLICATIONS FOR ANTITRUST

As noted above, a recent conference on behavioral approaches to areas of economics besides finance and macroeconomics (the two fields where the applications are relatively well established) did not include any discussion of industrial economics because there has been virtually no interest within modern industrial economics in applying behavioral approaches. In light of that observation, the interest in behavioral approaches to antitrust economics might seem puzzling.

In broadest terms, one of the benefits claimed for antitrust enforcement is that it obviates the need for more direct and onerous regulation. If competition is sufficient to ensure that firms act in consumers’ interests, direct government involvement is unnecessary. The literature on shrouding highlights ways in which competition may not be sufficient to protect all customers. Arguably, therefore, the shrouding literature provides insights into the boundary between the reliance on antitrust and regulation.

What implications the behavioral economics literature has for how antitrust laws should be enforced is less clear. Superficially at least, shrouding relates to

bundling and tying. However, the behavior that allows firms to exploit uninformed customers is typically unbundling. For example, a relatively recent trend in the airline industry is the unbundling of the checking of bags. Interestingly, Southwest Airlines has run a major advertising campaign touting that on Southwest, “Bags fly free,” meaning that Southwest ties the right to check bags to a passenger seat. If the tying were anticompetitive rather than a form of competition, it is unlikely that Southwest would be advertising the point.

One of the general reasons to be skeptical that the behavioral economics literature has important implications for antitrust is that it has focused on how individual behavior is irrational. Ironically, there is a much older behavioral economics literature, dating back to Herbert Simon’s work on bounded rationality and the 1962 publication of Cyert & March’s *A Behavioral Theory of the Firm*. While that literature remains influential in various business school disciplines (such as organizational behavior), industrial economists have largely ignored it for decades.⁴¹

A long-standing issue in the literature on firm behavior is whether mergers reflect managerial objectives rather than those of shareholders.⁴² In turn, the issue is part of a broader issue about the quality of corporate governance.⁴³ While recent events have perhaps provided more reason to doubt the effectiveness of corporate governance, the implications for merger review are not clear.⁴⁴ Under the assumption that management is constrained to operate in shareholders’ interests, the objective of merger review is to assess the relative likelihood of competitive and anticompetitive explanations for the increase in value anticipated from a merger. The possibility that a merger might destroy value arguably reduces the likelihood that it is intended to create monopoly power, but it also limits the concern about a merger challenge standing in the way of preventing efficiencies.

A LONG-STANDING ISSUE IN THE LITERATURE ON FIRM BEHAVIOR IS WHETHER MERGERS REFLECT MANAGERIAL OBJECTIVES RATHER THAN THOSE OF SHAREHOLDERS.

VI. Conclusions

My objective in this article has been to provide some background for policy makers who are wondering what behavioral economics is and what implications they should draw from it.

My guess is that most policy makers who were hoping for great insights will be disappointed.

The behavioral literature does not address the trade-off that consumer protection policy makers routinely confront. For economic policy in general and consumer protection policy in particular, the assumption of perfectly rational behavior is a straw man. Once one digs through the jargon used to describe them, the deviations from rationality that have been documented are not very surprising.

Real people are not geniuses. Real people sometimes act impulsively. Policy makers already know this. But real people are not idiots, they have mechanisms for dealing with their own impulsiveness, and government intervention is necessarily costly and imperfect. Consumer protection policy makers were struggling with

BUT REAL PEOPLE ARE NOT IDIOTS,
THEY HAVE MECHANISMS FOR
DEALING WITH THEIR OWN
IMPULSIVENESS, AND GOVERNMENT
INTERVENTION IS NECESSARILY
COSTLY AND IMPERFECT.

how to balance bounded rationality and imperfect government intervention long before behavioral economics became a hot topic. At this point, the literature is catching up to the policy issues rather than providing insights that can lead it.

With respect to antitrust, the focus of the new behavioral literature on individuals, not firms, severely limits the insights it can provide. As described in the previous section, there are perhaps a few implications that touch on antitrust, but they are not central to the major antitrust debates of the day. ▼

- 1 The field of behavioral economics is quite vast and a complete survey of the literature is far beyond the scope of this article. For those seeking a more extensive review, see C. Camerer, S. Issacharoff, G. Loewenstein, T. O'Donoghue, & M. Rabin, *Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism,"* 151 U. PA. L. REV. 1211 (2003); P. Diamond & H. Vartiainen (eds.) *BEHAVIORAL ECONOMICS AND ITS APPLICATIONS* (2007); and R. Thaler & C. Sunstein, *NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS* (2008). In the Diamond & Vartiainen volume, the chapter by Bernheim & Rangel provides a clear and succinct summary of ways in which individual behavior has been documented to deviate from rationality and the issues in how to incorporate such behavioral factors into economic models. See B.D. Bernheim & A. Rangel, *Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision-Makers*, *BEHAVIORAL ECONOMICS AND ITS APPLICATIONS* (P. Diamond & H. Vartiainen eds.) (2007).
- 2 The classic statement of this position is M. FRIEDMAN, *ESSAYS IN POSITIVE ECONOMICS* (1953).
- 3 Predictive accuracy was one of the justifications Cyert & March gave for their behavioral theory of a firm. See R.M. CYERT & J.G. MARCH, *A BEHAVIORAL THEORY OF THE FIRM* (1963).
- 4 See, H. SIMON, *MODELS OF MAN* (1957) and H. SIMON, *MODELS OF BOUNDED RATIONALITY*, VOLS. 1 AND 2 (1982).
- 5 The observation that the behavior is technically consistent with revealed preference and a strong taste for looking to the left is due to Bernheim & Rangel. See B.D. Bernheim & A. Rangel, *supra* note 1. Thaler & Sunstein also discuss the London pedestrian example. See R. THALER & C. SUNSTEIN, *NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS* 90 (2008).
- 6 Governments are still needed to provide public goods like defense and they can have a role in correcting externalities that arise when a transaction between two parties affects others that are not party to the transaction. The efficiency of market outcomes does not necessarily imply a distribution of income that society would consider ideal or even satisfactory, so there might be a role for government in redistributing income.
- 7 G.J. Stigler, *The Economics of Information*, 69, J. POL. ECON. 213 (1961).
- 8 G.R. Butters, *Equilibrium Distribution of Sales and Advertising Prices* 44 REV. ECON. STUD. 465 (1977).

- 9 P. Diamond & H. Vartiainen, *Introduction*, P. Diamond & H. Vartiainen (eds.) *supra* note 1 at 2.
- 10 For a more complete elaboration of deviations from rationality, see Thaler & Sunstein, *supra* note 1 and Bernheim & Rangel, *supra* note 5.
- 11 M. Allais, *Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine*, 21 *ECONOMETRICA* 503 (1953).
- 12 Although the probabilities do not match up exactly, suppose that the choice is made at the start of the Major League Baseball season in United States and Canada. Think of the 10 percent occurrence as the New York Yankees winning the World Series, the 1 percent chance being the Cincinnati Reds winning the World Series, and the 89 percent chance being any team besides the Yankees or Reds winning the World Series. In the first choice, the individual either takes \$1 million for sure or, alternatively, \$5 million if the Yankees win, \$0 if the Reds win, and \$1 million otherwise. The choice is between getting \$1 million if either the Yankees or Reds win the World Series as opposed to \$5 million if just the Yankees win. For both choices, the individual's decision only matters if the Reds or Yankees win. The amount one gets if someone other than the Yankees or Reds wins does not rationally affect this choice.
- 13 When behavioral economists say that "framing" affects decisions, they mean that people make different choices when the same information is presented in different ways.
- 14 Tversky died in 1996. Nobel Prizes are not awarded posthumously.
- 15 D. Kahneman & H. Tversky, *Judgment under Uncertainty: Heuristics and Biases*, 185 (New Series) *SCIENCE* 1124 (1974).
- 16 For example, they gave respondents a description of a person with characteristics that matched the stereotype of a librarian and asked for the relative probabilities that the person was farmer, physician, salesman, airline pilot, or librarian. Research has shown that on average, respondents give too little weight to the fact that some professions are far more common than others.
- 17 Thaler & Sunstein, *supra* note 1, at 23.
- 18 *Id.* at 97.
- 19 Status quo bias can be understood as a way in which framing matters.
- 20 D. Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 *Q. J. ECON.* 861 (1997).
- 21 W. Güth, R. Schmittberger, & B. Schwarze, *An Experimental Analysis of Ultimatum Bargaining*, 3 *J. ECON. BEHAV. & ORG.* 367 (1982).
- 22 The insight that humans sometimes give in to short-run temptation with disproportionately severe long-term consequences dates back at least to the story of the Garden of Eden. Restricting attention to people considered to be economists, consider the following quote: "The man who acts according to the rules of perfect prudence, of strict justice, and of proper benevolence, may be said to be perfectly virtuous. But the most perfect knowledge of these rules will not alone enable him to act in this manner: his own passions are very apt to mislead him; sometimes to drive him and sometimes to seduce him to violate all the rules which he himself, in all his sober and cool hours, approves of." See A. SMITH, *THE THEORY OF MORAL SENTIMENTS* (1976). (The reference is to a Clarendon Press edition. The book was originally published in 1759.)
- 23 The equity premium puzzle is that the average return to equities so far exceeded the return to bonds that it was hard to reconcile investors' willingness to hold bonds with plausible degrees of risk aver-

sion. See R. Mehra & E.C. Prescott, *The Equity Premium Puzzle*, 15 J. MONETARY ECON. 145 (1985). (With the S&P 500 now roughly 20 percent below its value a decade ago, the puzzle may not be as pronounced as it once was.)

- 24 See, for example, R.J. Schiller, *Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends*, 71 AM. ECON. REV. 421 (1981) and S.F. LeRoy & R.D. Porter, *The Present Value Relations: Tests Based on Implied Variance Bounds*, 49 ECONOMETRICA 551 (1981).
- 25 See comments of P. Diamond in P. Diamond & H. Vartiainen (eds.), *supra* note 1, at 303.
- 26 J.B. DeLong, A. Schleifer, L. Summers, & R.J. Waldmann, *Noise Trader Risk in Financial Markets*, 98 J. OF POL. ECON. 703 (1990).
- 27 S. Salop & J. Stiglitz, *Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion*, 44 REV. ECON. STUD. 493 (1977).
- 28 X. Gabaix & D. Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121 Q. J. ECON. 505 (2006).
- 29 The customers who end up paying the late fuel charges may not literally be uninformed. The group would include those with strong preferences not to refuel and little concern for the cost and those who, due to the unfolding of travel-day events, end up not having the time to refuel.
- 30 Sometimes they do. As is discussed below, Southwest Airlines advertises its competitors' shrouded charges.
- 31 *The FTC as National Nanny*, THE WASH. POST, March 1, 1978.
- 32 J.H. Beales, III, *The Federal Trade Commission's Use of Unfairness Authority: Its Rise, Fall, and Resurrection*, 22 J. PUB. POL. AND MKTG 192, (2003).
- 33 Camerer et al., *supra* note 1.
- 34 Camerer et al. define a successful asymmetrically paternalistic policy as one for which:
$$(p^*B) - (1 - p)C - I + \Delta\pi > 0,$$
where p is the fraction of boundedly rational individuals, $(1 - p)$ is the fraction of fully rational individuals, B is the benefit of the policy to the boundedly rational individuals, C is the cost to the fully rational consumers, I is the implementation costs, and π is the economic profits of companies supplying the market in question.
- 35 Cameron et al. consider disclosure requirements more intrusive than defaults because they might be more costly.
- 36 Thaler & Sunstein, *supra* note 1.
- 37 In 2007, the Bureau of Economics at the FTC held a conference on the application of behavioral economics to consumer protection regulation, *available at* <http://www.ftc.gov/be/consumerbehavior/index.shtml>. For a summary of the conference, see JOSEPH P. MULHOLLAND, SUMMARY REPORT ON THE FTC BEHAVIORAL ECONOMICS CONFERENCE, *available at* <http://www.ftc.gov/be/consumerbehavior/docs/070914mulhollandrpt.pdf>.
- 38 An important part of the explanation is that many of the FTC consumer protection actions entail fraudulent activities by "judgment-proof" parties. Liability is not a close call requiring a careful weigh-

ing of costs and benefits, and penalties are driven more by the financial resources of defendants than by a measure of economic damages.

- 39 Christine Jolls, *Behavioral Law and Economics*, Diamond and Varitainen (eds), *supra* note 1 at 115.
- 40 As noted above, the need for consumer protection may lie more in some consumers making some very bad decisions than in ways in which consumer decisions are wrong (but only slightly so) on average.
- 41 It is interesting to consider why there has been little interest in behavioral economics approaches in industrial economics. A possible explanation is that one of the factors that drives interest in behavioral approaches is anomalies that seem hard to explain with a model of rational behavior. With the decline in interest in cross-market empirical analysis in industrial economics, industrial economists have not done the sort of work that might be expected to yield anomalies. Anomalies may exist, however. For example, continued investment in the airline industry despite a history of massive losses should arguably be considered a major anomaly in the industrial economics literature.
- 42 R. Roll, *The Hubris Hypothesis of Corporate Takeovers*, 59 J. Bus. 197 (1986).
- 43 See, for example, A. Berle & G. Means, *THE MODERN CORPORATION AND PRIVATE PROPERTY* (1932) and W.J. BAUMOL, *BUSINESS BEHAVIOR, VALUE, AND GROWTH* (1959).
- 44 I do not mean to suggest that there are no implications. I am not aware of any analysis of them, however, and they are not transparent without analysis.

The Future of Behavioral Economics in Antitrust Jurisprudence

Douglas H. Ginsburg &
Derek W. Moore

U.S. Court of Appeals for the DC Circuit &
Cadwalader, Wickersham & Taft LLP

The Future of Behavioral Economics in Antitrust Jurisprudence

*Douglas H. Ginsburg & Derek W. Moore**

Neoclassical economics or “price theory” has had a profound effect upon antitrust analysis, first as practiced in academia and then as reflected in the jurisprudence of the Supreme Court of the United States. More recently, behavioral economics has had a large and growing influence upon legal scholarship generally. Still, behavioral economics has not yet affected judicial decisions in the United States in any substantive area of law. The question we address is whether that is likely to change in the foreseeable future, *i.e.*, whether the courts’ present embrace of price theory in antitrust cases portends the courts’ imminent acceptance of behavioral economics in either antitrust or consumer protection cases.

*Douglas H. Ginsburg is a judge on the United States Court of Appeals for the District of Columbia Circuit, where he served as Chief Judge from July 16, 2001 until February 10, 2008. Derek W. Moore is an associate in the Washington, D.C. office of Cadwalader, Wickersham & Taft, LLP.

I. Introduction

The influence of an academic movement upon judicial decision-making can be evaluated in various ways. One simple, perhaps simplistic, way is to count the citations in judicial opinions to scholarly works representative of that movement. By that metric, behavioral economics has not yet affected courts in the United States in any substantive area of law. The question we address is whether that is likely to change in the foreseeable future.

There is some reason to think judges will consult behavioral economics or literature influenced by behavioral economics with increasing regularity in the not-too-distant future; the movement is already influential and becoming more so within the legal academy. Legal scholars have begun to incorporate behavioral economics into their work in much the same way antitrust scholars began to incorporate neoclassical economics into their work in the 1950s and 1960s.¹ Since then, price theory has had a major effect upon legal analysis in general² and a profound effect upon antitrust law³ and the antitrust jurisprudence of the Supreme Court.⁴

Considering the outpouring of legal scholarship influenced by behavioral economics, we might expect future judges—taught by professors with a behavioral bent—to incorporate that learning into their decisions. Behavioral economics in the late 20th and early 21st centuries and neoclassical economics in the 1950s and 1960s are not perfect analogues; however, there are several reasons to think behavioral economics will not have nearly as significant an effect upon future judicial decisions as neoclassical economics has had upon courts in the last several decades.

In this paper we briefly trace the influence price theory has had upon the antitrust academy and the judiciary. Then we place behavioral economics upon an evolutionary timeline to highlight its growing importance within the legal academy, but also to demonstrate its failure thus far to influence judicial decision-making in any meaningful way. We conclude with reasons for thinking behavioral economics is unlikely to influence future judicial decision-making.

CONSIDERING THE OUTPOURING
OF LEGAL SCHOLARSHIP
INFLUENCED BY BEHAVIORAL
ECONOMICS, WE MIGHT EXPECT
FUTURE JUDGES—TAUGHT
BY PROFESSORS WITH A
BEHAVIORAL BENT—TO
INCORPORATE THAT LEARNING
INTO THEIR DECISIONS.

II. Price Theory and Antitrust

Industrial organization (“IO”) is a sub-discipline of microeconomics that focuses upon the structure and performance of markets and upon firms’ strategic behavior within those markets. Although IO today is a rigorous discipline using com-

plex mathematical models, game theory, and econometric analysis, Richard Posner has aptly described IO in the 1950s and early 1960s as

“tend[ing] to be untheoretical, descriptive, ‘institutional,’ and even metaphorical. Casual observation of business behavior, colorful characterizations (such as the term “barrier to entry”), eclectic forays into sociology and psychology, descriptive statistics, and verification by plausibility took the place of the careful definitions and parsimonious logical structure of economic theory.”⁵

In the 1940s and 1950s, however, economists such as Aaron Director and George Stigler had begun to apply the rigorous fundamentals of neoclassical economics, or “price theory,” to the study of industrial organization. These are the simple propositions “that demand curves slope downward, that an increase in the price of a product will reduce the demand for its complement, that resources gravitate to the areas where they will earn the highest return, etc.”⁶

Although the teachings of price theory naturally appeared first in the economics journals, they began to show up in law reviews in the mid-1950s,⁷ as antitrust scholars comfortable with economic analysis and economists willing to conform to the conventions of legal scholarship—often in collaboration—began to retail technical economics to a legal audience. Significantly, the interdisciplinary and (unlike law reviews) peer-reviewed *Journal of Law and Economics* debuted in 1958. The move from economics journals and textbooks to interdisciplinary journals, law reviews, and eventually casebooks was significant because most policymakers in government had a legal background and were not well-versed in the economic way of thinking, let alone able to read technical economic literature.

The influence of economic reasoning upon antitrust analysis, practice, and jurisprudence continued to grow throughout the following decades. In 1965 Donald Turner, an antitrust law scholar with a Ph.D. in economics,⁸ became the head of the Antitrust Division of the Department of Justice. In the 1970s Richard Posner and Robert Bork each applied neoclassical economics to antitrust as a whole.⁹ Rather than using economic analysis to address a specific problem or analyze a particular case, they viewed the whole field, as Posner later said, “through the lens of price theory.”¹⁰ In 1981 James Miller, Ph.D., an economist without legal training, became Chairman of the Federal Trade Commission and William Baxter, a professor of antitrust law who had unqualifiedly embraced economic analysis, was put in charge of the Antitrust Division. In 1985 his successor (one of the present authors) elevated the position of Chief

Economist to that of a Deputy Assistant Attorney General, thereby giving the economists in the Division standing equal to the lawyers.

The changes that had originated in the academy and then affected the enforcement agencies eventually transformed the antitrust jurisprudence of the United States Supreme Court, which by the mid-1970s had recognized the importance of sound economic analysis in antitrust law.¹¹ First, the range of conduct deemed unlawful *per se* narrowed markedly as economic analysis displaced free-ranging considerations of political economy in giving meaning to the Sherman Act. Indeed, the Supreme Court's antitrust opinions have increasingly relied expressly upon the work of leading academic economists and their co-authors and colleagues in the law schools.¹² As a result of this convergence upon IO to inform the law, the degree of agreement among the Justices in antitrust cases has increased markedly; indeed, most antitrust decisions are now decided by a super majority of six or more. Economic analysis is also responsible for the prominence of simply-stated legal norms in antitrust decisions.¹³

FIRST, THE RANGE OF CONDUCT
DEEMED UNLAWFUL *PER SE*
NARROWED MARKEDLY AS
ECONOMIC ANALYSIS DISPLACED
FREE-RANGING CONSIDERATIONS
OF POLITICAL ECONOMY
IN GIVING MEANING TO
THE SHERMAN ACT.

The Court's reliance upon price theory reflects the near-consensus among academics on the proper approach to antitrust analysis. There is now broad and non-partisan agreement in academia, the bar, and the courts regarding the importance of price theory in antitrust decision-making. And that has transformed the dialogue in the courts. Today, it is common to see briefs on both sides of a case making arguments based upon sophisticated economic literature.

Even if economic analysis does not indicate a uniquely correct result in every case, it significantly constrains the decision-making of the courts by narrowing the range of plausible outcomes. As a result, neoclassical economic analysis has promoted predictability and consistency in antitrust jurisprudence.

III. Behavioral Economics and the Legal Academy

The migration of behavioral economics from economics departments to law schools is, of course, a movement in the same direction as the earlier infiltration of the legal academy by price theory. There are some important differences, however, in the implications of the two phenomena.

A. BRIEF HISTORY OF BEHAVIORAL ECONOMICS

Neoclassical economics assumes that man, *homo economicus* if you will, has consistent preferences—if he prefers apples to oranges and oranges to nuts, then he prefers apples to nuts—and makes choices that maximize his utility at all times.¹⁴

Cognitive psychologists began to question this assumption and in the 1950s the polymathic Herbert Simon made prominent the idea that individuals have only limited or “bounded” rationality and therefore sometimes make choices that satisfy their preferences but do not maximize their utility.¹⁵ This idea flows from the observation that humans do not possess the cognitive capacity required to process all the information necessary to maximize utility at all times; instead, they use heuristics or “shortcuts” to make decisions that sometimes fail to jibe with the predictions of neoclassical economics.¹⁶

BUILDING UPON AND ADDING TO
BOUNDED RATIONALITY AND
PROSPECT THEORY, BEHAVIORAL
ECONOMISTS AND PSYCHOLOGISTS
IDENTIFIED SEVERAL HEURISTICS
OR BEHAVIORAL ANOMALIES.

In the 1970s psychologists Daniel Kahneman and Amos Tversky developed an alternative to the rational choice model of neoclassical economics, which they called “prospect theory.” They first suggested that, when making a financial decision, individuals assign greater weight to a loss than to a gain of the same amount.¹⁷ Building upon and adding to bounded rationality and prospect theory, behavioral economists and psychologists identified several heuristics or behavioral anomalies

which, in their view, demonstrated the rational choice model was flawed because individuals depart from it in systematic, predictable ways. One example is the “availability heuristic,” meaning people estimate “the frequency of some event . . . by judging how easy it is to recall other instances of this type (how ‘available’ such instances are).”¹⁸ Another is the “endowment effect,” which holds there is a gap between the price at which an individual will sell an item he owns and the lesser amount he would be willing to pay to purchase precisely the same item.¹⁹ In other words, an individual values a good more highly if he owns (*i.e.*, is “endowed” with) it.

B. FROM BE TO BLE

Fifty years after the pioneering work of Herbert Simon and 20 years after Kahneman and Tversky first developed prospect theory, behavioral economics began to make serious inroads into the scholarship published in law journals — and thus to create the field now denominated “behavioral law and economics (‘BLE’).” Milestones of BLE include: Christine Jolls, Cass Sunstein, & Richard Thaler’s *A Behavioral Approach to Law and Economics*, published in the 1998 *Stanford Law Review*, which outlines and advocates using behavioral economics to analyze and reform law; Sunstein’s 2000 anthology, *Behavioral Law and Economics*, a collection of scholarship largely co-authored by behavioral economists and legal academics; and Sunstein and Thaler’s 2008 book *Nudge*, which attempts to make behavioral economic analysis of law accessible to the public at large.

A Westlaw search for the term “behavioral economics (‘BE’)” in American law journals reveals some interesting trends.

Years	Number of BE appearances in text	Number of BE appearances in titles
1980-84	1	0
1985-89	14	0
1990-94	12	0
1995-99	103	9
2000-04	548	27
2005-09	917	26

First, there is almost no mention of BE in law reviews until the latter half of the 1980s. Second, almost no scholarship in law reviews focused upon BE—as evidenced by the absence of that term from the titles of articles—until the latter half of the 1990s. Third, and most telling for the future, there has been a significant increase in the amount of BLE scholarship in the past decade.

How does the ascent of behavioral economics in the legal academy compare with that of price theory in antitrust scholarship some decades earlier? On the one hand, one might be inclined to see the late 1990s, when BE first became a subject treated in law reviews, as analogous to the late 1950s when neoclassical economic analyses of antitrust first appeared in law reviews.²⁰ But the type of scholarship produced by the two schools of thought in those nascent periods is markedly different. The antitrust pieces were the first attempts to retail economic analyses of business practices to an audience of lawyers. Director & Levi focused upon a narrow question: How should one understand the competitive effects of particular business practices? By contrast, the seminal piece by Jolls, Sunstein, & Thaler takes a much broader view. Rather than identify one case or even one area of the law that would benefit from a behavioral approach, they “propos[ed] a systematic framework for a behavioral approach to economic analysis of law, and us[ed] behavioral insights to develop specific models and approaches addressing topics of abiding interest in law and economics.”²¹

HOW DOES THE ASCENT OF
BEHAVIORAL ECONOMICS IN THE
LEGAL ACADEMY COMPARE WITH
THAT OF PRICE THEORY IN
ANTITRUST SCHOLARSHIP SOME
DECADES EARLIER?

The ambitiousness of their work makes it more analogous to the treatises of the late 1970s, in which Bork and Posner sought to analyze the whole of antitrust law through the lens of price theory, but there is a major difference as well. When Jolls, Sunstein, & Thaler, published in 1998, BE was already influential and becoming more so in economics,²² but BLE had hardly begun its upward trajectory; indeed there had been only a half-decade at most of serious scholarship exploring its implications in the law journals. Bork and Posner, on the other hand, were build-

ing upon two decades of work in price theory scholarship as applied specifically to antitrust issues and vetted in the law journals—scholarship that itself built upon a century of progress in economics. Their project was to synthesize and extend, not to establish, an intellectual movement.

IV. Behavioral Economics and the Courts

Behavioral economics has since infiltrated deeply into the legal academy and, with Cass Sunstein now heading the office that reviews proposed regulations for consistency with the policies of the president,²³ BE is poised to make inroads

IF, HOWEVER, BE IS TO AFFECT
AN AREA OF LAW—WHETHER
ANTITRUST, CONSUMER
PROTECTION, OR ANY OTHER—
THEN LEGISLATURES AND COURTS
MUST TAKE NOTICE.
SO FAR THEY HAVE NOT.

within the executive branch of government. If, however, BE is to affect an area of law—whether antitrust, consumer protection, or any other—then legislatures and courts must take notice. So far they have not.²⁴

The Supreme Court has cited a BLE article only once.²⁵ In considering whether an award of punitive damages was excessive, the Court cited an article by Cass Sunstein, David Schkade, & Daniel Kahneman for the modest proposition

that “juries do not normally engage in . . . a finely tuned exercise of deterrence calibration when awarding punitive damages.”²⁶

BLE has played no more prominent a role in the lower federal courts.²⁷ Indeed, the term “behavioral economics” appears in only three reported cases. *Honorable v. Easy Life*²⁸ involved a claim of racial discrimination in the sale of residential real estate. The plaintiffs alleged the defendant real estate company exploited unsophisticated buyers.²⁹ The court cited several BLE articles,³⁰ and stated:

“[T]he economic theories that imply that market prices are efficient, thus beneficial for consumers, presuppose that consumers are informed, markets are competitive, and the costs of making transactions are not excessively burdensome. . . . But these assumptions must be relaxed, and perhaps, ultimately replaced, if economic theory is to have any application to what happens in actual markets.”³¹

A search of federal court decisions for citations to BLE literature, *i.e.*, without regard to whether the term “behavioral economics” appears, does not change the outcome. We could find only 13 other cases that cite any of this literature and

most of those cases either concern the issue whether a jury's award of punitive damages is excessive or cite a BLE work only in passing.³²

V. What Conclusions Can We Draw?

If BE is going to affect judicial decision making to a degree at all similar to the effect that price theory has had upon antitrust law, then we should see a significant increase in courts' reliance upon BLE scholarship in another 10 to 15 years. Why then? Because, as Herbert Hovenkamp has pointed out:

“Elite American judges generally absorb the thinking of elite American intellectuals. Classical constitutional doctrine followed after the political economy that prevailed in America's best universities. . . . This was the political economy taught in American universities in the 1870s and 1880s. . . . Like judges of every era, [American judges] drew their wisdom — particularly the wisdom they applied to public law — from outside. . . . When the dominant American economic ideology changed—not until the first three decades of the twentieth century—the legal ideology followed close behind.”³³

A similar lag occurred between the emergence of price theory in IO and its influence upon antitrust jurisprudence. To the extent BE is now infiltrating if not permeating undergraduate economics courses and becoming a more prevalent mode of analysis in American law schools, we may in due course see judges sympathetic to behavioral ideas they first encountered at a younger age.

There are several reasons to think we will not see courts relying upon BE scholarship in antitrust cases, however. First, from a judicial perspective, BE is almost the opposite of price theory, which narrows significantly the range of outcomes a court may reach in an antitrust case; that price theory ideally generates determinate results is its great virtue as an aspect of jurisprudence. By interpreting the Sherman Act to promote consumer welfare *qua* allocative efficiency,³⁴ first scholars and then the Supreme Court delegitimated and excluded from consideration the myriad factors that had influenced judges in the past, including such whimsical goals as preserving small, locally owned businesses³⁵ or avoiding aggressive price cutting that could drive out less efficient rivals.³⁶

THERE ARE SEVERAL REASONS TO THINK WE WILL NOT SEE COURTS RELYING UPON BE SCHOLARSHIP IN ANTITRUST CASES, HOWEVER. FIRST, FROM A JUDICIAL PERSPECTIVE, BE IS ALMOST THE OPPOSITE OF PRICE THEORY.

BE may provide some guidance in the odd case, but the central theme of the discipline—that in certain circumstances humans are not rational utility maximizers and their departure from the rational choice paradigm is systematic and predictable—rather than foreclosing possibilities, opens them up and thereby increases the degrees of freedom with which a court may pursue personal, idiosyncratic goals. BE does not—at least not yet—provide or even promise to provide a general standard by which to decide any particular type of case.³⁷ Perhaps this is why BLE scholars were so quick to promote a “systematic framework for a behavioral approach to economic analysis of law”³⁸ rather than apply BE to the narrow legal questions that are actually decided by courts. The greater degrees of freedom a court would have if it departed from the rational choice model might well appeal to a lower court judge but would be anathema to the Supreme Court, which has labored for more than 30 years to reign in and rationalize antitrust law.

ONE CAN CERTAINLY IMAGINE THE
FTC USING BEHAVIORAL
ECONOMICS TO GIVE CONTENT
TO ITS AUTHORITY TO
PROHIBIT “UNFAIR
METHODS OF COMPETITION.”

Second, courts are unlikely to embrace BLE on their own initiative. Indeed it is difficult to imagine ways in which behavioral economics can be made useful to a particular legal issue that must be decided by a court in the first instance. In reviewing the decision of an administrative agency, on the other hand, an adversely affected firm might argue it was arbitrary and capricious for the agency not to have taken into account the relevant BE literature. This might arise in a case objecting to the decision of a risk-regulating agency, such as the Food and Drug Administration or the Consumer Protection Safety Commission, either requiring or failing to require some disclosure about a product.

Alternatively, a private party or an agency charged with consumer protection, such as the Federal Trade Commission, might argue that a practice is deceptive because it exploits a widespread cognitive bias identified in the BE literature.³⁹ One can certainly imagine the FTC using behavioral economics to give content to its authority to prohibit “unfair methods of competition.”⁴⁰ But this is far different from a court using BE to interpret what is meant by that phrase, which seems unlikely again because BE could serve only to broaden, rather than to narrow the meaning of the term “unfair.”⁴¹

Third, we are not likely to see an increase in judicial reliance upon BLE literature because courts are constitutionally averse to broad principles, whether drawn from neoclassical or from behavioral economics.⁴² BE, by suggesting individuals tend to deviate from the rational-choice-utility-maximization paradigm in certain respects, weighs in favor of general rules, *e.g.*, that a certain type of conduct—say, advertising that something is “free”—is deceptive. Courts are more inclined to make circumscribed decisions that narrowly answer the question whether a particular practice is impermissible in a specific context and avoid speaking more broadly than is necessary to decide the case before them.

For example, in a recent article on “behavioral antitrust,” Avishalom Tor & William Rinner have argued “a behavioral analysis suggests that real-world, boundedly rational manufacturers are prone to overuse” resale price maintenance as a competitive strategy.⁴³ Even if true, this point is of limited utility to a court in an RPM case, where the issue is whether a particular instance of resale price maintenance is anticompetitive; that is an empirical question the answer to which does not depend upon the state of mind of the manufacturer. Tor & Rinner’s point is of greater relevance to a legislature considering whether to make resale price maintenance illegal *per se* or to an antitrust enforcement agency deciding whether to devote resources to RPM cases.

FINALLY, AND PERHAPS MOST
IMPORTANT, BE IS STILL
IN A NASCENT FORM.

Finally, and perhaps most important, BE is still in a nascent form. Although modern price theory was refined greatly in the mid-20th century, its seeds were sewn much earlier.⁴⁴ The first instance of recognizable BE dates back just 50 years and economists still debate whether certain cognitive biases, which are now commonly assumed, can be replicated in laboratory experiments.⁴⁵ Even when natural experiments have been observed, there is evidence the behavioral approach fails to predict consumer choice better than does the rational choice model.⁴⁶ Although legal scholars are quick to devise policy prescriptions upon the basis of some admittedly interesting conclusions that can be drawn from the work of behavioral economists, courts and even regulatory agencies are not likely to shape the law accordingly until there is greater agreement among economists doing behavioral work about whether particular findings are verifiable, replicable, and empirically supported.

This is not to say BE is doomed forever to be irrelevant to the law and to legal policy. The problems we have identified serve only to show the BLE literature—in its present state—is of little if any utility to a court. We think it highly unlikely, even in the long run, that courts will view any particular area of law—consumer protection and antitrust law included—let alone the law more generally, through the lens of BE. The executive and the legislature are better suited and more likely than the judiciary to incorporate the teachings of BE—if they are persuasive—into policy prescriptions. ▼

1 For clarity, we shall refer to legal scholarship influenced by neoclassical economics or “price theory” as “traditional law and economics.”

2 See, e.g., RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW*, 7th ed. (2007).

3 See ROBERT H. BORK, *THE ANTITRUST PARADOX* (1978); RICHARD A. POSNER, *ANTITRUST LAW: AN ECONOMIC PERSPECTIVE* (1976).

4 See Leah Brannon & Douglas H. Ginsburg, *Antitrust Decisions of the U.S. Supreme Court 1967 to 2007*, 3 *COMP. POL’Y INT’L* 1 (2007).

- 5 Richard A. Posner, *The Chicago School of Antitrust Analysis*, 127 U. PENN. L. REV. 925, 928-29 (1979) (hereinafter "Chicago School").
- 6 *Id.*, at 928.
- 7 See e.g., Ward Bowman, *Tying Arrangements and the Leverage Problem*, 67 YALE L.J. 19 (1957); Aaron Director & Edward Levi, *Law and the Future: Trade Regulation*, 51 NW. L. REV. 281 (1956); Robert H. Bork, *Vertical Integration and the Sherman Act: The Legal History of an Economic Misconception*, 22 U. CHI. L. REV. 157 (1954).
- 8 Turner himself was a member of the "Harvard School" of antitrust analysis, which was based upon a different strand of industrial organization and entailed a somewhat more interventionist approach. See William E. Kovacic, *The Intellectual DNA of Modern U.S. Competition Law For Dominant Firm Conduct: The Chicago/Harvard Double Helix*, 2007 COL. BUS. L. REV. 1, 30-31. Because he adhered to the "structure-conduct-performance" paradigm—which holds that the competitive performance of a particular industry or market is determined largely by the number and relative size of the incumbent firms—as reflected in the 1968 Merger Guidelines, Turner's appointment did not represent the triumph of price theory in antitrust policy. Still, Turner's appointment demonstrates the then-growing importance to antitrust analysis of being able to speak the language of economics.
- 9 See BORK, *supra* note 3; POSNER, *supra* note 3.
- 10 Posner, *Chicago School*, *supra* note 5, at 928.
- 11 See generally Brannon & Ginsburg, *supra* note 4.
- 12 Over the last 20 years more than 75 percent of the Court's antitrust opinions cite antitrust scholarship applying price theory. Brannon & Ginsburg, *supra* note 4.
- 13 See, e.g., *NCAA v. Bd. of Regents*, 468 U.S. 85, 107 (1984) ("Congress designed the Sherman Act as a consumer welfare prescription. A restraint that has the effect of reducing the importance of consumer preference in setting price and output is not consistent with this fundamental goal of antitrust law" (internal citations omitted)).
- 14 See GARY S. BECKER, *THE ECONOMIC APPROACH TO HUMAN BEHAVIOR* 14 (1976) ("[A]ll human behavior can be viewed as involving participants who maximize their utility from a stable set of preferences and accumulate an optimal amount of information and other inputs in a variety of markets").
- 15 HERBERT A. SIMON, *THE MODELS OF MAN* (1957); Herbert A. Simon, *A Behavioral Model of Rational Choice*, 69 Q.J. ECON. 99 (1955), reprinted in this journal at 241-58.
- 16 See Christine Jolls, Cass. R. Sunstein, & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1476-77 (1998).
- 17 Daniel Kahneman & Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974); Kahneman & Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 ECONOMETRICA 263 (1979).
- 18 Jolls, Sunstein, & Thaler, *supra* note 16, at 1477.
- 19 *Id.*, at 1484.
- 20 See sources in note 7, *supra*.
- 21 Jolls, Sunstein, & Thaler, *supra* note 16, at 1473-74.

- 22 The John Bates Clark Medal, then awarded biannually to a leading American economist under the age of 40, was consecutively awarded to behavioral economists Andrei Shleifer and Matthew Rabin in 1999 and 2001.
- 23 The Office of Information and Regulatory Affairs is a component of the Office of Management and Budget in the Executive Office of the President.
- 24 This paper is limited in scope to the question whether courts have been influenced by BLE; whether a legislature has adopted the teachings of BLE is outside that scope.
- 25 In *Exxon Shipping Co. v. Baker*, 128 S.Ct. 2605 (2008), the Court cited several BLE articles concerning jury awards of punitive damages but “[b]ecause the research was funded in part by Exxon,” the Court explicitly “decline[d] to rely on it.” *Id.* at 2626 n.17.
- 26 *Cooper Industries v. Leatherman Tool Group*, 532 U.S. 424, 439 (2001).
- 27 For the sake of brevity, we have limited our discussion to the federal courts. We have no reason to believe state courts have been more receptive to BLE.
- 28 *Honorable v. Easy Life Real Estate System*, 100 F.Supp. 2d 885 (N.D. Ill. 2000).
- 29 *Id.* at 888.
- 30 *Id.* at 892.
- 31 *Id.* at 888.
- 32 It is possible, though we think it unlikely, that a case recently argued before the United States Supreme Court could yield an opinion that relies upon certain BLE literature. In *Jones v. Harris Associates*, 527 F.3d 627 (7th Cir. 2008) reh’g denied 537 F.3d 728, cert. granted, 129 S.Ct. 1579 (2009), economists Robert Litan, Joseph Mason, & Ian Ayers have filed an amicus brief arguing behavioral economics is relevant to a dispute over whether a family of mutual funds overpaid its investment advisor:

[R]ecent discoveries in the behavioral finance literature reveal that the vast majority of investors cannot accurately assess the quality of the mutual funds in which they invest — that is, they do not behave rationally as predicted by economic models. Problems such as ‘misperceptions of chance,’ ‘sample-size neglect,’ ‘loss aversion,’ and ‘mental accounting’ render the assessment of mutual fund performance extremely difficult for investors. Accordingly, the proposition that mutual fund investors will simply ‘fire’ their advisers by ‘redeeming their shares and investing their assets elsewhere’ ignores these natural impediments to competition.
- See also Adam Liptak, *Supreme Court to Hear Case on Executive Pay*, NY TIMES, Aug. 18, 2009, A10.
- 33 HERBERT HOVENKAMP, *ENTERPRISE AND AMERICAN LAW, 1836-1937* (1991), 97-98, 182, 204.
- 34 See *Reiter v. Sonotone Corp.*, 442 U.S. 330, 343 (1979) (noting that legislative history “suggest[s] that Congress designed the Sherman Act as a ‘consumer welfare prescription’”) (quoting BORK, *supra* note 3, at 66); *Nat’l Collegiate Athletic Ass’n v. Bd. of Regents*, 468 U.S. 85, 107 (1984) (“Congress designed the Sherman Act as a consumer welfare prescription. A restraint that has the effect of reducing the importance of consumer preference in setting price and output is not consistent with this fundamental goal of antitrust law.”) (citation omitted); see also *GTE Sylvania Inc. v. Continental T.V. Inc.*, 537 F.2d 980, 1003 (9th Cir. 1976) (en banc), *aff’d* 433 U.S. 36 (1977) (“Since the legislative intent underlying the Sherman Act had as its goal the promotion of consumer welfare, we decline blindly to

condemn a business practice as illegal per se because it imposes a partial, though perhaps reasonable, limitation on intrabrand competition, when there is a significant possibility that its overall effect is to promote competition between brands.”) (citing *BORK*, *supra* note 3, at 11).

- 35 See *U.S. v. Trans-Missouri Freight Ass’n.*, 166 U.S. 290, 323 (1897) (trade may be “badly and unfortunately restrained by driving out of business the small dealers and worthy men”).
- 36 See *Utah Pie v. Continental Baking Co.*, 386 U.S. 685, 690 (1967) (“There was ample evidence to show that each of the respondents contributed to what proved to be a deteriorating price structure over the period covered by this suit, and each of the respondents in the course of the ongoing price competition sold frozen pies in the Salt Lake market at prices lower than it sold pies of like grade and quality in other markets considerably closer to its plants”).
- 37 Indeed in *Jones v. Harris*, *supra* note 32, behavioral economics seems to raise more questions than it answers: What behavioral anomalies might be present? Is there empirical proof they are present? How would the presence of such anomalies in a certain percentage of market actors affect the market overall? Even if a model predicated upon behavioral insights might more accurately describe the behavior of market actors — and it surely might — it is difficult to imagine the Supreme Court adopting behavioral insights without reason to believe these questions can be answered reliably in the crucible of litigation.
- 38 Jolls, Sunstein, & Thaler, *supra* note 16, at 1473-74.
- 39 This is similar to the argument made by the economists in *Jones v. Harris* that the Court should not rely upon market forces in a particular context because consumers’ cognitive biases prevent the market from functioning properly.
- 40 15 U.S.C. § 45.
- 41 Cf. Robert A. Lande, *Revitalizing Section 5 of the FTC Act Using “Consumer Choice” Analysis*, ANTITRUST SOURCE, Feb. 2009, available at <http://www.abanet.org/antitrust/at-source/09/02/Feb09-Lande2-26f.pdf> (suggesting § 5 must be interpreted “in a manner that is definite, predictable, principled, and clearly bounded” if its application is to survive judicial review); see also Michelle M. Burtis, *Modern Industrial Organization: A Comment*, 12 GEO. MASON L. REV. 39, 41-43 (2003) (arguing game theory-based economic models have not been prominently featured in antitrust litigation in part because models “suffer from a ‘multiplicity of equilibria’”); Dennis W. Carlton, *The Relevance for Antitrust Policy of Theoretical and Empirical Advances in Industrial Organization*, 12 GEO. MASON L. REV. 47, 50 (2003) (“[A]lthough there are some examples where game theory can improve our understanding of some antitrust issues (e.g., “agreement” in Section 1 cases), in general, game theory alone is unlikely to produce insights that are sufficiently robust to provide a clear guide to the implementation of antitrust policy”).
- 42 See Edward R. Becker, *The Uses of ‘Law and Economics’ by Judges*, 33 J. LEGAL EDUC. 306, 310 (1983) (“[L]aw and economics will not influence the judgments of courts if it consists of an abstruse methodology. You are not going to get very far in presenting a system or a methodology for all seasons”).
- 43 Tor & Rinner, *Behavioral Antitrust: A New Approach to the Rule of Reason After Leegin*, University of Haifa Faculty of Law Legal Studies Research Paper Series (December 1, 2009) available at <http://ssrn.com/abstract=1522948>.
- 44 See, e.g., ALFRED MARSHALL, *PRINCIPLES OF ECONOMICS* (1890); ANTOINE A. COURNOT, *RESEARCH INTO THE MATHEMATICAL PRINCIPLES OF THE THEORY OF WEALTH* (1838).
- 45 See, e.g., Charles R. Plott & Kathryn Zeiler, *Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?*, 97 AM. ECON. REV. 1449 (2007); Plott & Zeiler, *The*

Willingness to Pay-Willingness to Accept Gap, the Endowment Effect, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations, 95 AM. ECON. REV. 530 (2005).

- 46 Joshua D. Wright, *Behavioral Law and Economics, Paternalism, and Consumer Contracts: An Empirical Perspective*, 2 NYU J. LAW & LIBERTY 470 (2007).

The Role of Behavioral Economics in Competition Law: A Judicial Perspective

Vivien Rose

U.K. Competition Appeal Tribunal

The Role of Behavioral Economics in Competition Law: A Judicial Perspective

*Vivien Rose**

To date, the literature on the role of behavioral economics in the context of competition policy has largely focused on the development of the theory and practice of behavioral economics in the analysis of competition cases. It is also useful, however, to consider how arguments about behavioral economics are likely to be received in a judicial setting.

It will take longer for academic writings on behavioral economics to filter through into the arguments before and judgments of the Tribunal or the High Court than it does in the U.S. courts—indeed that might never happen. But, in fact, what courts have been doing all along may be closer to behavioral economics than to more conventional economic theories of rational behavior.

*The author is a Chairman of the Competition Appeal Tribunal in London. The views expressed in this article are entirely personal, do not necessarily reflect the views of my colleagues, and do not indicate how the Tribunal is likely to decide any cases whether currently pending or arising in the future.

I. Introduction

To date, the literature on the role of behavioral economics in the context of competition policy has largely focused on the development of the theory and practice of behavioral economics in the analysis of competition cases. It is also useful, however, to consider how arguments about behavioral economics are likely to be received in a judicial setting.

An important point to bear in mind concerning economics and litigation is that private law litigation in the United Kingdom and elsewhere in the European Union arises in the context of claims brought by businesses rather than by individual consumers. A business affected by anticompetitive conduct is more likely to suffer loss and damage on a scale which is sufficient to make it worthwhile litigating. The two most common kinds of claim are where a business seeks damages from cartel members for losses arising from the cartel or (most frequently) by competitors alleging that the behavior of an allegedly dominant firm is foreclosing the market and restricting market entry and expansion. Thus the most likely scenario to come before a court is not a consumer or group of consumers saying “This is what we did, this is how we behaved” but rather a business saying, “This is how my customers behave” or “This is how my competitors’ customers *would* behave if their behavior were not being distorted by the allegedly unlawful conduct of the other party to the action—i.e. they would in fact now be my customers, not his.”

AN IMPORTANT POINT TO BEAR IN MIND CONCERNING ECONOMICS AND LITIGATION IS THAT PRIVATE LAW LITIGATION IN THE UNITED KINGDOM AND ELSEWHERE IN THE EUROPEAN UNION ARISES IN THE CONTEXT OF CLAIMS BROUGHT BY BUSINESSES RATHER THAN BY INDIVIDUAL CONSUMERS.

The question how customers do behave within existing market conditions or would behave if those conditions were different can arise in a number of situations, such as defining the relevant market or considering the extent of foreclosure of the market caused by tying or loyalty rebates. These issues may arise in recently liberalized markets so that the defendant is a monopolist or quasi-monopolist in a market where there is no history of competitive conditions to provide evidence of actual consumer conduct under such conditions. The court is therefore being invited to engage in some crystal ball gazing when hypothesizing what customers would do if the market were competitive.

A second key point is that the citation of academic articles is much less prevalent in the English courts across the whole range of topics than it is in the United States. Until relatively recently, when an advocate cited an academic article in support of his or her submissions, it was a sign of a certain level of desperation—indicating there must be no “real” authorities in the form of case law from either domestic or Commonwealth courts to rely on. Things have moved on since then—particularly in the Supreme Court where citation of both text books and

articles is now commonplace. For example, at a recent meeting of the Competition Law Association in London, there was a panel discussion of the *Norris* decision in the House of Lords¹ concerning whether cartel activity could constitute the common law offence of conspiracy to defraud. There was discussion about the use made by their Lordships of an article that a panel member had co-authored. Another recent example in the Court of Appeal is the decision about whether bank charges are subject to the test of fairness. In *Abbey National v Office of Fair Trading*² the Court referred to a leading textbook on contract law and to an article written by a Professor of Law at Swansea University. The Supreme Court recently overturned that decision,³ and referred to an article

THE COURT IS THEREFORE BEING
INVITED TO ENGAGE IN SOME
CRYSTAL BALL GAZING
WHEN HYPOTHESIZING WHAT
CUSTOMERS WOULD DO IF THE
MARKET WERE COMPETITIVE.

written in 1994 by Professor Hugh Collins, *Good Faith in European Contract Law*.⁴ If reference to legal academic writings is fairly rare then reference to academic writings in other disciplines such as economics is even rarer.

In the Competition Appeal Tribunal where I sit, it is very common to hear evidence from expert economists and for these witnesses to be cross-examined extensively. But although such experts may include references to papers and articles in footnotes in their witness statements, these are rarely the focus of their evidence and are not usually picked up by either the advocates or the Tribunal. One exception to this was in the *GISC* case⁵ concerning whether Article 81 applied to the rules promulgated by the General Insurance Standards Council. The Director General of Fair Trading had found they were not covered by the prohibition. The CAT quashed this decision. Before the hearing of the action, the Tribunal alerted the parties to an article written in 1990 by John Kay and John Vickers called *Regulatory Reform: an Appraisal*⁶—an article of interest not just because of its intrinsic merit but because John Vickers was at the time of the appeal the Director General of Fair Trading and the article appeared to advance an argument contrary to the DGFT's case in the *GISC* appeal.

There have been a couple of instances where more complex economics has arisen in cases involving utilities regulation, for example in *Albion*⁷ where the Tribunal was concerned with the correct approach to access pricing or in *Mobile Call Termination Charges*⁸ as regards the merits of asymmetric price regulation in sectors characterized by a large incumbent provider and a smaller new entrant. But for more basic questions about relevant market definition or the likely effect of alleged anticompetitive conduct on customers, competitors, and consumers, the courts have tended to rely on legal authorities and on expert oral and written economic evidence.

That leads to the third key point that judges, even if they sit on a specialist tribunal like the Competition Appeal Tribunal, are lawyers first and economists second. They are unlikely to have formal training in economics. This has upsides

and downsides. The downside is that they are unlikely to be up to speed in the latest thinking in economic journals. What is second nature to economists will not be so to judges. The upside is that judges are likely to be receptive to anything that accords with a common sense approach to issues and which chimes with how they/we in our rather untutored way are likely to think that people behave. So if a party is seeking to establish that people (e.g. consumers) are likely to react in a particular way to the conduct under discussion, the judge is more likely to accept this if he/she thinks “that is how I would probably react.” This leads to the conclusion that if behavioral economics is a way of bringing economic theory more in line with a realistic idea of what people actually do then it is likely to be an attractive line of argument. Judges do tend to consider that they are good assessors of human nature.

II. Some Examples

It might assist to consider two examples from legal practice where the theory of behavioral economics was definitely not at the forefront of everyone’s mind but where some kind of academic framework would have been useful as a context in which to discuss the issues. Perhaps without knowing it, the counsel and judiciary were applying behavioral economics!

A. TEE SHIRTS WITH LOGOS OF HEAVY METAL ROCK BANDS

A straightforward claim for the price of goods sold and delivered was brought by a company that produced “heavy metal tee shirts.” The plaintiff wholesaler company had the exclusive right granted by the rock band’s promoter to use images connected with a particular rock band—or with a particular global tour of that band—for the purpose of creating merchandise in the form of tee shirts. The tee shirts sold were ordinary black tee shirts but with the name and logo or other design of the heavy metal rock band printed on them. The shirts were mainly sold to small-scale retailers who often sold them from stalls set up at the rock concert venue.

PERHAPS WITHOUT KNOWING
IT, THE COUNSEL AND
JUDICIARY WERE APPLYING
BEHAVIORAL ECONOMICS!

A batch of several dozen shirts was sold to a stallholder who failed to pay for them. In defense to a straightforward claim for the price, the retailer pleaded alleged breaches of Articles 81 and 82 EC [now Articles 101 and 102 TFEU] and counterclaimed for damages. In an application for summary judgment, the issue arose as to what was the relevant market within which these tee shirts competed. If the relevant market was “all tee shirts” then clearly the seller had a negligible market share and the allegations of dominance were doomed to failure. If the market was the market for tee shirts bearing the logo of this particular heavy metal band, the wholesaler as the exclusive licensee would have 100 percent of the market over the relevant period. There were a number of other markets posited: the market was a

market for heavy metal band tee shirts—or for tee shirts with any kind of music band logo or sign on them.

At the hearing, the plaintiff seller relied on the well known *United Brands* formula. In *United Brands* the Court of Justice had identified a group of people who were unable to eat other fruit and could only eat bananas. It is not enough to identify a group of people who like the taste of bananas but do not like the taste of other kinds of fruit—there needs to be something more substantial separating the group of banana-only eaters from the rest of the population in order to find that bananas occupy their own market. In the case of the tee shirts, there is no one who is capable of wearing a heavy metal rock band tee shirt but is physically unable to wear an ordinary black tee shirt or a tee shirt with any other kind of decoration. There was nothing therefore to separate out a group of heavy-metal-band-only tee shirt wearers from the general population and hence no separate market for these tee shirts.

The flaw in this argument was that the heavy metal tee shirt retailed at the concert venue for about £18 and a plain black tee shirt retailed in shops generally at £3. Applying a normal price elasticity test indicated therefore that these tee shirts did occupy separate relevant markets because they appeared to be

THIS BEGS THE FURTHER QUESTION—WAS IT FAIR TO DESCRIBE THE PURCHASERS AS ACTING “IRRATIONALLY” OR WERE THEY RATHER MAKING THEIR PURCHASING DECISION ON THE BASIS OF CRITERIA THAT WERE NOT SO EASY TO ASSESS IN QUANTITATIVE TERMS.

attractive to people at a greatly inflated price even though they were largely indistinguishable from the ordinary shirt. To recast this in terms of behavioral economics—it may be that consumers were acting totally irrationally in paying £18 for a £3 tee shirt simply because of the logo on it. But the fact is that a significant number of them were prepared to pay this inflated price—and a large industry in merchandising rights had grown up as a result. This begs the further question—was it fair to describe the purchasers as acting “irrationally” or were they rather making their purchasing decision on the

basis of criteria that were not so easy to assess in quantitative terms. Needless to say, summary judgment was refused and the case settled before coming to trial.

B. MATERNITY SAMPLE BAGS

The second case concerned the exclusive arrangements entered into by a company that creates bags containing free samples that are distributed by maternity ward staff to women who have just given birth in hospital. The bag compiler entered into exclusive arrangements with the maker of only one of the range of brands of each of the products to include in the bag—one brand of nappies, one brand of cotton wool buds, one brand of baby wipes, etc. The companies competing in the market of each of these products fought to have their product included in the bag. Why? It was perceived that there was a kind of “first mover”

advantage—if the mother got used to using one brand of the particular product she would stick with that—particularly if it had a tacit endorsement of the hospital by being included in the bag given on the ward.

A manufacturer whose product was not chosen for inclusion in the bag sought an injunction against the bag compiler to prevent it from distributing the bags. It was alleged that the exclusive arrangements with the competitor were contrary to Articles 101 and 102 TFEU. One issue in the case was whether the bag compiler was dominant in “the market of providing promotion through inclusion of [the product] in the hospital gift bags” or whether the gift bag was just one of a wide range of promotions that could be used by the manufacturers. This, in turn, depended on whether it was really true that women tended to stick thereafter with the product that they got in the bag. The bag compiler argued that there was no reason to suppose that this was the case. Women make hundreds of repeat purchases of these items during the early years of their child’s life. Was it really plausible to suppose that they would ignore all differences in price and other advertising during those years to stick with the brand they had been given in the bag? Further, there was evidence that the hospital instructed expectant mothers to bring their own supply of these items with them when they came to hospital so the women would have made their first purchases of these items before receiving the bag. Many mothers getting the bag were having their second or third child so may have been already fixed in their purchasing habits.

But the answer to all this was that the companies competing in this market were prepared to pay the bag compiler considerable sums of money for the privilege of being included in this bag. Either they were not acting rationally in considering that it was an important marketing tool or they knew from their own research that the mothers did not act rationally and were in fact unduly influenced by the presence of the particular brand in this bag of free goods. The case settled before the Court had to arrive at a solution. But behavioral economics might have provided a useful framework in which to have that discussion.

BUT IN FACT WHAT COURTS
HAVE BEEN DOING ALL ALONG
MAY BE CLOSER TO BEHAVIORAL
ECONOMICS THAN TO
MORE CONVENTIONAL
ECONOMIC THEORIES OF
RATIONAL BEHAVIOR.

III. Conclusion

In conclusion, it is certainly the case that it will take longer for academic writings on behavioral economics to filter through into the arguments before and judgments of the Tribunal or the High Court than it does in the U.S. courts—indeed that might never happen. But in fact what courts have been doing all along may be closer to behavioral economics than to more conventional economic theories of rational behavior. ▼

- 1 House of Lords ([2008] UKHL 16).
- 2 Abbey National v Office of Fair Trading [2009] EWCA Civ 116.
- 3 Office of Fair Trading v Abbey National [2009] UKSC 6.
- 4 Hugh Collins, *Good Faith in European Contract Law*, 14(2) OXFORD J. LEGAL STUD. 229 (1994).
- 5 GISC case [2001] CAT 4.
- 6 John Kay & John Vickers (1990), *Regulatory Reform*, DEREGULATION OR REREGULATION (G. Majone, ed.) (1990).
- 7 Albion [2006] CAT 23.
- 8 Mobile Call Termination charges [2008] CAT 11.

What Does Behavioral Economics Mean for Competition Policy?

Matthew Bennett, John Fingleton,
Amelia Fletcher, Liz Hurley, &
David Ruck

U.K. Office of Fair Trading

What Does Behavioral Economics Mean for Competition Policy?

*Matthew Bennett, John Fingleton,
Amelia Fletcher, Liz Hurley, & David Ruck**

This paper looks at whether behavioral economics fundamentally changes our understanding of competition policy. We argue that behavioral economics is an important incremental advance in our understanding, just as informational economics was before it. But this does not mean that all previous economic models of competition and markets are now irrelevant. For the most part, they still provide valid and valuable insights. Importantly, behavioral economics does not question our belief in competition policy as a tool for making markets work well for consumers.

Nevertheless, the existence of behavioral biases does have a number of implications for the way in which markets work. Behavioral biases on the consumer side emphasize the importance of the demand side in making markets work well, and the important synergies between consumer policy and competition policy. Behavioral biases may also have implications for anticompetitive behavior. In spite of this, behavioral economics does not necessarily imply more intervention. Markets can often solve their own problems and even where they can't, there are dangers inherent in over-paternalism limiting consumer choice. Behavioral economics also emphasizes the difficulties that authorities can have in trying to correct for such biases.

*The authors are all members of the Office of Fair Trading ("OFT") in the United Kingdom. We would also like to thank the many colleagues with whom we have discussed the ideas in this paper. Nevertheless, the opinions within the paper are those of the authors alone and not necessarily those of the OFT.

I. Introduction & Summary

Over the last ten years behavioral economics has become one of the rare subjects in economics to move beyond academia and into wider society. Interest is now so widespread that the sight of behavioral economics books on the best seller list is no longer incongruous. But what exactly is behavioral economics and what is it likely to mean for competition policy?

In very general terms, behavioral economics is an area of economics that combines insights from psychology with traditional economic models to more accurately reflect decision-making by consumers and other economic agents.¹

Traditional models of markets have assumed economic agents that can be classified as “Homo Economicus,” rather than Homo Sapiens. Homo Economicus has infinite capacity to take in and process information; is neutral to how things are presented; can anticipate and take the future into account; cares only about self-maximizing; and treats gains the same as losses.

Literature on behavioral economics, however, provides evidence of the various ways in which real economic actors differ from Homo Economicus. Homo Sapiens has limits to taking in information; is taken in by how things are presented; tends to be poor at anticipating the future; cares about other people and fairness; and cares more about losses than gains. In short, Homo Sapiens exhibits systematic biases in the way he views both the world and markets.²

IN SHORT, HOMO SAPIENS
EXHIBITS SYSTEMATIC BIASES
IN THE WAY HE VIEWS BOTH
THE WORLD AND MARKETS.

These are known as “behavioral biases.” While there are a large number of behavioral biases, the main ones involve a lack of processing power,³ the importance of framing,⁴ the existence of time inconsistency,⁵ and aversion to losses.⁶

In fairness, economists never really asserted that market participants were as ultra-rational as Homo Economicus. Rather, the assumption was used as a helpful shortcut. It enabled the development of sophisticated economic models of markets, which in turn appeared to reflect reality fairly well. The argument was that market participants did not actually have to be ultra-rational in order for markets to work as if they were.

The recent growth in interest in behavioral economics has, however, forced economists to question this thesis. Specifically, are there ways in which behavioral biases might lead to systematic biases in the models of markets and competition on which we have been relying? Do they have implications for the effective analysis of, and intervention in, markets by competition authorities? Or is behavioral economics just the latest economics fad, which does not really change anything, and will soon be buried at the bottom of the economics toolbox?

This paper provides some thoughts on these issues. Like many economic answers to complicated questions, our conclusions are nuanced. While behavioral economics does raise questions regarding current thinking, and may alter how competition authorities carry out some of their analysis or design their interventions, behavioral economics does not represent the fundamental shift some might fear.

First, behavioral economics does not mean that all previous economic models are negated. Much of the existing literature continues to provide us with important and useful insights. Arguably, behavioral biases simply take a place along-

SECOND, BEHAVIORAL ECONOMICS
DOES NOT CHANGE THE VIEW
THAT COMPETITION POLICY IS A
CRUCIAL TOOL FOR MAKING
MARKETS WORK WELL
FOR CONSUMERS. IT DOES,
HOWEVER, EMPHASIZE
THAT COMPETITION POLICY IS
ONLY ONE SIDE OF THE STORY.

side other well-known market failures, such as market power, asymmetric information, and externalities, the implications of which are now well understood and incorporated into economic thinking.

Second, behavioral economics does not change the view that competition policy is a crucial tool for making markets work well for consumers. It does, however, emphasize that competition policy is only one side of the story. To make markets work well we not only need competition policy, we also need engaged consumers, able to assess, access, and act on information. This highlights the crucial role that consumer policy, and other forms of demand-side intervention, plays in driving effective markets.

Third, behavioral economics does not alter the view that the market may find its own solutions to any problems. Reputation, learning effects, intermediaries, the media, and even firms themselves, can all help to solve market problems arising from consumer biases. Behavioral economics instead strengthens something we already knew: that we can not blindly assume the market will solve everything.

Finally, even to the extent that there are persistent problems in markets, behavioral economics also emphasizes the difficulties that competition (or consumer) authorities can face in trying to correct for such biases. For example, behavioral economics tells us that simply providing more information may not be a good solution when consumers have problems assessing such information. As such, it is far from clear that a greater emphasis on behavioral economics implies more intervention. It may well imply less.

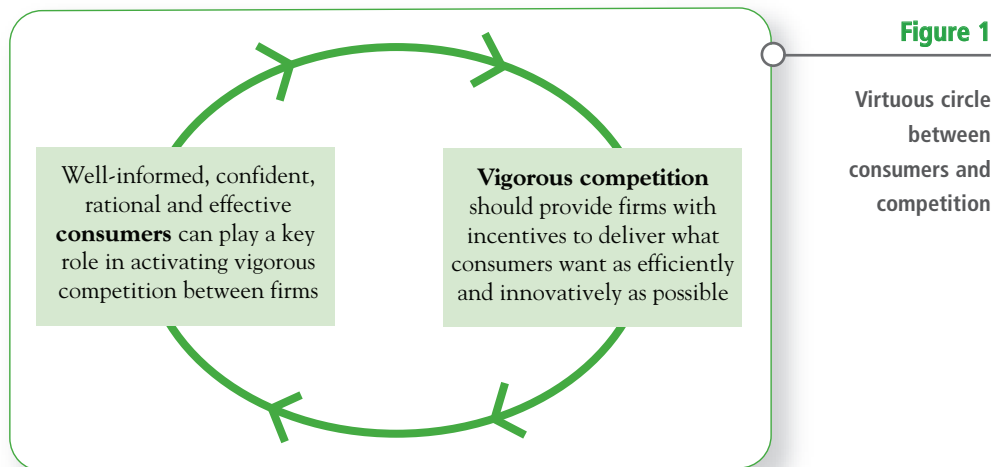
The second section of this paper examines the implications of behavioral economics for our understanding of competition in markets, with a particular focus on the factors that drive, or inhibit, competition. It looks first at behavioral biases on the demand side, and how these can potentially soften competition, especially when we take account of firms' incentives to exploit such demand-side biases.

The third section of the paper then looks at the implications of these consumer biases for antitrust enforcement, before considering the potential for behavioral biases on the supply side, and how these too could alter our thinking on antitrust questions. This section also examines briefly the role that consumer policy, and other demand-side interventions, can have in driving competition in markets, working alongside antitrust enforcement.

The fourth section of this paper examines the implications of behavioral economics for interventions in markets. It looks at the potential for the market to solve its own problems and at the unintended consequences of interventions. It concludes with some thoughts on the practical implications of behavioral economics for the empirical analysis involved in designing interventions.

II. Behavioral Economics and Competition in Markets

Markets work well when there are efficient interactions on both the demand (consumer) side and the supply (firm) side. On the demand side, confident consumers activate competition by making well-informed and well-reasoned decisions which reward those firms which best satisfy their needs. On the supply side, vigorous competition provides firms with incentives to deliver what consumers want as efficiently and innovatively as possible. When both sides function well, a virtuous circle is created between consumers and competition. This is illustrated in figure one below.



While active and rational consumers and vigorous competition work together in tandem to deliver consumer benefits, the failure of either side of the circle can harm the effectiveness of markets. Competition policy has traditionally been

focused on protecting the supply side—if competition between firms is diminished then consumers do not get what they want. However, the demand side is also crucial—if consumers are less engaged in the buying process, then firms may find it harder to win market share by providing what consumers most want. This will, in turn, reduce the incentive of firms to work towards that end, competition will be weakened, and less consumer benefit will be delivered by the market.

Behavioral biases clearly have a key role to play here. On the demand side, they can impact on the extent to which consumers play their active, effective, and rational part in this virtuous circle. On the supply side, they can impact on whether firms act on their apparent incentives in the profit-maximizing way that most economic models assume.

A. BEHAVIORAL BIASES: THE FOURTH MARKET FAILURE?

It is not newsworthy that there can be problems in markets. Economists have long been aware of three broad circumstances in which the market may fail, the implications of which are now well understood:

- a) Market power: Gives firms a reduced incentive to compete for customers, and a greater incentive to exploit them (or, alternatively, to have an easy monopolist life).
- b) Asymmetries in information between consumers and firms: Can hinder consumers' ability to effectively drive competition or firms' ability to target consumers effectively.
- c) Externalities not captured within consumers' preferences: Can mean that a market produces too much or too little of a good or service from a societal point of view.

ARGUABLY, BEHAVIORAL BIASES
CAN BE VIEWED SIMPLY AS A
FOURTH TYPE OF MARKET FAILURE.

Arguably, behavioral biases can be viewed simply as a fourth type of market failure, albeit one for which the economic literature is rather more nascent and the implications less well established.

B. CONSUMER BIASES AND THE ROLE OF CONSUMERS IN DRIVING COMPETITION

In order for consumers to drive competition in the way described above, they ideally need to:

- a) **access** information about the various offers available in the market;
- b) **assess** these offers in a well-reasoned way; and
- c) **act** on this information and analysis by purchasing the good or service that offers the best value to the customer.

When any of these three elements of the consumer decision-making process breaks down, consumers' ability to drive effective competition can be harmed.

The knowledge that consumers play a key role in driving competition is not new or indeed specific to behavioral economics. Moreover, the role that various consumer barriers can play in hindering the competitive process has been examined for several decades. For example, the key role of search costs in obstructing consumers' ability to access information, and the impact this has on competition, was shown nearly forty years ago by Diamond in his famous paradox in which one finds many firms charging monopoly prices.⁷ Likewise the role of switching costs in creating a barrier for consumers to act, and the potential this has to soften competition, has been well studied by academics and understood by competition authorities.⁸

Behavioral economics, however, provides several important insights that go beyond this existing literature. First, it highlights that consumers may find it hard to assess information and compare across products. This step in the consumer decision-making process had been largely overlooked in the standard literature. Second, it allows us to better understand the underlying causes of search costs (which affect access) and switching (which limits ability to act). This is important because understanding the underlying causes of these search and switching costs can be key in designing effective remedies to address them. Third, the prevalence of consumer biases may mean that existing problems within the consumer decision-making process are more entrenched than we had believed, and more prevalent.

THE PREVALENCE OF CONSUMER BIASES MAY MEAN THAT EXISTING PROBLEMS WITHIN THE CONSUMER DECISION-MAKING PROCESS ARE MORE ENTRENCHED THAT WE HAD BELIEVED, AND MORE PREVALENT.

The following examples illustrate how behavioral biases can affect each of the three steps in the consumer decision-making process.

- a) **Accessing information.** Behavioral biases may exacerbate existing problems for consumers in accessing information. For example consumers tend to look at relative costs rather than absolute search costs. This means a consumer may be willing to travel an hour across town for a half price offer on a £20 pen, but would not travel an hour across town for £10 off of a £500 television even though the amount saved (£10) would be the same. This may imply that search costs are more prevalent on large ticket items than small ticket items. Consumers may also fail to anticipate add-ons and search only on headline price, or consumers may forget previous experiences.⁹
- b) **Assessing offers.** Behavioral biases may create or exacerbate consumers' difficulties in assessing the best deal. For example, consumers' ability to assess which product would suit them best may be impaired by incorrectly anticipating risk, underestimating or overestimating

future use, or overweighting the present.¹⁰ Furthermore, when faced with more information than can be easily analyzed, they may only look at a sub-set of information and use rules of thumb to assess the information.¹¹ Finally they may be distracted by the way in which information is framed and presented.¹²

- c) **Acting on information and analysis.** Behavioral biases may give rise to, or exacerbate, consumers' difficulty in acting to get the best deals. For example, if consumers have overconfidence in their ability to act in the future, this can create inertia and a tendency to fail to act today.¹³

Thus, while barriers to "access, assess, and act" may exist without behavioral biases, behavioral biases provide both a greater understanding of why and when those barriers exist, how prevalent they are, and how severe a barrier they can present to active and reasoned consumer choice.¹⁴

C. FIRMS' REACTIONS TO CONSUMER BIASES

Such consumer biases are not simply relevant to understanding how consumers act in a market; they also have a bearing on firms' behavior. Where such biases exist, firms can act to exacerbate and exploit them, at every stage in the decision-making process. Indeed, forthcoming OFT research illustrates the way in

which a number of common practices by firms can have significant impacts on the extent of the biases exhibited by consumers.¹⁵

WHERE SUCH BIASES EXIST, FIRMS
CAN ACT TO EXACERBATE AND
EXPLOIT THEM, AT EVERY STAGE IN
THE DECISION-MAKING PROCESS.

This is not a new finding. Arguably, marketing experts have long known it. Moreover, the standard economic literature already indicates that firms may have an incentive to increase search or switching costs in order to increase the barriers.¹⁶ Introducing intuitions from behavioral economics, however, suggests that such behavior may be more prevalent and longer-lasting than initially thought:

- a) **Accessing information.** Firms can make it more difficult for consumers to perform optimal search. For example, behavioral economics shows that consumers do not tend to look at pricing terms that are not provided upfront. Firms may exploit this by putting more of the price into add-on services; restructuring their tariffs,¹⁷ adding clauses within the terms and conditions;¹⁸ or making price searching harder (for example, by drip pricing—only revealing the true price after the customer has spent some time choosing).¹⁹
- b) **Assessing offers.** Firms can make it more difficult for consumers to assess the best deal. Because behavioral economics indicates that consumers have difficulties comparing across differently structured offers, firms may exploit this by obfuscating their prices or increasing choice

or complexity. They may also use price promotions and framing to distract and distort decision-making.²⁰

- c) **Acting on information and analysis.** Firms can make it more difficult for consumers to act to get the best deals. Behavioral economics indicates that consumers may display more inertia than traditionally suggested, perhaps due to overconfidence in their capacity to improve things at a later time. Firms, knowing that consumers display this inertia, can increase switching costs (for example, making consumers use registered post to cancel). They can also use defaults and automatic enrollments, or use time limited offers to inhibit switching.

Of course, in many circumstances, firms' ability to exploit such biases in this way will be limited, for example, by the potential for new firms to enter and provide products which make a virtue out of not exploiting biases.²¹ Such market solutions to problems arising from behavioral biases are discussed in the last section. However, some of the recent behavioral literature suggests that there may be equilibria in which all firms exploit consumer biases and none of them has a unilateral incentive to correct this situation.²²

THERE MAY BE EQUILIBRIA IN WHICH ALL FIRMS EXPLOIT CONSUMER BIASES AND NONE OF THEM HAS A UNILATERAL INCENTIVE TO CORRECT THIS SITUATION.

The nature, prevalence, and self-awareness of consumer biases can also differ across markets, and this too can have an impact on how firms react. For example, in some markets there will be a proportion of consumers that know about their biases and correct for them (termed “sophisticated” in the literature) and a proportion who do not (termed “myopes”). In such markets, firms may have an incentive to exploit the myopes, but competition will force them to compete away some of the resulting rents on low upfront prices in order to entice them in the first place. This is competition from which the sophisticated gain. Effectively, the sophisticated get a better price than they would absent the exploitation of the myopes.

In such a situation, any firm that tried to stop exploiting the myopes would have to raise its initial price, which would, in turn, cause both types of customer to switch away. The myopes switch because they no longer see a cheap upfront price, and the sophisticated switch because they are no longer subsidized by myopes. The result is that under certain conditions no firm can profit from moving to a non-exploitative outcome unilaterally.

D. THE IMPLICATIONS OF CONSUMER BIASES FOR COMPETITION

So behavioral economics has established the existence of consumer biases, and added to our understanding of how these can be exacerbated by firms. What, though, are the implications of this for competition, and is there a difference between the short term or “static” and the longer term or “dynamic” effects?

As is clear from the virtuous circle discussed above, static competition will be affected by the failure of consumers to access, assess, and act on information. Passive consumers do not provide the same type of constraints on firms as active consumers do. In traditional economic terms, this reduction in price sensitivity is similar to a general reduction in both the product's absolute elasticity, and its substitutability (or cross-elasticity) with other products.¹³ Such reductions in substitutability can translate into a lessening of the intensity of competition—a softening of competition—and, as a result, higher prices for consumers.²⁴ Again, while this potential for harm was discussed within the traditional switching and search costs literature, behavioral economics has brought insights as to its underlying causes, its prevalence, and its magnitude.

For example, in the context of add-on pricing or aftermarkets, the static harm from softening competition manifests itself in two ways. First, there is a direct loss in consumer welfare from overall higher prices—although some of this harm may be competed away in the primary market depending on the level of competition. This harm will tend to be greater when there are more myopic consumers who are unaware of their biases, and less competition. Second, there is a loss in allocative efficiency that exists even if there is perfect competition. This results from over-consumption of the subsidized primary product and under-consumption of the expensive add-on by the sophisticates.²⁵

Dynamic competition may also be affected by consumer biases within the market. One of the key benefits of competition is the role it has in ensuring that those firms that provide the best value continue in the market while those that provide poor value exit. Over time this evolutionary role of competition implies that the average efficiency of the market increases for all consumers. This role is diminished when consumers no longer reward those firms that provide them with what they really want but, instead, reward those that best play on their biases. For example, if consumers perform only limited search, then firms might compete on, and be rewarded for, being the first to attract consumers, rather than

offering the best deal. This potentially implies overuse of resources on advertising or paying for the prominence of their product on an internet search site rather than providing a lower-priced or higher-quality product.

THIS ROLE IS DIMINISHED WHEN
CONSUMERS NO LONGER REWARD
THOSE FIRMS THAT PROVIDE THEM
WITH WHAT THEY REALLY WANT
BUT, INSTEAD, REWARD THOSE
THAT BEST PLAY ON THEIR BIASES.

The other key dynamic role that competition plays is as an efficient framework to promote product and process innovation. In general, competition among innovators increases the intensity of innovation and development.²⁶ When consumers have behavioral biases this may reduce the ability of firms with innovative products to win customers. This, in turn, may reduce firms' incentives to invest in the research or innovation needed to generate new products. Such reductions in the dynamic role of competition may be far greater than any static effects on competition given their long-term nature.²⁷

Finally, to the extent that these biases facilitate firms' ability to foreclose markets—as will be discussed in the next section—behavioral biases may also retard competition and dynamic efficiency.

The types of static and dynamic concerns outlined above have played a role in several recent OFT market studies, including Sale and Rent Back²⁸ and the Personal Current Accounts (“PCA”) market study. For example, within the PCA market study, a combination of consumer biases and firm behavior led the OFT to conclude that the market was not working efficiently.

“...A significant number of consumers do not know how much they will effectively pay in bank fees or how individual elements in the charging structure will be implemented, either before or after they are incurred [...] this means banks have less incentive to provide better offers on insufficient funds and interests. Without better offers from banks, however, consumers have little incentive to switch. [...] The OFT believes that the market may be stuck in an equilibrium that does not work well for many consumers.”²⁹

In summary, behavioral economics has provided practitioners with greater insights into how consumer biases may create new distortions or accentuate existing distortions in competition.

III. Behavioral Biases and Antitrust Enforcement

What, then, are the implications of behavioral biases for the realm of antitrust enforcement, which covers mergers, abuse of dominance, and anticompetitive agreements?

The first point to highlight is that standard antitrust policy is not necessarily well designed to address the demand-side effect of consumer biases described above—that of distorting or weakening the virtuous circle of competition. However, antitrust enforcement is just one tool in the toolbox available for solving market problems. Other tools include consumer policy, market studies, investigations (in the UK at least), and even the potential for authorities to advocate legislation in a particular market.

This choice requires consumer and competition policy to work closely together; picking the best tool to fix the problem and not simply thinking about which has traditionally been used. In this regard, the UK is in a relatively unique situation in having a third type of instrument that sits between pure antitrust instruments and consumer instruments—market studies and investigations.³⁰ These

are designed to examine features of the market that distort competition, arguably precisely the type of concerns that consumer behavioral biases may create.

Nevertheless, even within standard antitrust enforcement, there are a number of ways in which behavioral biases can potentially have an impact. This is an underdeveloped area, and thinking is at an early stage, but in this section we put forward a few initial ideas that may merit further development.

A. THE IMPACT OF CONSUMER BIASES ON ANTITRUST ENFORCEMENT

Our focus initially is on the consumer biases discussed above. We examine how these could potentially play into bundling and tying cases, aftermarket cases, and collusion cases.

1. Consumer Biases and Bundling and Tying

Tying and bundling practices, carried out by a dominant firm, can be anticompetitive if they significantly raise the cost to competitors of competing, and thereby foreclose the market. A key piece of evidence in such a foreclosure story is whether tying creates a significant switching cost for customers in switching to rival products. Behavioral biases can clearly be relevant here. In general, one might assume that where customers face only nominal costs to switch products,

HOWEVER, BEHAVIORAL
ECONOMICS SUGGESTS THAT EVEN
SMALL SWITCHING COSTS CAN
HAVE SIGNIFICANT EFFECTS ON
CONSUMER BEHAVIOR
IN THE PRESENCE OF CONSUMER
INERTIA, ENDOWMENT EFFECTS,
AND DEFAULT BIAS.

then tying or bundling is unlikely to be able to foreclose. However, behavioral economics suggests that even small switching costs can have significant effects on consumer behavior in the presence of consumer inertia, endowment effects, and default bias.³¹ This can, in turn, make foreclosure more likely to occur through tying and bundling.

Arguably, such a bias played a role in the Media Player element of the EU Microsoft case.³² Microsoft set its own product, Media Player, as the default program to play all media when consumers bought a PC. Setting this as a default when viewed through a rational lens may be unimportant since consumers were easily able to download other media players for free, and at only a minimal cost to them in terms of time. However, when viewed through a behavioral lens, it becomes clear that consumers are significantly less likely to switch from the preloaded Microsoft settings than might otherwise be expected. In this setting, a strategy to foreclose could move from being unlikely to being much more plausible.

It is worth adding, however, that behavioral biases can also help to provide an efficiency rationale for tying and bundling. If consumers find it difficult to make complex choices, they may value being provided with a fully bundled product

where these choices are effectively made for them. Such efficiency rationales need to be weighed against any potential for anticompetitive effect.

2. Consumer Biases and Aftermarkets

Much of the past economic literature and commentary in aftermarkets has assumed that consumers are rational and that they can access, assess, and act on the necessary information. Under these assumptions, any attempt to monopolize the aftermarket will simply result in consumers acting to switch primary products; hence, such conduct is unprofitable. This is one of the main tenets underpinning the Chicago School's view in this area—if monopolizing the aftermarket through refusal to supply is unprofitable, then firms must do it for efficiency reasons.

However, recent developments in the economic literature, including the behavioral literature, have questioned the Chicago critique in several different ways. First, the fact that firms can make greater profits from more confused consumers may provide firms with an incentive to exacerbate the impact of consumer biases.³³ Second, the level of profits competed away in the secondary market depends on the degree of competition in the primary market, with only perfect competition leading to all the profits being competed away in the primary market.³⁴ Finally, as discussed in the previous section, it cannot be assumed that firms have a unilateral incentive to educate consumers to their biases.³⁴

THE CENTRAL THRUST OF THIS LITERATURE IS THAT, CONTRARY TO THE CHICAGO CRITIQUE, THERE MAY EFFECTIVELY BE MORE THAN “ONE MONOPOLY PROFIT.”

The central thrust of this literature is that, contrary to the Chicago critique, there may effectively be more than “one monopoly profit.” Hence, one cannot assume that a firm has no incentive to foreclose the aftermarket. The observation that the “one monopoly profit” theory only holds under certain assumptions is not new. There is, by now, a well-established “post-Chicago” literature examining circumstances in which the Chicago critique does not hold. Behavioral economics simply highlights that consumer rationality is another key assumption underlying this critique.

3. Consumer Biases and Collusion

Finally, the behavioral literature points to the possibility that firms can potentially increase their joint profits by agreeing to exploit consumer biases and thereby soften competition; for example, by restricting or obfuscating the information provided to customers. There would seem to be little reason why such an agreement should not be viewed as anticompetitive, even if there is no agreement relating specifically to prices or volumes.

A nice example of such a concern is the U.S. Federal Trade Commission's (“FTC”) case on bulletproof vests.³⁶ In this case, the association of bulletproof

vest manufacturers adopted a rule that prevented comparative advertising. Under this rule, no member could represent that another member's vests had failed certification testing, even if the advertising claim were true. The FTC determined that the rule had the impact of restraining quality competition, obviously a key competitive parameter in bulletproof vests!

B. SUPPLY-SIDE BIASES AND ANTITRUST³⁷

The focus of this paper has so far been on consumer (or demand-side) biases. Firms on the supply side have been assumed to act rationally, and in a profit-maximizing manner. But is this assumption always correct?

There are several reasons to expect that firms will, in fact, tend to be more rational than consumers when making decisions. First, firms can benefit from economies of scale in decision-making. Consider firms as purchasers. Whereas a consumer might buy a product just once, for a relatively small amount of money, firms' purchasing decisions will tend to involve repeated spend of large amounts of money. As such, firms will have a stronger incentive to make the investment required to get this purchasing decision right.

Second, the market might be expected to discipline firms that make mistakes or depart significantly from profit-maximization behavior. Competition can drive poor performers (bad decision makers) out of the business and reward the better ones. In contrast, consumers who make mistakes will not (typically) be driven out of the market in quite the same way.

Third, firms will tend to hire in analytical expertise, and will have departments who specialize in making business decisions and are evaluated on their efficiency in doing so. Workers and managers that are inefficient, or do not learn from their mistakes, may be less likely to climb the promotion ladder to positions of control.

THERE IS A GROWING EMPIRICAL
LITERATURE THAT PROVIDES
EVIDENCE TO SUPPORT
THE NOTION OF NON-RATIONAL
BEHAVIOR BY FIRMS.

Nevertheless, despite these arguments, there is a growing empirical literature that provides evidence to support the notion of non-rational behavior by firms.³⁸ Explanations relate in particular to the facts that: firms operate in complex environments and need to solve complex problems, and thus tend to resort to short-cuts and rules of thumb just as consumers do; they typically function on the basis of group decision-making, which can itself lead to biases; and the nature of recruitment, promotion, and entrepreneurialism implies that the people who run firms will often have specific characteristics such as over-confidence and willingness to take risks.

Because the literature is relatively nascent there are no strong conclusions to draw. Indeed, in a recent paper commissioned by the OFT and also published within this journal,³⁹ Armstrong & Huck suggest that the implications of firms'

behavioral biases on competitive outcomes are ambiguous and depend on circumstances. In some situations, supply-side biases might help to promote competitive outcomes, while in others they might hinder.

1. Supply-Side Biases and Collusion

An example of these mixed findings can be seen in the impact of firms' behavioral biases on the likelihood of collusion. Firms colluding to keep prices high typically face a short-term incentive to cheat on the agreement (since they will gain market share) but a long-term incentive to sustain a high price. When firms have behavioral biases, the ease of sustaining collusion could increase or decrease. For example:

- a) **Collusion could be harder to sustain** if one assumes that firms have a desire to maximize relative profits (rather than absolute profits). Under this assumption the benefit to deviating from the collusive price increases, since firms will not only gain from the increased profit of cheating (lowering price to steal business from a rival) as they do in the rational models, but will also generate utility from reducing a rival's profit (relative to their own).
- b) However, **collusion could also be easier to sustain**. For example, there is evidence that personal friendship and trust can play an important role in sustaining collusion, with cartel members often investing a lot of time and effort in individual relationships.⁴⁰

This latter point has several interesting implications. First, one may view private information exchanges more cautiously given that the communication often associated with such exchanges may facilitate the trust.⁴¹ Second, this need to create a relationship implies that although cartels may be more costly to set up and, hence, less common than might be thought, they may also be more stable to shocks due to the relationships and last longer.

2. Supply-Side Biases and Mergers

Even if firms were found not to act in a fully profit-maximizing manner, it is far from obvious that this would significantly affect much of merger analysis, since a merger might typically be expected to change the incentives of the merging parties in a similar way, irrespective of what those firms are maximizing.

However, there may still be some subtle implications of supply-side biases for mergers which merit further consideration. For example, there is some evidence in the behavioral literature that firms give greater weight to fixed- and sunk- costs in their pricing decisions than might be expected by standard economic theory.⁴² Could this have any

EVEN IF FIRMS WERE FOUND NOT TO ACT IN A FULLY PROFIT-MAXIMIZING MANNER, IT IS FAR FROM OBVIOUS THAT THIS WOULD SIGNIFICANTLY AFFECT MUCH OF MERGER ANALYSIS.

implications for the treatment of fixed/sunk cost efficiencies arising from mergers? Such efficiencies are not typically accepted as relevant in merger cases, since they are not expected to feed through to consumers as lower prices. But does this reasoning hold true in a situation where there is reason to believe that firms use pricing rules that do reflect fixed- and sunk- costs?

3. Supply-Side Biases and Entry

Supply-side biases may also have implications for the way entry is assessed in antitrust cases. There is significant empirical evidence to show that firms are often overconfident when it comes to predicting their success on a market.⁴³ A very large percentage of firms that enter into a new market fail within a short period of time. Does this mean that it is right to be more cautious than we would otherwise be when considering the possibility of entry as a countervailing force for the creation or abuse of market power?

In summary, the discussion in this section has provided some thoughts on how behavioral biases could potentially affect antitrust cases. It is intended to provide a flavor of early thinking on this issue, rather than a complete assessment or a statement of how competition authorities will be altering their analysis going forward. In the end, whether the existence of behavioral biases is likely to impact on any given antitrust case will need to be analyzed on a case-by-case basis.

IV. Implications for Interventions in Markets

Some elements of the preceding discussion might seem to point to behavioral biases creating more need for intervention in markets. However, this would be too strong a policy conclusion to draw. While behavioral biases can exacerbate problems in markets, it is important not to throw out all we have learned with

WHILE BEHAVIORAL BIASES CAN EXACERBATE PROBLEMS IN MARKETS, IT IS IMPORTANT NOT TO THROW OUT ALL WE HAVE LEARNED WITH REGARD TO WHEN ONE SHOULD INTERVENE.

regard to when one should intervene. In most circumstances, the pricing, marketing, and advertising practices of firms can still be viewed as benign with no need for action, even where they undoubtedly seek to exploit behavioral biases.

More generally, markets can be self-correcting and interventions can potentially do more harm than good. It will, therefore, typically be important to make a careful assessment of interventions before they are put in place.

A. PROBLEMS IN MARKETS CAN BE SELF-CORRECTING

It will be unnecessary to intervene, and could indeed have negative unintended consequences, where the problems in the market would otherwise be self-cor-

recting. This will be true, for example, where there are market solutions, where consumers may learn, or where firms can self-regulate.

1. Market Solutions

In many instances, the problems arising from behavioral biases will be solved by the actions of market participants themselves. For example, in response to consumer self-control problems in terms of attendance at the gym, we observe some gyms providing options for consumers to pay for blocks of visits.

This is part of a broader class of situations where firms may not have an incentive to exploit consumers' biases. For example, in the models of add-on prices discussed in the previous section, the proportion of myopes in the population can play a key role in determining whether firms reveal their add-on prices or exploit these consumer biases by keeping these prices hidden.⁴⁴ As the proportion of myopes declines, there are too few of them for the firms to base their price structure on; hence, firms choose to reveal the add-on price. An interesting implication from this result is that the market may require a catalyst in order to change from an equilibrium in which all firms want to exploit consumer biases to an equilibrium in which all firms want to help consumers by revealing their prices.

AN INTERESTING IMPLICATION FROM THIS RESULT IS THAT THE MARKET MAY REQUIRE A CATALYST IN ORDER TO CHANGE FROM AN EQUILIBRIUM IN WHICH ALL FIRMS WANT TO EXPLOIT CONSUMER BIASES TO AN EQUILIBRIUM IN WHICH ALL FIRMS WANT TO HELP CONSUMERS BY REVEALING THEIR PRICES.

The media can play this catalyst role by making consumers aware of their biases, or at least aware of the tariff structures that exploit their biases. This may result in a virtuous circle in which the more consumers understand about the situation, the less firms try to exploit their biases. For example, in personal banking in the UK, the OFT recently lost an appeal relating to its proposed intervention on unauthorized overdraft fees.⁴⁵ Nevertheless, the substantial publicity surrounding this case may well have been a factor in a variety of changes in the market. Some smaller banks are positioning themselves explicitly as offering a simple deal, while some larger ones are promoting new tariffs without overdraft charges or have restructured their tariffs to include daily charges rather than usage charges. While it is too early to tell whether these will be successful in driving better outcomes for consumers, these examples illustrate how information can work alongside competition to provide incentives for firms to overcome market failings.

Advisors and intermediaries can also play a catalyst role in improving consumer decision-making where there are consumer biases. Consumer organizations, such as Which? in the UK, advise consumers of potential pitfalls (i.e. hidden terms or prices) and make recommendations to help reduce complexity.

Arguably, the reach and effectiveness of intermediaries have been greatly extended with the advent of the internet and the ability to compare prices and terms across different sellers.

Of course, intermediaries do not necessarily provide a panacea. Indeed, there may be cases when incentives of the intermediaries are not always aligned with consumers. For example, when firms pay intermediaries for their advice to consumers then their impartiality may well be questioned. However, there are many cases where market solutions, of one sort or another, will work well. And where they do, this will typically be preferable to intervention, given that the market is typically better placed to devise solutions than a competition authority.

2. The Power of Learning

Even if firms have an incentive to mislead consumers this may not be possible (for long) if consumers learn from their mistakes. For example, a deodorant advert in which consumers are told the deodorant will make them irresistible to women is, no doubt, attempting to exploit a bias. Despite this, it is unlikely to raise significant consumer or competition concerns primarily because consumers can easily guess it is simply not true and even if they don't, they will soon learn! A consumer who learns can switch supplier or purchase more intelligently.⁴⁶ This will mean those firms who have a reputation for dealing fairly with consumers will thrive, while those that treat them poorly will gain a poor reputation and exit.

There are clearly limits to learning. In markets where consumers make frequent purchases (or can benefit from the learning of others via word of mouth) learning may be expected. By contrast, when purchases are infrequent or large value (for example, when entering into a sale and rent back arrangement), then learning may not provide the constraint required. Similarly, there will be circumstances where biases are hardwired (for example, limits to computation can not

be overcome) or where consumers cannot learn from others (for example, word of mouth may be limited for products, such as sale and rent back, which consumers are embarrassed to admit they have bought).

HOWEVER AN INTERESTING RESULT
OF THE BEHAVIORAL LITERATURE IS
THAT IT MAY NOT BE ESSENTIAL
FOR CONSUMERS TO "CORRECT"
THEIR BEHAVIORAL BIASES.

However an interesting result of the behavioral literature is that it may not be essential for consumers to "correct" their behavioral biases.⁴⁷ As long as consumers learn that they have the bias, then they will make allowances for this in their behavior. This will, in turn, limit the extent to which firms can exploit such biases. The implication is that educating consumers about their biases, even if this does not change them, may be sufficient to remove much of the associated consumer detriment.

Moreover, there are wider benefits of allowing consumers to make mistakes and learn from them. Such experiences will teach consumers market skills that are transferable across many of their day-to-day decisions in markets. This may, in turn, enhance consumers' active involvement in markets.

3. Self-regulation

Finally, self-regulation can also play an important role in improving consumer decision-making or ensuring firms do not exploit consumer biases. Self-regulation occurs where firms opt to join schemes that require them to behave in particular ways. This can be helpful where firms do not have a unilateral incentive to improve market outcomes but might have a collective industry incentive to do so.⁴⁸ For example, if reducing price complexity could increase industry-wide demand, by making consumers more confident to enter the market, then this may be something that could be achievable through self-regulation.

SELF-REGULATION CAN ALSO
PLAY AN IMPORTANT ROLE IN
IMPROVING CONSUMER DECISION-
MAKING OR ENSURING FIRMS DO
NOT EXPLOIT CONSUMER BIASES.

A more general example of a self-regulatory body is the UK Advertising Standards Authority ("ASA"). Part of the ASA's responsibility is to adjudicate over claims of false, or misleading, advertising. In doing so it ensures that firms do not unduly attempt to play on consumer behavioral biases through such techniques.

B. INTERVENTION CAN POTENTIALLY DO MORE HARM THAN GOOD

Where markets would otherwise self-correct, intervention can clearly be unnecessary or even harmful. Moreover, even if markets cannot self-correct, care must be taken when intervening because it is not always clear that interventions will improve outcomes for consumers. This is nothing new, having been recognized by John Stuart Mill over 150 years ago:

"All errors which [man] is likely to commit against advice and warning, are far outweighed by the evil of allowing others to constrain him to what they deem his good."⁴⁹

This quote highlights two overarching issues that often overshadow intervention. First, on a principled level, we want solutions that solve the problem, but we do not want to remove consumer choice. This has been described as a "liberal paternalist approach."⁵⁰ Second, there is no guarantee that authorities will necessarily improve the market or not create unforeseen consequences else-

where. Asymmetries in information are inherent in intervention. Firms may have incentives to manipulate the information they provide to authorities in order to gain more favorable outcomes. More simply, it may be that authorities

IT IS WORTH NOTING THAT
ALTHOUGH THIS PAPER ONLY
DISCUSSES CONSUMER AND FIRM
BEHAVIORAL BIASES, THIS DOES
NOT PRECLUDE THE POSSIBILITY
OF AUTHORITIES HAVING
BEHAVIORAL BIASES AS WELL!

simply do not have the level of expertise required to make delicate interventions. In such situations an authority would be wise to be conscious of its own limitations.

Finally, it is worth noting that although this paper only discusses consumer and firm behavioral biases, this does not preclude the possibility of authorities having behavioral biases as well! This, and the previous two points, all cau-

tion us against being too paternalistic even when behavioral biases point to problems within the market.

In summary, we cannot assume that behavioral economics implies more intervention. Markets still can, and may, provide solutions to problems.

C. LESSONS FOR DESIGN OF REMEDIES

While behavioral economics may not necessitate more intervention, there will always be times—just as there has always been—when intervention is necessary.

Under pure antitrust enforcement, intervention will tend to take the form of penalties for infringement, rather than more proactive remedies in the market place. However, this is not necessarily always the case, as shown by the significant number of Article 9 Commitment decisions within EC Article 102 and 101 cases in the last year.⁵¹ While such remedies have typically been based on the supply side in the past, there is little reason why they should not be based on the demand side, if consumer behavior were found to be an important driver of problems in the market.⁵²

Moreover, as mentioned above, there are other tools than antitrust enforcement available for solving market problems arising from behavioral biases. These include consumer enforcement, consumer education, and (in the UK at least) market studies and investigations. There is also potential for authorities to advocate legislation in a particular market.

Where proactive remedies are feasible, they should ideally fit with the liberal paternalistic approach to intervention discussed above. For example, where one outcome is clearly superior to another, it may be possible to design an intervention that defaults consumer behavior to the superior outcome, but without restricting the ability of consumers to make an alternative choice if they so wish. An example of such a remedy is the use of automatic enrollments in pensions to overcome inertia in pension savings. Automatic enrollment nudges those con-

sumers who are “accidentally” under-saving towards a better outcome, while still preserving choice (via an opt out).

Another example is the recent Article 9 remedy in the EU Microsoft Internet Explorer case. As agreed between the Commission and Microsoft, computers in Europe downloaded an update to the Windows operating system.⁵³ Once the update was downloaded the user was presented with a screen providing, in a random order, several different choices of internet browsers (including Microsoft Internet Explorer). By ensuring the user makes an active choice, the intervention was designed to cut the tie between Windows and Internet Explorer. However, in maintaining the choice, the intervention may be described as a remedy in the liberal paternalist vein.

A final example of a positive intervention in this regard may be obligations on firms to require them to help consumers make decisions. For example, rather than centering on directly reducing market power, recent OFT work in Personal Current Accounts in banking has highlighted clarity, transparency, and consumer empowerment as keys to making the market function effectively. This, in turn, may mean that banks in the UK will need to change what information they provide and how they provide it.

However, behavioral economics also tells us that it is important to consider how consumers will react to market interventions. For example, we know that consumers can face behavioral barriers to assessing information. Indeed, it is well documented that consumers do not always read and understand the information provided to them.⁵⁴ This can mean that an intervention that simply improves the information available to consumers will be ineffective in solving market problems.

Moreover, such interventions can even have negative consequences in terms of increasing consumer confusion. For example, a study by the FTC found that revealing to consumers the compensation that mortgage brokers would receive on loans led to consumers placing too much focus on the compensation payments and less on whether or not the loan in question was good value; which, in turn, actually led to them paying more for their loans. It was also found that this placed brokers at a disadvantage to direct lenders and might have led to less competition and higher costs for all mortgage customers.⁵⁵

MOREOVER, SUCH INTERVENTIONS CAN EVEN HAVE NEGATIVE CONSEQUENCES IN TERMS OF INCREASING CONSUMER CONFUSION.

Behavioral economics therefore shows us the importance of making use of “smarter information”—thinking carefully about its framing, the context in which information is read, and the ability of consumers to understand it. A report for the OFT highlighted the positive story around “traffic light” information in food labeling. Here, simple (and consequently less detailed) information

on nutritional value of food led to better outcomes than a full list of food ingredients and nutritional value.⁵⁶

A further concern that can arise around interventions to solve problems associated with consumer biases is that such interventions can be inherently redistributive. In many markets, the gains that firms make from exploiting consumer biases will be to some extent passed back, through the competitive process, to customers who do not exhibit those biases. In this case, there is effectively a form of cross-subsidy between customers, and this may be unwound with intervention. This does not imply that such interventions should not be made, but it is important to be aware that there can be losers as well as winners in such situations.

D. PRACTICAL IMPLICATIONS FOR THE ASSESSMENT OF INTERVENTIONS

One very clear lesson from behavioral economics is that it can be difficult to predict how consumers will react. There are many potential biases at play and theorizing will only get you so far. Empirical evidence can be crucial.

However, behavioral biases can also impact on the design and use of surveys and other empirical techniques. For example, if framing matters, then it is important to design surveys carefully so the way in which questions are framed does not distort the responses provided. While we have always recognized the potential differences between a survey response and what consumers and firms really do, behavioral economics reinforces this message.⁵⁷ Likewise, if framing matters to consumer purchasing decisions, econometric analysis may sometimes need to

ONE VERY CLEAR LESSON FROM
BEHAVIORAL ECONOMICS IS THAT
IT CAN BE DIFFICULT TO PREDICT
HOW CONSUMERS WILL REACT.

incorporate information on the context in which consumers' decisions were made if its results are to provide a full picture of consumer behavior.

Empirical evidence is especially important when designing remedies that are intended to alter consumer behavior and thereby improve competition. One recent example of the importance of empirical evidence is the Market Investigation by the UK Competition Commission ("CC") into Payment Protection Insurance ("PPI"). Here the firms investigated were found to have effective point of sale monopolies of PPI, as they're typically sold as a follow-on product alongside other financial products. As a result, prices were found to be very high.⁵⁸

The CC proposed a package of measures (which affected the supply and demand side) to bring competition into the market. This included a prohibition on the sale of PPI during the sale of the credit product and for seven days afterwards. However, the CC's appeal body, the Competition Appeal Tribunal ("CAT"), rejected this remedy and sent it back to the CC for further consideration.⁵⁹ The CAT contended that the CC had not provided sufficient evidence

regarding how consumers would actually respond to the remedy and whether it would benefit them.

In practice, our ability to design appropriate remedies is likely to be enhanced by empirical research (often experimental) that can capture how consumers will really respond to mooted solutions. The importance of road testing remedies is outlined in a 2009 OFT report.⁶⁰

IN PRACTICE, OUR ABILITY TO
DESIGN APPROPRIATE REMEDIES
IS LIKELY TO BE ENHANCED
BY EMPIRICAL RESEARCH
(OFTEN EXPERIMENTAL) THAT
CAN CAPTURE HOW CONSUMERS
WILL REALLY RESPOND
TO MOOTED SOLUTIONS.

V. Conclusion

This paper asked whether behavioral economics questions the foundations of antitrust. The answer, like many answers to economic questions, is nuanced. Behavioral economics does question some of our current thinking and it may alter how we carry out some of our analysis. However, it does not represent the fundamental shift some would advocate (and some would fear).

Behavioral economics is an incremental advance in our understanding just as informational economics was to the basic competition model. Economics is an evolving science, changing all the time, and economists are used to this. Two-sided market theory was a relatively recent incremental change in our understanding of how some markets work. This led to direct changes in how we understand these markets and when and how we intervene. The same will be true of behavioral economics, but it is important to resist any claim that behavioral economics means everything must change.

This highlights one last point—where behavioral biases appear to be creating problem, some may advocate abandoning competition for regulation. We discussed above the dangers of over-paternalism and limiting choice. Competition authorities have a key role in reminding government of the benefits that competition and choice bring.⁶¹ In doing so, however, they need to be cognisant of the available evidence on behavioral economics and its implications. We hope that this paper contributes to that goal. ▼

- 1 For a general review of this literature, see S. DellaVigna, *Psychology and Economics: Evidence from the Field*, 47 J. ECON. LITERATURE 315-372 (June 2009).
- 2 Note that behavioral economics does not describe a chaotic world in which consumers make random decisions. In general, the behavioral biases exhibited by consumers are systematic and are often boundedly rational.
- 3 The processing power biases include: choice overload (consumers make choices on sets of information); representational biases (consumers use visible value as a good indicator of hidden value); and rules of thumbs (consumers imitate what other consumers do rather than make their own decisions).

- 4 The framing biases include: relative utility (a consumer's utility is affected by reference points such as past actions); default biases (consumers adopt the default option); and placement biases (consumers' choices depend on where goods are placed on a list—e.g. a tendency to choose the first).
- 5 The time inconsistency biases include: projection bias (consumers expect that they will feel the same tomorrow as they do today); over optimism (consumers over estimate how much they will use a good, or underestimate how much it will cost them); and hyperbolic discount biases (consumers value today disproportionately greater than tomorrow).
- 6 The loss aversion biases include endowment biases (consumers value something more once they have owned it more than before they own it).
- 7 If there are search costs, Diamond found that consumers may not search the market but simply choose a firm randomly. The best response of firms is then to charge a monopoly price to these consumers. P. Diamond, *A Model of Price Adjustment*, 3(2) J. ECON. THEORY 156-58 (1971).
- 8 P.D. Klemperer, *Markets with Consumer Switching Costs*, 102(2) Q. J. ECON. 375-394 (1987) showed that, in the context of a single period model, switching costs could be thought of as a form of artificial product differentiation, reducing the intensity of competition between competitors. In the two period model in J. Farrell & C. Shapiro, *Dynamic Competition with Switching Costs*, 19(1) RAND 123-137 (1988) the result is more complicated as firms may compete more intensely in order to exploit their established base in the second period, thus creating a "bargain and rip-off" cycle. See also Office of Fair Trading, *Switching Costs*, Economic Discussion Paper 5 (April 2003), for a review of the literature.
- 9 For example, S. Agarwal, J. C. Driscoll, X. Gabaix, & D. Laibson, *Learning in the Credit Card Market*, Working paper series (2008) investigate learning in the credit card market. They find that although consumers learn (through negative feedback), this hard-earned knowledge does not fully persist (i.e. knowledge depreciates).
- 10 For example, S. DellaVigna & U. Malmendier, *Paying not to go to the gym*, 96(3) AMER. ECON. REV. 694-719 (2006), using data from three U.S. health clubs find that consumers frequently choose contracts that appear sub-optimal given their attendance frequency. Members who choose a contract with a flat monthly fee pay a price per expected visit of more than \$17, even though they could pay \$10 per visit using a 10 visit pass. They suggest this could be driven by consumer overconfidence about gym attendance.
- 11 For example, V.G. Morwitz, E.A. Greenleaf, & E.J. Johnson, *Divide and prosper: Consumers' reactions to partitioned prices*, 35 J. MARKETING RES. 453-463 (1998) find that, when prices are presented in parts, consumers' ability to recall the entire price for the good is diminished and demand is increased. This might suggest that consumers may be anchoring to the first piece of information seen (generally the base price) and then attributing less importance to later pieces of information (i.e. surcharges or add-ons). Similarly, T. Hossain & J. Morgan, *Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiment on eBay*, ADVANCES IN ECON. ANALYSIS & POL'Y (2005) suggest that consumers treat the base price separately from the handling fee in a natural field experiment they conducted using eBay auctions. The authors found that charging a low reserve price compared to the retail price of the good and high shipping and handling costs resulted in a higher total sales price than the reverse situation (low shipping and handling but high reserve price). This result may also be driven by consumers ignoring or missing the additional costs, although there are alternative behavioral explanations such as endowment bias.
- 12 For example, M. Baye, J. Morgan, & P. Scholten, *Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site*, 52(4) J. INDUS. ECON. 463-496 (2004), using a UK data set of consumer click throughs from Kelkoo.co.uk for 2003 to 2004, found that even though Kelkoo does not order results by price by default, a firm listed first on a search results page still benefited from 17.5 percent higher demand on average than when it was listed second. This is despite the ease with which the consumer can usually reorder the results by lowest price. More generally, A. Tversky & D. Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 (44810) Sci.

- 453-458 (1981) show how psychological principles govern the perception of decision problems and the evaluation of options.
- 13 For example, as well as finding consumers are overconfident about gym use (as suggested above) DellaVigna & Malmendier (2006), *supra* note 10, suggest that consumers might overestimate their propensity to cancel automatically renewed contracts.
 - 14 For a discussion of these and other consumer issues which can arise, see M. Armstrong, *Interactions between competition and consumer policy: A report prepared for the OFT*, OFT991 (April 2008).
 - 15 The research uses a controlled economic experiment to test five pricing frames, whereby the true price is provided in a complex way. The pricing frames investigated are drip pricing, "sales," complex pricing, bait pricing, and time limited offers. Drip pricing is where the consumers see only part of the full price up front and price increments are dripped through the buying process. "Sales" occur where a sale price is given and a pre-sale price is also given as a reference to the consumer, for example "was £2 is now £1." Complex pricing is where the unit price may be difficult to determine, for example "3 for the price of 2." Bait pricing is where sellers may promote a special price but there are only a limited number of goods actually available at that price. Time limited offers are where a price is advertised as only being available for a pre-defined short period of time. The report found that all of these pricing practices have some adverse effect on consumer choice and that most of them do significantly impact on consumer welfare. It suggests that the root of the errors can be found in the existence of the behavioral biases, largely the endowment effect and cognitive errors. Office of Fair Trading, *The Impact of Price Frames on Consumer Decision Making*, Economic Discussion paper, (Forthcoming, April 2010).
 - 16 For example, P. Klemperer, *Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade*, REV ECON. STUDIES 515-539 (1995) finds that in a simple multi-period model with two firms both firms are able to maintain higher prices and earn higher profits in the presence of switching costs than without switching costs. See also A. Banerjee & L. Summers, *On frequent flyer programs and other loyalty-inducing arrangements*, H.I.E.R. DP no. 1337 (1987).
 - 17 For example DellaVigna & Malmendier (2006), *supra* note 10, argue that a reason why gyms favor the use of term contracts with upfront payments may be to lock in overconfident consumers. See also M. Grubb, *Selling to Overconfident Consumers*, 99(5) AMER. ECON. REV. 1770-1807 (2009) who analyzed U.S. mobile phone data to investigate whether the three part tariffs seen within the U.S. mobile phone industry were developed as a means of capturing consumers' overconfidence. He found this was the most plausible of different explanations for the tariff structure. Grubb argues that the model can be reinterpreted more widely to explain the use of flat rates and late fees in rental markets, and teaser rates on loans.
 - 18 See, for example, M. Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47(2) STAN. L. REV. 211-59 (1995).
 - 19 Several commentators have argued that the low-cost airlines are particularly effective in using drip pricing to exploit the fact that consumers are more likely to buy the product after they have invested time in it, see, for example, D. Milmo, *Ryanair Scraps Airport Check-in Desks*, GUARDIAN (Sept. 30, 2009). See also G. Ellison & S.F. Ellison, *Search, Obfuscation, and Price Elasticities on the Internet*, 77(2) ECONOMETRICA 427-452, 03 (2009) who argue that economists should think about firms' active incentives to obfuscate as well as consumers' incentives to search.
 - 20 Indeed one strategy consultancy advertises courses on how to minimize banking competition by increasing the difficulties for consumers to compare across products, stating that: "The likelihood that banks continually try to undersell one another is greater if their price structures make it easy for customers to compare offers. In order to prevent easy comparisons, a bank should create price structures that are clearly distinguishable from those of its rivals. Price systems with several price components are especially effective." G. Wuebker & J. Baumgarten, *Strategies against Price Wars in the Financial Service Industry*, Simon-Kucher & Partners.

- 21 See, for example, C. Shapiro, *Aftermarkets and Consumer Welfare: Making Sense of Kodak*, 63(2) ANTITRUST L. J. 496 (1995). See also H. Beales, R. Craswell & S. Salop, *The Efficient Regulation of Consumer Information*, 24(3) J. L. & ECON. 491-539, §1B (1981).
- 22 See X. Gabaix & D. Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121(2) Q.J. ECON. 505-40 (2006).
- 23 This point is discussed in respect of switching costs within Klemperer (1987), *supra* note 8. But the point is more general.
- 24 This result can be derived from either the switching/search literature (*supra* note 8) or the behavioral literature. With regards to the behavioral literature, Ellison and Ellison (2009), *supra* note 19, examine price data for internet retailers. They show that some retailers engage in obfuscation in order to frustrate consumer search, thus resulting in much less price sensitivity on other products. At the extreme, R. Spiegler, *Competition Over Agents with Boundedly Rational Expectations*, 1(2) THEORETICAL ECON. 207-31 (2006) showed that under certain circumstances firms' prices may be entirely independent of competition.
- 25 In standard competition issues, prices above the competitive level result in underconsumption of the product (the Harberger triangle). In aftermarket there may be two distortions: the underconsumption of the secondary product, but also overconsumption of the primary product sold below cost. In large markets the resulting allocative loss may be significant. For example, in the UK the Competition Commission estimated that the cross subsidy from the price of insurance on loans to loans resulted in an allocative inefficiency in excess of £200m. See Competition Commission, *Market investigation into payment protection insurance, Final Report*, ¶10.494 (January 2009).
- 26 While the relationship between competition and innovation is complicated; in general, competition before the innovation takes place drives faster innovation. For a discussion of the links among competition, productivity, and innovation see *Productivity and Competition, an OFT Perspective on Productivity Debate*, OFT877 (2007).
- 27 See, for example, J. Hausman & G. Leonard, *The Competitive Effects of a New Product Introduction: A Case Study*, 50(3) J. INDUS. ECON. 237-63 (2002). See also A Petrin, *Quantifying the Benefits of New Products: The Case of the Minivan*, 110 J. POL. ECON. 705 (2002).
- 28 Sale and rent back is a product which allows consumers who are in difficulties in making mortgage payments to sell their home to a company and then rent it back from them. *Sale and rent back: An OFT Market Study*, OFT (October, 2008).
- 29 *Personal Current Accounts in the UK: An OFT Market Study*, OFT (July 2008).
- 30 For example, the UK market investigation reference into payment protection insurance ("PPI") is discussed later in this paper. In this case, the OFT could potentially have attempted to bring an Article 102 abuse case against each of the suppliers of credit and PPI. Alternatively, action under consumer legislation could have been used. However, the OFT took the view that a reference to the Competition Commission was most appropriate, since the CC could consider all aspects in considering whether a problem existed and, if so, how best to remedy the problem both from the competition and consumer perspective. See, *Payment protection insurance: Report on the market study and proposed decision to make a market investigation reference*, OFT869 (October 2006).
- 31 It is noteworthy that the Vertical Restraints Block Exemption Guidelines explicitly mention that tying may lead to supra-competitive prices, especially "in the case of long-term contracts or in the case of aftermarket with original equipment with a long replacement time [as] it becomes difficult for customers to calculate the consequence of the tying." *Commission notice of 13 October 2000: Guidelines on Vertical Restraints*, COM(2000/C 291/01), Official Journal C 291 ¶217 (October, 2000).

- 32 Case COMP/C-3/37.792 Microsoft.
- 33 See Gabaix & Laibson (2006) *supra* note 22. See also S. DellaVigna & U. Malmendier, *Contract Design and Self-Control: Theory and Evidence*, Q. J. Econ. 119: 353-402 (2004) who describe a model in which consumers make their purchase decisions based on naïvely low estimates for their use of the secondary good. Once consumers are locked into the primary good they find they use the secondary good more often than initially estimated. Firms can exploit this naivety by extracting the consumers' rents upfront, but then extracting an additional rent when consumers are locked in. The result is more than one monopoly rent.
- 34 There are a few exceptions to this. First, in Spiegler's model firms respond to a greater number of competitors by hiding their prices even more, *supra* note 24. Second, in other general models involving product differentiation all profits are competed away only when the market is fully covered—that is when industry demand is independent of price. When this unrealistic assumption is relaxed, the amount of profits the firm can retain in exploiting the secondary market is proportional to the degree of primary competition. See, for example, C. Genakos & T. Valletti, *Testing the 'Waterbed' Effect in Mobile Telephony*, CEIS Working Paper No. 110 (January 2008).
- 35 See Gabaix & Laibson (2006), *supra* note 22.
- 36 Federal Trade Commission, Personal Protective Armor Ass'n, 59 Fed.Reg. 19,019 (1994).
- 37 This section draws from a recent paper by Armstrong & Huck. See M. Armstrong & S. Huck, *Behavioral Economics as Applied to Firms: A Primer*, 6(1) COMPETITION POL'Y INT'L 3-45 (Spring, 2010), also published as M. Armstrong & S. Huck, *Behavioral Economics as Applied to Firms: A report prepared for the OFT*, OFT 1213.
- 38 For a discussion of the literature see Armstrong & Huck, *Id.*
- 39 *Id.*
- 40 See M. Levenstein & V. Suslow, *Cartel bargaining and monitoring: The role of information sharing*, THE PROS AND CONS OF INFORMATION SHARING (Mats Bergman ed.), Swedish Competition Authority (2006).
- 41 This may also address some of the criticism of the "cheap talk" literature. This literature argues that there is no point in exchanging information in order to monitor an agreement if it can't be verified, since no firm is going to cheat on the agreement but still provide the true information that reveals they have cheated. However, the role of private information exchange may not be to monitor the cartel but rather to create and sustain the trust needed to maintain the cartel. See J. Farrell & M. Rabin, *Cheap Talk*, 10 J. ECON. PERSPECTIVES 103-118 (1996).
- 42 N. Al-Najjar, S. Baliga, & D. Besanko, *Market forces meet behavioural biases: Cost misallocation and irrational pricing*, 39 RAND 214-23 (2008) describe recent physiological and experimental literature suggesting that firms may confuse fixed, sunk, and variable costs. They show that even, if pricing strategies that increase profits do better in the market, there may be equilibrium in which there are non-standard pricing practices.
- 43 See, for example, U. Malmendier & G. A. Tate, *Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction*, AFA 2004 San Diego Meetings (March 15, 2003).
- 44 See Gabaix & Laibson (2006), *supra* note 22.
- 45 See *Judgment: The Office of Fair Trading v Abbey National plc & Others*, UK Supreme Court, UKSC 6 (2009).
- 46 Of course where a mistake leads to irrevocable and significant detriment (for example, buying the wrong pension) learning may not be sufficient to prevent serious harm.

- 47 See DellaVigna (2009), *supra* note 1.
- 48 Self-regulation may also create competition concerns as it may provide opportunities for anticompetitive practices, such as foreclosure or price-fixing. For more details, see Office of Fair Trading, *The Economics of Self Regulation in Solving Consumer Quality Issues*, Economic Discussion Paper OFT1059 (March 2009).
- 49 JOHN STUART MILL, ON LIBERTY (1859).
- 50 For a popular explanation of liberal paternalism see R. THALER & C.R. SUNSTEIN, *NUDGE: IMPROVING DECISION ABOUT HEALTH, WEALTH AND HAPPINESS* (2008).
- 51 Recent Article 9 decisions include Rambus, RWE, Eon, Ship Classification, and Microsoft. See, *Commission accepts commitments from Rambus lowering memory chip royalty rates*, Rambus IP/09/1897 Brussels, (December 2009); *Commission opens German gas market to competition by accepting commitments from RWE to divest transmission network*, RWE, IP/09/410 Brussels, (March 2009); *Commission opens German electricity market to competition*, EON, IP/08/1774 Brussels, (November 2008); *Commission accepts commitments by GDF Suez to boost competition in French gas market*, GDF Suez, MEMO/09/536 Brussels (December 2009); *Commission paves way for more competition in ship classification market by making IACS' commitments legally binding*, Ship Classification, IP/09/1513 (October 2009); *Commission welcomes Microsoft's roll-out of web browser choice*, Microsoft, IP/10/216 Brussels, (March 2010).
- 52 Such an approach has been articulated in the context of enforcing exploitative abuses. See A. Fletcher & A. Jardine, *Towards an Appropriate Policy for Excessive Pricing*, 12th Annual Competition Law and Policy Workshop (June 2007).
- 53 *Commission welcomes Microsoft's roll-out of web browser choice*, *supra* note 51.
- 54 For example, see *Warning: Too much information can harm, A final report by the Better Regulation Executive and National Consumer council on maximising the positive impact of regulated information for consumers and markets*, (November 2007).
- 55 See Federal Trade Commission Bureau of Economics Staff Report, *The Effect of Mortgage Broker Compensation Disclosures on Consumers and Competition: A Controlled Experiment*, (February 2004).
- 56 See Office of Fair Trading, *Assessing the effectiveness of potential remedies in consumer markets*, Economic Discussion Paper OFT994, ¶¶ 4.39 – 4.42 (April 2008).
- 57 The concepts of stated and revealed preference are well understood in competition analysis. For a discussion of the implications of these consequences, see Office of Fair Trading / Competition Commission, *Good practice in the design and presentation of consumer survey evidence in merger inquiries*, Forthcoming for consultation March 2010.
- 58 See Competition Commission, *Market investigation into payment protection insurance*, ¶4.97 (January 2009).
- 59 *Barclays Bank PLC v Competition Commission*, 2009, Competition Appeal Tribunal, Case: 1109/6/8/09.
- 60 See Competition Commission and Office of Fair Trading, *Road testing of consumer remedies*, Economic Discussion Paper OFT1099 (July 2009).
- 61 For a recent report that provides a toolkit for introducing choice and competition into public service markets, drawing on the 'access, assess, act' framework set out above, see Office of Fair Trading, *Choice and Competition in Public Service Markets*, Economic Discussion Paper OFT (March 2010).

Behavioral Economics and Merger Analysis

Alison Oldale

U.K. Competition Commission

Behavioral Economics and Merger Analysis

*Alison Oldale**

The papers in this volume by Eliana Garces, and Matthew Bennett, John Fingleton, Amelia Fletcher, Liz Hurley & David Ruck provide a very clear overview of behavioral economics and its application to competition policy generally. In this note I will comment on some implications of what they have to say for merger analysis.

*Chief Economist, Competition Commission (CC). All views expressed are personal to the author and do not necessarily represent the views of the CC.

I. Introduction & Summary

The papers in this volume by Eliana Garces (“Garces”),¹ and Matthew Bennett, John Fingleton, Amelia Fletcher, Liz Hurley & David Ruck (Bennett *et al.*)² provide a very clear overview of behavioral economics and its application to competition policy generally. In this note I will comment, in particular, on some implications of what they have to say for merger analysis.

II. Consumer Biases

Bennett *et al.* draw a useful distinction between consumer biases and supply side biases, and I start my comments by looking at the implications of consumer biases for merger control. (These comments are relevant only for mergers where the customers are final consumers, and the following should be read as being limited to consumer-facing mergers, even if this is not made explicit.)

In short, there are some intriguing possibilities, some of which have already begun to have a modest impact. However, the field needs further development if it is to have substantial effects. Interestingly, it is unclear whether taking on board the lessons from behavioral economics would lead to more—or less—enforcement.

ANALYZING THE REASONS
WHY CUSTOMERS BEHAVE AS
THEY DO CAN PROVIDE USEFUL
CORROBORATION FOR OTHER
EVIDENCE ABOUT THE
DEMAND FUNCTION.

A. CONSUMER DEMAND

The most obvious area where behavioral economics could influence the practice of merger analysis is in the understanding of consumer demand. The response of customers to changes in the terms of sale is at the heart of all merger analysis. It is often the most important constraint on the actions of firms in a market, and is the focus of market definition analysis.

However, merger control is, in the end, concerned with the impact of a change in market structure. The response of customers to changes in terms affects this but, in many cases, the analysis can proceed taking this response as a given. There is, therefore, an argument that there is no need to understand why customers behave as they do as long as the demand function has been correctly measured. There is some weight to this argument, though it should not be pushed too far: It may sometimes be useful to understand what underlies the demand function.

First, analyzing the reasons why customers behave as they do can provide useful corroboration for other evidence about the demand function. In this respect, the three “A”s introduced in Bennett *et al.* combine to provide a helpful organizing principle for some factors affecting the willingness of customers to switch. The three “A”s are: information about how well consumers Access information

(including search costs), how well they Assess that information, and the degree of their freedom to Act on it (including switching costs). However, it should not be forgotten that information about the extent to which products are, or are perceived to be, differentiated will often be more important.

Second, considering the reasons why customer responses are what they are could highlight possible ways in which the merger might affect the demand function itself, and so suggest reasons why demand should not be treated as a given. This is because firms can sometimes choose to act in ways that exploit consumer biases, causing consumers to make worse choices. Interestingly, however, it is not clear from the papers whether mergers that reduce competition mitigate or exacerbate this problem.

Both papers note that competition can sometimes reduce the extent to which firms benefit from exploiting consumer biases. This can happen if firms can take customers from their rivals by making a virtue of their plain dealing. In these cases, a merger that reduces competition could potentially lead to worse outcomes than would be expected from simply taking the existing demand function as a given.

Both papers also report, however, that the opposite can happen: Theoretical and empirical papers show that obfuscation and complexity can increase with competition, so that mergers that reduce competition could, in principle, make consumers better off by reducing the incentives to obfuscate.

In addition, Garces notes that advertising and branding are close cousins of some of the ways that firms can exploit consumer biases. These are areas where there is a lot of uncertainty about the relationship between the extent of competition and the level of investment by firms, and about the impact on consumers of any investments made.

THEY BOTH ALSO SUGGEST
THAT OUR UNDERSTANDING OF
THE IMPLICATIONS OF CONSUMER
BIASES ON THE EFFECTS
OF A MERGER IS STILL
IN ITS EARLY STAGES.

In sum, although both papers provide reasons why it may be useful to understand—as well as measure—demand, they both also suggest that our understanding of the implications of consumer biases on the effects of a merger is still in its early stages.

B. ENTRY AND THE EVOLUTIONARY ROLE OF COMPETITION

The papers suggest that understanding more about consumer demand could affect the analysis of entry costs. For example, Bennett *et al.* note that if search is costly then firms can invest in ways to attract consumers other than by offering the best deal (by paying to be prominent on search engines, through advertising, etc). This would make entry more costly. Whether this creates a barrier to entry, however, is unclear.

C. PRICE DISCRIMINATION MARKETS

One area where behavioral economics could have an early impact on merger analysis is if poorly informed consumers form a vulnerable group. Competition authorities reviewing mergers will often pay particular attention to the potential effects on groups of customers that can be targeted with discriminatory prices. In some cases such vulnerable groups will be defined as separate markets. The papers suggest that “myopes” might form a vulnerable group, distinguished by their limited ability to gather and process information, or to correct for their biases.

D. MATERIALITY

Finally, Garces especially emphasizes the problems of understanding the welfare effects of changes to a market when consumer choices are not a reliable guide to consumer welfare. This can complicate the already difficult task of deciding whether any loss of competition from a particular merger is material or not.

III. Supply Side Biases

Although the focus of behavioral economics to date has been on the behavior of consumers, Bennett *et al.* discuss recent developments in considering the implications for the behavior of firms. This line of research could potentially have far reaching consequences for merger control, though it is too early to say what these consequences might be. Here I point to some possibilities based on the early indications in the papers.

At a basic level, the reason why research that affects our understanding about how firms behave could affect merger control is, simply, because merger control is all about predicting firm behavior. These predictions are usually based on thinking about how a profit-maximizing firm would behave. If research into behavioral economics uncovers better ways to understand how firms behave, this could obviously affect the analysis.

Bennett *et al.* stress that research in this area is in its infancy and, to date, it has done more to highlight possibilities than to generate concrete improvements in predicting firm behavior. They also note that what matters most is the change in incentives following a change in market structure, not the details of the firm’s objective function. But still it is possible to highlight some areas to watch.

IF RESEARCH INTO BEHAVIORAL ECONOMICS UNCOVERS BETTER WAYS TO UNDERSTAND HOW FIRMS BEHAVE, THIS COULD OBVIOUSLY AFFECT THE ANALYSIS.

A. ENTRY

One area where behavioral economics might have an early impact on merger analysis involves the evaluation of evidence about entry. As Bennett *et al.* note,

it is apparent that many firms enter markets when their chances of success are very slim. This has two implications for evaluating evidence on entry:

1. Attempts at entry may be more likely than would be suggested by evaluating the barriers to entry and considering whether entry would be profitable.
2. Successful entry may be less likely than would be suggested by reviewing the plans of third parties.

B. COORDINATED EFFECTS

Another area highlighted by Bennett *et al.* is the analysis of coordinated effects. This is an area where traditional approaches to understanding how a change in market structure could affect firm behavior have not always been especially helpful to competition authorities seeking to distinguish between good and bad mergers. If behavioral economics could add to what we know, it could make an important contribution. So far Bennett *et al.* report some additional insights. Notably, coordination is generally more likely if there is a degree of trust, fostered by contact and personal relationships. However, this is not yet the clear diagnostic test that competition authorities would like.

NOTABLY, COORDINATION IS GENERALLY MORE LIKELY IF THERE IS A DEGREE OF TRUST, FOSTERED BY CONTACT AND PERSONAL RELATIONSHIPS. HOWEVER, THIS IS NOT YET THE CLEAR DIAGNOSTIC TEST THAT COMPETITION AUTHORITIES WOULD LIKE.

So far Bennett *et al.* report some additional insights. Notably, coordination is generally more likely if there is a degree of trust, fostered by contact and personal relationships. However, this is not yet the clear diagnostic test that competition authorities would like.

C. EFFICIENCIES

Finally, behavioral economics could affect the assessment of efficiencies. On the one hand, as Bennett *et al.* note, there is some evidence that

fixed and sunk cost efficiencies matter more for pricing decisions than traditional models of firm behavior would suggest. This could mean that some fixed and sunk cost efficiencies benefit consumers, whereas competition authorities often assume they could not.

On the other hand, behavioral economics reinforces what competition authorities always suspected about claims that a merger will generate efficiencies: that these should be treated with a degree of healthy scepticism. On top of the problem that efficiencies are inherently easy to claim and hard to prove, there is evidence that firms may believe in the existence of efficiencies, but be wrong. ▼

1 Eliana Garces, *The Impact of Behavioral Economics on Consumer and Competition Policies*, 6(1) COMPETITION POL'Y INT'L 145-52 (Spring, 2010).

2 Matthew Bennett, John Fingleton, Amelia Fletcher, Liz Hurley, & David Ruck, *Behavioral Economics and Competition Policy: Some Potential Implications for the Analysis of Markets and Interventions?* 6(1) COMPETITION POL'Y INT'L 111-37 (Spring, 2010).



CPI COMPETITION POLICY
INTERNATIONAL

Volume 6 | Number 1 | Spring 2010

The Impact of Behavioral Economics on Consumer and Competition Policies

Eliana Garcés
European Commission

The Impact of Behavioral Economics on Consumer and Competition Policies

*Eliana Garcés**

Interesting questions are being asked about the policy implications of relaxing commonly held assumptions about how people make decisions. If consumers are not always rationally maximizing some kind of utility function, can we still claim that their decisions are always in their own best interest? And should this be a policy concern at all? We commonly rely on the competitive process to produce the market outcomes that are the most favorable to consumers. In a model of rational behavior, firms in a competitive environment compete mostly on the merits and the market outcome is efficient and welfare-maximizing. Does this result continue to hold when the rationality assumption about consumer behavior is relaxed?

*Eliana Garcés is currently a member of the Cabinet of the European Commissioner for Competition Policy. The views expressed in this piece are personal and in no way represent the views of the European Commission.

Behavioral Economics is gradually becoming mainstream microeconomics and can no longer be considered the fringe research field that was largely disregarded by neo-classical economists and game theorists until a decade ago. But, notwithstanding its increasing success, behavioral economics is struggling to establish itself as a well-developed tool to analyze markets and provide policy conclusions. This is partly because behavioral economics has done a better job at questioning well-established assumptions underlying traditional models than at providing an alternative framework with a similar analytical predictive power. But one fact, which is often forgotten, is true: a lot of our analytical conclusions about efficiency and welfare rely on the interpretation we give to consumer decisions. These interpretations are directly derived from the “rationality” assumptions at the foundation of the neo-classical supply and demand model. Assumptions of rational and profit-maximizing behavior seem to have been an overall satisfactory approach. But Behavioral Economics is currently investigating the limits to the generalization of this economic rationality framework. It is gathering evidence on its empirical relevance and testing the explanatory power of alternative behavioral hypotheses. More interestingly, applications of behavioral economics to the field of industrial organization examine the consequence of different behavioral hypotheses on the predicted efficiency and welfare outcome of markets.

MORE INTERESTINGLY,
APPLICATIONS OF BEHAVIORAL
ECONOMICS TO THE FIELD OF
INDUSTRIAL ORGANIZATION
EXAMINE THE CONSEQUENCE OF
DIFFERENT BEHAVIORAL
HYPOTHESES ON THE PREDICTED
EFFICIENCY AND WELFARE
OUTCOME OF MARKETS.

A new framework has not yet emerged. But interesting questions are being asked about the policy implications of relaxing commonly held assumptions about how people make decisions. If consumers are not always rationally maximizing some kind of utility function, can we still claim that their decisions are always in their own best interest? And should this be a policy concern at all? We commonly rely on the competitive process to produce the market outcomes that are the most favorable to consumers. In a model of rational behavior, firms in a competitive environment compete mostly on the merits and the market outcome is efficient and welfare-maximizing. Does this result continue to hold when the rationality assumption about consumer behavior is relaxed?

Let us start this discussion with a brief definition of the rational individual. The decision-making process underlying both neoclassical economics and game theoretic models relies on the following assumption: people have pre-existing, well-ordered, and complete preferences. This means they can assign a given value to everything under any contingency. For example, a rational agent currently knows how much she would be willing to pay to rent a bike during a stay in Bangkok next summer. She is in a position to sign a contract today that she will not regret next summer. Factors such as the average price of bikes that she

will observe in Bangkok when she gets there are external to her utility and are not supposed to have an effect on her enjoyment of the contract. Individuals can, therefore, rank at any time all possible choices according to inherent tastes. There is also some stability and continuity in those tastes so that a small change in the characteristics of the product will not dramatically change the value that the individual assigns to it.

Given the value assigned to each available product and their initial resources, individuals maximize the level of welfare they can achieve in economic exchanges. These assumptions about the nature and structure of preferences and the utility maximizing decision rule form the basis of the “rational behavior” of the “homo economicus.” The rational behavior assumption is what allows us to interpret the price paid by an individual as a direct manifestation of his or her

BEHAVIORAL ECONOMICS HAS
PRODUCED EMPIRICAL EVIDENCE
INDICATING THAT INDIVIDUALS
MAY NOT BEHAVE LIKE “HOMO
ECONOMICUS” BUT RATHER USE
THEIR BRAINS IN RICHER
AND MORE COMPLEX—IF NOT
ALWAYS BETTER—WAYS.

preferences and welfare. The “rational” behavioral assumption has been so widely accepted that any departure from it was, for a long time, considered to be out of bounds for mainstream economics.

Behavioral economics has produced empirical evidence indicating that individuals may not behave like “homo economicus” but rather use their brains in richer and more complex—if not always better—ways. This is partly due to the fact that individuals suffer from cognitive

limitations, systematic misperceptions, and emotional reactions to their environment that affect their decision making.¹ One proposition is that people use rules of thumb or they adopt second-best behaviors when faced with either complex decisions or perceived high search costs. Proponents of contingent preferences also argue that preferences and willingness to pay are significantly affected by personal expectations, which are not a fixed concept but rather a function of recent experience and a person’s particular environment at a point in time.² Willingness to pay changes not only with inherent taste, but also with the state a person is in and the information carried by the environment.

Another strand of literature examines how the choices people make are influenced by how the choice is presented to them. The role of default choices and the effect of framing choices have been extensively documented.³ Choice overload has been found in some instances to paralyze and upset people although the literature in this field produces contradicting results depending on the situation.⁴ Too much or too complex information will also drive people to take shortcuts or fixate on a particular dimension of the choice while ignoring the others, thereby making suboptimal decisions. People can also drastically overreact to seemingly anodyne changes in the product description.⁵ A strand of literature has illustrated the time inconsistency of people’s preferences and the tendency to make errors when forecasting future preferences and choices.⁶ In particular, people

have been found to underestimate future needs in favor of present ones and to overestimate how much future preferences will look like current ones, i.e. to underestimate how their needs and circumstances will change.

What is the practical consequence of accepting these new behavioral features in decision making? Information about consumer choices is used to construct the market demand function. In competition policy, this demand function is used to establish the market demand elasticity, an important component for market definition and the analysis of firms' pricing power. Demand functions are also commonly used to calculate consumer welfare under particular market outcomes.

One can argue that once we empirically observe a demand, the underlying mechanism by which this demand is formed is of secondary importance. Observed choices will give us the information we need for market and competition policy analysis. But several issues arise. First, if choices are the results of imperfect or second-best decision-making rules, then the willingness to pay may not be perfectly correlated with the welfare ultimately obtained by the consumer. Second, if preferences and willingness to pay are not exogenous or consistent or stable, then firms may be able to manipulate them. Neoclassical economics already recognizes the possibility that a firm invests in advertisement to increase the demand of its product. In these models, advertisement increases the utility derived from consuming the product and therefore the willingness to pay.⁷ Behavioral economics introduces the theoretical possibility that firms invest in "manipulating" consumer choice in order to increase willingness to pay without necessarily increasing the utility derived from consuming the product.

BEHAVIORAL ECONOMICS
INTRODUCES THE THEORETICAL
POSSIBILITY THAT FIRMS INVEST
IN "MANIPULATING" CONSUMER
CHOICE IN ORDER TO INCREASE
WILLINGNESS TO PAY WITHOUT
NECESSARILY INCREASING THE
UTILITY DERIVED FROM
CONSUMING THE PRODUCT.

Several papers have studied pricing strategies that might be able to produce higher prices and potentially higher profits for firms without increasing consumer utility. The most obvious one is complex pricing or price obfuscation. This is a strategy whereby firms adopt multidimensional tariffs or multiply the described characteristics of a particular product for the sole purpose of decreasing comparability of offers. Excessive personalization of the offer can have a similar effect. For this to be a strategy that exploits behavioral biases, the "complication" of the product or of the terms of the offer must offer no utility to the consumer but just reduce the consumer's willingness or ability to compare alternatives. Ellison and Ellison (2009) provide empirical evidence of offer complication over the internet for the sole purpose of decreasing comparability on internet price search engines.⁸ Pricing strategies in the online sale of airline tickets are another example of price obfuscation through the slicing of the transaction into sequential acquisitions of options and sequential payments of fees. The idea that firms introduce obfuscation to avoid the

potential price pressure of competition is formalized in Spiegler (2006).⁹ This and other papers predict endogenous contract complexity and search costs that increase with competition.

Although the literature has argued that obfuscation strategies have negative effects on consumer welfare and efficiency, there are fewer conclusions on the effect of such strategies on firm profits. Literature on search costs already argued that search costs and differences in search abilities led to prices above the competitive solution.¹⁰ In behavioral models, firms voluntarily increase search costs to protect themselves from competition. Equilibrium with higher profits depends on the lack of incentives for firms to deviate from this obfuscation strategy. Such disincentives may exist either if consumer heterogeneity causes firms providing transparent information to attract less profitable customers, or if the market is susceptible to significant shrinkage if most consumers are made aware of the full cost of the good or service when they make their purchase decision.¹¹ Evidence of price-increasing obfuscation strategies in competitive environments has been presented in the case of retail financial services.¹²

Other pricing strategies play into the inability of consumers to accurately forecast the actual usage of the good or service they buy. Tariffs that offer an entry fee that includes a specified usage and charge high marginal cost for any additional usage can be understood as screening for under- and over-confident customers.¹³ These three-part tariff plans are optimal for those customers who can accurately forecast their use. But consumers who under- or over-estimate their usage pay higher unit prices than foreseen at the time of transaction.¹⁴ Three-part

tariffs contracts are profit maximizing for industries with low marginal costs such as telecommunications or financial services. Most literature on such tariffs does not show supra-competitive profits in competitive settings. But the consumer welfare of over- and under-confident users is nonetheless decreased.¹⁵

BEHAVIORAL ECONOMICS
INTRODUCES THE POSSIBILITY OF
PRICE DISCRIMINATION BETWEEN
CONSUMERS BASED ON THEIR
COGNITIVE ABILITIES, THEIR
INFORMATION PROCESSING
ABILITIES, AND THEIR ABILITY
TO FORECAST FUTURE NEEDS.

It is known that firms implement price discrimination that caters to customers with different tastes. Behavioral economics introduces the possibility of price discrimination between consumers

based on their cognitive abilities, their information processing abilities, and their ability to forecast future needs. Most of these models result in inefficient outcomes for a segment of consumers and, in some cases, there are cross subsidies from one segment of consumers to another based on factors over which the subsidizing users have limited awareness and control. Competitive markets do not resolve the inefficiencies present for these consumers because, as in the complex pricing case, there may be disincentives to do so for any single firm. For example, if some users underestimate the likelihood of running into credit, a credit card company may not have an incentive to compete on the average inter-

est rate. If it did, it would attract the less solvent and riskier customers who are the most sensitive to interest rates while, at the same time, decreasing its profits on those customers who are insensitive to the interest rate and run a balance which they did not foresee. Ausubel (1991) uses this argument to explain interest rate stickiness and evidence of supra-competitive profits in the competitive credit card industry.¹⁶

In addition to pricing strategies, firms may also adopt marketing strategies that exploit the different ways in which consumers deviate from “rational” decision making. Firms spend money on positioning, search engine management, sales effort, and advertisement content to influence consumers’ perceptions of the relevant alternatives for a given choice.¹⁷ Firms may also encourage consumers to focus on irrelevant aspects of the product when making a decision. Examples include advertising investment funds based on past performance and the low marketing importance given to providing management fee information. By doing so, firms exploit a natural tendency of people to extrapolate from the past even when the past carries no information about the future.¹⁸ Firms might be able to use persuasive strategies that lead to decisions that are not welfare-maximizing for consumers.

What have we learned? Behavioral models suggest that consumers might make inefficient decisions in competitive markets and might be overcharged with no resolution by the competitive process. Firms can adopt pricing and marketing strategies that allow them to soften price competition even in competitive environments and in ways that reduce consumer welfare. In other words, competitive market outcomes might not, in every single case, be efficient and welfare-maximizing.

IN OTHER WORDS, COMPETITIVE MARKET OUTCOMES MIGHT NOT, IN EVERY SINGLE CASE, BE EFFICIENT AND WELFARE-MAXIMIZING.

The role of consumer policy in neoclassical economics is to fight fraud and resolve information issues. Behavioral considerations have, nevertheless, already played a role in regulatory initiatives such as the imposition of a cooling off period in certain purchases. The idea that firms might be adopting strategies that create inefficiencies for consumers raises the question of whether such commercial behavior falls under the remit of consumer policy. The increasing ability and incentives to price discriminate against individuals with reduced cognitive or predictive abilities, issues of self control, or high levels of distress also calls for a judgment call on how much we want to care about issues of fairness. There is certainly a case to be made for consumer protection intervention in those markets, such as financial markets, where the consumer risk in the case of a suboptimal decision can be very high. But any remedial intervention will have to be targeted at the particular problem to be solved and must not generate additional inefficiencies. Market remedies should be carefully designed and should aim at improving the conditions for optimal decision-making by consumers. Restricting particular commercial practices might be efficiency- and wel-

fare-enhancing but only if it can be shown that the sole purpose of the practice is to distort the consumer decision.

Competition policy currently relies on market entry and lack of competitor foreclosure to elicit efficient outcomes from markets. Behavioral economics suggests that, in some particular cases, this might be overoptimistic. But competition policy is currently quite powerless in the face of collective “exploitative”

COMPETITION POLICY CURRENTLY
RELIES ON MARKET ENTRY
AND LACK OF COMPETITOR
FORECLOSURE TO ELICIT EFFICIENT
OUTCOMES FROM MARKETS.
BEHAVIORAL ECONOMICS SUGGESTS
THAT, IN SOME PARTICULAR CASES,
THIS MIGHT BE OVEROPTIMISTIC.

strategies adopted by firms in a competitive market. Running successful excessive pricing cases in the face of competition or collective dominance cases is notoriously difficult since it is hard to explain within the traditional framework why competition will not eliminate traders that do not serve consumers well.

The current models cannot be sufficiently generalized yet to provide an operational framework for policy. In most markets, preserving a competitive environment will, in fact, be sufficient for efficiency and welfare maximization. But behavioral models show that one has to be careful of inferring too much from the competitive environment in those markets where behavioral biases and commercial strategies that exploit these biases are likely to play a big role in the transaction decision. In such cases one must be cautious regarding conclusions on the market efficiency and consumer welfare outcome.

But behavioral models show that one has to be careful of inferring too much from the competitive environment in those markets where behavioral biases and commercial strategies that exploit these biases are likely to play a big role in the transaction decision. In such cases one must be cautious regarding conclusions on the market efficiency and consumer welfare outcome.

Behavioral economics is a field that will develop further in the next years. Without a doubt our understanding of how markets work will increase. Meanwhile, rapid technological change is providing both consumers and firms with increased market access and massive amounts of information and data. This will generate radical changes in commercial practices in the years to come. Behavioral economics will play an important role in explaining what is likely to become an increasingly complex and sophisticated commercial environment. ▼

- 1 For an interesting neuroscience perspective, see Camerer, Loewenstein, & Prelec, *Neuroeconomics: How Neuroscience Can Inform Economics*, 43 J. ECON. LITERATURE 9-64 (2005).
- 2 Koszegi & Rabin, *A Model of Reference Dependent Preferences*, 121 Q.J. ECON. 1133-1165 (2006).
- 3 For example, see Johnson & Goldstein, *Do Defaults Save Lives?* 302 SCI. 1338-1339 (2003) on the role of default choices in the rates of organ donations; Benartzi & Thaler, *Heuristics and Biases in Retirement Savings Behavior*, 21(3) J. ECON. PERSP. 81-104 (2007) on the effect of defaults and framing of retirement options on retirement savings decisions.
- 4 Iyengar & Lepper, *When choice is demotivating: Can one desire too much of a good thing?*, 79(6) J. PERSONALITY & SOC. PSYCHOL. 995-1006 (2000); Fasolo, Hertwig, Huber & Ludwig, *Size, entropy, and density: What is the difference that makes the difference between small and large real-world assortments?*, 26(3) PSYCHOL. & MARKETING 254 - 279 (2009).

- 5 Bertrand, Karlan, Mullainathan, Shafir, & Zinman, *What's Psychology Worth? A Field Experiment in the Consumer Credit Market*, Yale University Economic Growth Center Discussion Paper No. 918 (2005), available at SSRN: <http://ssrn.com/abstract=770389>.
- 6 Loewenstein, O'Donoghue, & Rabin, *Projection Bias in Predicting Future Utility*, 118 Q. J. ECON. 1209-1248 (2003); Della Vigna & Malmendier, *Paying not to go to the gym*, 96(3) AMER. ECON. REV. 694-719 (2005).
- 7 See, for example, Becker & Murphy, *A Simple Theory of Advertisement as a Good or Bad*, 108 Q.J. ECON. 942-64 (1993).
- 8 Glenn Ellison & Sara Fisher Ellison, *Search, Obfuscation, and Price Elasticities on the Internet*, 77(2) ECONOMETRICA, 427-452,03 (2009).
- 9 R. Spiegler, *Competition over agents with boundedly rational expectations*, 1 THEORETICAL ECON. 207-231 (2006).
- 10 Salop & Stiglitz, *Bargain and Rip-offs: A Model of Monopolistically Competitive Price Dispersion*, 44 REV. ECON. STUDIES 493-510 (1977).
- 11 See Gabaix & Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121(2) Q. J. ECON. 505-540 (2006); Also Ellison, *A Model of Add-on Pricing*, 120(2) Q. J. ECON. 585-637 (2005).
- 12 Carlin, *Strategic Price Complexity in Retail Financial Markets*, 91(3) J. FIN. ECON. 278-287 (2009).
- 13 Grubb, *Selling to Overconfident Consumers*, Stanford Institute for Economic Policy Research Discussion Paper No. 0618 (2006).
- 14 In this literature, over-optimism, projection bias, and issues of self-control can be sources of systematic bias in some people's forecasts of usage.
- 15 See, for example, Della Vigna & Malmendier, *supra* note 6 at 694-719. This paper also argues that one cannot make inferences on the consumer consumption preferences and consumer welfare from observed contracts.
- 16 Ausubel, *The Failure of Competition in the Credit Card Market*, 81(1) AMER. ECON. REV. 50-81(1991).
- 17 Eliaz & Spiegler, *Consideration Sets and Competitive Marketing* (2009) available at http://www.econ.brown.edu/fac/Kfir_Eliaz/considersept2009.pdf.
- 18 Choi, Laibson, & Madrian, *Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds*, Yale ICF Working Paper N. 08-14 (2008). Available at SSRN: <http://ssrn.com/abstract=1125023>.

The Failed Resurrection of the Single Monopoly Profit Theory

Einer Elhauge
Harvard Law School

The Failed Resurrection of the Single Monopoly Profit Theory

*Einer Elhauge**

I. INTRODUCTORY SUMMARY

I am very grateful to CPI for generously holding a symposium on my article, *Tying, Bundled Discounts, and the Death of the Single Monopoly Profit Theory*.¹ Of course, the downside of having a bunch of academics invited to critique your article is that typically all of them will disagree, and often one will become disagreeable. The Comments in the CPI symposium are no exception to this norm. Luckily, the Comments all disagree with my article in largely different ways, so I can simply address them Comment by Comment.

Professor Paul Seabright claims that an absence of empirical proof supports the single monopoly profit theory. This claim fails because the single monopoly profit theory is an impossibility theorem. It also fails because my recommended exception applies to whatever empirical extent the necessary conditions for the single monopoly profit theory actually exist.

Seabright likewise claims that a lack of empirical proof favors critics of current tying doctrine. This claim fails because it is the critics that favor a categorical rule that requires empirical proof across the category: namely critics favor categorical legality either for all ties or for all ties that lack substantial foreclosure. In contrast, current tying doctrine uses no categorical rule, but rather weighs efficiencies against anticompetitive effects in each case and permits ties to whatever extent it turns out to be empirically true that the efficiencies outweigh the anticompetitive effects. Current tying doctrine is thus preferable to the critics' recommended alternatives whether the standard is consumer welfare or total welfare, and whether one thinks most ties flunk that standard or not.

*Einer Elhauge is the Petrie Professor of Law at Harvard Law School, author of *U.S. Antitrust Law & Economics*, co-author of *Global Antitrust Law & Economics*, and editor of the forthcoming *Research Handbook on the Economics of Antitrust Law*.

Seabright also makes the more minor claim that, absent empirical proof that most ties harm welfare, the law should shift the burden of proof on efficiencies away from defendants. But this claim fails for four reasons. First, the burden of empirical proof on legal issues is on those who want to overrule precedent. Second, the fact that defendants have better access to evidence on tying efficiencies favors giving defendants the burden to prove those efficiencies, regardless of what one assumes about the welfare effects of most ties. Third, in allocating this burden of proof, the relevant set of ties are those for which defendants would have the burden to prove efficiencies, which is not all ties, but rather is only ties of separate products with tying market power where my recommended exception does not apply. The relevant category thus excludes: (1) ties of items deemed a single product because they are routinely bundled in competitive markets, (2) ties without market power, *and* (3) ties without a substantial foreclosure share that bundle products lacking separate utility in a fixed ratio. Fourth, even without general empirical proof, theoretical considerations indicate that ties in the relevant set will usually reduce both consumer welfare (the actual antitrust standard) and ex ante total welfare.

Professors Daniel Crane and Joshua Wright claim that bundled discounts cannot credibly threaten unbundled prices that exceed but-for prices. This claim conflicts with the fact that firms demonstrably can credibly threaten the refusal to sell at any price that is necessary to get buyers to agree to tying and monopoly pricing. This claim also ignores the fact that, in markets with many buyers, buyers have collective action problems that make them price takers.

Professor Barry Nalebuff offers models on ties that achieve intra-product price discrimination by metering use of the tying product that confirm my model's conclusions on that subset of ties. To the extent our models diverge on some details, I think it is more accurate to model metering ties by assuming that buyers purchase a whole number of tied units, rather than infinitely divisible fractions of tied units (as he assumes). I also think it is more accurate to assume that buyers have varying valuations, rather than the same valuation for tied product usage over the relevant range (as some of his models assume).

My legal conclusions are also generally confirmed by the conclusions that Professor Harry First reaches with a multi-goal approach. However, I prefer a welfarist analysis because I find that the multi-goal approach and its non-welfarist components are conclusory and unpersuasive when they conflict with welfare.

II. The Seabright Attack

A. SEABRIGHT IS WRONG TO CLAIM THAT A LACK OF EMPIRICAL PROOF UNDERMINES MY ANALYSIS AND SUPPORTS THE CRITICS OF CURRENT TYING DOCTRINE

Seabright's main argument is that the single monopoly profit theory is "undead" because I have not empirically proven how often the conditions that invalidate

it apply.² This is an odd defense because the single monopoly profit theory is an *impossibility* theorem: it claims that a firm with market power cannot possibly increase monopoly profits with tying because there is only a single monopoly profit the firm can get. Given that Seabright acknowledges that many market conditions do invalidate the single monopoly profit theory,³ his argument does not resurrect this impossibility theorem, but rather adds a few more spadefuls of dirt on top of its grave.

Seabright does not dispute that I have correctly identified the market conditions that invalidate the single monopoly profit theory. While prior work has shown that the single monopoly profit theory is invalid under particular conditions, it remains undisputed that my synthesis of the literature shows that the

IN SHORT, SEABRIGHT OFFERS NO
GROUNDS TO THINK THAT I HAVE
NOT CORRECTLY SPECIFIED THE
CONDITIONS UNDER WHICH THE
SINGLE MONOPOLY PROFIT THEORY
HOLDS, NOR ANY REASON TO
THINK THAT THE THEORY'S RULE
OF *PER SE* LEGALITY SHOULD
EXTEND BEYOND THE SET OF CASES
WHERE THOSE CONDITIONS OBTAIN.

single monopoly profit theory does not hold with or without a fixed ratio, with or without a strong positive demand correlation, and with or without a substantial foreclosure share.⁴ The conditions under which it does not hold thus clearly seem much broader than had previously been appreciated or than is suggested by Seabright's begrudging concession that "the Single Monopoly Profit theory is not true always and everywhere."⁵

Indeed, it remains undisputed that my analysis shows that the single monopoly profit theory holds *only* when there is a combination of a

fixed ratio, a strong positive demand correlation, *and* no substantial foreclosure share.⁶ If Seabright wishes to argue simply that there are *some* circumstances under which a tying firm can obtain only a single monopoly profit, then he agrees with me, but this argument does not resurrect the original single monopoly profit theory—at best it gives birth to a new baby single monopoly profit theory. Nor can this baby single monopoly profit theory justify the sweeping rule of *per se* legality for all tying that the Chicago School had advocated based on the original theory. Instead, the baby theory justifies only what I advocated in my article: a limited rule of *per se* legality applicable only to ties satisfying three conditions: (1) a fixed ratio, (2) a strong positive demand correlation (inferred from a lack of separate utility), *and* (3) no substantial foreclosure share.⁷

In short, Seabright offers no grounds to think that I have not correctly specified the conditions under which the single monopoly profit theory holds, nor any reason to think that the theory's rule of *per se* legality should extend beyond the set of cases where those conditions obtain. Nor is he right that my policy arguments depend on any empirical assumptions about how often those market conditions hold. To whatever empirical extent those conditions happen to hold, my proposed exception would apply a rule of *per se* legality.⁸ If those conditions usually hold, then my approach would usually apply a rule of *per se* legality. But if

those conditions usually don't hold, then my approach would usually not apply a rule of *per se* legality. My approach thus requires no empirical assumptions about the frequency with which those market conditions hold; it rather makes the legal results depend on empirical assessments of whether the conditions are present in actual cases.

In contrast, a rule of *per se* legality would require strong empirical evidence because it makes a categorical judgment that the single monopoly profit theory holds for all tying cases, even though Seabright himself admits that it actually does not hold in some tying cases.⁹ Such a categorical judgment would make sense only if one empirically believed *both* (1) that the conditions necessary for the single monopoly profit theory would apply in the vast bulk of cases covered by a tying doctrine with my exception *and* (2) that courts are incapable of distinguishing cases where those conditions do not apply from those where they do. Seabright offers no empirical evidence for the first conclusion, which seems implausible not only because of the limited conditions under which the theory is valid, but also because my exception would exclude ties that meet those conditions. Nor does Seabright offer any empirical evidence for the second conclusion, which again seems implausible because it does not seem especially difficult to determine when there are fixed ratios, separate utility, and a substantial foreclosure share.

In addition to mistakenly defending the single monopoly profit theory, Seabright criticizes my defense of the current *quasi-per se* rule because I have not provided empirical proof that ties usually harm welfare.¹⁰ But he is wrong that my defense of current tying doctrine depends on any such empirical premise. As I pointed out, calling current doctrine a *quasi-per se* rule is actually a misnomer.¹¹ Instead, current tying doctrine applies a particular form of rule of reason analysis that requires tying market power and then considers on a case-by-case basis whether any harmful effects are outweighed by offsetting efficiencies.¹² My defense of current tying doctrine thus does not depend on any empirical premise that welfare is usually worsened by the set of all ties with tying market power, but instead depends only on the claim that welfare is generally harmed by the *subset* of such ties that *lack offsetting output-increasing efficiencies*.

MY DEFENSE OF CURRENT TYING DOCTRINE THUS DOES NOT DEPEND ON ANY EMPIRICAL PREMISE THAT WELFARE IS USUALLY WORSENERD BY THE SET OF ALL TIES WITH TYING MARKET POWER, BUT INSTEAD DEPENDS ONLY ON THE CLAIM THAT WELFARE IS GENERALLY HARMED BY THE SUBSET OF SUCH TIES THAT LACK OFFSETTING OUTPUT-INCREASING EFFICIENCIES.

In his text, Seabright asserts that my claim is "both unjustified as science and impractical as policy."¹³ But the only support he provides for his condemnation simply ignores the fact that my claim was explicitly limited to ties without offsetting output-increasing efficiencies: he argues that when one considers the set of all price discrimination ties, they could conceivably increase or decrease wel-

fare.¹⁴ Buried in his endnotes, he admits that I am “correct” to claim that “the economic literature proves that price discrimination *always* decreases total welfare *unless it affirmatively increases output*.”¹⁵ It is hard to see how Seabright can say that the economic literature on price discrimination shows I am unjustified in claiming a proposition is “generally” true, when he acknowledges in the footnotes that this literature actually proves my proposition is “always” true.

Seabright nonetheless dismisses this economic literature on the ground that my above-quoted description of it was “phrased in such a way as to imply that increasing output is an unusual thing for price discrimination to do.”¹⁶ But there is nothing in my phrasing that implied any such thing, nor does justifying the current doctrine require such an empirical premise. If it empirically turns out to be true that defendants can usually prove an offsetting output-increasing efficiency, then current tying doctrine would usually not impose liability.¹⁷ But to

IN CONTRAST, THE QUASI-CHICAGO POSITION THAT TYING SHOULD NEVER BE ILLEGAL WITHOUT PROOF OF A SUBSTANTIAL FORECLOSURE SHARE DOES DEPEND ON A STRONG EMPIRICAL PREMISE BECAUSE IT MAKES A CATEGORICAL JUDGMENT THAT ALL TIES WITHOUT A SUBSTANTIAL FORECLOSURE SHARE SHOULD BE *PER SE* LEGAL.

the extent it is empirically the case that offsetting output-increasing efficiencies cannot be shown, then current tying doctrine would correctly impose liability. Once again, my position does not depend on any empirical premise; it rather makes the legal results turn on empirical assessments in each tying case of whether welfare-increasing effects actually exist.

In contrast, the quasi-Chicago position that tying should never be illegal without proof of a substantial foreclosure share *does* depend on a strong empirical premise because it makes a categorical judgment that all ties without a substantial foreclosure share should be *per se* legal.¹⁸

Although often described as the rule of reason position, this quasi-Chicago position really amounts to a rule of *quasi-per se* **legality** that mandates non-liability for all ties without a substantial foreclosure share. Seabright provides no explanation for why the law should *categorically* deem all ties without substantial foreclosure to be welfare-enhancing when he himself concedes that such ties can decrease welfare when tying market power exists.¹⁹ Justifying this *quasi-per se* legality position would require strong empirical evidence both (1) that ties with market power and no substantial foreclosure share almost always enhance welfare *and* (2) that courts cannot distinguish when such ties do or do not enhance welfare. Seabright provides no empirical evidence on either point. In contrast, the current tying doctrine that I defend requires no strong empirical premise because it empirically assesses whether welfare-increasing effects actually exist in challenged cases.

In short, the Chicago School position requires empirical evidence about all ties to justify its *per se* legality rule of categorical non-liability for all ties, and the quasi-Chicago School position requires empirical evidence about the set of all

ties without a substantial foreclosure share to justify its *quasi-per se* legality rule of categorical non-liability for all ties without a substantial foreclosure share. So critics of current tying doctrine certainly need powerful empirical evidence, which Seabright admits is lacking,²⁰ to justify their demands for radical changes to current tying doctrine. But defenders of current tying doctrine need no such empirical evidence because current tying doctrine, unlike its critics, makes no categorical judgment of liability or non-liability for ties with market power. Instead, current doctrine just requires case-by-case empirical assessment of the possibility that ties can harm welfare when there is tying market power. Given that Seabright himself admits that possibility is real even without a substantial foreclosure share, it is hard to fathom his objection to allowing courts to consider that possibility.

Moreover, to the extent the critics' proposals for radical change to current tying doctrine did turn on an assessment of the empirical evidence, it would be strange to say those proposals should be adopted even though (as Seabright admits), there is no empirical evidence to support them. Absent empirical evidence, standard law on *stare decisis* requires sticking with existing precedent. The burden of proof is on the critics who advocate changing existing law, not on those who favor adhering to long-established precedent. The critics' burden cannot be met by simply assuming they are empirically right until someone provides empirical evidence to the contrary.

ABSENT EMPIRICAL EVIDENCE,
STANDARD LAW ON STARE
DECISIS REQUIRES STICKING
WITH EXISTING PRECEDENT.
THE BURDEN OF PROOF IS
ON THE CRITICS WHO ADVOCATE
CHANGING EXISTING LAW,
NOT ON THOSE WHO FAVOR
ADHERING TO LONG-
ESTABLISHED PRECEDENT.

Given all this, what could possibly be Seabright's basis for asserting that my position hinges on an empirical claim? Other than simply ignoring the fact that my claim was limited to ties without offsetting output-increasing efficiencies, Seabright relies on two moves.

First, Seabright asserts that I claim "the support of the economic literature for the conclusion that 'imperfect price discrimination likely decreases consumer welfare,'" which he calls "a travesty of what the literature says."²¹ Now, while academics sometimes get disagreeable, it is not every day that an academic gets quite so disagreeable that he accuses another academic of committing a "travesty." Even when feeling impolite, no reasonable academic would do so unless he is absolutely sure his position is unassailable. But when one examines the full language of what I said in the passage that Seabright selectively quotes, it turns out that it did not even make the empirical claim that Seabright asserts it made. Instead, it stated that: "*The critics' analogy to perfect discrimination means that imperfect price discrimination likely decreases consumer welfare.*"²² My point was that critics of current tying doctrine were using an analogy to make a claim that, because perfect price discrimination increased total welfare, imperfect price discrimination was likely to increase total welfare as well, and that if one applied

that same analogy evenhandedly, it meant that, because perfect price discrimination reduces consumer welfare, imperfect price discrimination is likely to reduce consumer welfare as well.²³ Which part of my actual proposition does Seabright find objectionable? Does he claim that perfect price discrimination doesn't reduce consumer welfare? If so, he claims a position that no competent economist holds. Or does he claim that analogies should not be applied evenhandedly? If so, he has a very self-serving view about analogical reasoning, for which he

provides no support. The only travesty here is that, because of his regrettable misquotation, Seabright provides no response to what I actually said.

THE CRUCIAL ANALYTICAL POINT,
WHICH SEABRIGHT APPEARS
TO HAVE MISSED, IS THAT IT IS
ONLY THE CRITICS WHO HAVE TO
MAKE A CATEGORICAL WELFARE
CLAIM BECAUSE THEY ARE
THE ONLY ONES ARGUING
FOR A CATEGORICAL RULE.

None of which means that it would be a travesty to claim that economic theory indicates that tying-induced price discrimination is more likely to reduce consumer welfare than increase it. I shall offer reasons below to think that it does, and Seabright offers nothing to the contrary other than bald assertion. But that isn't

the claim I was making, nor is it a claim that is required to defend current tying doctrine. The crucial analytical point, which Seabright appears to have missed, is that it is only the critics who have to make a categorical welfare claim because they are the only ones arguing for a categorical rule. What my analysis showed was that the analogy the critics relied on for their categorical welfare claim undermines their position because that analogy indicates that price discrimination ties are categorically likely to reduce consumer welfare, and antitrust law embraces a consumer welfare standard, rather than the *ex post* total welfare standard used by the critics.²⁴ This demonstration that the critics' own analogy undermines their position does not mean that current tying doctrine requires relying on this same analogy or on a contrary categorical welfare claim; it doesn't because current doctrine makes no categorical liability claim. In other words, a conclusion that the analogy is persuasive favors current tying doctrine (because the correct standard is consumer welfare), and a conclusion that the analogy is unpersuasive also favors current tying doctrine (because critics rely on it to make a categorical welfare claim that is necessary to their position, whereas current doctrine requires no such categorical claim).

Second, Seabright argues that my position must rest on an empirical claim that ties generally harm welfare because current tying doctrine (which I defend) gives defendants the burden of proving an offsetting output-increasing efficiency.²⁵ But Seabright cites no support for his premise that, absent empirical evidence on whether a proposition is usually true, the burden of proof must favor the defendant. There are many reasons to allocate a burden of proof other than using the pro-defendant bias that Seabright favors. One simple reason is adhering to precedent, which in tying cases has long put the burden of proving efficiencies on a

defendant with tying market power. His claim that we lack empirical evidence in either direction hardly provides a compelling reason to deviate from *stare decisis*. Another reason favors putting the burden of proof on the party that has the best access to evidence on the relevant issue, because that is more likely to lead to accurate resolutions. Even if the ties covered by current doctrine generally have efficiencies, defendants clearly have better access to evidence about the efficiencies of their own ties than others can have. Finally, even if we didn't have those two compelling reasons, one might reasonably conclude that, absent empirical evidence on the issue, one should allocate the burden of proof based on theoretical considerations about which welfare effect is more likely across the set of cases covered by current tying doctrine.

B. ALLOCATING THE BURDEN OF PROOF BASED ON LIKELY WELFARE EFFECTS UNDER THE CURRENT CONSUMER WELFARE STANDARD

Suppose, just for argument's sake, that we put aside the first two grounds for allocating the burden of proof, and decide to allocate it instead based on theoretical considerations about whether consumer welfare was likely to increase or decrease for the set of ties covered by current tying doctrine. I begin with consumer welfare because it is the governing legal standard, but in the following sections I consider (and reject) Seabright's argument that antitrust and competition law should change to an *ex post* total welfare standard and show that in any event such a change in legal standard would require little change to the analysis of tying doctrine.

FIRST, THERE IS A POWER EFFECT THAT SEABRIGHT STUDIOUSLY IGNORES: TYING CAN CREATE INTER-PRODUCT PRICE DISCRIMINATION ACROSS THE BUNDLED PRODUCTS.

Because tying doctrine does not even apply unless the defendant ties separate products together, the relevant set of cases obviously would exclude any bundles that constitute a single product. When two bundled items are a single product, we have no tie that triggers tying doctrine at all, but rather have only the sale of a single product. Nonetheless, with no basis whatsoever, Seabright asserts that I would apply tying doctrine to the sale of bundled items that are plainly a single product under current law, such as guitars with strings, cameras with memory cards, and airplanes with toilets.²⁶ I have written over 100 pages elaborating single product tests and explaining their importance in screening out bundles whose efficiency can be inferred from market tests.²⁷ Nowhere in my tying article is there any suggestion that, having so carefully elaborated these single product tests, I now favor abandoning the separate products element that must be satisfied to show a tie at all. Perhaps Seabright is simply unaware of the well-known separate products element of tying doctrine, but whatever the explanation, he is simply mistaken in asserting that I would require defendants to show efficiencies for the sale of many single products just because he can imagine describing them as bundles of two items.

The relevant set of cases thus clearly includes only ties of separate products. Further, given the doctrine I am defending, it also includes only ties with tying market power and where the exception for products in a fixed ratio that lack separate utility does not apply. In such cases, there are, as I showed in my article, three relevant power effects.

1. Inter-Product Price Discrimination

First, there is a power effect that Seabright studiously ignores: tying can create inter-product price discrimination across the bundled products.²⁸ One can see why Seabright prefers to ignore this power effect. As I pointed out, the economic literature proves that “assuming a normal distribution of buyer valuations, [such] tying *always* decreases consumer welfare absent perfect positive demand correlation.”²⁹ But this proven result cannot properly be ignored if one wishes to accurately assess the likely effects of tying with market power. A normal bell-shaped distribution is a common assumption in economic analysis, and it seems quite reasonable to assume that usually there are more buyers with moderate valuations of a product than with extreme valuations. A perfect positive demand correlation also seems unlikely, especially in cases where the products have separate utility, which is the relevant set for my analysis given that this power effect assumes fixed ratio bundling and my exception would apply if the products also lacked separate utility. Absent empirical evidence to the contrary, it thus seems

entirely reasonable to think that ties with this power effect usually reduce consumer welfare.

SECOND, THERE IS A POWER EFFECT
THAT SEABRIGHT ADMITS REDUCES
CONSUMER WELFARE AND CAN
LEVERAGE ONE MONOPOLY PROFIT
INTO TWO MONOPOLY PROFITS:
TYING CAN EXTRACT INDIVIDUAL
CONSUMER SURPLUS.

2. Extracting Individual Consumer Surplus

Second, there is a power effect that Seabright admits reduces consumer welfare and can leverage one monopoly profit into two monopoly profits: tying can extract individual consumer surplus.³⁰ This effect Seabright dismisses with

the combination of a theoretical claim, an empirical claim, and a conceptual claim, each of which is necessary to his argument, and each of which is unsupported and demonstrably false. His theoretical claim is that this power effect necessitates a requirements tie, which is a tie that obligates the buyer to make all its purchases of the tied product from the defendant.³¹ For this theoretical claim, he relies on my description of one illustration I gave, which did assume such a requirements tie, but he neglects to acknowledge that on the very next page I explicitly stated that: “extracting individual consumer surplus does not necessitate a requirements tie that forbids buying the tied product from rivals. . . . A firm could achieve the same effect by requiring buyers to buy some fixed quantity of the tied product at a supracompetitive price (say 200 scanners at \$400) if they want to make purchases of the tying product at the monopoly price.”³² Other

economic literature agrees with me that extracting individual consumer surplus does not necessitate a requirements tie.³³

Even if Seabright were right on his theoretical claim, his admission that requirements tying can have this adverse welfare effect means that he needs to couple his theoretical claim with an empirical claim, and it is a doozy. Seabright asserts that requirements tying is something “the world has rarely seen outside of gangster life.”³⁴ Now that is quite an empirical assertion, and remarkably he provides zero empirical support for it, despite having spent his entire paper repeatedly chastising me for my alleged lack of empirical support. It is also an assertion that reflects a charming naïveté about the actual world of law and business. He bases his assertion on a claim that no requirements tie “could possibly be enforced without illegal coercion” unless “the monopoly good is technologically complementary to the competitively supplied good in such a way as to make useless (or more generally to lower the value of) any version of the latter supplied by a competitor.”³⁵ But single-product exclusive dealing and requirements contracts are in fact commonplace, and by definition their enforcement cannot depend on rivals’ technological incompatibility with a tying product. Indeed, requirements and exclusive dealing contracts are so common that they get their own sections under both contracts and antitrust statutes.³⁶ Nor is it at all uncommon to attach such exclusive conditions to a tying agreement. At least seven Supreme Court cases have involved requirements ties, and in none of those cases was it true that the rival tied product was technologically incompatible with the defendant’s tying product.³⁷ Indeed, the description of the requirements clauses in three of these cases indicates that the ties were not even limited to tied products that were used with the defendant’s tying product.³⁸

SEABRIGHT ASSERTS THAT REQUIREMENTS TYING IS SOMETHING “THE WORLD HAS RARELY SEEN OUTSIDE OF GANGSTER LIFE.” NOW THAT IS QUITE AN EMPIRICAL ASSERTION, AND REMARKABLY HE PROVIDES ZERO EMPIRICAL SUPPORT FOR IT.

Although these seven Supreme Court cases involved requirements ties of the sort that Seabright claims are rare for non-gangsters to impose, they did involve tied products that were needed to get value out of the tying product, so were likely metering ties, rather than ties that extract individual consumer surplus. However, bundled loyalty discounts frequently involve products with no strong positive demand correlation, for which extracting individual consumer surplus is possible. For example, in *LePage’s*, the defendant used bundled loyalty rebates “conditioned on purchases spanning six of 3M’s diverse product lines” and that also covered both brand name and private-label tape which, if anything, have negatively related demand.³⁹ Likewise, in *Masimo*, the Ninth Circuit found that: “Tyco’s bundling contracts gave customers a price discount for purchasing a number of *unrelated* products together, one being pulse oximetry. However, receipt of the discount was conditioned upon customers purchasing 90-95% of *their requirements of those products* from Tyco.”⁴⁰ Indeed, bun-

dled loyalty rebates spanning unrelated products are rampant in U.S. health-care.⁴¹ I am sure the firms that use them would be surprised to learn that Seabright equates all of them with gangsters.

Nor is there any great mystery why buyers comply with exclusivity obligations even when gangsters are not around to enforce them. Sellers can require contractual promises (which most businesses honor voluntarily)⁴² or buyer self-reporting (even fewer business are willing to lie and commit fraud) or rely on simple observation,⁴³ followed by threats to enforce the contract, withhold bundled rebates, or cut off supply of the tying product when buyers are noncompliant. Seabright asserts this would amount to “illegal coercion” that is “unenforceable . . . in law,” but cites no law to support his legal conclusion.⁴⁴ The irony, of course, is that such exclusive tying conditions would be illegal only if tying doctrine continues to make them illegal, which is precisely the doctrine that Seabright criticizes.

Seabright’s conceptual claim is that, if the conditions for this power effect are rare, the law should ignore it rather than focusing the doctrine on cases when those conditions hold. He provides no basis for this claim. Even if Seabright were right in his empirical assertion that ties usually involve fixed ratios,⁴⁵ the more logical response would be (as I advocate) precluding litigation of this power effect when the tie actually involves a fixed ratio, but allowing it to be litigated when the right to buy the tying product is tied to an obligation to buy a tied product without a fixed ratio.⁴⁶ This does not mean that all fixed ratio ties should be *per se* legal because the power effect that Seabright studiously ignores—inter-product price discrimination—remains possible with a fixed ratio.⁴⁷ However, a

fixed ratio does preclude the other two power effects, and thus should preclude the *quasi-per se* rule entirely if coupled with evidence of a lack of separate utility that indicates the sort of strong positive demand correlation that makes inter-product price discrimination impossible as well.⁴⁸ Because such cases fit within my exception, they are not within the relevant set of cases whose likely welfare effects are at issue.

THERE IS NO BASIS FOR SEABRIGHT’S THEORETICAL CLAIM THAT EXTRACTING CONSUMER SURPLUS NECESSITATES REQUIREMENTS TIES, NO BASIS FOR HIS EMPIRICAL CLAIM THAT SUCH REQUIREMENTS TIES ARE RARE FOR NON-GANGSTERS, AND NO BASIS FOR HIS CONCEPTUAL CLAIM THAT, IF THE CONDITIONS ARE RARE, LITIGATION SHOULD BE PROHIBITED.

In short, there is no basis for Seabright’s theoretical claim that extracting consumer surplus necessitates requirements ties, no basis for his empirical claim that such requirements ties are rare for non-gangsters, and no basis for his conceptual claim that, if the conditions are rare, litigation should be prohibited rather than focus-

ing the doctrine on cases when the conditions are met. His concession that this power effect harms consumer welfare thus helps support a conclusion that theoretical considerations indicate consumer welfare is likely to be harmed in the rel-

evant set of cases, which are ties with market power that lie outside the exception for products in a fixed ratio that lack separate utility.

3. Intra-Product Price Discrimination

Third, there is the one power effect that Seabright does address on the merits: the fact that metering ties can enable intra-product price discrimination. It is true that, other than proving that imperfect price discrimination cannot increase welfare unless it produces output-increasing efficiencies that offset its misallocation inefficiency, the economic literature before my article tended to simply say that imperfect price discrimination might or might not do so, and then pronounce the issue ambiguous. But this is rather unhelpful if one needs to know the general tendency across a range of cases, as tying critics need to know to support categorical non-liability, and as Seabright asserts we need to know in order to allocate the burden of proof under current tying doctrine. After all, tomorrow you may live or die, so I suppose we could say the issue is “ambiguous,” but that doesn’t mean you should assume the two are equally likely when planning your calendar.

To fill this gap in the literature, I offered my own model of the welfare effects of metering ties that create imperfect intra-product price discrimination, and mathematically proved that metering ties reduce consumer welfare significantly in that linear model, with the reduction converging for large numbers of tied items on a 18.85% loss of the consumer welfare that would be enjoyed without price discrimination.⁴⁹ Using a linear model that assumes the number of tied products is continuous rather than discrete, Professor Nalebuff’s Comment on my article reaches the similar conclusion that metering ties reduce consumer welfare by 18.75%.⁵⁰ Seabright complains that we cannot be sure my results will be the same without linear demand.⁵¹ That is true, but my article used my linear model only to help rebut the assertion by critics that the welfare effects were likely to be categorically positive, not to make any claim of categorical liability.⁵² Further, linear models are commonly used in antitrust economics and, indeed, were commonly used by the Chicago school to develop many of their propositions. Absent other models that can tell us the likely effects, linear models appear to be the best we have.

Moreover, there remains the fact that perfect discrimination clearly does lower consumer welfare, no matter what one assumes about the shape of the demand curve. To be sure, reasoning by analogy is less satisfactory than having a formal model for every possible demand curve. But no matter what the demand curve is, we know that the overall move from uniform pricing to perfect discrimination

IN SHORT, OF THE THREE POWER EFFECTS PRODUCED BY TIES WITH MARKET POWER THAT DON’T FIT INTO MY EXCEPTION, THEORETICAL CONSIDERATIONS INDICATE THAT TWO OF THOSE POWER EFFECTS ALMOST SURELY REDUCE CONSUMER WELFARE AND THE THIRD LIKELY DOES SO AS WELL.

lowers consumer welfare. We also know that the entire point of metering ties is to increase the perfection of price discrimination. While some movements that increase the perfection of price discrimination might not reduce consumer welfare, the sum of all movements that increase the perfection of price discrimination must, in aggregate, reduce consumer welfare, just like we know the driver whose final destination is east must go east more often than west, even though some of the movements in his trip might not. It thus seems likely that metering ties that increase the perfection of price discrimination on average reduce consumer welfare. Given that this analogy points in the same direction as the linear demand model, theoretical considerations certainly provide more reason to think that metering ties reduce consumer welfare than increase it. This conclusion gets even stronger if we include the costs of implementing and monitoring a tying scheme, which I omitted from my analysis to be conservative but which others have stressed.⁵³

4. Summary

In short, of the three power effects produced by ties with market power that don't fit into my exception, theoretical considerations indicate that two of those power effects almost surely reduce consumer welfare and the third likely does so as well. This more than suffices to conclude that theoretical considerations favor putting the burden of proof on the defendant. Moreover, even if theoretical considerations were too ambiguous, as Seabright insists, the burden should still be put on defendants because they have the best access to information about the efficiency and output effects of their ties. Finally, even if both those factors were

ambiguous, we would have no more grounds to put the burden on the plaintiff than on the defendant, so it would make more sense to allocate the burden by sticking to *stare decisis*.

OF THE FIVE ARGUMENTS I PUT
FORTH FOR THE CONSUMER
WELFARE STANDARD, ONLY ONE OF
THEM EVEN ARGUABLY IMPLIES
THAT PRODUCER SURPLUS SHOULD
BE GIVEN A LOWER WEIGHT, AND
EVEN FOR THIS ONE THE
IMPLICATION IS INCORRECT.

C. SEABRIGHT'S MISTAKEN ARGUMENTS AGAINST THE CURRENT CONSUMER WELFARE STANDARD

I argue that antitrust law correctly embraces a consumer welfare (i.e., consumer surplus) standard rather than a total welfare (i.e., total sur-

plus) standard. Seabright raises various arguments against my use of a consumer welfare standard. None are valid. First, he argues that: "Professor Elhauge claims that producer surplus should essentially be given zero weight in social welfare, even though most of the arguments he gives for this conclusion (such as the higher average income of shareholders when compared to consumers) imply that they should be given a lower weight but still one greater than zero."⁵⁴ However, his characterization of my arguments is false. Of the five arguments I put forth for the consumer welfare standard, only one of them even arguably implies that

producer surplus should be given a lower weight, and even for this one the implication is incorrect.

The only one of my arguments that even arguably has this implication is the one that Seabright mentions: the argument that, given average incomes, a consumer welfare standard has beneficial distributive effects compared to a total welfare standard. But, in fact, even this argument does not imply that courts should give producer surplus some weight between zero and 1, nor that one should (as Seabright suggests in a footnote) vary the weight depending on the income of particular consumers and producers.⁵⁵ Varying the weight of producer surplus is judicially inadministrable. Doing so depending on the wealth of particular consumers and producers is even worse because the economic literature proves that liability rules that vary with litigant income are less efficient at redistribution than income taxation. The reason is that while income taxes inefficiently discourage income creation, liability rules that vary with each litigant's income discourage not only income creation, but also some efficient conduct regulated by the rule.⁵⁶ Thus, even if (contrary to fact) the distributive point were the only argument, it would not imply that the law should change to adopt Seabright's weighing approach. Instead, antitrust law should continue furthering distributive goals with a general rule that uses a consumer welfare standard, assuming one agrees, as Seabright admits is "probably correct," that consumers generally have lower income than shareholders.⁵⁷ A consumer welfare standard is not only far more administrable than Seabright's weighing approach, but also does not penalize income creation because it does not vary the liability rule with each litigant's income, whereas his weighing approach would. In any event, the point is moot because the other four arguments for a consumer-welfare standard obviate any need to weigh producer surplus at all.

Seabright simply ignores three of the other four arguments for a consumer welfare standard. (1) He does not dispute the point that antitrust law in fact requires a consumer welfare standard, which is true both in the United States and the EU.⁵⁸ This point makes his arguments legally irrelevant to the issue of how courts should interpret tying doctrine. (2) Seabright also does not dispute my point that any conduct that truly enhances total welfare can generally be restructured to shift enough of the gain to consumers to advance consumer welfare while still profiting the producer.⁵⁹ This point means that a consumer welfare test does not in fact require avoiding conduct that increases total welfare, but instead helps verify that the relevant conduct really does produce a net gain to total welfare by forcing producers to

A CONSUMER WELFARE
STANDARD IS NOT ONLY
FAR MORE ADMINISTRABLE
THAN SEABRIGHT'S WEIGHING
APPROACH, BUT ALSO DOES NOT
PENALIZE INCOME CREATION
BECAUSE IT DOES NOT VARY
THE LIABILITY RULE WITH
EACH LITIGANT'S INCOME,
WHEREAS HIS WEIGHING
APPROACH WOULD.

put their money where their mouth is regarding the claimed size of efficiency gains.⁶⁰ (3) While Seabright suggests that adjudicators should simply give different weights to consumer and producer welfare, he does not dispute my point that, in a world of concurrent antitrust jurisdiction, only a pure consumer welfare standard gives optimal enforcement incentives to the decisive regulator.⁶¹ This point seems confirmed by the fact that the only court I know of to try a weighing approach approved a merger that it acknowledged would increase prices by 11% because most of the merging firms' product was exported to foreign consumers, whose interests the court gave zero weight.⁶²

WHAT I POINTED OUT WAS THAT
THOSE WHO ARGUED THAT TYING'S
POWER EFFECTS USUALLY
INCREASED TOTAL WELFARE WERE
REALLY ONLY POINTING OUT
SITUATIONS WHERE THEY WOULD
INCREASE *EX POST* TOTAL WELFARE,
WHICH IS NOT AT ALL THE SAME
AS OVERALL TOTAL WELFARE.

Seabright addresses my fifth argument for a consumer-welfare standard, but misunderstands it. What I pointed out was that those who argued that tying's power effects usually increased total welfare were really only pointing out situations where they would increase *ex post* total welfare, which is *not* at all the same as overall total welfare.⁶³ The reason is that some or all of the additional monopoly profits created by the power effects would be dissipated by

ex ante costs.⁶⁴ The cases of interest are those where the difference in standards leads to different results, namely those where tying's power effects reduce consumer welfare but increase *ex post* total welfare, which means cases where the *ex post* monopoly profit increase outweighs the consumer welfare harm. If some or all of that *ex post* monopoly profit increase is dissipated by *ex ante* costs, then the overall monopoly profit increase may well be smaller than the consumer welfare harm, in which case the tying power effect lowers total welfare even though it increases *ex post* total welfare.

For example, let's take the power effect that provides the best case for critics (and thus, not surprisingly, the one that many critics prefer to discuss to the exclusion of others): metering ties that create intra-product price discrimination. Suppose a market with linear demand in which, at a uniform monopoly price, consumer surplus is \$100 million, monopoly profits are \$200 million, and thus total welfare is \$300 million. The models by myself and Nalebuff indicate that, at least for high numbers of tied items, allowing metering ties that create intra-product price discrimination would reduce consumer welfare by 19% and increase total welfare by 5%.⁶⁵ This means that allowing metering ties would reduce consumer surplus by \$19 million, increase *ex post* total welfare by \$15 million, and thus increase *ex post* monopoly profits by \$34 million. If more than \$15 million (or 44%) of those additional *ex post* monopoly profits would be dissipated by the *ex ante* costs of *all* firms' efforts to acquire that monopoly position, then a rule allowing such metering ties would result in an overall producer profit increase of less than \$19 million, which is smaller than the consumer harm, and

thus would reduce total welfare even though it would increase *ex post* total welfare. The degree of dissipation required is even smaller if we consider metering ties with fewer tied items (because the ratio of consumer welfare loss to *ex post* total welfare gain is usually larger for them⁶⁶) or ties with the other two power effects (because their *ex post* total welfare effects are more ambiguous compared to their consumer welfare harm).⁶⁷

Seabright offers various responses to this argument, none of which are valid. First, he argues that this argument amounts to a claim that producer surplus “should be given a lower weight” than consumer surplus.⁶⁸ But that is not what this argument shows. Instead, this argument shows that some or all of the nominal producer surplus increase is a mirage caused by failing to consider *ex ante* producer costs. This point is no different than saying that consumer surplus measurements should subtract the costs that consumers paid for the products. The point is about accurately measuring overall producer surplus, not about weighing it differently than consumer surplus.

SOME OR ALL OF THE NOMINAL
PRODUCER SURPLUS INCREASE
IS A MIRAGE CAUSED BY
FAILING TO CONSIDER
EX ANTE PRODUCER COSTS.

Second, Seabright complains that I lack “any empirical backing” for my claim that the additional monopoly profits produced by tying will be dissipated by *ex ante* costs.⁶⁹ But it is hard to see what he is complaining about because he admits: “That there is some such dissipation is not seriously disputed by economists.”⁷⁰ His implication appears to be that I asserted that monopoly profits would always be *completely* dissipated. But that is not what my article says. To the contrary, I explicitly stated: “There are thus two possible cases. In cases where Judge Posner is right that 100% of monopoly profits are dissipated, then any *ex post* increase in monopoly profits effectively washes out *ex ante*, which means that the consumer welfare effects actually determine the overall total welfare effects.... In cases where Fisher is right, then some share less than 100% of monopoly profits are dissipated, which still means that tying that increases *ex post* total welfare will often decrease overall total welfare. It will do so whenever the consumer welfare harm exceeds the non-dissipated share of the monopoly profit gain.”⁷¹ Because Seabright never confronts my actual argument, he never explains what, if anything, he deems wrong about this analysis. Instead, he oddly complains that I don’t provide empirical proof for a proposition he admits is not seriously disputed.

Third, Seabright claims that my analysis is somehow rebutted by the fact that “there are also beneficial effects on innovation of competition to obtain market power, as is recognized in the patent system.”⁷² But this fact is perfectly consistent with my analysis. Indeed, I affirmatively base my analysis on it. What I pointed out, and Seabright ignores, is that the patent system has already considered this beneficial effect and set patent lengths on the assumption that: “Patent holders are entitled to the normal monopoly profits they make by selling their

patented goods, but are not currently entitled to extract more than those profits through tying.”⁷³ Changing current tying doctrine to allow firms to use tying to get more than normal monopoly profits thus gives them more than what patent law determined was the optimal reward for their innovation.

CHANGING CURRENT TYING
DOCTRINE TO ALLOW FIRMS TO USE
TYING TO GET MORE THAN
NORMAL MONOPOLY PROFITS THUS
GIVES THEM MORE THAN WHAT
PATENT LAW DETERMINED WAS
THE OPTIMAL REWARD
FOR THEIR INNOVATION.

Relatedly, Seabright suggests that my analysis is somehow in tension with the fact that “several scholars have found ‘U-shaped’ results, with some degree of market power being more beneficial to innovation and growth than either complete monopoly or a high degree of competition.”⁷⁴ The literature that he cites is actually about an entirely different issue—the extent to which *existing* market power fosters more or less

ex post innovation—whereas the relevant issue at hand is about whether the *future prospect* of gaining more than normal monopoly profits from successful innovations is necessary to induce optimal *ex ante* investments in innovation.⁷⁵ Moreover, on the relevant issue, my analysis affirmatively *relied* on an inverted U-shaped result proven by other economic literature that models competition to obtain patents.⁷⁶ As I stated, this literature proves that there is a socially optimal fraction of the total surplus created by an innovation that the firm winning a patent should get in order to induce socially optimal investment in innovation. If the patent holder captures all of this total surplus with perfect price discrimination, then that would lead to socially excessive investments. Thus, as I said, “What keeps that fraction from being exceeded is precisely the fact that part of the total surplus is instead enjoyed by consumers, as the consumer surplus they earn at a uniform monopoly price.”⁷⁷ On the other hand, if the patent holder received no fraction of this total surplus because no patent was recognized, then there would be socially insufficient incentives to invest.

Whether patent holders get the socially optimal fraction of overall total surplus will turn both on the patent length and the share of total surplus they get during the patent term. Currently, patent law attempts to achieve this optimal fraction by setting the length of patents on the assumption that, during the patent term, the share of total surplus received by patent holders will reflect normal monopoly profits, but will not reflect any additional profits that could be earned by using tying to extract the consumer surplus that buyers would earn at normal monopoly prices.⁷⁸ Thus, the efforts by patent law to award the optimal fraction would be undermined if tying law were changed to allow patent holders to extract more than their normal monopoly profits during the patent term.⁷⁹ This does not mean that anything that decreases the fraction earned by patent holders is desirable. To the contrary, I have equally objected to other proposals to deprive patent holders of some of those normal monopoly profits because such proposals would reduce their share of total surplus during the patent term below

the share that patent law assumed when setting the patent length.⁸⁰ If we assume patent law has set patent lengths to achieve the optimal fraction, then changes that try to increase or decrease that fraction will necessarily move us away from the social optimum; i.e., away from the apex of the inverted U-curve. If patent law has not been set to achieve the optimal fraction, then the correct solution is to reform patent law systematically, rather than change antitrust or other laws to allow certain *ad hoc* deviations from the normal monopoly profits that patent holders are entitled to get during the patent term.⁸¹

In short, even if we care only about total welfare, we have no basis to favor changing tying doctrine to allow ties that increase *ex post* total welfare by giving firms more than their normal monopoly profits at the expense of consumer welfare. If we assume patent law has already set the socially optimal patent terms, then allowing such ties will give patent holders more than the socially optimal fraction of total surplus, and thus will affirmatively reduce total welfare. In addition, if Judge Posner is right that 100% of any additional monopoly profits would be dissipated by *ex ante* costs, then allowing any tie that harms consumer welfare will, once again, reduce total welfare even if it creates enough additional monopoly profits to increase *ex post* total welfare. Finally, in cases where Professor Fisher is right that less than 100% of additional monopoly profits will be dissipated, then ties that increase *ex post* total welfare but reduce consumer welfare will still reduce total welfare unless the former effect sufficiently outweighs the latter, which is unlikely because the consumer welfare harm is generally stronger and less ambiguous than any *ex post* total welfare gain. Overall, then, when judging ties that allow firms to reap more than normal monopoly profits from their market power (i.e., ties that have one of the three power effects I identified), using a consumer welfare standard is more likely to further total welfare than using an *ex post* total welfare standard would.

IN SHORT, EVEN IF WE CARE ONLY ABOUT TOTAL WELFARE, WE HAVE NO BASIS TO FAVOR CHANGING TYING DOCTRINE TO ALLOW TIES THAT INCREASE EX POST TOTAL WELFARE BY GIVING FIRMS MORE THAN THEIR NORMAL MONOPOLY PROFITS AT THE EXPENSE OF CONSUMER WELFARE.

D. CHANGING TO AN EX POST TOTAL WELFARE STANDARD WOULD HAVE LITTLE IMPACT ON PROPER TYING DOCTRINE

Even if one believed (despite all the above) that we should change antitrust law to adopt an *ex post* total welfare standard, it is striking what little difference that would make to proper tying doctrine. Because even critics like Seabright admit that ties with power effects can reduce *ex post* total welfare, there would still be no sound basis for any categorical rule of non-liability for either all ties (the Chicago view) or all ties without a substantial foreclosure share (the quasi-Chicago view). Instead, it would remain the case that court should stick with current tying doctrine, which balances power effects against efficiencies under

the misnamed *quasi-per se* rule. The only clear difference would be that the defendant could win by proving that the tradeoff resulted in a net improvement to *ex post* total welfare, rather than (as under current law) having to prove that the tradeoff resulted in a net improvement in consumer welfare. I already pointed this out in my initial article.⁸² The only other arguable difference would be

EVEN IF ONE BELIEVED (DESPITE ALL THE ABOVE) THAT WE SHOULD CHANGE ANTITRUST LAW TO ADOPT AN *EX POST* TOTAL WELFARE STANDARD, IT IS STRIKING WHAT LITTLE DIFFERENCE THAT WOULD MAKE TO PROPER TYING DOCTRINE.

that, under an *ex post* total welfare standard, the argument for changing the current burden of proof on efficiencies would be not be quite as weak as it is now under the consumer welfare standard used by current antitrust law.

Although changing antitrust law to adopt an *ex post* total welfare standard would make the argument for changing the current burden of proof somewhat less weak, this does not mean it

would make that argument strong. Consider the three power effects that ties can have. Ties that extract individual consumer surplus would reduce *ex post* total welfare in the typical tying case where spending or valuation is significantly higher for the tying product than the tied product.⁸³ Ties that achieve inter-product price discrimination across both products increase *ex post* total welfare only if demand strength relative to cost is high, and otherwise decrease *ex post* total welfare.⁸⁴ Ties that achieve intra-product price discrimination on the tying product generally increase *ex post* total welfare unless the number of tied items is small or the buyers are intermediaries.⁸⁵ The last power effect is the only one favorable to critics, which is why it is the one they like to focus on, but even this effect is smaller and somewhat more mixed than the decrease in consumer welfare, and does not apply when the buyers are intermediaries, which is actually typical in most tying cases.⁸⁶ Considering the three power effects as a group, we have no reason to think that ties with market power are more likely to increase *ex post* total welfare than decrease it, and thus these predicted effects provide no reason to allocate the burden to plaintiffs even under a pure *ex post* total welfare standard. The reasons are even weaker if the *ex post* total welfare standard is advocated based only on the mistaken belief that it provides a closer proxy to overall total welfare than a consumer welfare standard. Given that the predicted consumer welfare decrease is stronger and more uniform and that, in at least some cases, an *ex post* total welfare increase will mean a decrease in overall total welfare given monopoly profit dissipation, the predicted total welfare effects if anything suggest the burden should be put on defendants.

Even if we ignore *ex ante* effects, the ambiguous *ex post* total welfare effects provide no reason to reallocate the burden of proof to plaintiffs for all ties with market power. Thus, even under a pure *ex post* total welfare standard, the proper burden allocation would turn on factors other than predicted effects. Those other two factors—*stare decisis* and allocating the burden to the party with the best access to the relevant evidence—support putting the burden of proof on the

defendant to show that the output-increasing efficiencies do offset the anticompetitive effects.

The most one can say is this: *If* antitrust law chooses to change to a pure *ex post* total welfare standard that ignores *ex ante* effects, *and* decides to change the burden of proof in tying cases to reflect *only* the most likely effects across a set of cases (without considering precedent or access to evidence), *and* elects to have different burdens of proof for different power effects, *then* in cases involving ties that are alleged to *only* increase intra-product price discrimination among final consumers, it makes sense to reallocate the burden of proof to plaintiffs on the issue of whether offsetting efficiencies outweighed anticompetitive effects enough to produce a net increase in *ex post* total welfare. But *even* in this case, the initial burden of *production* to show such offsetting efficiencies should be on the defendant because that burden should always be allocated based on who has the best access to the evidence. We would be switching the burden of proof only after that burden of production had been met, only for a limited set of tying cases based on a contestable view about how to allocate burdens of proof, and only if antitrust law first wrongly changed to a pure *ex post* total welfare standard.

EVEN IF WE IGNORE *EX ANTE* EFFECTS, THE AMBIGUOUS *EX POST* TOTAL WELFARE EFFECTS PROVIDE NO REASON TO REALLOCATE THE BURDEN OF PROOF TO PLAINTIFFS FOR ALL TIES WITH MARKET POWER.

In short, even if we spot the critics an undesirable change in both the antitrust welfare standard and the standards for allocating burdens of proof, the economic literature shows that the only change to tying doctrine that could possibly be warranted would be changing the burden of proof (but not production) on one subset of ties with market power. If this is the only doctrinal change that could even arguably be justified under the critics' own (quite dubious) standards, then the once-mighty single monopoly profit theory is down to a minor quibble indeed.

III. The Crane-Wright Challenge

Although Crane and Wright mainly focus on a theoretical claim about bundled discounts, their Comment starts with some assertions about the relevant empirics and welfare standards. I thus address those assertions briefly before moving on to their theoretical challenge.

A. THE EMPIRICAL EVIDENCE DOES NOT SUPPORT CHANGING CURRENT TYING DOCTRINE

Unlike Seabright, who argues that the necessary empirical evidence does not exist, Crane and Wright make an affirmative claim that the empirical evidence shows that very few ties harm consumer welfare, stating: "the best available empirical evidence suggests the frequency of instances of bundled discounts and

tying arrangements resulting in harm to consumers as compared to those arrangements improving consumer welfare is very low.”⁸⁷ They cite two articles for this empirical claim, but neither supports it. The cited articles do assert that bundling is usually efficient, but do so based on what one of the articles admits is “casual empiricism” rather than any rigorous empirical study.⁸⁸ More important, the relevant question is not whether bundling is generally efficient, but whether the

MORE IMPORTANT, THE RELEVANT
QUESTION IS NOT WHETHER
BUNDLING IS GENERALLY EFFICIENT,
BUT WHETHER THE TIES
CONDEMNED BY CURRENT TYING
LAW ARE GENERALLY EFFICIENT.

ties condemned by current tying law are generally efficient. After all, as I have noted, it would be improper to conclude that, because driving is generally desirable, the drunk driving condemned by law is generally desirable as well.⁸⁹

Currently, tying law condemns only ties that (1) involve separate products, (2) have tying market power, and (3) lack any proven offsetting efficiency, and the cited articles provide no evidence that ties meeting those three conditions are generally efficient. To the contrary, their argument that bundling is generally efficient is based largely on bundles that exist on competitive markets,⁹⁰ which would flunk not only the market power requirement, but also the separate products element necessary to have a tie at all, because two items are deemed a single product if they are routinely bundled in a competitive market under a test I elaborated in my portion of an antitrust treatise.⁹¹ The rest of their argument is based on the possibility of various efficiencies,⁹² which I fully acknowledge, but if offsetting efficiencies exist, the tie would not be condemned under current tying doctrine. No empirical evidence is presented in either cited article that the ties that are actually condemned under current doctrine generally benefit consumer welfare. To the contrary, one of the cited articles expressly acknowledges that we do not have empirical studies of the effects of antitrust actions that condemned ties.⁹³

Moreover, even if Crane and Wright were right that most ties are efficient and that the conditions necessary for ties to be anticompetitive are rare,⁹⁴ that is no reason to change current tying doctrine to adopt either the Chicago view that *all* ties should be categorically legal or the quasi-Chicago view that *all* ties without a substantial foreclosure share should be categorically legal. After all, Crane and Wright themselves explicitly acknowledge that ties and bundled discounts can create monopoly leverage, impose efficiency-reducing price discrimination, exclude competitors, and harm consumers.⁹⁵ Thus, rather than adopt a categorical rule that denies the possibility of what they admit is possible, it is better to have a doctrine that makes case-by-case determinations of whether the necessary conditions exist for anticompetitive effects and whether they are offset by output-increasing efficiencies, which is precisely what occurs under current tying doctrine and is the approach I advocated.

Of course, one could argue that, although offsetting efficiencies often exist, it is hard for defendants to prove them.⁹⁶ But the empirical evidence cited to sup-

port this claimed defendant inability is weak: it consists of the fact that, lacking access to “internal cost information,” two scholars were unable to establish cost savings in 2 of 3 case studies where they felt confident cost savings must explain the tie.⁹⁷ But defendants would have access to precisely the internal cost information that these scholars lacked, and findings from 2 of 3 handpicked case studies is hardly sufficient to draw general empirical conclusions. Moreover, a balanced analysis would have to acknowledge it is also hard for plaintiffs to prove the absence of efficiencies and the existence of anticompetitive effects, so the relative difficulty of proof may not favor defendants.

In any event, even if one thinks that most cases covered by current tying doctrine involve efficiencies and that defendants have much more difficulty proving those efficiencies than plaintiffs have proving their absence, that would at most justify shifting the burden of proof on efficiencies to plaintiffs.⁹⁸ It would not justify the categorical non-liability rules advocated by the Chicago or quasi-Chicago view. Nor has any rigorous empirical evidence been provided for the premises necessary to justify a change in the litigation burden of proof. Given that the policy burden of proof is on those who want to overrule decades of *stare decisis*, that burden has clearly not been met.

B. THE CRANE-WRIGHT ARGUMENT AGAINST A CONSUMER WELFARE STANDARD

Crane and Wright also argue against judging ties with market power under a consumer-welfare standard. They state that they object: “to Professor Elhauge’s claim that antitrust law has committed to a course that would require it to micromanage markets to identify and sanction instances of tying, bundling, and bundled discounts that reduce static consumer welfare. We believe such a policy would be counterproductive for consumers, unadministrable, and run afoul of antitrust law’s tolerance of simple monopoly pricing (which obviously reduces static welfare), and would be inconsistent with the Supreme Court’s antitrust jurisprudence.”⁹⁹

Their claim that my position conflicts with antitrust tolerance of monopoly pricing is quite mistaken. I explicitly noted that tying that merely extracts more profits out of monopoly power, rather than extending that monopoly power by excluding rivals, cannot be condemned as monopolization.¹⁰⁰ Because monopoly pricing does not exclude rivals, it also cannot be condemned as monopolization, and because it involves no agreement or conditioned sale, it cannot be condemned under other antitrust provisions. In contrast, tying and

THE CRANE-WRIGHT ARGUMENT
THUS AMOUNTS TO A CLAIM
THAT, IF WE ALLOW MONOPOLY
PRICING THAT HAS ADVERSE
WELFARE EFFECTS, WE MUST
ALLOW AGREEMENTS THAT
HAVE SIMILAR ADVERSE
WELFARE EFFECTS. THAT CLAIM
OBVIOUSLY CONFLICTS WITH
ANTITRUST LAW NOT JUST ON
TYING BUT ON ALL AGREEMENTS
IN RESTRAINT OF TRADE.

bundled discounts do involve agreements and conditioned sales and thus can be judged under doctrines other than monopolization. The Crane-Wright argument thus amounts to a claim that, if we allow monopoly pricing that has adverse welfare effects, we must allow agreements that have similar adverse welfare effects. That claim obviously conflicts with antitrust law not just on tying but on all agreements in restraint of trade, including horizontal price-fixing. Nor does their argument bear on the choice between a consumer or total welfare standard because their mistaken analogy to monopoly pricing would apply no matter which welfare standard were used.

The Crane-Wright claim that my approach would be unadministrable and contrary to antitrust jurisprudence is hard to square with the fact that current tying doctrine clearly does weigh any efficiencies of a tie against its anticompetitive effects, as does the rule of reason for all agreements in restraint of trade. Further, while the authorities I collected clearly establish that consumer welfare is the legal metric for making such a trade off, the Crane-Wright objection to the administrability of case-by-case rule of reason analysis would be equally applicable if total welfare were the metric. This argument thus also fails to bear on the choice of welfare standard. Instead, its logic amounts to a radical claim that all agreements in restraint of trade should be judged either *per se* legal or illegal, with no case-by-case rule-of-reason analysis under any welfare metric.

C. THE CRANE-WRIGHT ARGUMENT ON BUNDLED DISCOUNTS IGNORES THE FACT THAT BUYER COLLECTIVE ACTION PROBLEMS MAKE THEM PRICE-TAKERS

Crane and Wright mainly argue that bundled discounts cannot have the same power effects as tying.¹⁰¹ They reason that a firm cannot credibly threaten to charge an unbundled price that exceeds the monopoly price to buyers who refuse a bundle because carrying out that threat against noncompliant buyers would be

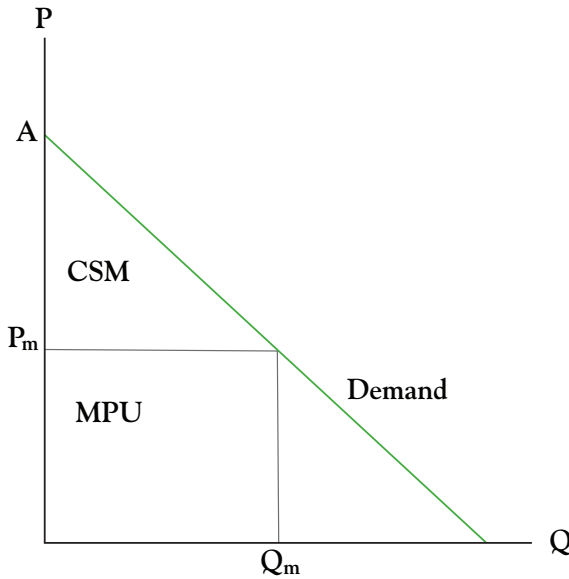
IT IS EASY TO SHOW THAT A
THREAT TO CHARGE AN
UNBUNDLED PRICE THAT EXCEEDS
THE MONOPOLY PRICE HAS
AS MUCH (IF NOT MORE)
CREDIBILITY THAN THE
CONVENTIONAL TYING THREAT
NOT TO SELL AN UNBUNDLED
PRODUCT AT ANY PRICE.

less profitable to the firm than lowering its unbundled price to the monopoly level.¹⁰² Their claim here repeats Crane's critique that a prior article by Nalebuff was invalid because it assumed that the seller could threaten an unbundled price that exceeded the monopoly price, which Crane asserted was not a credible threat for the same reason.¹⁰³

However, it is easy to show that a threat to charge an unbundled price that exceeds the monopoly price has as much (if not more) credibility than the conventional tying threat not to sell an unbundled product at any price, and that the Crane-Wright argument would thus imply that tying is also impossible, which is clearly untrue. To see why, let's take the simple case of a market where each buyer has linear demand

of $A - P$ and costs are zero. The monopoly price will thus be $P_m = A/2$, each buyer will purchase a quantity Q_m , and the consumer surplus for each buyer at this monopoly price will be the triangle marked “CSM.”

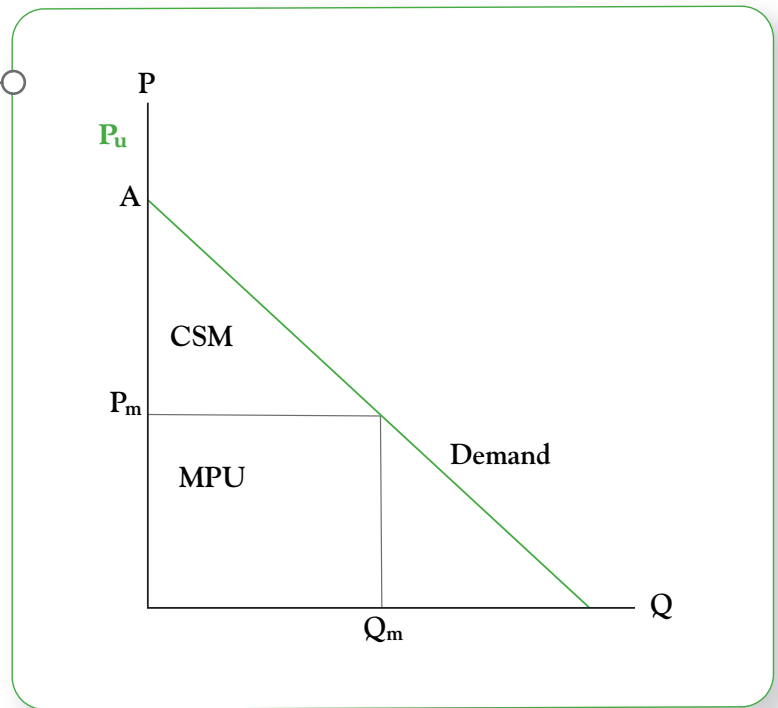
Figure 1



Suppose this monopolist used a conventional tie where it refused to sell this monopoly product at the monopoly price unless the buyer buys the tied product from it at a supra-competitive price. Then standard economic analysis finds that the buyer will accept the tie if CSM exceeds the consumer surplus lost (“CSL”) on the tied product from having to buy it at a supra-competitive price.¹⁰⁴ The buyer will do so because the buyer gets more surplus by accepting the tie than by rejecting it. Thus, through tying, the monopolist can increase its profits per buyer from MPU (its monopoly profit at a uniform price) up to $MPU + CSL$.

Now suppose the monopolist instead imposes a bundled discount where the unbundled price P_u exceeds A, the lowest price that chokes off demand, but the monopolist gives a “discount” of $P_u - P_m$ on the monopoly product to buyers who purchase the tied product at the same supra-competitive price as in the conventional tying case.

Figure 2



Then, just as in the conventional tying case, standard economic analysis finds that the buyer would accept the bundle if $CSM > CSL$ because the buyer gets more surplus by accepting the bundle than by rejecting it, and the monopolist will thus increase its monopoly profit to $CSM + CSL$.¹⁰⁵ Under the Crane-Wright logic, the buyers would instead reject the bundle because the seller's threat to charge P_u to a non-compliant buyer is not credible given that the monopolist would make more money by caving to a rejecting buyer and selling the monopoly product at the monopoly price. But by that logic, one could equally say that buyers would reject any tie because the seller's threat not to sell the tying product at any price to a non-compliant buyer is not credible given that the monopolist would make more money by caving to a rejecting buyer and selling the tying product at the monopoly price. In either the tying or bundled discount case, the seller would lose the profits on selling to this buyer at a uniform monopoly price (MPU) by carrying out its threat. The cases are economically indistinguishable. Yet we know that tying threats can be sufficiently credible to induce buyers to accept ties, which Crane and Wright do not deny. Thus, bundled discounts must be equally credible when they make a threat that is economically indistinguishable from the tying threat.

What is the flaw in the Crane-Wright logic? It is that they have one-sidedly focused on the credibility of only the seller's threat, without considering the

credibility of their assumed buyer threat to reject the bundle. They simply assumed that all buyers have credibly rejected the bundle, so that the seller's only choice is to sell at the unbundled price (which here results in no sales) or cave and sell at the monopoly price with the bundle. But, in fact, the buyer threat to reject the bundle is not credible because each buyer would lose CSM - CSL in consumer surplus if the buyer rejected the bundle. If the market had only one seller and one buyer, then one could imagine a bargaining game of chicken with unclear resolution, but in a typical market the seller faces many buyers who have a collective action problem. It is better for each buyer to accept than reject, and no single buyer's rejection would cause the seller to deviate from a bundling strategy that increases seller profits by CSL to all other buyers. The seller is a unitary actor, but the buyers have a collective action problem. Thus, each buyer would accept the bundle and the seller need never carry out the threat or sacrifice any profits, as Nalebuff correctly concluded in prior work.¹⁰⁶

Indeed, the seller's threat has as much, if not more, credibility as conventional monopoly pricing itself. In the standard monopoly pricing case, the seller's threat is to refuse to sell the product at any price unless buyers agree to pay the monopoly price. Under the Crane-Wright logic, the seller threat under monopoly pricing would not be credible because, if the buyer threatened not to buy the product unless the monopolist lowered the price below the monopoly price to some above-cost level, the monopolist would find it more profitable to sell at that above-cost price than to forego sales and lose all profits to that buyer. Thus, the Crane-Wright logic would imply that monopoly pricing itself is impossible, which again conflicts with commonplace observation. Instead, standard economics finds that monopoly pricing works because collective action problems among many buyers make them price takers.

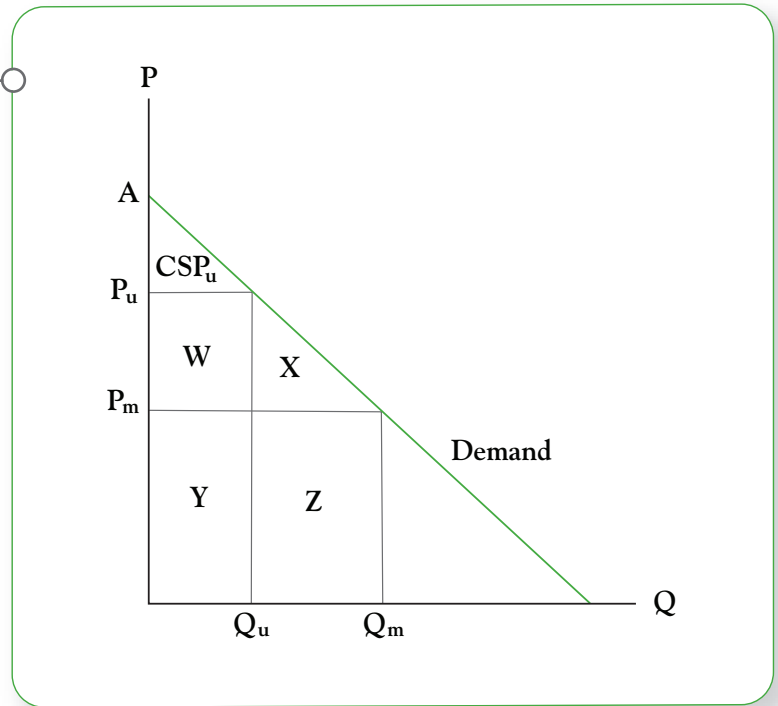
Crane and Wright's contrary logic thus conflicts with the standard economic observation that buyers are price takers in any typical market with many buyers. If we instead stick to this standard price-taker observation, then, in tying and bundled discount cases, buyers will accept because they prefer accepting the tied or bundled terms to doing without the tying product, just like buyers pay the monopoly price because they prefer paying it to doing without the product.

Now consider the case where the unbundled price is below the choke price, A . If a buyer rejected the bundle, it would not lose all of CSM, because rejecting buyers would buy some quantity Q_u at the unbundled price P_u and thus get their consumer surplus at the unbundled price, which is CSP_u . But if they are price takers, all buyers would accept the bundle as long as the difference between CSM and CSP_u , which in Figure 3 is $W + X$, exceeds the consumer surplus lost by purchasing the linked product at supra-competitive prices. The dynamic on the buyer side is precisely the same as the conventional tying case where buyers com-

THE CRANE-WRIGHT LOGIC
WOULD IMPLY THAT MONOPOLY
PRICING ITSELF IS IMPOSSIBLE,
WHICH AGAIN CONFLICTS WITH
COMMONPLACE OBSERVATION.

pare CSM to CSL, with the only difference being that here the buyers compare $W + X$ to CSL.

Figure 3



On the seller's side, the threat to charge P_u to noncompliant buyers is no less credible than the seller's threat in a conventional tying case. To the contrary, it is *more* credible. In the conventional tying case, carrying out the seller's threat means not selling the tying product at all and thus sacrificing all of $Y + Z$. In the bundled discount case where $P_u < A$, carrying out the seller's threat means selling the monopoly product at P_u and getting Q_u in sales, thus earning $W + Y$ rather than $Y + Z$. The profits that would be sacrificed if this threat ever had to be carried out are thus just $Z - W$, which is much smaller than $Z + Y$. Thus, if a buyer rejected a bundled discount where the unbundled price was lower than the choke price, carrying out the seller threat would require much less of a profit sacrifice than carrying out a conventional tying threat, making the bundled discount threat, if anything, more credible. Accordingly, if one thought (like Crane and Wright) that the credibility of the seller threat mattered, then the threat to charge an unbundled price that exceeds the monopoly price (but is below the choke price) is clearly more credible than the threat under conventional tying (or monopoly pricing) not to sell the product at any price. In fact, buyers have a collective action problem that makes them price takers, so buyers in either case will accept the bundle, and the seller will never have to carry out the threat.

The rest of Crane and Wright's arguments about the credibility of an unbundled price that exceeds the monopoly price all rest on their mistaken premise that buyers would respond by rejecting the bundle, thus forcing the seller to sacrifice profits and lose sales to rivals or substitute products.¹⁰⁷ In reality, no profit sacrifice is required because price-taking buyers will accept the bundle, which increases seller profits. Nor does the above analysis change if we assume that the seller has rivals in the tying market or that the tying product has substitutes. The existence of rivals and product substitutes will simply affect the shape of the seller's firm-specific demand curve. As long as that firm-specific demand curve has a downward slope—i.e., as long as the seller has tying market power—then buyers who buy from that firm will get some consumer surplus at the profit-maximizing price for the tying product, and all the analysis above will continue to hold. (I already pointed this out in analysis that Crane and Wright do not address.¹⁰⁸) Accordingly, the seller can offer a tie or bundled discount that all its buyers will accept because the consumer surplus that each buyer would lose by rejecting the bundle exceeds the consumer surplus that each buyer would lose by accepting it. The bundling seller with market power thus need not sacrifice any profits nor lose any sales to its rivals or substitute products.

THE BUNDLING SELLER WITH MARKET POWER THUS NEED NOT SACRIFICE ANY PROFITS NOR LOSE ANY SALES TO ITS RIVALS OR SUBSTITUTE PRODUCTS.

The above focuses on bundled discounts that extract individual consumer surplus, but we can say much the same about the credibility of bundled discounts that cause the other two power effects. For bundled discounts that cause inter-product price discrimination, the economic literature has already mathematically proven that bundled discounts are *more* profitable for the seller than a pure tie *and* that sellers will maximize profits by setting the unbundled price above the but-for price for any product over which it has market power.¹⁰⁹ Crane and Wright offer no rebuttal to these mathematical proofs.

For bundled discounts that create intra-product price discrimination, if the unbundled price exceeds the choke price, then the bundled discount is economically equivalent to a tie. Thus, the threat to charge the unbundled price to buyers who refuse the bundle has precisely the same credibility as the conventional tying threat of refusing to sell the tying product at any price to buyers who do not accept the bundle. If the unbundled price is lower than the choke price, then the price discrimination effects are the same as tying for any buyers who value the tying product less than the unbundled price. Consider Figure 3 again, but with Q now meaning the number of buyers who purchase the tying product, and assume each buyer purchases only one unit of a tying product whose value correlates with usage of a tied product. This bundled discount could not price discriminate among the buyers from 0 to Q_u because those buyers could always avoid any effort to extract the portion of their valuation above P_u by just purchasing the tying product at P_u . But this bundled discount could achieve precise-

ly the same profitable price discrimination effects as tying for the buyers who value the tying product less than P_u .¹¹⁰ Again, this threat is, if anything, more credible than conventional tying, because if buyers were to reject the bundled discount, the seller only loses $Z - W$, whereas if buyers were to reject a conventional tie, the seller would lose $Y + Z$. However, because buyers are price takers in any market with many buyers, in fact buyers who value the tying product less than P_u would accept the bundle as long as the surcharge on the tied product did not exceed the consumer surplus each buyer enjoyed on the tying product, just as they would with a conventional tie.

D. BUYER-INITIATION DOES NOT DISPROVE ANTICOMPETITIVE EFFECTS

Crane and Wright also argue that buyers may initiate bundled loyalty discounts that create efficiencies.¹¹¹ However, buyer initiation of bundled or unbundled loyalty conditions does not disprove anticompetitive effects because such conditions can raise market-wide prices when they cover a sufficient share of the market, and the lions' share of that market-wide price increase is externalized onto other buyers in the market.¹¹² Even more of that market-wide price increase is externalized if the buyers are intermediaries who pass most or all of the price increase on to downstream buyers.¹¹³ Because of that externality, entering into a loyalty agreement in exchange for side-payments or some trivial discount from the elevated market price will be individually profitable for each buyer, even though the externality means it is harmful to buyers collectively. Each buyer thus

EACH BUYER THUS HAS
INDIVIDUAL INCENTIVES TO ENTER
INTO LOYALTY AGREEMENTS EVEN
THOUGH THE RESULT OF ALL OF
THEM FOLLOWING THOSE
INDIVIDUAL INCENTIVES IS THAT
ALL BUYERS ARE HARMED.

has individual incentives to enter into loyalty agreements even though the result of all of them following those individual incentives is that all buyers are harmed.

Whether buyers initiate such a loyalty agreement is thus irrelevant because the same externality problem that makes it individually profitable for buyers to accept an anticompetitive loyalty condition also makes it individually profitable for buyers to initiate an anticompetitive loyalty condition that harms all buyers collectively.¹¹⁴ Buyer initiation is thus no more relevant than voluntary action is in any other situation where externalities exist. For example, in the classic tragedy of the commons, each cow herder initiates bringing too many cows to the commons because each considers only the individual benefit of doing so and ignores the harm to other cow herders, but this does not alter the inefficiency of them doing so. Likewise, individuals may initiate littering because they ignore the effects of their littering on others, but this does not alter the desirability of laws against littering to prevent everyone from initiating littering that collectively harms everyone.

Crane and Wright argue that this externality problem is inapplicable to bundled loyalty discounts procured by Group Purchasing Organizations ("GPOs") or

Pharmacy Benefit Managers (“PBMs”) because those groups can solve the collective action problem among their members.¹¹⁵ But their argument falters on two scores. First, even Crane and Wright admit that intermediate buyers may initiate anticompetitive loyalty agreements because they pass the price increase on to downstream consumers.¹¹⁶ GPOs and PBMs have even more incentive than intermediate buyers to initiate anticompetitive loyalty agreements because GPOs and PBMs don’t purchase the product at all, but rather serve as brokers who earn a percentage of the purchase price, and thus have affirmative incentives to agree to loyalty conditions that increase market prices.¹¹⁷ Second, even if one thought that GPOs and PBMs perfectly represented their downstream purchasers, each GPO or PBM would still externalize most of its agreement’s adverse effect on market prices onto other groups and downstream purchasers. Under U.S. guidelines, each GPO must keep its share of purchases in any market below 35% to avoid possible challenge for being an illegal horizontal combination.¹¹⁸ Thus, each GPO externalizes 65% or more of the market harm caused by its agreement to an anticompetitive loyalty agreement. The largest PBM has a smaller market share than the largest GPO and thus would externalize even more of the market harm that would be caused if it agreed to an anticompetitive loyalty agreement.

EACH GPO EXTERNALIZES
65% OR MORE OF THE MARKET
HARM CAUSED BY ITS AGREEMENT
TO AN ANTICOMPETITIVE
LOYALTY AGREEMENT.

Crane and Wright also assert that, although customer-initiated bundled discounts can harm consumers, they can do so only if they create predatory below-cost discounts that exclude rivals.¹¹⁹ However, as the economic literature shows, and as I explained in my article using concrete illustrations, above-cost bundled loyalty discounts can harm consumer and total welfare by raising the costs of equally efficient rivals or by excluding less efficient rivals who would otherwise constrain market prices.¹²⁰ Further, the economic literature also shows that above-cost bundled loyalty discounts can—without excluding rivals or reducing rival efficiency—reduce the incentives of firms and their rivals to compete on price, which rivals may have no incentive to undo because it is profitable for them.¹²¹ Crane and Wright simply provide no substantive response to this economic literature.

Finally, Crane and Wright rely heavily on a recent article by Professors Benjamin Klein and Kevin Murphy that argues that retailers may have incentives to initiate exclusive dealing agreements in differentiated product markets.¹²² In essence, Klein and Murphy argue that, in such a differentiated market, bidding for an exclusive contract with a retailer can increase the relevant demand elasticity by combining downstream buyers with high and low valuations for the seller’s product.¹²³ This, they argue, will cause sellers to price at cost and result in a gain in consumer surplus that outweighs the lost product variety.¹²⁴ Crane and Wright argue that this analysis can be extended by analogy to bundled loyalty

discounts that are initiated by buyers.¹²⁵ But there are several problems with this line of argument.

First, the Klein-Murphy model is problematic. Under their model, the two sellers in a differentiated market would sell at cost and earn zero profits if they used exclusive contracts, but would sell at prices that were double their cost if they did not.¹²⁶ Given that premise, it is hard to see why the sellers would be willing to bid on an exclusive basis, let alone why, as Klein and Murphy assert, sellers would have “the exact same motivation” as retailers to initiate exclusive bidding.¹²⁷ Under their model, exclusive contracts harm the sellers and thus any seller with market power would avoid them. A seller who agrees to bid on an exclusive contract would earn zero profits and thus earn just as much by not bidding.¹²⁸ The seller can thus costlessly threaten not to bid on an exclusive contract, and the retailer cannot credibly respond by insisting on an exclusive contract because doing so would mean buying exclusively from the other seller at a monopoly price (given the resulting lack of competitive bidding). Further, while the sellers have market power, the retailers are plentiful and will suffer from a collective action problem that makes them price takers, not entities who can insist that sellers with market power bid on the basis that is most advantageous to retailers. The Klein-Murphy model seems to oddly flip the assumption about who the price taker is when a seller has market power in a market with many buyers.

Even if we posit that, for some reason, retailers can credibly threaten not to buy from a seller with market power unless the seller bids on an exclusive basis that results in zero seller profit, such retailers could, with equal credibility, threaten not to buy from the seller unless it bids at cost on a *non-exclusive* basis. Each seller would, in this scenario, sell half of the retailer’s demand at cost, but if Murphy and Klein are right that each seller would prefer to sell all of the retailer’s demand at cost rather than not sell to the retailer at all, then each seller would also prefer to sell half of the retailer’s demand at cost rather than not sell to the retailer at all.¹²⁹ Retailers would be better off buying at cost on a nonexclusive basis because that increases the satisfaction of their consumer’s varying

IF RETAILERS HAD THE ABILITY
TO CREDIBLY INSIST ON BIDS
THAT LED TO SELLER PRICES
THAT EQUALED COST, RETAILERS
WOULD BE BETTER OFF DOING SO
WITHOUT ANY EXCLUSIVITY.

brand preferences compared to buying at cost on an exclusive basis.¹³⁰ Thus, if retailers had the ability to credibly insist on bids that led to seller prices that equaled cost, retailers would be better off doing so without any exclusivity.

Second, even if the Klein-Murphy model were convincing on single-product exclusive contracts, one cannot simply extend it by analogy to bundled loyalty discounts. Other models

that have analyzed bundled loyalty discounts in differentiated markets find that they produce an inefficient product mix and excessive bundling.¹³¹

Third, even if the Klein-Murphy model were convincing and applicable to bundled loyalty discounts, it shows that retailer-initiated exclusive contracts can lead to efficiencies only under very particular assumptions about market differentiation and costs, not that exclusive contracts always or usually do so, let alone that any efficiencies always or usually outweigh any anticompetitive effects. If a particular bundled loyalty discount actually did create such efficiencies, then the test I propose would fully consider them.¹³² The Crane-Wright analogy to the Klein-Murphy model thus provides no reason to deviate from my suggested test for bundled discounts.

IV. THE NALEBUFF MODELS

Nalebuff's Comment makes a major contribution to modeling imperfect price discrimination created by metering ties. This is just the sort of article I searched for when I wrote the section of my article on that power effect, and if it had existed earlier, it could have saved me a lot of time. However, the Nalebuff Comment does proceed on a misapprehension about my claim regarding metering ties. Correcting that misapprehension shows that his models support my actual position. To the extent our models diverge on some details, I think my model better captures the imperfect price discrimination produced by real metering ties by assuming that: (1) buyers purchase a whole number of tied units, rather than infinitely divisible fractions of tied units (as he assumes), and (2) buyers have varying valuations, rather than the same valuation for tied product usage over the relevant range (as some of his models assume).

A. THE MISAPPREHENSION

As Nalebuff correctly observes, metering ties are just one of the three power effects that I considered in assessing the overall effects of ties, and I argued that we should focus on consumer welfare, or at least total welfare, rather than on *ex post* total welfare.¹³³ Thus, my defense of current tying doctrine holds on these grounds whether or not metering ties usually increase *ex post* total welfare.

EVEN IF *EX POST* TOTAL WELFARE WERE THE RIGHT STANDARD, A CONCLUSION THAT METERING TIES USUALLY INCREASE *EX POST* TOTAL WELFARE WOULD NOT JUSTIFY REPLACING CURRENT DOCTRINE WITH A CATEGORICAL RULE OF LEGALITY FOR METERING TIES, BECAUSE SUCH A CATEGORICAL RULE WOULD INSTEAD WRONGLY ASSUME THAT METERING TIES ALWAYS INCREASE *EX POST* TOTAL WELFARE.

However, Nalebuff incorrectly states that I also claimed that the imperfect price discrimination produced by metering ties usually reduces *ex post* total welfare.¹³⁴ That is not what I said. My claim was that: "Imperfect intraproduct price discrimination actually reduces *ex post* total welfare by misallocating output, *unless that inefficiency is offset by an output-increasing efficiency*."¹³⁵ Although I pointed out cases when an offsetting output-increasing efficiency would not

exist, I did not deny that they can or usually exist. Quite the opposite, I found that (assuming linear demand and equal-sized groups) a metering tie “lowers *ex post* total welfare for 2 or 3 tied units, but increases it for 4 or more units,” with the *ex post* total welfare gains “ranging from 0.4% to 9% and converging on 4.85% for large numbers of tied units.”¹³⁶ My argument was not based on a claim that metering ties generally reduce *ex post* total welfare, but was rather that: “in those cases where tying-induced price discrimination does increase *ex post* total welfare, the defendant should be able to prove an output-increasing efficiency. . . . Indeed, if (by hypothesis) the critics were right that the relevant legal welfare standard is *ex post* total welfare, then that would be the standard the quasi-per se rule applies to determine whether the efficiency offsets the harm, and the quasi-per se rule would never condemn a tie that increased *ex post* total welfare.”¹³⁷

Thus, even if *ex post* total welfare were the right standard, a conclusion that metering ties *usually* increase *ex post* total welfare would not justify replacing current doctrine with a categorical rule of legality for metering ties, because such a categorical rule would instead wrongly assume that metering ties *always* increase *ex post* total welfare.¹³⁸ Even less would such a conclusion justify replacing current doctrine with a categorical rule of legality for all ties with power effects,

given that the other two power effects are less likely to have positive effects on *ex post* total welfare.¹³⁹

EVEN IF THE PROPER STANDARD WERE TOTAL WELFARE, THERE IS NO REASON TO FIXATE ON *EX POST* TOTAL WELFARE, WHICH IN TYING CASES PROBABLY CORRELATES LESS WELL TO TOTAL WELFARE THAN CONSUMER WELFARE DOES.

Further, I pointed out that consumer welfare is actually the right standard, and that the same theoretical considerations that suggest metering ties might usually increase *ex post* total welfare mean they are even more likely to reduce consumer welfare.¹⁴⁰ Finally, I showed that even if the proper standard were total welfare, there is

no reason to fixate on *ex post* total welfare, which in tying cases probably correlates less well to total welfare than consumer welfare does.¹⁴¹

That was the policy argument, no part of which relied on a claim that metering ties usually reduce *ex post* total welfare. With my actual policy argument in mind, let’s consider Nalebuff’s three models.

B. NALEBUFF’S BASELINE MODEL

In his baseline model, Nalebuff assumes that each buyer values the tying product in direct proportion to the number of tied units they use and that each buyer puts the same value as other buyers on each usage.¹⁴² Given these assumptions, a tie that prices those tied units at that value amounts to perfect price discrimination. Nalebuff correctly acknowledges this and that, in reality, “price discrimination is usually imperfect. . . .”¹⁴³ However, he argues that this baseline model provides

some intuition for the claim that metering ties “will typically increase [*ex post*] total welfare and decrease consumer welfare.”¹⁴⁴ Nalebuff mistakenly thinks I disagree with this claim,¹⁴⁵ but in fact I confirmed it in my own model of metering ties. My argument was instead that: (1) this intuitive analogy did not mean that the metering ties that are actually condemned by current tying doctrine would usually increase *ex post* total welfare, because that doctrine permits metering ties that have offsetting efficiencies,¹⁴⁶ and (2) this intuitive analogy did not support the critic’s claim that the consumer welfare effects of metering ties were more ambiguous than the *ex post* total welfare effects.¹⁴⁷ Nalebuff does not address the first argument, but supports me on the second because he affirms that metering ties “typically . . . decrease consumer welfare.”¹⁴⁸

Nalebuff also argues that this baseline model does a surprisingly good job of describing the tie of printer heads to ink that was at issue in *Illinois Tool Works v. Independent Ink*.¹⁴⁹ I’m not sure about that; it seems to me quite plausible that different customers would use printer heads to print different amounts and value what they printed differently. Indeed, in his amicus brief in *Illinois Tool Works*, Nalebuff argued against the metering tie in that case based partly on his conclusion that, given customer variation, the tie would produce only imperfect price discrimination that could not be assumed to increase efficiency.¹⁵⁰

But suppose Nalebuff is now right in his characterization of *Illinois Tool Works*: what are the implications? One implication is that such a perfect metering tie totally eliminates all consumer surplus.¹⁵¹ Because consumer welfare is the actual legal standard, that implication resolves the economics that are relevant to the law. Although Nalebuff’s Comment suggests it might be better to use some weighed sum of producer profits and consumer welfare,¹⁵² such an approach would raise the problems already detailed above in Part I. Further, Nalebuff’s amicus brief in *Illinois Tool Works* agreed with me that consumer welfare is actually the correct standard as a matter of both law and policy.¹⁵³

TO THE EXTENT THAT METERING
TIES LIKE IN *ILLINOIS TOOL
WORKS* DO PRODUCE PERFECT
PRICE DISCRIMINATION, THEY
WILL REDUCE TOTAL WELFARE,
EVEN THOUGH THEY MAXIMIZE
EX POST TOTAL WELFARE.

Another implication is that a perfect metering tie reduces total welfare, even though it increases *ex post* total welfare. The reason is that, while there is some debate about precisely what fraction of total surplus to give innovators in order to maximize total welfare, we know that giving 100% of total surplus to the successful innovator produces excessive investment and reduces total welfare.¹⁵⁴ That is, we know, as discussed above, that the curve does not constantly increase up to 100% but is instead an inverted-U.¹⁵⁵ Thus, to the extent that metering ties like in *Illinois Tool Works* do produce perfect price discrimination, they will reduce total welfare, even though they maximize *ex post* total welfare.¹⁵⁶

Nalebuff’s baseline model thus provides no basis to conclude that metering ties likely increase total welfare. A superficial reading of Nalebuff might suggest otherwise, but he is careful to explain that he is using the term “total welfare” to refer only to “*ex post* total welfare,” and that he has not considered *ex ante* effects.¹⁵⁷ Moreover, his baseline model also shows a clear decline in the consumer welfare that his prior work acknowledged is the correct antitrust standard.

C. NALEBUFF’S MODEL I

The model that Nalebuff’s paper discusses the most is his Model I, which is very similar to my own model of metering ties with the exception that Nalebuff assumes buyers make a continuous choice about how many tied units to buy, whereas I assumed buyers make discrete choices.¹⁵⁸ That is, whereas I assumed buyers can buy 1, 2, 3 or some other whole number of cartridges, Nalebuff assumes buyers can also buy 1.1 or 2.26 cartridges or any other infinitely divisible fraction of cartridges. This permits Nalebuff to offer a more mathematically powerful proof than I could. However, it also means his model deviates more from reality because in fact buyers cannot buy fractions of cartridges. Nor can buyers purchase fractions of other tied product units; if they could, then by definition whatever minimum fraction they could buy would be the tied “unit” used in my model.

Nalebuff’s model comes in two flavors, both of which confirm my own conclusions about the likely welfare effects of metering ties. In one version, Nalebuff assumes that buyers can buy any fraction of tied units, even less than one tied unit. Because, in his model, buyers are basically choosing among an infinite number of tied unit choices, his results are, not surprisingly, quite similar to my findings when the number of tied units is very large, as the following table shows. Nalebuff’s conclusions thus strongly confirm my own for large numbers of tied units. In particular, Nalebuff and I both find that metering ties reduce consumer welfare by almost 19%, which supports presumptive condemnation under the consumer welfare standard used by antitrust law.

Table 1

	Elhaug Model Conclusion ¹⁵⁹ (for large numbers of tied units)	Nalebuff Model I Conclusion ¹⁶⁰ (for infinitely divisible tied units)
Consumer Welfare	−18.85%	−18.75%
Total Welfare	+4.85%	+4.88%
Tying Product Output	+39.79%	+39.66%
Tied Product Output	−2.29%	−2.29%

In the other version of Nalebuff Model I, he assumes that buyers have to buy at least one tied unit but can buy any fraction of units above one. Here, he finds that *ex post* total welfare effects are positive only if the number of tied units exceeds 4.58.¹⁶¹ I found that the *ex post* total welfare effects are positive only if the number of tied units is 4 or higher.¹⁶² Thus, Nalebuff's latter model is quite consistent with my findings and indicates that, if anything, my model is slightly conservative about when metering ties are likely to reduce *ex post* total welfare.

Nalebuff and I both find that, even if metering ties increase *ex post* total welfare and tying product output, they decrease the number of tied products used. As Nalebuff notes, this was a surprising result, and I am glad his analysis confirms it.¹⁶³ However, because usage of the tied product is what correlates with actual productive output in a metering tie, this result does cut against metering ties for those who think that antitrust should focus on the extent to which restraints increase or decrease productive output. Further, in the real world (unlike in our model) there are real costs to making the tying product, so that productive efficiency seems likely to be adversely affected to the extent that metering ties result in the increased creation of costly tying products that are utilized less often.

NALEBUFF AND I BOTH FIND
THAT, EVEN IF METERING TIES
INCREASE EX POST TOTAL
WELFARE AND TYING PRODUCT
OUTPUT, THEY DECREASE
THE NUMBER OF TIED
PRODUCTS USED.

I should caution also that both of our models depend on the assumption that, at a competitive tied product price, the number of tying product buyers who would use a low amount of tied units equals the number of buyers who would use a medium or high amount. Although this assumption is a useful heuristic, it often may not hold. One may reasonably think that buyers who would use many tied units would be more enthusiastic about the tying product and that there would thus be more of them. If so, then that will increase the size of the groups that use many tied units, which means that metering ties will have worse effects on consumer welfare and total welfare. Or one might think that buyers are likely to reflect a normal bell-shaped distribution where buyers who use a medium amount of tied units are more likely than buyers at either extreme. In that case, I conjecture (but have not proven) that the welfare effects of metering ties would be worse because a uniform price would generally result in sales to the medium buyers and there would be relatively fewer low unit buyers picked up by metering ties.

This last paragraph doesn't mean that one can assume that metering ties will usually decrease *ex post* total welfare. It simply means that, even if one thought that *ex post* total welfare were the correct standard, one should not over-read our models as showing that metering ties always increase *ex post* total welfare unless the number of tied units is fairly small. Metering ties may well often or usually decrease *ex post* total welfare under different assumptions about the distribution of buyers who use low, medium, and high amounts of tied units. Thus, even if the law were to switch to an *ex post* total welfare standard, the law should stick to

judging metering ties under current tying doctrine, which makes case-by-case judgments that can reflect varying buyer distributions.

D. NALEBUFF MODEL II

In his final model, Nalebuff assumes that each customer has a declining marginal value for usage of the tied product such that each “customer of type a values the q th copy at $a - q$.”¹⁶⁴ This assumption allows customers to have a range of valuations for the first tied unit they buy. But by assuming that buyers keep buying units of the tied product until the valuation of the last unit they buy equals the tied product price, this model assumes that all buyers have precisely the same valuation for the last tied unit they buy, as well as the same valuation for the penultimate unit, and so on until we get up to the value

EVEN IF THE LAW WERE TO SWITCH
TO AN EX POST TOTAL WELFARE
STANDARD, THE LAW SHOULD
STICK TO JUDGING METERING TIES
UNDER CURRENT TYING DOCTRINE,
WHICH MAKES CASE-BY-CASE
JUDGMENTS THAT CAN REFLECT
VARYING BUYER DISTRIBUTIONS.

of the first unit (that is, a) for any customer group. Thus, for small increases in the tied product price, the model effectively assumes that all buyers will have the same valuation for the marginal tied product usage affected.

Nalebuff Model II accordingly assumes a lot more uniformity about valuation than my model or Nalebuff Model I, which assumed that buyers within and across groups had different valuations for usage of the tied product. Instead,

Nalebuff Model II assumes uniformity in buyer valuation for small increases in the tied product price, which is the relevant price range considered in this model. His Model II thus effectively assumes a form of quasi-perfect price discrimination that comes close to Nalebuff’s baseline model. Not surprisingly, Model II thus leads to the similar result that small increases in price discrimination via metering ties always increase *ex post* total welfare.¹⁶⁵ Although Nalebuff also says that the effect on consumer welfare is ambiguous in his Model II, his Theorem 5 and Appendix do not claim to have proven the consumer-welfare effect is ambiguous.¹⁶⁶ Instead, he infers this ambiguity in consumer-welfare effect from the fact that the tying product price decreases while the tied product price increases.¹⁶⁷ But that is also true for tying-tied product pricing under the perfect price discrimination produced by metering ties with constant valuation per tied unit, and we know that such ties *unambiguously reduce* consumer welfare, so parallel conclusions about pricing for metering ties that imperfectly price discriminate do not suffice to prove that their consumer-welfare effects are ambiguous.

The reader will have to judge for himself or herself the plausibility of the Nalebuff Model II assumption that consumers keep using printer cartridges up until the point when the value of printing equals the cartridge price, so that all consumers value the last thing they print precisely the same. My own sense is to the contrary, that I (and those I know) value the last thing we print at way more than the marginal price of printing, and we stop printing instead because we

have no use for an additional unit. To be sure, there must be some marginal buyers in the market who value the last thing they print at the marginal price of printing, or else the cartridge price would increase. However, I suspect most of us are infra-marginal (like most of us in most markets) and enjoy consumer surplus even on the last cartridge we use. Further, I suspect that the amount of consumer surplus we enjoy on that last cartridge varies considerably. If so, that makes a model like mine or Nalebuff Model I more appropriate than Nalebuff Model II.

The situation might be different when the buyers are intermediaries whose usage of the tied product creates a downstream output whose valuation largely reflects a common downstream market price. In those cases, each buyer might keep expanding usage/output until valuation reflects the marginal tied product price. But when the buyers are intermediaries rather than consumers, then there are other reasons (not considered by Nalebuff or any of the other Comments) to conclude that tying-induced price discrimination is likely to reduce both consumer welfare and *ex post* total welfare.¹⁶⁸

V. THE FIRST FUNDAMENTAL QUESTION

Professor First raises a more fundamental question: should the goal of antitrust be limited to enhancing welfare at all? Instead, he argues for considering multiple goals, including: (1) consumer welfare, (2) producer welfare, (3) preserving the competitive process, (4) consumer choice, (5) innovation efficiency, (6) preventing firms from getting “too big to fail,” and (7) other distributive concerns.¹⁶⁹ Applying these goals, he concludes that I am right to defend the current *quasi-per se* rule, but wrong to recognize an exception to it.¹⁷⁰ I am glad to have his support for my major conclusion, but find myself in disagreement with his multi-goals approach and with his rejection of my exception.

A. THE MULTI-GOAL APPROACH

I disagree with First’s multi-goal approach at both the wholesale and retail levels. My wholesale objection is that using such a *mélange* of goals makes the analysis entirely indeterminate. One person might apply this set of goals to ties and reach one conclusion, another might reach the opposite conclusion, and there would be no real way to choose between them. Nor is the problem limited to the fact that different people would reach different judgments. Even if we imagined only a single adjudicator, the rejection of any overarching goal means we would have no common metric for weighing each of the multiple goals, making them incommensurable. Making tradeoffs among such incommensurable goals is like asking whether a car is bluer than it is fast; the question has no real answer (unless we made the characteristics commensurable by measuring their contribution to an overarching goal like consumer preference

MAKING TRADEOFFS AMONG
SUCH INCOMMENSURABLE GOALS
IS LIKE ASKING WHETHER A CAR
IS BLUER THAN IT IS FAST; THE
QUESTION HAS NO REAL ANSWER.

satisfaction). The sheer multiplicity of goals thus means the goals will provide no real guidance in resolving doctrinal issues. We will instead be back to making conclusory judgments based on raw intuitions about whether tying or other conduct seems good or bad.

At the retail level, the problem is that each of the stated goals beyond consumer welfare is unhelpful because each is unpersuasive when it conflicts with consumer welfare. Let me address each of First's additional goals in turn.

1. Producer Welfare

First argues "we can't be completely indifferent to what happens to producer surplus. How else to understand antitrust's continuing concern for efficiencies?"¹⁷¹ This question is easy to answer. We should understand antitrust as being concerned about efficiencies only to the extent they are passed on to consumers to a sufficient extent that they improve consumer welfare. This is precisely what antitrust law provides.¹⁷² There are also several sound policy reasons not to weigh

producer surplus against consumer surplus, which I detailed in Part I.

THE COMPETITIVE PROCESS GOAL
IS VACUOUS AT ITS CORE BECAUSE
SOMETIMES DECREASING THE
NUMBER OF COMPETITORS AND
INCREASING COLLABORATION
AMONG THEM IS TREATED AS
WORSENING THE COMPETITIVE
PROCESS AND SOMETIMES IT IS
TREATED AS IMPROVING IT.

2. The Competitive Process

As First acknowledges, the goal of preserving the competitive process is "poorly defined."¹⁷³ But the problem is not merely vagueness at the edges. The competitive process goal is vacuous at its core because sometimes decreasing the number of competitors and increasing collaboration among them is treated as worsening the competitive process and sometimes it is treated

as improving it.¹⁷⁴ The only way to make sense of this pattern is to realize that what drives the results is not some freestanding notion of process, which would indicate that all those cases should be condemned because they reduce the process of competition. Instead, the results turn on whether the relevant conduct likely increases or decreases consumer welfare. The "competitive process" conclusion is simply a label applied to signal whether a court has concluded the conduct seems likely to increase consumer welfare or not. It thus adds nothing useful to a consumer-welfare standard. Indeed, the vacuity of the competitive process standard for judging issues of tying doctrine seems neatly illustrated by the fact that, while First apparently concludes it favors retaining the current *quasi-per se* rule, precisely the opposite conclusion is reached by Gregory Werden, the main current champion of the competitive process standard.¹⁷⁵

In response, First does not so much defend the competitive process standard as cite Werden's arguments that consumer welfare is also poorly defined, mainly because Werden claims that a consumer-welfare standard is inconsistent with the

fact that antitrust law condemns buyer cartels.¹⁷⁶ But as I already explained: “Condemnation in such cases is perfectly consistent with a consumer welfare standard because, if such conduct affects consumer welfare at all, the effect can only be negative. Allowing the anticompetitive ... creation of upstream market power could only reduce output and market choices in the downstream consumer market not only currently, but also in the future by making firms less willing to enter such markets.”¹⁷⁷ That is, buyer cartels lead to subcompetitive upstream prices which lower upstream output to subcompetitive levels. That reduced upstream output will be passed on downstream, because one cannot sell output that does not exist or make it from inputs that don’t exist. But the reduced upstream price will *not* be passed on downstream because the downstream price will be determined by the lower downstream output, which will raise downstream prices.

So, even though it seems counterintuitive, upstream monopsony power that reduces upstream prices will increase downstream prices to the extent it has a downstream effect. This effect could certainly be muted to the extent that the firms in the upstream buyer cartel lack downstream market power as sellers. But it would not be entirely eliminated unless downstream rivals of the cartel members really have infinitely elastic supply, which is rare. In any event, even if the effect can be muted and sometimes eliminated by downstream rival expansion, the direction of any effect is bad for downstream consumers. That is, the upstream buyer cartel either harms downstream consumers or has no discernable effect on them, but it doesn’t ever benefit downstream consumers. Because the only possible effect on downstream consumers is negative, it makes perfect sense to condemn the conduct under a consumer-welfare standard.

Even if we imagine some product for which there is no new output—like some set of famous old paintings—allowing buyer cartels could only reduce the willingness of other artists to produce new paintings. The reason is that the prospect that future buyer cartels would be allowed to suppress resale prices (once the new paintings become old) would reduce the new paintings’ expected value and thus would reduce the initial price the first buyer would be willing to pay for any new painting. Thus, allowing buyer cartels that were nominally directed only at old paintings could only decrease the output of new paintings and harm consumer welfare.

As the above suggests, a consumer-welfare standard does not require proving a harm to consumer welfare in each case. Sometimes antitrust uses rules rather than standards, and given the possible harm to consumer welfare and lack of any possible benefit to it, there is nothing wrong with a *per se* rule that condemns all buyer cartels without requiring proof in each particular case of a harm to consumer welfare. Using such a rule does not alter the fact that “consumer welfare is

AS THE ABOVE SUGGESTS,
A CONSUMER-WELFARE STANDARD
DOES NOT REQUIRE PROVING
A HARM TO CONSUMER
WELFARE IN EACH CASE.

the ultimate metric used to design antitrust laws, whether they take the form of rules or standards.”¹⁷⁸

In the end, I am not sure First disagrees with me on this point because he ultimately acknowledges that by “the competitive process” he means “processes that are likely to achieve the results that consumer surplus tries to measure.”¹⁷⁹ That appears to agree with my conclusion that “courts judge whether conduct worsens the competitive process by whether it produces a process that is likely to harm consumer welfare.”¹⁸⁰ But if one agrees with that, then it seems to me that the competitive process notion is not an independent goal and does no useful work. To the contrary, it just obscures the ultimate welfare question. One might as well proceed directly to analyze whether the challenged conduct or class of conduct seems to be the sort that is likely to harm consumer welfare, and whether a standard that looks at each case or a rule that applies to a category of cases seems the best approach for advancing consumer welfare.

3. Consumer Choice

Consumer choice is an important goal, but only because it bears on consumer welfare. If conduct reduces consumer choice, then absent some offsetting benefit, that will tend to reduce the satisfaction of consumer preferences and thus lower consumer welfare. To the extent the consumer choice goal is meant to be a corrective to the view that the only way to harm consumer welfare is by raising prices, then I think it is all to the good. But I don’t think this means consumer choice should be pursued as a goal even when it conflicts with consumer welfare. Instead, the consumer choice goal is just a factor that should be consid-

ered only to the extent it affects consumer welfare. While consumer choice is certainly relevant to consumer welfare, the latter remains the ultimate standard.

WHILE CONSUMER CHOICE IS
CERTAINLY RELEVANT TO
CONSUMER WELFARE, THE LATTER
REMAINS THE ULTIMATE STANDARD.

The scholars that First cites for the consumer choice goal, Neil Averitt and Robert Lande, seem to agree with me about its subordinate relevance because they do not claim it is a goal

distinct from consumer welfare. Instead, they argue that often “[t]here is no good way to assess consumer welfare ... without considering the non-price choice issues.”¹⁸¹ This leads them to conclude that: “The consumer choice model of antitrust ... explains ..., better than the price or efficiency models can, why antitrust is good for consumer welfare.”¹⁸² Moreover, one of those scholars, Robert Lande, has argued strongly for a consumer-welfare standard.¹⁸³

To test whether consumer choice should be a freestanding goal, rather than a subordinate factor relevant to consumer welfare, the cases of interest are those where the goals conflict. In particular, consider a tie that reduces consumers’ ability to chose the tying and tied products separately, but also creates some efficiency that is sufficiently passed on to consumers that it enhances consumer wel-

fare. First apparently believes that in such a case the consumer choice goal would be thwarted because consumers are “denied a choice they might prefer in the tied product market.”¹⁸⁴ His conclusion seems right if we define the consumer choice goal to be violated by anything that reduces the number of consumer choices. But why should we condemn a tie that would give consumers an alternative choice that makes them better off? There seems little reason to expand the number of consumer choices when that harms consumers. Alternatively, one might instead conclude that the consumer choice factor is ambiguous in such a case because, although the tie deprives consumers of the choice of picking the products separately, condemning the tie deprives consumers of a choice too—the ability to choose a tie that they would prefer to either of the separate choices. Because the ability to choose a tie makes consumers better off, we could say that allowing the tie furthers the consumer choice goal. But then we are really making decisions based on consumer welfare rather than on some freestanding notion of consumer choice. In short, either the consumer choice goal is undesirable (if defined in a way that allows it to conflict with consumer welfare) or subordinate (if defined to be consistent with consumer welfare).

4. Innovation Efficiency

Innovation efficiency is another goal that is important, but only as a means to the end of improving welfare. Nor does the goal offer much independent guidance when assessing ties because people have countervailing intuitions on whether ties advance or worsen innovation.¹⁸⁵ First objects to ties on the grounds that “innovation in the tied product market might be dampened or suppressed.”¹⁸⁶ Crane, Wright, and others favor ties that extract more than normal monopoly profits because they think that will increase incentives to innovate.¹⁸⁷ In the end, as discussed in Part I, economics favors an inverted-U approach, where we maximize innovation efficiency by allowing firms to reap all their normal monopoly profits from having created a market option that is preferable to other options, but do not allow firms to also use ties to extract the consumer surplus that consumers enjoy at normal monopoly prices. Thus, maximizing innovation efficiency is entirely consistent with prohibiting ties that reduce consumer welfare, and considering innovation efficiency separately does nothing to clarify the analysis.

INNOVATION EFFICIENCY IS ANOTHER GOAL THAT IS IMPORTANT, BUT ONLY AS A MEANS TO THE END OF IMPROVING WELFARE.

5. Preventing Firms From Becoming “Too Big to Fail”

As I understand it, the concern with firms becoming “too big to fail” is that their failure would create too many systemic problems in the economy, so the government must bail them out if they do fail. The prospect of these bailouts then gives these large firms incentives to engage in excessively risky transactions because

they externalize much of the downside costs onto taxpayers. This is a legitimate concern that, if valid, would justify some form of legal regulation.

But it seems to me the law should directly target the distorted incentive by regulating, taxing, or requiring insurance premiums for the excessively risky transactions engaged in by firms that are too big to fail. One solution would be to say that, when a firm reaches such a size, the implicit government insurance should be made explicit and an insurance premium should be charged that reflects the level of risk the firm incurs. That would deter inefficient risk taking, protect tax-

payers, and prevent economic dislocation because failure would result in a pre-defined insurance payment.

IT IS THUS PREFERABLE TO KEEP
ANTITRUST FOCUSED ON THE TASK
OF PROTECTING CONSUMER
WELFARE, AND LET OTHER
REGULATORY STRATEGIES PROTECT
TAXPAYERS AND DEAL DIRECTLY
WITH THE EXTERNALITIES CAUSED
BY IMPLICIT GOVERNMENT
INSURANCE FOR FIRMS THAT
ARE TOO BIG TO FAIL.

In contrast, simply blocking mergers that would produce firms that are too big to fail seems a poor remedy to the problem. Blocking such mergers would be of no help if firms grow to be too big to fail through internal expansion rather than through mergers. Nor would blocking such mergers be necessary if we had a regulation, tax, or mandatory insurance that directly addressed the risk-taking externality. Indeed, blocking such mergers seems affirmatively

undesirable if the merger would lower costs after considering any increased tax or premium costs, because then the merger would lower prices in a way that benefits consumer welfare without imposing any uncompensated externalities. Blocking such a merger would harm consumer welfare, but be unnecessary to protect taxpayers or prevent the inefficient distortion that prompts the concern. It is thus preferable to keep antitrust focused on the task of protecting consumer welfare, and let other regulatory strategies protect taxpayers and deal directly with the externalities caused by implicit government insurance for firms that are too big to fail.

In any event, it is hard to see how “too big to fail” concerns are likely to have much relevance in a tying case. Ties rarely have any bearing on whether a firm becomes too big to fail. So even if this were a valid independent goal for antitrust, it would have little impact on tying doctrine.

6. Other Distributive Concerns

Finally, First suggests that antitrust should consider “distributive concerns in more specific cases where business practices may have uncertain effects on the welfare of infra-marginal customers but substantial effects on customers who are priced out of the market.”¹⁸⁸ But if, as in the examples First cites, a restraint raises market prices in a way that prices out some consumers, then that does harm

consumer welfare. Such consumer harm could hardly be outweighed by ambiguous effects on infra-marginal customers.

Perhaps First has in mind the claim that, if conduct benefits some consumers and harms other consumers, antitrust should consider the income of the particular consumers at issue rather than decide cases based on the aggregate effect. But such an approach would be judicially inadministrable and, to my knowledge, no U.S. court has been willing to engage in it. Moreover, even if administrable, it would be theoretically flawed because taxation is a more efficient means of achieving redistribution than varying liability rules with the income of the affected parties. Although income taxation inefficiently discourages income creation, varying conduct liability with income not only discourages income creation to the same degree as income taxation, but also adds a discouragement of welfare-enhancing conduct.¹⁸⁹ In contrast, antitrust rules that protect consumer welfare do not create this distortion (even though they have favorable distributional effects) because such rules do not make conduct liability vary with party income and thus does not discourage income creation. Moreover, banning agreements that lessen overall consumer welfare: (1) is consistent with precedent, (2) is more administrable because it does not require consumer-by-consumer analysis, (3) helps coordinate global enforcement; (4) does not prevent efficiency-increasing conduct because compensating payments can be made; and (5) optimizes investment in innovation and improves *ex ante* total welfare.¹⁹⁰

B. DEFENDING THE EXCEPTION

Because the single monopoly profit theory does hold for “ties that involve a fixed ratio, no separate utility, *and* no substantial foreclosure share or effect,” I would recognize a rule of *per se* legality for such ties.¹⁹¹ First objects on several grounds, but with all respect I do not think any of his objections is persuasive.

One objection he raises is that such a tie might still harm consumer choice and innovation efficiency. Thus, he asks: “Why not stick with the presumption of illegality and shift the burden to the defendant to show an efficiency justification for refusing to sell the products unbundled?”¹⁹² The answer is simple. In the limited conditions when the single monopoly profit theory does hold, we know the firm could profit from imposing the tie only if it has some efficiency justification.¹⁹³ Thus, proving those conditions itself rebuts any presumption by showing there must be some efficiency justification. Such efficiencies will, to some extent, be passed on to consumers and the tie cannot otherwise harm consumers, so such ties should benefit consumer welfare. For reasons discussed above, this suffices to allow the

BECAUSE THE SINGLE MONOPOLY
PROFIT THEORY DOES HOLD FOR
“TIES THAT INVOLVE A FIXED
RATIO, NO SEPARATE UTILITY,
AND NO SUBSTANTIAL
FORECLOSURE SHARE OR EFFECT,”
I WOULD RECOGNIZE A RULE OF
PER SE LEGALITY FOR SUCH TIES.

tie even if it reduces notions of consumer choice that conflict with consumer welfare. Nor, as also discussed above, do we have any reason to think such a tie would reduce innovation efficiency. To the contrary, developing such a tie would itself be an efficient innovation.

Another objection he raises is that the firm might impose such a tie “to impede or deter entrants in the tied product market that might grow to challenge its monopoly position in the tying product market.”¹⁹⁴ That is a valid concern, but as I showed, that anticompetitive effect requires a substantial foreclosure share or effect in the tied market.¹⁹⁵ If a substantial foreclosure share or effect has not been shown, then this concern is invalid and we know efficiencies must motivate the tie; thus, my approach reaches the right result with *per se* legality. If a substantial

foreclosure share or effect has been shown, then my rule of *per se* legality would not apply. My exception to the *quasi-per se* rule for ties involving a fixed ratio and lack of separate utility would instead trigger a traditional rule of reason analysis, under which showing a substantial foreclosure share or effect would (as First desires) shift the burden to the defendant to show an efficiency justification.¹⁹⁶

Relatedly, First objects to my conclusion that the *Microsoft* case was right to recognize an exception to the *quasi-per se* rule because the tie there involved a fixed ratio and lack of separate utility.¹⁹⁷ He reasons that many customers didn’t want to use the browser at all and varied in how often they upgraded browsers and operating systems, so that their proportions were not truly

fixed.¹⁹⁸ But to defeat the possibility of power effects for products that lack separate utility, the products need only be “used or tied in fixed ratios,” so it suffices that “the ties ... involve a fixed ratio.”¹⁹⁹ Even if buyers might want to use the products in varying proportions, the fact that the tie bundles them in a fixed proportion suffices to mean that “buyers would experience any tied product price increase as an increase in the marginal price of buying the tying product.”²⁰⁰ In *Microsoft*, regardless of whether buyers might desire varying proportions, the challenge was to conduct that did bundle the operating system and browser in a fixed ratio, and (assuming the browser lacked separate utility) such a fixed bundle cannot have the power effects that justify the *quasi-per se* rule.²⁰¹ Instead, the real anticompetitive concern was, as First correctly recognizes, that a substantial foreclosure share or effect in the browser market could help preserve market power in the operating system market.²⁰² But focusing on that inquiry is precisely what is correctly achieved by recognizing the exception to the *quasi-per se* rule.

THE COMMENTS ALL AGREE WITH ME THAT TIES WITH MARKET POWER CAN REDUCE CONSUMER WELFARE AND TOTAL WELFARE EVEN WITHOUT A SUBSTANTIAL FORECLOSURE SHARE. THAT CONCLUSION IS ALL WE NEED TO REJECT NOT ONLY THE SINGLE MONOPOLY PROFIT THEORY BUT ALSO A CATEGORICAL RULE OF *PER SE* LEGALITY FOR EITHER ALL TIES OR ALL TIES WITHOUT A SUBSTANTIAL FORECLOSURE SHARE.

VI. CONCLUSION

The Comments all agree with me that ties with market power can reduce consumer welfare and total welfare even without a substantial foreclosure share. That conclusion is all we need to reject not only the single monopoly profit theory but also a categorical rule of *per se* legality for either all ties or all ties without a substantial foreclosure share. Because the critics of current doctrine advocate one of those categorical rules, this conclusion thus suffices to reject their legal position whether one thinks the proper standard is consumer welfare or total welfare.

In fact, the correct standard is consumer welfare as a matter of both law and policy. Consumer welfare should thus be the standard used when judging whether, under the current *quasi-per se* rule, a particular tie with market power has output-increasing efficiencies that offset any harmful anticompetitive effects. Allocating the burden of proof on those efficiencies to defendants remains supported by precedent, access to evidence, and the fact that theoretical considerations indicate that ties with market power will generally reduce consumer welfare. Even if we instead think that total welfare should be the standard, there is no good reason to fixate on *ex post* total welfare, and judging ties based on their consumer welfare effects is likely to correlate better to overall total welfare.

However, when the tie involves no substantial foreclosure share or effect and the bundled products lack separate utility and are used or tied in fixed proportions, then the tie cannot harm consumer or total welfare even with tying market power. While the old single monopoly profit theory is dead, a new baby single monopoly profit theory does apply to such ties, and thus they should be *per se* legal. ▼

1 Einer Elhauge, *Tying, Bundled Discounts, and the Death of the Single Monopoly Profit Theory*, 123 HARV. L. REV. 397 (2009).

2 Paul Seabright, *The Undead? A Comment on Professor Elhauge's Paper*, 5(2) COMPETITION POL'Y INT'L 243, 243-44 (2009).

3 *Id.* at 243, 246-47.

4 Elhauge, 123 HARV. L. REV. at 400.

5 Seabright, *supra* note 2, at 243.

6 Elhauge, 123 HARV. L. REV. at 400-01.

7 *Id.* at 402.

8 *Id.*

9 Seabright, *supra* note 2, at 243, 246-47.

10 *Id.* at 244-45.

11 Elhauge, 123 HARV. L. REV., at 425-26.

12 *Id.*

13 Seabright, *supra* note 2, at 244.

14 *Id.* at 244-45.

15 *Id.* at 250 n.8 (emphasis added) (quoting my article).

16 *Id.*

17 Elhauge, 123 HARV. L. REV. at 401, 427.

18 *Id.*

19 Seabright, *supra* note 2, at 243-47.

20 *Id.* at 245.

21 *Id.* at 245.

22 Elhauge, 123 HARV. L. REV. at 401. Seabright quotes to page 2 of my September 30th working paper rather than to my final article, see Seabright, *supra* note 2, at 250 n.2, but the passage he selectively quotes was identical in that working paper with the trivial difference that it said "their analogy" rather than saying "the critics' analogy."

23 Elhauge, 123 HARV. L. REV. at 401, 430, 435-36.

24 *Id.* at 435-42.

25 Seabright, *supra* note 2, at 248, 250 n.15.

26 *Id.* at 248-49.

27 X AREEDA, ELHAUGE, & HOVENKAMP, ANTITRUST LAW 175-279 (1996).

28 See Elhauge, 123 HARV. L. REV. at 405-07.

29 *Id.* at 406 (emphasis added).

30 Seabright, *supra* note 2, at 246-28.

31 See *Id.* at 247.

32 See Elhauge, 123 HARV. L. REV. at 409.

33 See Jose Carbajo et al., *A Strategic Motivation for Commodity Bundling*, 38 J. INDUS. ECON. 283, 284 (1990).

34 Seabright, *supra* note 2, at 247.

- 35 *Id.* at 247.
- 36 See U.C.C. §2-306; Clayton Act §3.
- 37 *Illinois Tool Works v. Independent Ink*, 547 U.S. 28, 32 (2006); *Eastman Kodak Co. v. Image Technical Servs., Inc.*, 504 U.S. 451, 458 (1992); *Jefferson Parish Hosp. Dist. No. 2 v. Hyde*, 466 U.S. 2 (1984); *Int'l Salt Co., Inc. v. United States*, 332 U.S. 392 (1947); *Int'l Bus. Machs. Corp. v. United States*, 298 U.S. 131 (1936); *United Shoe Mach. Co. v. United States*, 258 U.S. 451, 456 (1922); *Motion Picture Patents Co. v. Universal Film Mfg. Co.*, 243 U.S. 502 (1917).
- 38 *Illinois Tool*, 547 U.S. at 32 (buyers who used any of the seller's printheads had to "agree that they will purchase their ink exclusively from petitioners."); *Kodak*, 504 U.S. at 458 (seller refused to sell parts to buyers who bought any service from Kodak's rivals); *United Shoe*, 258 U.S. at 456 (lessees of shoe machinery had to agree to "purchase supplies exclusively from the lessor.")
- 39 *LePage's Inc. v. 3M*, 324 F.3d 141, 154 (3d Cir. 2003) (en banc).
- 40 *Masimo v. Tyco*, 2009 WL 3451725, at *1 (9th Cir. 2009) (emphasis added). I was an expert witness for Masimo in the liability trial, and have been an expert witness for both plaintiffs and defendants in other cases involving bundled or loyalty discounts.
- 41 See Elhauge, *The Exclusion of Competition for Hospital Sales Through Group Purchasing Organizations*, at 7 (2002) (report to the U.S. Senate on behalf of the Medical Device Manufacturer's Association).
- 42 Most contractual enforcement is not by law, but by reputational sanctions. See Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, 113 YALE L. J. 541, at 557 (2003).
- 43 In cases involving hospitals, such observation can be accomplished by sending salespersons to the hospitals or by having distributors and GPOs track the products each hospital purchases through them.
- 44 Seabright, *supra* note 2, at 247.
- 45 *Id.*
- 46 See Elhauge, 123 HARV. L. REV. at 402, 409, 443.
- 47 See *id.* at 405-07.
- 48 *Id.* at 402, 443.
- 49 See Elhauge, 123 HARV. L. REV. at 433-434, 479-481.
- 50 Professor Nalebuff states the percentage as 18.7% with rounding, but his proof indicates that the figure to two decimal places is 18.75%. See Barry Nalebuff, *Price Discrimination and Welfare*, 5(2) COMPETITION POL'Y INT'L 221, 227, 236 (2009).
- 51 Seabright, *supra* note 2, at 245.
- 52 See Elhauge, 123 HARV. L. REV. at 432-34.
- 53 See Nalebuff, *supra* note 50, at 232.

54 Seabright, *supra* note 2, at 246.

55 *Id.* at 250 n.9.

56 See Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. LEGAL STUD. 667 (1994).

57 Seabright, *supra* note 2, at 250 n.9.

58 This is true not only under U.S. antitrust law, see Elhauge, 123 HARV. L. REV. at 436-38, but also under EU competition law, see ELHAUGE & GERADIN, GLOBAL ANTITRUST LAW & ECONOMICS 69-70 (Foundation Press 2007).

59 See Elhauge, 123 HARV. L. REV. at 438.

60 *Id.*

61 *Id.*

62 Commissioner of Competition v. Superior Propane Inc., 2000 Canada Comp. Trib. 16 (April 4, 2002).

63 See Elhauge, 123 HARV. L. REV. at 439-432.

64 *Id.* at 441-42.

65 *Id.* at 479-81; Nalebuff, *supra* note, at 225, 227.

66 See Elhauge, 123 HARV. L. REV. at 481.

67 *Id.* at 434-35.

68 Seabright, *supra* note 2, at 246.

69 *Id.*

70 *Id.*

71 See Elhauge, 123 HARV. L. REV. at 442.

72 Seabright, *supra* note 2, at 246.

73 Elhauge, 123 HARV. L. REV. at 440.

74 Seabright, *supra* note 2, at 246.

75 Elhauge, *Defining Better Monopolization Standards*, 56 STAN. L. REV. 253, 298-300 (2003) (explaining the distinction between these two issues).

76 See Elhauge, 123 HARV. L. REV. at 440 (relying on literature finding such an invested U-shaped result).

77 *Id.*

78 *Id.*

- 79 *Id.* at 440-441.
- 80 Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMPETITION L. & ECON. 535, 543 (2008).
- 81 See *Id.* at 545; Elhauge, 123 HARV. L. REV. at 440-41.
- 82 Elhauge, 123 HARV. L. REV. at 434.
- 83 *Id.* at 412, 435.
- 84 *Id.* at 406-07, 434-35.
- 85 *Id.* at 433-34, 479-481.
- 86 *Id.* at 432-34, 479-81.
- 87 Daniel A. Crane & Joshua D. Wright, *Can Bundled Discounting Increase Consumer Prices Without Excluding Rivals?*, 5(2) COMPETITION POL'Y INT'L 209, 210 (Autumn 2009).
- 88 David S. Evans & Michael Salinger, *Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law*, 22 YALE J. REG. 37, 40-41, 52 (2005) (acknowledging that it relies only on "casual empiricism" and that "there has been essentially no empirical research into efficiencies from bundling and tying products together"); Kobayashi, *Does Economics Provide a Reliable Guide to Regulating Commodity Bundling by Firms? A Survey of the Economic Literature*, 1 J. COMPETITION L. & ECON. 707, 708 (2005).
- 89 Elhauge, 123 HARV. L. REV. at 462.
- 90 Evans & Salinger, *supra* note 88, at 38-41, 43-44, 65-84; Kobayashi, *supra* note 88, at 741-43.
- 91 See AREEDA, ELHAUGE, & HOVENKAMP, *supra* note 27, ¶ 1744-45.
- 92 Evans & Salinger, *supra* note 88, at 41-42, 52-65; Kobayashi, *supra* note 88, at 708, 742-43.
- 93 Kobayashi, *supra* note 88, at 744.
- 94 Crane & Wright, *supra* note 87, at 209-210.
- 95 *Id.*
- 96 Evans & Salinger, *supra* note 88, at 42, 44, 83-84, 85-86.
- 97 *Id.* at 83-84.
- 98 *Id.* at 42, 44, 86.
- 99 Crane & Wright, *supra* note 87, at 218 n.4.
- 100 Elhauge, 123 HARV. L. REV. at 439 n.112. Likewise, under EC law, it could not be an exclusionary abuse of dominance, but could be an exploitative abuse.
- 101 Crane & Wright, *supra* note 87, at 209.

- 102 *Id.* at 210, 212-13.
- 103 See Daniel A. Crane, *Mixed Bundling, Profit Sacrifice, and Consumer Welfare*, 55 EMORY L.J. 423, 461-462 (2006).
- 104 Elhauge, 123 HARV. L. REV. at 407-413.
- 105 *Id.* at 451-454.
- 106 See Barry Nalebuff, *Exclusionary Bundling*, 50 ANTITRUST BULL. 321, 326 (2005) ("Although the a la carte price of A is above the monopoly price, there is no loss to the firm, as it does not expect to make any sales at the inflated price. In equilibrium, all customers buy their B from the firm and thus are able to buy A at the profit-maximizing price of m.")
- 107 Crane & Wright, *supra* note 87, at 213-215.
- 108 Elhauge, 123 HARV. L. REV. at 412-13.
- 109 *Id.* at 455 (collecting literature).
- 110 *Id.* at 454.
- 111 Crane & Wright, *supra* note 87, at 215-17.
- 112 Elhauge, 123 HARV. L. REV. at 456 (summarizing literature); Einer Elhauge & Abraham L. Wickelgren, *Robust Exclusion Through Loyalty Discounts* (2010), available at <http://ssrn.com/abstract=1544008>; Elhauge, *Defining Better Monopolization Standards*, 56 STAN. L. REV. 253, 284-288 (2003).
- 113 Elhauge, 123 HARV. L. REV. at 456 (summarizing literature); Elhauge, *Defining Better*, *supra* note 112, at 288-292.
- 114 Elhauge, 123 HARV. L. REV. at 457; Elhauge, *Defining Better*, *supra* note 112, at 340.
- 115 Crane & Wright, *supra* note 87, at 217.
- 116 *Id.*
- 117 See Elhauge, *The Exclusion of Competition for Hospital Sales*, *supra* note 41, at 30.
- 118 Department of Justice and Federal Trade Commission Statements of Antitrust Enforcement Policy in Health Care Statement 7 (1996).
- 119 Crane & Wright, *supra* note 87, at 209, 217. In a footnote, Crane and Wright also assert that even below-cost bundled discounts will rarely be exclusionary because of Judge Easterbrook's argument that excluded rivals can always organize buyers to defeat it. *Id.* at 220 n.38. However, this argument fails for a number of reasons. First, for reasons that were discussed above, if the buyers are intermediaries who pass on most or all of the price increase onto downstream consumers, they can affirmatively profit from creating *supra*-competitive profits that they split with the seller, in which case the buyers would have little incentive to enter into such an agreement with the rival. Second, for reasons noted in the text following this footnote, if the loyalty agreement can raise market prices by reducing rival incentives to engage in price competition, then it can increase rival profits and eliminate any incentive for it to try to undo the loyalty agreements. Third, even when both rivals and buyers have the right incentives, the buyers will have collective action problems in joining a rival scheme to undo market exclusion because, while the buyers collectively benefit from that scheme,

individual buyers benefit even more if it occurs without their involvement. See Elhauge, *Why Above-Cost Price Cuts to Drive out Entrants Do Not Signal Predation or Even Market Power – and the Implications for Defining Costs*, 112 YALE L. J. 681, 760-61 (2003).

120 Elhauge, 123 HARV. L. REV. at 456-58, 461-464.

121 *Id.* at 414, 459-461, 463; Elhauge, *How Loyalty Discounts Can Perversely Discourage Discounting*, 5 J. COMPETITION LAW & ECON. 189 (2009); Elhauge & Wickelgren, *supra* note 112.

122 Crane & Wright, *supra* note 87, at 216-217.

123 Benjamin Klein & Kevin M. Murphy, *Exclusive Dealing Intensifies Competition for Distribution*, 75 ANTITRUST L.J. 433, 445-447 (2008).

124 *Id.* at 445-447.

125 Crane & Wright, *supra* note 87, at 216.

126 Klein & Murphy, *supra* note 123, at 447-448.

127 *Id.* at 448.

128 See Daniel Flores, *Exclusive Dealing Intensifies Competition for Distribution: Comment* at 4 (January 2010), available at <http://ssrn.com/abstract=1542695>. Murphy & Klein are also incorrect in asserting that in their model exclusive retail contracts make “all consumers on net better off.” Klein & Murphy, *supra* note 123, at 451-52. Instead, while all the consumers who preferred the brand that wins the exclusive contract will be better off, half the consumers who prefer the other brand will be better off and half will be worse off. Flores, *supra*, at 7-8.

129 Flores, *supra* note 128, at 5.

130 *Id.* at 8.

131 Elhauge, 123 HARV. L. REV. at 475-476.

132 *Id.* at 403, 451, 468, 478.

133 Nalebuff, *supra* note 50, at 221.

134 *Id.* at 221.

135 Elhauge, 123 HARV. L. REV. at 427 (emphasis added); see also *Id.* at 430, 432, 434 (repeating the point with the same caveat).

136 *Id.* at 433, 481.

137 *Id.* at 434.

138 *Id.* at 427 (“To the extent ties empirically have efficiencies that offset adverse power effects, the quasi-per se rule allows defendants to prove them. In contrast, eliminating the quasi-per se rule would make ties without substantial foreclosure shares per se legal, even when their adverse power effects exceed any efficiencies.”)

139 *Id.* at 427, 434-435.

- 140 *Id.* at 427, 433-39.
- 141 *Id.* at 427, 439-442.
- 142 Nalebuff, *supra* note 50, at 224.
- 143 *Id.* at 224.
- 144 *Id.* at 224. Throughout the body of his analysis, Nalebuff uses total welfare to refer to *ex post* total welfare. *Id.* at 239 n.2.
- 145 *Id.* at 224.
- 146 Elhauge, 123 HARV. L. REV. at 427, 430, 434.
- 147 *Id.* at 401, 426-427, 433-435.
- 148 Nalebuff, *supra* note 50, at 224.
- 149 *Id.* at 224-225.
- 150 Brief of Professors Barry Nalebuff, Ian Ayres, Lawrence Sullivan as Amici Curiae in Support of Respondent, *Illinois Tool Works v. Independent Ink*, 2005 WL 2427646, at *3, *19- 21.
- 151 Nalebuff, *supra* note 50, at 224.
- 152 *Id.* at 233.
- 153 Amicus Brief of Nalebuff, Ayres, & Sullivan, *supra* note 150, at 19.
- 154 Elhauge, 123 HARV. L. REV. at 440.
- 155 *Id.*
- 156 If all patent holders can engage in perfect price discrimination, then in theory the patent term could be shortened to provide the optimal fraction because consumers would enjoy some of the total surplus after the patent expires. But it is unrealistic to assume that all patent holders can perfectly price discriminate, and thus there is no reason to think patent terms have been set in this fashion. Further, market power often reflects other property rights that are not so term limited.
- 157 Nalebuff, *supra* note 50, at 222, 239 n.2.
- 158 *Id.* at 226.
- 159 Elhauge, 123 HARV. L. REV. at 433, 481.
- 160 Although Nalebuff states these percentage with rounding as negative 18.7%, positive 4.9%, positive 40%, and negative 2%, see Nalebuff, *supra* note 50, at 226-227, his proofs indicate that the figures to two decimal places are negative 18.75%, positive 4.88%, positive 39.66%, and negative 2.29%, *Id.* at 235-236.
- 161 *Id.* at 229.

- 162 Elhauge, HARV. L. REV. at 433, 481.
- 163 Nalebuff, *supra* note 50, at 222, 226-227.
- 164 *Id.* at 230.
- 165 *Id.* at 231-232.
- 166 *Id.* at 231-232, 237-239.
- 167 *Id.* at 232.
- 168 Elhauge, 123 HARV. L. REV. at 434.
- 169 Harry First, *No Single Monopoly Profit, No Single Policy Prescription?*, 5 COMPETITION POL'Y INT'L 199, 201-204 (2009).
- 170 *Id.* at 199, 204-06.
- 171 *Id.* at 202.
- 172 Elhauge, 123 HARV. L. REV. at 436-437.
- 173 First, *supra* note 169, at 201.
- 174 Elhauge, 123 HARV. L. REV. at 436 n.104; Elhauge, 56 STAN. L. REV. at 255, 260, 265-266.
- 175 Compare First, *supra* note 169, at 199, 205, with Gregory J. Werden, *Next Steps in the Evolution of Antitrust Law: What to Expect from the Roberts Court*, 5 J. COMPETITION L. & ECON. 49, 49 (2009).
- 176 First, *supra* note 169, at 201-202.
- 177 Elhauge, 123 HARV. L. REV. at 437 n.104.
- 178 *Id.* at 437 n.104.
- 179 First, *supra* note 169, at 202.
- 180 Elhauge, 123 HARV. L. REV. at 437 n.104.
- 181 Neil W. Averitt & Robert H. Lande, *Using The "Consumer Choice" Approach To Antitrust Law*, 74 ANTITRUST L. J. 175, 176 (2007).
- 182 *Id.* at 262.
- 183 John B. Kirkwood & Robert H. Lande, *The Fundamental Goal of Antitrust: Protecting Consumers, Not Increasing Efficiency*, 84 NOTRE DAME L. REV. 191 (2008).
- 184 First, *supra* note 169, at 205.
- 185 Innovation efficiencies also do not offer very useful guidance on whether to allow agreements or mergers that create market power because market power has mixed effects in that it: (1) decreases incentives to create innovations protected by intellectual property law but increases incentives to

create innovations that do not enjoy such protection, and (2) decreases incentives to create drastic innovations but increases incentives to create non-drastic innovations. Elhauge, 56 STAN. L. REV. at 298-299 & n.141; Elhauge, 112 YALE L. J. at 781 & n.266.

186 *Id.* at 205.

187 Crane & Wright, *supra* note 87, at 210; Dennis W. Carlton & Ken Heyer, *Extraction v. Extension: The Basis for Formulating Antitrust Policy Towards Single-Firm Conduct*, 4(2) COMPETITION POL'Y INT'L 285, 285, 290-92 (2008).

188 First, *supra* note 169, at 203.

189 See Kaplow & Shavell, *supra* note 56.

190 See *supra* Part I.

191 Elhauge, 123 HARV. L. REV. at 402.

192 First, *supra* note 169, at 205.

193 Elhauge, 123 HARV. L. REV. at 404

194 First, *supra* note 169, at 205.

195 Elhauge, 123 HARV. L. REV. at 417-419.

196 *Id.* at 402, 443, 469-470, 472.

197 *Id.* at 446-47.

198 First, *supra* note 169, at 205.

199 Elhauge, 123 HARV. L. REV. at 402, 409 (emphasis added); see also *Id.* at 416, 443 ("used or bundled in a fixed ratio").

200 *Id.* at 409; see also *id.* at 416.

201 *Id.* at 446.

202 First, *supra* note 169, at 205-206.

Clarifying the Scope of Judicial Review in Competition Inquiries: The Saga of PPI

Antonio Bavasso & Mark Friend
Allen & Overy

Clarifying the Scope of Judicial Review in Competition Inquiries: The Saga of PPI

*Antonio Bavasso and Mark Friend**

While early reports suggest that 2009 may well prove to be a good vintage for winemakers in Bordeaux, the same cannot unfortunately be said for the U.K. Competition Commission (“CC”). Indeed, 2009 is likely to be remembered as something of an *annus horribilis* for the CC, as the year in which it suffered an unprecedented succession of high-profile defeats before the Competition Appeal Tribunal (“CAT”).

The cases involved in these decisions raise several questions about the institutional balance between competition authorities and reviewing courts, all of which have far-reaching implications for the robustness of the entire system of competition enforcement. How far is it open to a court or tribunal to review a competition authority’s findings of fact, as opposed to the conclusions drawn by the authority from findings and judgments made in the light of those conclusions? Where is the line to be drawn between judicial review of the decision-making process and a full appeal on the merits? Has the CAT struck the right balance between allowing the CC, as a specialist competition authority, to exercise its judgment and intervene in markets where adverse effects on competition have been identified, while at the same time holding the CC to account? Is it realistic to require a competition authority such as the CC to conduct a detailed cost/benefit analysis of each element of its remedy proposals in the context of what is necessarily an imprecise and forward-looking exercise, involving qualitative judgments?

* Antonio Bavasso is Partner, Allen & Overy LLP; Visiting Professor at University College London; Co-founder and Executive Director of the Jevons Institute for Competition Law and Economics at University College London, antonio.bavasso@allenoverly.com. Mark Friend is Partner, Allen & Overy LLP. mark.friend@allenoverly.com.

I. Introduction

The *annus horribilis* year of 2009 all began with Tesco's challenge to the CC's remedy proposals in the groceries inquiry, relating to the introduction of a "competition test" into the planning regime.¹ The CAT agreed with Tesco that the CC had failed to take into account the economic costs associated with its remedy proposal and, specifically, that it had failed to assess the risk that the imposition of a competition test might produce adverse effects for consumers, resulting in unmet demand for grocery retailing. Things took a turn for the worse when the CC's remedy proposals in the payment protection insurance ("PPI") inquiry, relating to the imposition of a point of sale prohibition ("POSP"), were quashed following a successful challenge by Barclays, relying on similar arguments to those upheld in the *Tesco* case.² Essentially the CAT found that the CC had failed to take adequate account of the risk of a radical remedy producing adverse consequences to consumers due to reduced take-up of PPI. The next major challenge to the CC concerned an allegation of apparent bias on the part of one member of the CC panel investigating the supply of airport services by BAA; this led ultimately to substantial portions of the CC's report being quashed.³ The last case in the series was heard shortly before the BAA judgment was handed down: here, the CC suffered a more modest defeat at the hands of Sports Direct over the CC's refusal to provide Sports Direct, as a main party in a merger inquiry, with un-redacted versions of the CC's working papers. On the preliminary issue of whether the application by Sports Direct was premature, the CAT ruled that it was not, following which the CC withdrew its decision.⁴

While the BAA case will undoubtedly prove to be of major importance for the way in which the CC selects panel members in future inquiries, and while the *Sports Direct* case may have wider implications for other inquiries in which parties seek to challenge provisional decisions by the CC, both of these cases involved discrete and self-contained legal issues. By contrast, the *Tesco* and *Barclays* cases go to the heart of the CC's powers to impose or recommend remedial action, raising fundamental questions about the nature of the CAT's role in reviewing such decisions and, more generally, about the intensity of judicial review of decisions of competition authorities.

These questions about the institutional balance between competition authorities and reviewing courts have far-reaching implications for the robustness of the entire system of competition enforcement. How far is it open to a court or tribunal to review a competition authority's findings of fact, as opposed to the conclusions drawn by the authority from those findings and judgments made in the light of those conclusions? Where is the line to be drawn between judicial review of the decision-making process and a full appeal on the merits? Has the CAT struck the right balance between allowing the CC, as a specialist competition authority, to exercise its judgment and intervene in markets where adverse effects on competition have been identified, while at the same time holding the CC to account? Is it realistic to require a competition authority such as the CC

to conduct a detailed cost/benefit analysis of each element of its remedy proposals in the context of what is necessarily an imprecise and forward-looking exercise, involving qualitative judgments?

Of course the answer may vary from jurisdiction to jurisdiction. In this article we will focus on the judicial review principles applicable to the CAT in the United Kingdom with particular reference to the CAT's judgment in *Barclays*. Some of those principles are interesting as a point of reference for other jurisdictions

THESE QUESTIONS ABOUT
THE INSTITUTIONAL BALANCE
BETWEEN COMPETITION
AUTHORITIES AND REVIEWING
COURTS HAVE FAR-REACHING
IMPLICATIONS FOR THE
ROBUSTNESS OF THE ENTIRE
SYSTEM OF COMPETITION
ENFORCEMENT.

and indeed other areas of judicial review within the United Kingdom. The approach of the CAT (and, more generally, of the English Courts) in relation to competition enforcement is particularly interesting because this is one of the areas that is more naturally influenced by principles developed by the European Courts.

We will start by examining some of the general principles of judicial review and considering whether these principles allow for gradations of judicial scrutiny, depending on the nature of the decision being reviewed. We will

then discuss the background to the PPI inquiry, the basis for Barclays' appeal, and the grounds on which the CAT reached its decision. Finally, we will discuss the wider implications of the *Barclays* judgment, viewed against the background of the case law on the intensity of judicial review.

II. Statutory Background and General Principles of Judicial Review in the United Kingdom

In the United Kingdom, the key provisions of the Enterprise Act 2002 ("EA02") that set out the basis for review of the CC are sections 179 and 120.

Section 179 EA02 states:

"(1) Any person aggrieved by a decision of the OFT, the appropriate Minister, the Secretary of State or the Commission in connection with a reference or possible reference under this Part may apply to the Competition Appeal Tribunal for a review of that decision."

Then, crucially:

“(4) In determining such an application the Competition Appeal Tribunal shall apply the same principles as would be applied by a court on an application for judicial review.”

In relation to mergers section 120 EA02 follows an identical approach.⁵

It is well recognized that, since the United Kingdom’s accession to the European Communities (now the EU), principles of European law—developed in particular by the European Court of Justice (“ECJ”)—have had an increasing influence over national courts.⁶ This influence is most obvious in areas where U.K. legislation mirrors European provisions (e.g. the Human Rights Act 1998 and European Convention of Human Rights) or where—under U.K. legislation—courts are required to deal with questions of interpretation of national law consistently with EU (formerly Community) law (e.g. section 60 of the Competition Act 1998).

However, even when such a direct statutory link does not arise (as is the case under the EA02), principles developed under EU law (notably the principle of proportionality), increasingly find their way into competition enforcement policy and, as a result, into the jurisprudence of courts and tribunals reviewing such enforcement decisions.

Proportionality in relation to remedial action is a prime example of this. The EA02 requires the CC to “have regard to the need to achieve as comprehensive a solution as is reasonable and practicable to” either (a) the adverse effect on competition and any detrimental effects on customers so far as resulting from it (sections 134(6) and 138(2)(b) in the context of a market investigation), or (b) the substantial lessening of competition and any adverse resulting from it (sections 35(4) and 36(3) in the context of a merger investigation). While the EA02 only refers to “reasonable and practicable” the Guidelines adopted by the CC (pursuant to section 171(3) EA02 for market investigations and section 106(1) EA02 for mergers) refer to a proportionality standard.⁷ How does the proportionality standard in remedial action fit with general principles of judicial review?

In the United Kingdom the traditional grounds of judicial review can be broadly categorized as follows: legality, fairness, and reasonableness. Indeed, Lord Cooke of Thorndon has said that principles of judicial review could be summarized in three adverbs, namely, that a public body has to act lawfully, reasonably, and fairly.⁸ In this article we will focus principally on reasonableness.

These judicial review principles are not static but have evolved over time, influenced in part by principles derived from the case law of the European Courts. The importation of the principle of proportionality is an example.

Proportionality as a test for judicial review in English law was first suggested by Lord Diplock in 1985 in the *GCHQ* case.⁹ It is based on a principle that is widely known in European legal systems and is, in fact, now regarded as a “general principle of law” by the ECJ. The concepts of proportionality and reasonableness are closely linked,¹⁰ because the proportionality test, which seeks to prevent unduly oppressive decision-making, essentially requires the decision maker to

achieve a fair balance of relevant considerations, and thus bears similarities to (and arguably can fit within) the concept of reasonableness in English administrative law.

INDEED, LORD COOKE OF
THORNDON HAS SAID THAT
PRINCIPLES OF JUDICIAL REVIEW
COULD BE SUMMARIZED IN THREE
ADVERBS, NAMELY, THAT A PUBLIC
BODY HAS TO ACT LAWFULLY,
REASONABLY, AND FAIRLY.

Decisions by competition authorities, particularly decisions to impose remedies in the context of mergers and market investigations, often involve complex economic assessments relying on a mixture of economic theory, forensic examination of factual evidence, and qualitative

forward-looking assessments based on judgment and experience. When reviewing such decisions, the courts will assess the adequacy of the evidence relied upon to prove the competitive harm; regulatory intervention by a competition authority without a proper evidential basis will, in principle, be unlawful. Under traditional judicial review principles, inadequacy of evidence has to reach the standard of unreasonableness. However, it is recognized that the “no evidence” standard of judicial review “does not mean a total dearth of evidence. It extends to any case where the evidence, taken as a whole, is not reasonably capable of supporting that finding; or where, in other words, no tribunal could reasonably reach that conclusion on that evidence.”¹¹ The CAT has also recognized that in the “no evidence principle” is the principle that perverse or unreasonable action is unauthorized or *ultra vires*.¹² In this sense, a rationality review approaches the legality standard.¹³

How far do these general principles of judicial review allow for different levels of scrutiny in their application? Following that question, what is the appropriate level of scrutiny for decisions by competition authorities?

III. Different Levels of Scrutiny?

There can be little doubt that general principles of judicial review permit different levels of scrutiny in their application. Indeed, one of the leading U.K. textbooks of judicial review (De Smith)¹⁴ graphically depicts the different categories of review, ranging (in decreasing order) from “full intensity” (correctness review for abuse of power), to “structured proportionality” (where the burden of justification is on the decision-maker), “anxious scrutiny” (where the burden is again on the decision-maker), “standard *Wednesbury* unreasonableness” (where the burden is on the claimant), “light touch unreasonableness,” (again, where the

burden is on the claimant) and “non-justiciability” (where the court will also require an adequate justification).

How far, if at all, does the specialist nature of the CAT enter into the equation? Formally it appears that it does not. As the Court of Appeal put it in *Sky*, a court will apply its own specialized knowledge and experience, which enables it to “perform its task with a better understanding, and more efficiently.”¹⁵ However, the Court of Appeal also made clear that the possession of that knowledge and experience does not, in any way, alter the nature of the task at hand. The task of the CAT is to apply normal principles of judicial review, although as the Court of Appeal also noted, following its earlier decision in *IBA Health*, the CAT will have a better understanding of the issues at stake than a non-specialist court or tribunal.¹⁶

Related to this is the question of how the threshold for judicial intervention is affected by the specialist nature of the decision maker (i.e. the competition authorities whose decisions are under review). As the editors of De Smith put it: “[t]he threshold of intervention is particularly influenced by the respective institutional competence of the decision-maker and the court.”¹⁷ The greater the degree of specialism and “institutional” competence of the decision-maker, the greater the need for self-restraint on the part of the reviewing court or tribunal. The logical corollary of self-restraint where an adjudicative body lacks institutional capacity is that, where that body enjoys enhanced institutional capacity, it should exercise more intensive scrutiny. The CAT itself acknowledged that this may result in its being “a more demanding and/or less deferential tribunal than might otherwise be the case where a court is called upon to review a decision of a specialist regulator.”¹⁸ The ordinary principles of judicial review give the CAT sufficient latitude to do so.

THE GREATER THE DEGREE
OF SPECIALISM AND
“INSTITUTIONAL” COMPETENCE
OF THE DECISION-MAKER, THE
GREATER THE NEED FOR SELF-
RESTRAINT ON THE PART OF THE
REVIEWING COURT OR TRIBUNAL.

Leaving aside the CAT’s institutional competence we suggest that there are three other reasons that may call for enhanced scrutiny, all of which are equally applicable to non-specialist courts and, indeed, should influence the standards applied by the primary decision-maker, the competition authority.

The first reason relates to the nature of the competition authority’s intervention and, thus, the nature of the rights affected by the decision under review. In the most intrusive forms of intervention, which occur more frequently in the context of merger control (particularly in merger control regimes such as the United Kingdom which allow mergers to be completed without prior approval from the competition authority), but are also relevant in market investigations (the BAA example being a case in point), a divestment remedy engages a fundamental property right and the principles of Article 1 of Protocol 1 of the

European Convention on Human Rights. What is the appropriate standard of review in such cases? The case law in this area suggests a number of different answers to this question, which may to some extent depend on the nature of the interference with fundamental rights.¹⁹ The classic pre-Human Rights Act 1998 formulation of Sir Thomas Bingham MR in *Ex p. Smith*²⁰ (“the more substantial the interference with human rights, the more the court will require by justification before it is satisfied that the decision is reasonable”) can be regarded as an example of “anxious scrutiny.”²¹ More recent cases suggest a higher intensity of review, while continuing to draw a distinction between a proportionality review and a review on the merits: see the speech of Lord Bingham of Cornhill in the *Denbigh High School* case:²²

“... There is no shift to a merits review, but the intensity of review is greater than was previously appropriate, and greater even than the heightened scrutiny test adopted by the Court of Appeal in *R v Ministry of Defence, Ex p Smith* [1996] QB 517, 554.”

However, the distinction is not always clear-cut. As Rose points out, in the *Denbigh* case the court emphasized that it was concerned not with the decision-making process, but with the correctness of the decision (which implies that the reviewing court is substituting its own view for that of the decision-maker); while in cases involving the detention of mentally-ill patients, the case law indicates that the courts will conduct a full merits-based review.²³

The second reason relates to the nature of the theory relied upon by the competition authority. This can be conveniently referred to as the *Tetra Laval* doctrine, derived from ECJ case law.²⁴ In that case, the Commission appealed against a judgment of the CFI annulling its decision to prohibit a merger based on concerns about conglomerate effects. The ECJ referred to the need for a careful prospective analysis of the likely consequences of the merger, given that it would entail: “a prediction of events which are more or less likely to occur in future if a decision prohibiting the planned concentration or laying down the conditions for it is not adopted.” The ECJ noted that, in a conglomerate merger, the period in the future which would need to be examined is lengthy, and that the chains of cause and effect are “dimly discernible, uncertain and difficult to establish.” Therefore, the quality of the evidence relied on by the Commission to establish grounds for a prohibition is “particularly important” since it needs to support the conclusion that, in the absence of a prohibition, the economic development envisaged would be plausible. Although this case was concerned with the examination of mergers under the EC Merger Regulation, its relevance in U.K. pro-

ceedings under the EA02 was accepted by the Court of Appeal in *Sky*²⁵ (albeit found to be inapplicable in the circumstances of that case).

In essence, under the *Tetra Laval* doctrine, without departing from the balance of probability standard the expectation is that the decision-maker must exercise particular care; the standard of evidence required will be particularly high in relation to matters that are inherently less likely to give rise to competitive harm. This can be compared with the well-known Lord Hoffmann caveat about the standard required to prove a sighting of a lioness (rather than an Alsatian) in Regent's Park.²⁶

THE STANDARD OF EVIDENCE
REQUIRED WILL BE PARTICULARLY
HIGH IN RELATION TO MATTERS
THAT ARE INHERENTLY LESS
LIKELY TO GIVE RISE TO
COMPETITIVE HARM.

The third reason, which is closely related to both the first and second points, relates to the importance or the gravity of the issue reviewed in the context of the general context of the task entrusted to the competition authority, and applies when the remedy chosen by the competition authority is particularly intrusive, uncertain in its effects, or wide-ranging. This approach is not new but is now becoming known as “double proportionality.” It is a principle derived originally from ECJ jurisprudence in the *Fedesa* case,²⁷ where the ECJ described it in the following terms:

“... the lawfulness of the prohibition of an economic activity is subject to the condition that the prohibitory measures are appropriate and necessary in order to achieve the objectives legitimately pursued by the legislation in question; where there is a choice between several appropriate measures recourse must be had to the least onerous, and the disadvantages caused must not be disproportionate to the aims pursued.”

This principle has recently been applied in two successful challenges to the CC remedy proposals in *Tesco* and in *Barclays*. The principle was described by the CAT in *Tesco* in the following terms:²⁸

“... the application of these principles is not an exact science: many questions of judgment and appraisal are likely to arise at each stage of the Commission's consideration of these matters. This is most obviously the case when it comes to the balancing exercise between the (achievable) aims of the proposed measure on the one side, and any adverse effects it may produce on

the other side. In resolving these questions the Commission clearly has a wide margin of appreciation, with the exercise of which a court will be very slow to interfere in an application for judicial review. The margin of appreciation extends to the methodology which the Commission decides to use in order to investigate and estimate the various factors which fall to be considered in a proportionality analysis... The Commission can tailor its investigation of any specific factor to the circumstances of the case and follow such procedures as it considers appropriate. In this regard, it may well be sensible for the Commission to apply a 'double proportionality' approach: for example, the more important a particular factor seems likely to be in the overall proportionality assessment, or the more intrusive, uncertain in its effect, or wide-ranging a proposed remedy is likely to prove, the more detailed or deeper the investigation of the factor in question may need to be.”

Explaining the application of this test in *Barclays*, the CAT commented:²⁹

“double proportionality [is] simply a convenient label for the common sense proposition that, within a wide margin of appreciation, the depth and sophistication of analysis called for in relation to any particular relevant aspect of the inquiry needs to be tailored to the importance or gravity of the issue within the general context of the Commission’s task.”

We now turn to consider the *Barclays* judgment in more detail to see how these principles were approached by the CAT.

IV. PPI—The Background

The PPI investigation began life as an OFT market study, conducted using the OFT’s broad powers under section 5 EA02. Having identified competition concerns the OFT made a market investigation reference to the CC on February 7, 2007, pursuant to sections 131 and 133 EA02. PPI is a type of insurance that is intended to protect the borrower against certain defined risks (accident, illness, unemployment, etc.) that might otherwise prevent the borrower from being able to repay his or her loan. The main categories of PPI are personal loan PPI (“PLPPI”), mortgage PPI (“MPPI”), second mortgage PPI (“SMPPI”), credit card PPI (“CCPPI”), retail PPI, and motor finance PPI. However, 90 percent of all PPI sales in 2007 consisted of PLPPI, MPPI, and CCPPI.³⁰

The CC's report was published on January 29, 2009 and concluded that there were features of the PPI market which prevented, restricted, or distorted competition, resulting in an adverse effect on competition ("AEC") for the purposes of section 134(2) EA02. The report concluded:

- 1) Suppliers of PPI were found to face little competition when selling PPI in combination with the underlying credit product.
- 2) Consumers were found to face higher prices and less choice than would be the case in a competitive market.

Accordingly, providers of PPI were earning substantial excess profits, albeit the CC noted that some of these profits were used to cross-subsidize underlying credit prices.³¹

PROVIDERS OF PPI WERE EARNING SUBSTANTIAL EXCESS PROFITS, ALBEIT THE CC NOTED THAT SOME OF THESE PROFITS WERE USED TO CROSS-SUBSIDIZE UNDERLYING CREDIT PRICES.

Deficiencies in the competitive process for selling PPI were also identified. These included:

- 1) a failure by distributors and intermediaries actively to compete for customers on price or quality;
- 2) barriers to searching for consumers who wished to compare PPI policies, whether or not combined with credit; and
- 3) barriers to switching resulting in part from the excessive costs of switching out of single premium policies.

At the heart of the CC's concerns was the point of sale advantage ("POSA"), i.e. that PPI was generally sold by lenders at the point of sale of the underlying credit product, which was said further to restrict the extent to which other providers could compete effectively.³²

The CC's package of remedies included a series of measures designed to improve the level of information provided to consumers in order to facilitate searching, along with a prohibition on selling single premium policies in order to facilitate switching. But, most controversially, the package also included a POSP—in other words, a prohibition on distributors and intermediaries from selling PPI to their credit customers within seven days of a credit sale, unless the customer had proactively returned to the seller at least 24 hours after the credit sale. This was designed to address at least some of the incumbency advantages enjoyed by distributors selling PPI at the point of sale of the underlying credit product.

Explaining its decision to introduce the POSP, the CC analyzed a series of potential risks associated with this remedy that had been flagged by various providers during the course of the inquiry. The first was that the POSP would not be fully effective in that it would fail to remove all aspects of this incumbency advantage:³³

“We agree that this remedy will not entirely remove all aspects of the incumbency advantage enjoyed by distributors. However, we do not think that we need to remove all incumbency advantages of distributors in order effectively to remedy this aspect of the AEC.

[...]

We acknowledge that—as with any intervention aimed at enhancing competition—there is a risk that this element of the remedies package will not generate the changes in behaviour necessary fully to address the AEC.”

The second risk considered by the CC was that the POSP would lead to a reduced take-up of PPI, due to the inconvenience of no longer being able to buy PPI at the point of sale when taking out a loan. The CC’s view was that this would be partially or fully off-set by reduced PPI prices which would result from increased competition, and that the ability of consumers to initiate PPI purchases by telephone or internet communication 24 hours after the credit sale would significantly reduce this risk.³⁴

The CC also considered the risk of reduced consumer choice, but concluded that its remedies package would actually stimulate competition, increasing consumer choice.³⁵ On the other hand, the CC

THE CC ALSO CONSIDERED
THE RISK OF REDUCED
CONSUMER CHOICE, BUT
CONCLUDED THAT ITS REMEDIES
PACKAGE WOULD ACTUALLY
STIMULATE COMPETITION,
INCREASING CONSUMER CHOICE.

accepted that the POSP would lead to additional costs for distributors, and factored those costs into its assessment of whether the POSP was a proportionate remedy. It concluded that the POSP was a necessary part of its overall remedies package, which would lead to a “new, more competitive, market structure.”³⁶

It will be recalled that the CC has a discretion, when considering the need for remedial action to address an AEC in a market investigation reference, to take account of “relevant customer benefits” that would be jeopardized by the imposition of the remedy. Relevant customer benefits are defined by section 134(8) EA02 as benefits to existing or future customers in the form of:

- 1) lower prices, higher quality or greater choice of goods or services in any market in the United Kingdom (whether or not the market to which the feature or features concerned relate); or
- 2) greater innovation in relation to such goods or services.

During the CC inquiry it was argued by a number of PPI providers that the current market structure resulted in PPI prices cross-subsidizing lower credit prices,

and that this would be jeopardized by the introduction of the POSP. The CC accepted that lower credit prices were a direct result of the features of PPI that led to an AEC in the markets for PPI (in other words, that lower credit prices qualified as a relevant customer benefit),³⁷ but concluded that it would not be appropriate to modify its remedies package on that account. In the CC's view, the benefits of intervention would outweigh the benefit of lower credit prices.³⁸ The CC did not feel able to quantify all aspects of the benefits of intervention, but assessed one element as being in excess of £200 million. Against this, the CC assessed the costs of remedial action as involving one-off set-up costs of £100 million and ongoing annual costs of £50-60 million. Summarizing its assessment of the likely impact of its remedies package, the CC noted:³⁹

“We decided that the package of remedies we have set out will provide a comprehensive, reasonable and practicable solution to the AEC that we have identified in a timely manner.

As with any set of competition-enhancing remedies, we cannot predict exactly how the market will develop. However, we concluded that our remedies will remove barriers for searching and switching and lead to a larger stand-alone market whilst still enabling distributors to offer combinations of credit and PPI and to compete on the terms of the combination as well as of its component parts. We considered that the package of remedies will lead to more active competition for PPI consumers: through more active marketing before the credit sale; in response to increased consumer search just after the credit point of sale; and by encouraging the switching during the life of the credit product. This competition will manifest itself through more PPI advertising and lower prices.

[...]

We decided that the remedies set out in this decision document represent as comprehensive a solution to the AEC and resultant consumer detriment that we have identified as is reasonable and practicable, and that this package should not be modified to take account of credit prices being lower than they otherwise might be.”

V. The Basis for Barclays' Challenge

The summary of Barclays' notice of application, as it appeared on the CAT's web site,⁴⁰ identified four grounds of appeal, which to some extent overlapped. First, it was argued that the CC had failed to take into account considerations that were relevant to the proportionality of the POSP, in particular the benefits that

would arise from the remedies package and the incremental benefits from including the POSP in the package. Second, it was argued that there was no proper evidential basis for the CC's conclusion that the POSP was justified, and that the CC had incorrectly concluded that the POSP was a more reasonable, effective, and proportionate remedy than Barclays' own proposal involving informational remedies and an increased cooling-off period. Third, it was argued that the CC had failed to take into account relevant considerations (or, alternatively, had taken into account irrelevant ones) in its analysis of the consumer detriment arising from the AEC and whether the benefits of its intervention would outweigh the loss of relevant consumer benefits. Fourth, it was argued that the CC had failed to take into account relevant considerations in its analysis of the relevant market(s) and the extent of competition problems. In particular, it was said that the CC's market definition was too narrow, that the CC should have updated its earlier findings in the report in light of new information, and that it had failed to take into account changes in the market since its financial analysis was based on data only through the end of 2006.

VI. The CAT's Approach

Grounds 1-3 of Barclays' application were all essentially aimed at challenging the decision to impose the POSP. Ground 4, however, was a separate challenge to the CC's findings on market definition. As the CAT pointed out, this challenge, if well-founded, would have undermined the whole of the CC's findings on the

AEC, and the CAT therefore began its analysis with a consideration of these arguments.⁴¹

ALTHOUGH PRESENTED UNDER THE
GUISE OF THE TRADITIONAL
HEADINGS OF JUDICIAL REVIEW,
NAMELY A FAILURE TO TAKE
ACCOUNT OF RELEVANT
CONSIDERATIONS, IT SEEMS CLEAR
THAT ASPECTS OF BARCLAYS'
CASE WERE, IN SUBSTANCE,
A CHALLENGE ON THE MERITS.

A. GROUND 4

Although presented under the guise of the traditional headings of judicial review, namely a failure to take account of relevant considerations, it seems clear that aspects of Barclays' case were, in substance, a challenge on the merits, as the CAT pointed out at various points in its judgment.⁴² First, however, the CAT rejected Barclays' argument that the CC should have

adopted a broader market definition, as it had done in its 2003 report on extended warranties on domestic electrical goods. The mere fact that the CC had on different (albeit, loosely analogous) facts reached a different view in another investigation was not relevant to the lawfulness of its analysis in this particular case.⁴³

Arguments that the CC had failed to take into account more recent developments in the market were also rejected. The CAT found that the CC had, in fact, taken into account reduced profitability, falling penetration rates, and increased claims rates (and that its reasons for not taking into account substantial increas-

es in late 2008 were adequately explained and were well within the CC's margin of appreciation).⁴⁴ As to the overall decline in the PPI market in 2007/2008, the CAT considered that the CC had treated this as a consequence of the AEC; in the CAT's view, Barclays' complaint that the CC had taken no account of the decline was in substance a merits complaint.⁴⁵

“In other words, Barclays simply disagree with the Commission's conclusions as to the reasons for that decline. But that, in the absence of irrationality (and none is alleged in this respect), was a matter for the Commission to decide having, as we find, properly considered the evidence.”

Likewise, arguments that the CC had failed to take into account regulatory changes were also rejected, either on the basis that the CC had, in fact considered them, or on the basis that they were not relevant to the CC's assessment in the first place.⁴⁶

“It is we think a non-sequitur to suppose that the omission of a decision-making body to mention something which it clearly knew about as being irrelevant to its analysis means that the possible relevance of it went unconsidered.”

Also were rejected were the arguments about the CC's failure to carry out a proper analysis of market definition, including an incorrect application of the SSNIP test, the “cellophane fallacy,” and the CC's alleged failure to take account of evidence suggestive of a wider market definition. In concluding on Ground 4, the CAT noted that even if it had been persuaded that one or more of the arguments were well-founded for judicial review purposes, it would have been reluctant to conclude that they were relevant to the CC's findings as to the AEC.⁴⁷

“First, we were impressed by the breadth of analysis and verification underlying the Commission's market definition, and by the number of separate conclusions which all pointed to the same outcome. Secondly, we were equally impressed by the evident determination of the Commission not to be

enslaved by any particular market definition, but rather to assess the competition problems arising in the sale of PPI on an empirical rather than overly theoretical basis which, while no doubt influenced by market definition, was by no means controlled or dominated by it.”

B. GROUND 1

Barclays’ challenge under Ground 1—based essentially on a failure to take account of considerations relevant to a proportionality assessment—relied on a number of separate elements. First, it was claimed that the CC had failed to ana-

THIS FOCUS ON THE NEED TO
BALANCE THE EXTENT OF THE
CONSUMER DETRIMENT RESULTING
FROM THE AEC AGAINST THE
EXPECTED BENEFITS OF
THE REMEDIES PACKAGE HAS ITS
ROOTS IN THE *TESCO* JUDGMENT.

lyze or identify the extent of the benefits that would accrue from its remedies package; second, that the CC had failed to give consideration to the inevitable time lag before its remedies package would take effect; and third, that the CC had failed to assess the incremental benefit of adding the POSP to the remainder of the remedies package. This focus on the need to balance the extent of the consumer detriment resulting from the AEC against the expected

benefits of the remedies package has its roots in the *Tesco* judgment, where the CAT had expressed itself in the following terms:⁴⁸

“... it is necessary to know what the measure is expected to be able to achieve in terms of an aim, before one can sensibly assess whether that aim is proportionate to any adverse effects of the measure. The proportionality of a measure cannot be assessed by reference to an aim which the measure is not able to achieve.”

Indeed, Barclays went further than this and argued that the CC had repeated exactly the same mistake as in the *Tesco* case, by simply comparing the detriment associated with the AEC against the cost of implementing the remedies package. The CC’s rebuttal was that it had concluded that its remedies package would be fully effective to remedy the AEC, such that the consumer benefit from the remedies package was equivalent to the consumer detriment associated with the AEC. The issue was that, according to Barclays and the interveners (Lloyds and Shop Direct), the CC had nowhere in its report said anything about the degree of effectiveness of the remedies package. The issue therefore boiled down to a question

of how to interpret the CC's report. In answering this question, the CAT followed the approach in *Tesco*, accepted by all parties, that a CC market investigation report should be interpreted on the basis of "a fair and generous reading of the Report as a whole," rather than "word by word as a statute might be."⁴⁹

The debate nonetheless centered on whether the CC's analysis was predicated on the remedies being fully effective, or less than 100 percent effective. While the CC had recognized the risk that its remedies package might not be fully effective, this was not, in the CAT's view, inconsistent with a judgment that it probably would be fully effective.⁵⁰ Just because the remedies package would not completely remove the point of sale advantage this did not mean that the PPI market would not be a properly functioning market (in contrast to an ideal market, with every competitor on a completely level playing field).⁵¹ Thus, although on a dictionary definition, the CC's references to the word "effective" did not equate to "fully effective," viewed in the round, it was clear that the CC believed its remedies package would be "fully, or rather substantially, effective."⁵²

THE CAT WAS, HOWEVER, MORE SYMPATHETIC TO THE ARGUMENT THAT THE CC HAD FAILED TO INDICATE A TIMESCALE IN WHICH IT EXPECTED ITS REMEDIES TO TAKE EFFECT.

Having rejected these arguments, the CAT was, however, more sympathetic to the argument that the CC had failed to indicate a timescale in which it expected its remedies to take effect. The point was put pithily by Barclays that the CC had identified start-up implementation costs of £100 million and ongoing annual costs of £50-60 million, yet had said nothing about the timescale for the remedies to take effect save that they would do so in a "timely" manner. This was a problem, because this did not describe any measurement of time in an objective sense,⁵³ suggesting that the CC could not logically have carried out a systematic proportionality analysis. Curiously, the CAT then noted that it would not have regarded this failing on its own as sufficient to justify quashing the decision to impose the POSP, because it appeared that the CC had considered the issue, but had simply not spelled it out in the report.⁵⁴

The CAT was distinctly unsympathetic to Barclays' attempts to challenge the CC's proportionality analysis by reason of its failure to address the incremental impact, in terms of both costs and benefits, of adding the POSP to the remaining package of remedies. Indeed, had the CAT decided otherwise, it would probably have made the CC's task in devising a package of remedies practically unworkable.

C. GROUND 2

Ground 2, on the other hand, proved to be more troublesome for the CC. The essence of the argument was that the CC had decided upon the POSP without any proper evidential basis. During the CC's investigation, many of the PPI

providers had urged the CC to consider the loss of convenience that would arise from imposing the POSP, and the risk that this would lead to a contraction in the market. The CC dismissed these concerns in its report, noting:⁵⁵

“While we acknowledge that this element of the remedies package reduces the convenience of purchasing PPI at the credit point of sale, we consider that the potential reduction in PPI sales has been overestimated by some parties. By increasing competition and thereby reducing price, we expect our remedies package to lead to an increase in PPI sales that would partially or fully offset a decline from a reduction in convenience.”

The problem, according to Barclays and the interveners, was that the CC was proposing a radical and unprecedented remedy, yet had failed to reach a considered judgment on the extent of the reduction in take-up rates, and the extent to which this reduction would be off-set by any increase in PPI demand attributable to lower prices. The CAT agreed, noting that while the CC was entitled to decide how much weight to give to the evidence on loss of convenience, it should have made clear which evidence on reduced take-up it was discounting or rejecting:⁵⁶

“The potential for such a radical remedy to cause disadvantageous side-effects called for rigorous investigation and analysis of its potentially adverse consequences...

It was, of course, for the Commission to give such weight to that evidence as it reasonably thought fit, having regard in particular to the fact that most of it was tendered by parties with commercial reasons to be opposed to the imposition of the POSP. In that respect, we can identify no basis upon which the Commission’s decision to discount part of that evidence can be challenged.

In our view, however, it is unfortunate that the Commission did not identify which of the evidence that the loss of convenience would lead to a reduced take-up [of] PPI it discounted or rejected. This is particularly unfortunate because we have found it impossible to discern, from the conclusion at [para 10.50 of the CC report] that increased sales due to lower prices would ‘partially or fully off-set’ any reduced take-up, a sufficiently clear judgment either as to the extent to which the Commission considered that the convenience argument was established by the evidence, or as to the extent to which a decline in convenience would be offset by increased demand due to lower prices.”

Although the CAT agreed that the CC was right not to treat convenience as a relevant customer benefit for the purposes of section 134(8) EA02, given the very narrow definition in the statute, it did not, in the CAT's view, follow that the loss of convenience could not be a relevant disadvantage to be taken into account in the proportionality analysis. The likelihood of inconvenience leading to reduced take-up rates should, according to the CAT, have been weighed in the balance in the proportionality analysis.⁵⁷

“It could hardly be doubted that a remedies package which produced a theoretically perfectly competitive market for PPI, but at the expense of driving a majority of potential purchasers from the market place, would not be reasonable, proportionate, or for that matter, effective.”

It was the CC's failure to give any consideration to reduced take-up rates stemming from the loss of convenience that, in the CAT's view, amounted to a failure to take account of relevant considerations. Unless satisfied that this would not have affected the eventual outcome of the CC's report, this failure would justify quashing the decision to impose the POSP. After what it described as “anxious consideration” the CAT said that it was “not so satisfied.”⁵⁸

D. GROUND 3

Ground 3 was, in essence, a series of challenges to the CC's methodology for quantifying the consumer detriment caused by the AEC. First, it was claimed that the CC had modeled theoretical remedies packages rather than the actual package it was proposing. Further, the model was specifically challenged in several regards: it took no account of costs; was based on unjustified assumptions that the remedies would completely remove all excess PPI profits; took no account of the negative effect on PPI sales due to the loss of convenience of being able to buy PPI at point of sale; and was based on out-of-date information. The last challenge argued that the CC had failed to calculate the proper elasticity of demand.⁵⁹

The debate about modeling theoretical rather than actual remedies packages arose because the CC had identified and modeled both a “system” remedy (which was intended to increase information such that all consumers could search effectively for both credit and PPI before arriving at the point of sale), and a “non-system” remedy (in which prices would be reduced but there would be no increase in the amount of searching for PPI before the point of sale). However, the CAT noted that the primary purpose of the CC's modeling exercise was not to quantify the consumer benefits likely to flow from the proposed remedies package, but to identify any possible modification of the remedies package that

would preserve the relevant customer benefit of lower credit prices (which would be lost if the remedies package had the expected effect of reducing PPI prices).⁶⁰

The CAT accepted that the CC had indeed modeled theoretical remedies packages rather than the actual remedies, but considered that this was well within its margin of appreciation when considering whether its proposed remedies should be amended to preserve relevant customer benefits. Although the CAT expressed doubts about the use of such a methodology purely for quantifying the expected consumer benefits associated with the remedies package, in this particular case, it was only the secondary purpose of the modeling exercise. The use of the modeling exercise for its primary purpose led the CC to anticipate a net consumer benefit; since its actual remedies package lay further along the spectrum towards the more efficient end, the CC could confidently expect the benefits from its actual remedies package to be at least as great.⁶¹

THE CAT HAD MORE SERIOUS CONCERNS ABOUT THE OMISSION OF SET-UP AND ONGOING IMPLEMENTATION COSTS FROM THE MODEL.

The CAT had more serious concerns about the omission of set-up and ongoing implementation costs from the model. Implementation costs would be borne by PPI distributors and therefore should have been taken into account in considering the expected reduction in PPI prices. Although the CC had taken these costs into account when considering the proportionality of its remedies package, the CAT's concern was that the modeling did not fully reflect the impact of reduced PPI prices on higher credit prices. Also of concern was that the CC had failed to consider the increased costs of marketing PPI that would flow from any system remedy, such as the CC's proposed package of remedies. These were material facts that the CC ought to have taken into account. While on its own this omission would not have been sufficient to justify quashing the report, when coupled with the other defects, the CAT considered that this was something that the CC should re-consider as part of the overall conclusion that the CC's decision to impose the POSP should be quashed.⁶²

The next element of Ground 3, namely the claimed assumption by the CC that its remedies would be fully effective and would reduce excess PPI profits to zero, was rejected as in substance the same as Ground 1, which had already been rejected.⁶³

The claimed failure of the CC's modeling to take into account the adverse consequences of the remedies package (i.e. the loss of convenience associated with the imposition of the POSP) had already been considered under Ground 2 as sufficient to justify quashing the decision and, therefore, the CAT did not need to consider the point in detail again.⁶⁴

As to the claim that the model had used out-of-date (2006) information, the CAT noted that it might have been possible to use later figures but this did not

amount to a reviewable error in methodology. Just as the CC had had regard to data subsequent to 2006 when conducting its market definition analysis, and had been entitled to conclude that the more recent data did not undermine that analysis, so it was unnecessary for the CC to repeat the modeling analysis using more recent data than 2006. Although, in an ideal world, it might have been better if the CC had explained why it was content to rely on 2006 figures, the CAT, somewhat surprisingly, expressed itself satisfied that an explanation would have been forthcoming which would not have affected the CC's decision.⁶⁵

The final element of Ground 3 concerned an alleged error in the calculation of the elasticity of demand when estimating the likely effect of the remedies package on sales volumes. The CC conceded that its calculations assumed a price change by a single distributor rather than by all players in the market, but the CAT had little desire to delve into the merits of such a highly technical issue. On the face of it, the CAT considered that the use of an inappropriate elasticity of demand factor was a reviewable error, but on the facts of this case, the CAT was not persuaded that it was material to the decision to impose the POSP. Nonetheless, having identified other failings in the CC's analysis, this was another element that the CC would be required to reconsider as a result of the quashing of its decision.⁶⁶

E. RETAIL PPI

Finally, the CAT dealt briefly with the intervention by Shop Direct, a provider of retail PPI. Shop Direct's role as an intervener in the proceedings was necessarily limited to supporting Barclays' case that the decision to impose the POSP as a remedy for all types of PPI should be quashed. However, the main thrust of Shop Direct's case related to retail PPI, which, as the CAT put it,⁶⁷ "was a case that could only have been advanced under a separate application, rather than by way of intervention." Nonetheless, the CAT gave a very strong hint that, had such a case been brought independently by Shop Direct, it might well have succeeded.

The practical problem for stand-alone providers of retail PPI is that they cannot ascertain the level of credit being extended by the retailer, which makes it difficult to tailor stand-alone PPI policies to fluctuating amounts owed by the consumer. Shop Direct's argument was that the CC's remedies package "contained no solution to this conundrum, so that it could not therefore rationally be expected effectively to remedy the AEC in relation to retail PPI."⁶⁸ Although the CAT was understandably reluctant to express a view on the merits of this submission in the absence of full argument, it did nonetheless express concern about the point, noting that it had been unable to dismiss it as "obviously wrong," and inviting the CC to bear it in mind in its reconsideration of the POSP remedy, in order to avoid a further challenge being made to the CC's subsequent decision.⁶⁹

VII. Conclusion

Where does the CAT's judgment in *Barclays* leave the CC and, more generally, the scope for judicial review in competition inquiries? As we have shown, there are a number of circumstances where general principles of judicial review allow heightened scrutiny (or even more intense review); some of these apply to competition inquiries. We have identified three main circumstances where this may be relevant: (1) cases where the nature of the rights affected by a competition authority's action engages fundamental rights; (2) cases where the theory relied upon by the competition authority involves "uncertain" outcomes and when the causal links between action and effect are inherently difficult to establish; and

WILL THIS INDUCE THE CC
TO BE MORE CAREFUL IN ITS
ANALYSIS IN FUTURE CASES,
AND/OR IN ITS CHOICE OF
REMEDIES? PROBABLY BOTH.

(3) specific issues within the competition authority's overall analysis which—in light of their importance or gravity—require particularly careful consideration.

There is an important link between the level of judicial scrutiny over a competition authority's decisions and the standards to be applied by the primary decision-maker in its decision-making process: the "double proportionality" test can be seen as relating to the intensity of review, but it is in essence directed towards the nature of the balancing exercise to be conducted by the primary decision-maker. In *Barclays* the CAT was keen to show some deference to the CC's margin of appreciation and avoid a merits-based appeal under the guise of judicial review. It was even prepared to give the CC some latitude as to the erroneous use of evidence (e.g. the use of an inappropriate elasticity of demand factor) when it was not persuaded that the error would have made a material difference to the decision. In addition, the CAT was unsympathetic to arguments (the CC's failure to address the incremental impact, in terms of both costs and benefits, of adding the POSP to the remaining package of remedies) that would have made the CC's overall task practically unworkable. Nevertheless, the CAT was prepared to quash the CC's decision on the basis of defects in the proportionality assessment of such a radical remedy. Will this induce the CC to be more careful in its analysis in future cases, and/or in its choice of remedies? Probably both. ▼

- 1 Tesco plc v. Competition Commission [2009] CAT 6.
- 2 Barclays Bank plc v. Competition Commission [2009] CAT 27.
- 3 BAA Limited v. Competition Commission [2009] CAT 35 (main judgment); and [2010] CAT 9 (judgment on relief and permission to appeal).
- 4 Sports Direct International PLC v. Competition Commission [2009] CAT 32.
- 5 "(1) Any person aggrieved by a decision of the OFT, [OFCOM,] the Secretary of State or the Commission under this Part in connection with a reference or possible reference in relation to a rele-

vant merger situation or a special merger situation may apply to the Competition Appeal Tribunal for a review of that decision." Then, "(4) In determining such an application the Competition Appeal Tribunal shall apply the same principles as would be applied by a court on an application for judicial review."

- 6 See *generally*, DE SMITH'S JUDICIAL REVIEW 3 (H. Woolf, J. Jowell, & A. Le Sueur eds.) (6th ed. 2007).
- 7 See CC3 Market Investigations References: Competition Commission Guidelines (June 2003), ¶¶ 4.10-4.12 and CC8 Merger Remedies: Competition Commission Guidelines (November 2008), ¶¶ 1.9-1.13 (together, the Guidelines).
- 8 R. Cooke, *The Struggle for Simplicity in Administrative Law*, JUDICIAL REVIEW OF ADMINISTRATIVE ACTION IN THE 80s 3 (M. Taggart ed.) (1986).
- 9 Council for Civil Service Unions v. Minister of State for the Civil Service [1985] 1 A.C. 374 at 410.
- 10 See *generally* DE SMITH'S JUDICIAL REVIEW, *supra* note 6 at 590.
- 11 British Sky Broadcasting Group plc v. Competition Commission [2008] CAT 25, at ¶¶ 272-273.
- 12 *Id.*
- 13 Under English law an administrative decision is vitiated by illegality if it: contravenes or exceeds the terms of the power which authorises the making of the decision; pursues objectives other than those for which the power was conferred; is not authorised by any power; or contravenes or fails to implement a public duty. Compare this with the approach taken by the ECJ in competition cases. In Case C-413/06 P. Bertelsmann and Sony Corporation of America v. Impala [2008] ECR I-4951, at ¶ 133, the issue before the ECJ was the way in which the CFI had evaluated the role that market transparency should play in the context of a finding of collective dominance. The ECJ concluded that "in misconstruing the principles which should have guided its analysis of the arguments raised before it concerning market transparency in the context of an allegation of collective dominant position, the Court of First Instance committed an error of law."
- 14 *Supra* note 6 at 592.
- 15 British Sky Broadcasting Group plc v. Competition Commission [2010] EWCA Civ 2, at ¶ 37.
- 16 Office of Fair Trading & Ors v. IBA Health Ltd [2004] EWCA Civ 142.
- 17 *Supra* note 6 at 591.
- 18 *Supra* note 11 at ¶ 61.
- 19 For a review of the recent case law in this area see Vivien Rose, *Margins of Appreciation: Changing Contours in Community and Domestic Case Law*, 5(1) COMPETITION POL'Y INT'L 3 (Spring 2009).
- 20 R v. Ministry of Defence, Ex parte Smith [1996] QB 517, at 554.
- 21 See also R (Ross) v. West Sussex Primary Care Trust [2008] EWHC B15, cited by Rose (*supra* note 19) where this formulation was recently restated.
- 22 R (Begum) v. Denbigh High School [2006] UKHL 15.
- 23 *Supra* note 19, at 15, citing R (Wilkinson) v. Broadmoor [2001] EWCA 1545.

- 24 Case C-12/03 Commission v. Tetra-Laval B.V. [2005] ECR I-987, at ¶¶ 41-44.
- 25 *Supra* note 15.
- 26 “[S]ome things are inherently more likely than others. It would need more cogent evidence to satisfy one that the creature seen walking in Regent’s Park was more likely than not to have been a lioness than to be satisfied to the same standard of probability that it was an Alsatian.” Lord Hoffmann in Secretary of State for the Home Department v. Rehman [2001] UKHL 47, at ¶ 55.
- 27 Case C-331/88 R v. Minister of Agriculture, Fisheries and Food and Secretary of State for Health, ex parte Fedesa and others [1990] ECR I-4023, at ¶ 13.
- 28 *Supra* note 1, at ¶¶ 138-139.
- 29 *Supra* note 2, at ¶ 21.
- 30 *Supra* note 2, at ¶ 1.
- 31 *Id.*, at ¶ 3.
- 32 *Id.*, at ¶ 77.
- 33 *Id.*, at ¶¶ 84-85, citing the CC report at ¶¶ 10.41 and 10.43.
- 34 *Id.*, at ¶ 86.
- 35 *Id.*, at ¶ 87.
- 36 *Id.*, at ¶ 91, citing the CC report at ¶ 10.72.
- 37 *Id.*, at ¶ 93.
- 38 *Id.*, at ¶ 96.
- 39 *Id.*, at ¶ 99, citing ¶¶ 10.509, 10.513 and 10.514 of the CC report.
- 40 A summary of the notice of application is available at http://www.catribunal.org.uk/files/Notice_1109_Barclays_02.04.09.pdf.
- 41 *Supra* note 2, at ¶ 31.
- 42 *Id.*, at ¶¶ 43, 52.
- 43 *Id.*, at ¶ 36.
- 44 *Id.*, at ¶¶ 37-42.
- 45 *Id.*, at ¶ 43.
- 46 *Id.*, at ¶ 47.
- 47 *Id.*, at ¶ 65.

48 *Supra* note 1, at ¶ 143.

49 *Id.*, at ¶ 79; *supra* note 2, at ¶ 76.

50 *Id.*, at ¶ 103.

51 *Id.*, at ¶ 104.

52 *Id.*, at ¶ 105; see also ¶ 106.

53 *Id.*, at ¶ 115.

54 *Id.*, at ¶ 116.

55 CC report, at ¶ 10.50.

56 *Supra* note 2, at ¶¶ 128-130.

57 *Id.*, at ¶ 136.

58 *Id.*, at ¶ 140.

59 *Id.*, at ¶ 142.

60 *Id.*, at ¶ 145.

61 *Id.*, at ¶¶ 153-155.

62 *Id.*, at ¶¶ 158-163.

63 *Id.*, at ¶¶ 164-165.

64 *Id.*, at ¶¶ 166-167.

65 *Id.*, at ¶¶ 168-170.

66 *Id.*, at ¶¶ 171-175.

67 *Id.*, at ¶ 176.

68 *Id.*, at ¶ 178.

69 *Id.*, at ¶ 179.

A Behavioral Model of Rational Choice

Herbert A. Simon
with an introduction by
Lindsay W. McSweeney
Competition Policy International

Introduction to A Behavioral Model of Rational Choice

*Lindsay W. McSweeney**

Introduction

Herbert Alexander Simon (1916–2001) was a pioneer researcher in multiple scientific and economic fields, including economics, organizational theory, and artificial intelligence. After graduating from the University of Chicago he spent most of his career teaching at Carnegie Technical Institute in both the psychology and computer science departments.

Simon's early investigations in economics form the basis for much of his life's work, earning him the 1978 Nobel Memorial Prize in Economics "for his pioneering research into the decision-making process within economic organizations." To test his research, Simon also created new methodologies. He determined the best approach was through computer simulation modeling, a process that led to the computer science studies that would make him one of the founding fathers of Artificial Intelligence.¹

Simon's concept of "bounded rationality" was a precursor to many of the later concepts in behavioral economics. He was one of the first scientists to analyze uncertainty in organizational decision making, focusing on identifying the constraints under which decisions are made. As Salinger defines it, "Bounded rationality means that individuals (or firms) act purposefully, but not necessarily as if they are both fully informed and perfectly rational."² Simon identified several constraints that decision makers face, including: (1) only limited, often unreliable, information regarding possible alternatives and their consequences, (2) the human mind's limited ability to evaluate and process available information, and (3) time constraints. (An inability to deal with these constraints leads to the management ailment known as "analysis paralysis.")

I *Senior Editor, Competition Policy International and the CPI Antitrust Chronicle.

Simon first proposed the ideas behind bounded rationality in a 1953 study for the RAND Corporation called *A Behavioral Model of Rational Choice*. The first line in the paper states that it's "A model for the description of rational choice by organisms of limited computational ability."³ He further developed these ideas while working with the Cowles Commission for Research in Economics (an economic research institute then located at the University of Chicago) supported by a contract from the Office of Naval Research and a grant from the Ford Foundation. In 1955 Simon published his polished ideas in a similarly named paper in the *Quarterly Journal of Economics*. He doesn't use the actual term "bounded rationality" in this paper; the term first appeared in 1957, when he combined his papers in the book, *Models of Man*. However, Simon does refer to "approximate rationality" and also defines an aspiration level as the "boundary between what is considered satisfactory and unsatisfactory." In 1984, he claimed the phrase definitively, using it as the title in Volumes 1 and 2 of *Models of Bounded Rationality*, which compile his papers on classical and neo-classical economic theory. He published a third volume in 1997 to include his subsequent economic papers.

A Behavioral Model of Rational Choice is reprinted here as it was published in 1955. In this paper, Simon writes that "the distance is so great between our present psychological knowledge of the learning and choice processes and the kinds of knowledge needed for economic administrative theory that a marking stone placed halfway between might help travelers from both directions to keep to their courses."⁴ In his career Simon did far more than keep travelers in behavioral economics and organizational theory on course; he designed and engineered much of the course(s) they would follow. ▼

1 In 1955, Simon and Allen Newell created the "Logic Theorist," a computer program that would eventually prove 38 of the first 52 theorems in Russell and Whitehead's *Principia Mathematica* and presented this paper at the 1956 Dartmouth Conference that founded, defined, and named the new field of Artificial Intelligence. Simon was perhaps a bit too optimistic about AI; in 1965, he announced that "machines will be capable, within twenty years, of doing any work a man can do."

2 Michael Salinger, *Behavioral Economics, Consumer Protection, and Antitrust*, 6(1) Competition Pol'y Int'l at 71 (Spring, 2010).

3 This report is available online at www.rand.org/pubs/papers/2005/P365.pdf, last viewed on 03/05/10.

4 H. Simon, *A Behavioral Model of Rational Choice*, 6(1) Competition Pol'y Int'l at 241 (Spring, 2010).

A Behavioral Model of Rational Choice¹

*Herbert A. Simon*²

Traditional economic theory postulates an “economic man,” who, in the course of being “economic” is also “rational.” This man is assumed to have knowledge of the relevant aspects of his environment which, if not absolutely complete, is at least impressively clear and voluminous. He is assumed also to have a well-organized and stable system of preferences, and a skill in computation that enables him to calculate, for the alternative courses of action that are available to him, which of these will permit him to reach the highest attainable point of his preference scale.

Recent developments in economics, and particularly in the theory of the business firm, have raised great doubts as to whether this schematized model of economic man provides a suitable foundation on which to erect a theory—whether it be a theory of how firms *do* behave, or of how they “should” rationally behave. It is not the purpose of this paper to discuss these doubts, or to determine whether they are justified. Rather, I shall assume that the concept of “economic man” (and, I might add, of his brother “administrative man”) is in need of fairly drastic revision, and shall put forth some suggestions as to the direction the revision might take.

Broadly stated, the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist. One is tempted to turn to the literature of psychology for the answer. Psychologists have certainly been concerned with rational behavior, particularly in their interest in learning phenomena. But the distance is so great between our present psychological knowledge of the learning and choice processes and the kinds of knowledge needed for economic and administrative theory that a marking stone placed halfway between might help travelers from both directions to keep to their courses.

Lacking the kinds of empirical knowledge of the decisional processes that will be required for a definitive theory, the hard facts of the actual world can, at the present stage, enter the theory only in a relatively unsystematic and unrigorous way. But none of us is completely innocent of acquaintance with the gross characteristics of human choice, or of the broad features of the environment in which this choice takes place. I shall feel free to call on this common experience as a source of the hypotheses needed for the theory about the nature of man and his world.

The problem can be approached initially by inquiring into the properties of the choosing organism, or by inquiring into the environment of choice. In this paper, I shall take the former approach. I propose, in a sequel, to deal with the characteristics of the environment and the interrelations of environment and organism.

The present paper, then, attempts to include explicitly some of the properties of the choosing organism as elements in defining what is meant by rational behavior in specific situations and in selecting a rational behavior in terms of such a definition. In part, this involves making more explicit what is already implicit in some of the recent work on the problem—that the state of information may as well be regarded as a characteristic of the decision-maker as a characteristic of his environment. In part, it involves some new considerations—in particular taking into account the simplifications the choosing organism may deliberately introduce into its model of the situation in order to bring the model within the range of its computing capacity.

I. Some General Features of Rational Choice

The “flavor” of various models of rational choice stems primarily from the specific kinds of assumptions that are introduced as to the “givens” or constraints within which rational adaptation must take place. Among the common constraints—which are not themselves the objects of rational calculation—are (1) the set of alternatives open to choice, (2) the relationships that determine the pay-offs (“satisfactions,” “goal attainment”) as a function of the alternative that is chosen, and (3) the preference-orderings among pay-offs. The selections of particular constraints and the rejection of others for incorporation in the model of rational behavior involves implicit assumptions as to what variables the rational organism “controls”—and hence can “optimize” as a means to rational adaptation—and what variables it must take as fixed. It also involves assumptions as to the character of the variables that are fixed. For example, by making different assumptions about the amount of information the organism has with respect to the relations between alternatives and pay-offs, optimization might involve selection of a certain maximum, of an expected value, or a minimax.

Another way of characterizing the givens and the behavior variables is to say that the latter refer to the organism itself, the former to its environment. But if we adapt this viewpoint, we must be prepared to accept the possibility that what we

call “the environment” may lie, in part, within the skin of the biological organism. That is, some of the constraints that must be taken as givens in an optimization problem may be physiological and psychological limitations of the organism (biologically defined) itself. For example, the maximum speed at which an organism can move establishes a boundary on the set of its available behavior alternatives. Similarly, limits on computational capacity may be important constraints entering into the definition of rational choice under particular circumstances. We shall explore possible ways of formulating the process of rational choice in situations where we wish to take explicit account of the “internal” as well as the “external” constraints that define the problem of rationality for the organism.

Whether our interests lie in the normative or in the descriptive aspects of rational choice, the construction of models of this kind should prove instructive. Because of the psychological limits of the organism (particularly with respect to computational and predictive ability), actual human rationality-striving can at best be an extremely crude and simplified approximation to the kind of global rationality that is implied, for example, by game-theoretical models. While the approximations that organisms employ may not be the best—even at the levels of computational complexity they are able to handle—it is probable that a great deal can be learned about possible mechanisms from an examination of the schemes of approximation that are actually employed by human and other organisms.

In describing the proposed model, we shall begin with elements it has in common with the more global models, and then proceed to introduce simplifying assumptions and (what is the same thing) approximating procedures.

A. PRIMITIVE TERMS AND DEFINITIONS

Models of rational behavior—both the global kinds usually constructed, and the more limited kinds to be discussed here—generally require some or all of the following elements:

1. A set of *behavior alternatives* (alternatives of choice or decision). In a mathematical model, these can be represented by a point set, A .
2. The subset of *behavior alternatives that the organism “considers” or “perceives.”* That is, the organism may make its choice within a set of alternatives more limited than the whole range objectively available to it. The “considered” subset can be represented by a point set \hat{A} , with \hat{A} included in A ($\hat{A} \subset A$).
3. *The possible future state of affairs*, or outcomes of choice, represented by a point set, S . (For the moment it is not necessary to distinguish between actual and perceived outcomes.)
4. A “pay-off” function, representing the “value” or “utility” placed by the organism upon each of the possible outcomes of choice. The pay-off may be represented by a real function, $V(s)$ defined for all ele-

ments, s , of S . For many purposes there is needed only an ordering relation on pairs of elements of S —i.e., a relation that states that s_1 is preferred to s_2 or vice versa—but to avoid unnecessary complications in the present discussion, we will assume that a cardinal utility, $V(s)$, has been defined.

5. *Information as to which outcomes in S will actually occur* if a particular alternative, a , in A (or in \hat{A}) is chosen. This information may be incomplete—that is, there may be more than one possible outcome, s , for each behavior alternative, a . We represent the information, then, by a mapping of each element, a , in A upon a subset, S_a —the set of outcomes that may ensue if a is the chosen behavior alternative.
5. *Information as to the probability that a particular outcome will ensue* if a particular behavior alternative is chosen. This is a more precise kind of information than that postulated in (5), for it associates with each element, s , in the set S_a , a probability, $P_a(s)$ —the probability that s will occur if a is chosen. The probability $P_a(s)$ is a real, non-negative function with $\sum_{s \in S_a} P_a(s) = 1$.

Attention is directed to the threefold distinction drawn by the definitions among the set of behavior alternatives, A , the set of outcomes or future states of affairs, S , and the pay-off, V . In the ordinary representation of a game, in reduced form, by its pay-off matrix, the set S corresponds to the cell of the matrix, the set A to the strategies of the first player, and the function V to the values in the cells. The set S_a is then the set of cells in the a th row. By keeping in mind this interpretation, the reader may compare the present formulation with “classical” game theory.

B. “CLASSICAL” CONCEPTS OF RATIONALITY

With these elements, we can define procedures of rational choice corresponding to the ordinary game-theoretical and probabilistic models.³

- A. *Max-min Rule*. Assume that whatever alternative is chosen, the worst possible outcome will ensue—the smallest $V(s)$ for s in S_a will be realized. Then select that alternative, a , for which this worst pay-off is as large as possible:

$$\hat{V}(\hat{a}) = \underset{s \in S_{\hat{a}}}{\text{Min}} V(s) = \underset{a \in A}{\text{Max}} \underset{s \in S_a}{\text{Min}} V(s)$$

Instead of the maximum with respect to the set, A , of actual alternatives, we can substitute the maximum with respect to the set, \hat{A} , of “considered” alternatives. The probability distribution of outcomes, (6) does not play any role in the max-min rule.

- B. *Probabilistic Rule*. Maximize the expected value of $V(s)$ for the (assumed known) probability distribution, $P_a(s)$.

$$\hat{V}(\hat{a}) = \sum_{s \in S_{\hat{a}}} V(s) P_{\hat{a}}(s) = \underset{a \in A}{\text{Max}} \sum_{s \in S_a} V(s) P_a(s)$$

- C. *Certainty Rule*. Given the information that each a in A (or in \hat{A}) maps upon a specified s_a in S , select the behavior alternative whose outcome has the largest pay-off.

$$\hat{V}(\hat{a}) = V(S_a) = \text{Max}_{a \in A} V(S_a)$$

II. The Essential Simplifications

If we examine closely the “classical” concept of rationality outlined above, we see immediately what severe demands they make upon the choosing organism. The organism must be able to attach definite pay-offs (or at least a definite range of pay-offs) to each possible outcome. This, of course, involves also the ability to specify the exact nature of the outcomes—there is no room in the scheme for “unanticipated consequences.” The pay-offs must be completely ordered—it must always be possible to specify, in a consistent way, that one outcome is better than, as good as, or worse than any other. And, if the certainty or probabilistic rules are employed, either the outcomes of particular alternatives must be known with certainty, or at least it must be possible to attach definite probabilities to outcomes.

My first empirical proposition is that there is a complete lack of evidence that, in actual human choice situations of any complexity, these computations can be, or are in fact, performed. The introspective evidence is certainly clear enough, but we cannot, of course, rule out the possibility that the unconscious is a better decision-maker than the conscious. Nevertheless, in the absence of evidence that the classical concepts do describe the decision-making process, it seems reasonable to examine the possibility that the actual process is quite different from the ones the rules describe.

Our procedure will be to introduce some modifications that appear (on the basis of casual empiricism) to correspond to observed behavior processes in humans, and that lead to substantial computational simplifications in the making of a choice. There is no implication that human beings use all of these modifications and simplifications all the time. Nor is this the place to attempt the formidable empirical task of determining the extent to which, and the circumstances under which humans actually employ these simplifications. The point is rather that these are procedures which appear often to be employed by human beings in complex choice situations to find an approximate model of manageable proportions.

A. “SIMPLE” PAY-OFF FUNCTIONS

One route to implications is to assume that $V(s)$ necessarily assumes one of two values, (1,0), or of three values, (1,0,-1), for all s in S . Depending on the circumstances, we might want to interpret these values, as (a) (satisfactory or unsatisfactory), or (b) (win, draw or lose).

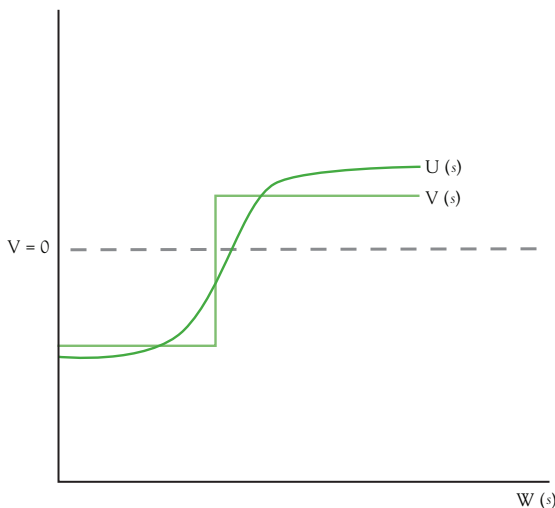
As an example of (b), let S represent the possible positions in a chess game at White's 20th move. Then a (+1) position is one in which White possesses a strategy leading to a win whatever Black does. A (0) position is one in which White can enforce a draw, but not a win. A (-1) position is one in which Black can force a win.

As an example of (a) let S represent possible prices for a house an individual is selling. He may regard \$15,000 as an "acceptable" price, anything over this amount as "satisfactory," anything less as "unsatisfactory." In psychological theory we would fix the boundary at the "aspiration level"; in economic theory we would fix the boundary at the price which evokes indifference between selling and not selling (an opportunity cost concept).

The objection may be raised that, although \$16,000 and \$25,000 are both "very satisfactory" prices for the house, a rational individual would prefer to sell at the higher price, and hence, that the simple pay-off function is an inadequate representation of the choice situation. The objection may be answered in several different ways, each answer corresponding to a class of situations in which the simple function might be appropriate.

First, the individual may not be confronted simultaneously with

Figure 1



a number of buyers offering to purchase the house at different prices, but may receive a sequence of offers, and may have to decide to accept or reject each one before he receives the next. (Or, more generally, he may receive a sequence of pairs or triplets or n -tuples of offers, and may have to decide whether to accept the highest of an n -tuple before the next n -tuple is received.) Then, if the elements S correspond to n -tuples of offers, $V(s)$ would be 1 whenever the highest offer in the n -tuple exceeded the "acceptance price" the seller had determined upon at that time. We can then raise the further question of what would be a rational process for determining the acceptance price.⁴

Second, even if there were a more general pay-off function, $W(s)$, capable of assuming more than two different values, the simplified $V(s)$ might be a satisfactory approximation to $W(s)$. Suppose, for example, that there were some way of introducing a cardinal utility function, defined over S , say $U(s)$. Suppose further that $U(W)$ is a monotonic increasing function with a strongly negative second derivative (decreasing marginal utility). Then $V(s) = V\{W(s)\}$ might be the approximation as shown on page 248.

When a simple $V(s)$, assuming only the values $(+1,0)$ is admissible, under the circumstances just discussed or under other circumstances, then a (fourth) rational decision-process could be defined as follows:

- D. (i)** Search for a set of possible outcomes (a subset, S' in S) such that the pay-off is satisfactory ($V(s) = 1$) for all these possible outcomes (for all s in S').
- (ii)** Search for a behavior alternative (an a in \hat{A}) whose possible outcomes all are in S' (such that a maps upon a set, S_a , that is contained in S').

If a behavior alternative can be found by this procedure, then a satisfactory outcome is assured. The procedure does not, of course, guarantee the existence of uniqueness of an a with the desired properties.

B. INFORMATION GATHERING

One element of realism we may wish to introduce is that, while $V(s)$ may be known in advance, the mapping of A on subsets of S may not. In the extreme case, at the outset each element, a , may be mapped on the whole set, S . We may then introduce into the decision-making process information-gathering steps that produce a more precise mapping of the various elements of A on nonidentical subsets of S . If the information-gathering process is not costless, then one element in the decision will be the determination of how far the mapping is to be refined.

Now in the case of the simple pay-off functions, $(+1,0)$, the information-gathering process can be streamlined in an important respect. First, we suppose that the individual has initially a very coarse mapping of A on S . Second, he looks for

an S' in S such that $V(s) = 1$ for s in S' . Third, he gathers information to refine that part of the mapping on A on S in which elements of S' are involved. Fourth, having refined the mapping, he looks for an a that maps on to a subset of S' .

Under favorable circumstances, this procedure may require the individual to gather only a small amount of information—an insignificant part of the whole mapping of elements of A on individual elements of S . If the search for an a having the desirable properties is successful, he is certain that he cannot better his choice by securing additional information.⁵

It appears that the decision process just described is one of the important means employed by chess players to select a move in the middle and end game. Let A be the set of moves available to White on his 20th move. Let S be a set of positions that might be reached, say, by the 30th move. Let S' be some subset of S that consists of clearly “won” positions. From a very rough knowledge of the mapping of A on S , White tentatively selects a move, a , that (if Black plays in a certain way) maps on S' . By then considering alternative replies for Black, White “explores” the whole mapping of a . His exploration may lead to points, s , that are not in S' , but which are now recognized also as winning positions. These can be adjoined to S' . On the other hand, a sequence may be discovered that permits Black to bring about a position that is clearly not “won” for White. Then White may reject the original point, a , and try another.

Whether this procedure leads to any essential simplification of the computation depends on certain empirical facts about the game. Clearly all positions can be categorized as “won,” “lost,” or “drawn” in an objective sense. But from the standpoint of the player, positions may be categorized as “clearly won,” “clearly lost,” “clearly drawn,” “won or drawn,” “drawn or lost,” and so forth—depending on the adequacy of this mapping. If the “clearly won” positions represent a significant subset of the objectively “won” positions, then the combinatorics involved in seeing whether a position can be transformed into a clearly won position, for all possible replies by Black, may not be unmanageable.⁶ The advantage of this procedure over the more common notion (which may, however, be applicable in the opening) of a general valuation function for positions, taking on values from -1 to 1 , is that it implies much less complex and subtle evaluation criteria. All that is required is that the evaluation function be reasonably sensitive in detecting when a position in one of the three states—won, lost, or drawn—has been transformed into a position in another state. The player, instead of seeking for a “best” move needs only to look for a “good” move.

We see that, by the introduction of a simple pay-off function and of a process for gradually improving the mapping of behavior alternatives upon possible outcomes, the process of reaching a rational decision may be drastically simplified from a computational standpoint. In the theory and practice of linear program-

ming, the distinction is commonly drawn between computations to determine the *feasibility* of a program, and computations to discover the *optimal* program. Feasibility testing consists in determining whether a program satisfies certain linear inequalities that are given at the outset. For example, a mobilization plan may take as given the maximum work force and the steel-making capacity of the economy. Then a feasible program is one that does not require a work force or steel-making facilities exceeding the given limits.

An optimal program is that one of the feasible programs which maximizes a given pay-off function. If, instead of requiring that the pay-off be maximized, we require only that the pay-off exceed some given amount, then we can find a program that satisfies this requirement by the usual methods of feasibility testing. The pay-off requirement is represented simply by an additional linear inequality that must be satisfied. Once this requirement is met, it is not necessary to determine whether there exists an alternative plan with a still higher pay-off.

For all practical purposes, this procedure may represent a sufficient approach to optimization, provided the minimum required pay-off can be set “reasonably.” In later sections of this paper we will discuss how this might be done, and we shall show also how the scheme can be extended to vector pay-off functions with multiple components (Optimization requires, of course, a complete ordering of pay-offs).

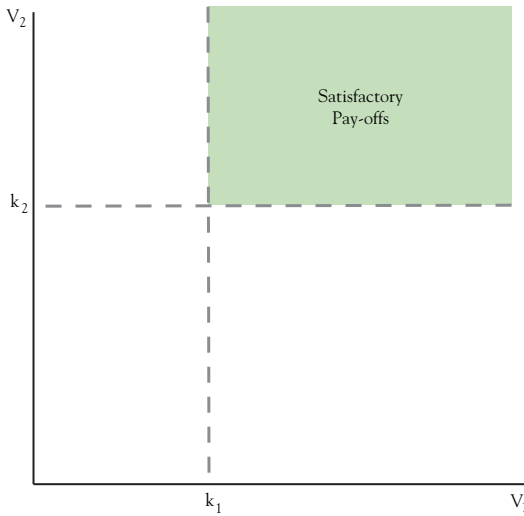
C. PARTIAL ORDERING OF PAY-OFFS

The classical theory does not tolerate the incomparability of oranges and apples. It requires a scalar pay-off function, that is, a complete ordering of pay-offs. Instead of a scalar pay-off function, $V(s)$, we might have a vector function, $V(s)$; where V has the components V_1, V_2, \dots . A vector pay-off function may be introduced to handle a number of situations:

1. In the case of a decision to be made by a *group of persons*, components may represent the pay-off functions of the individual members of the group. What is preferred by one may not be preferred by the others.
2. In the case of an individual, he may be trying to implement a number of *values that do not have a common denominator*—e.g., he compares two jobs in terms of salary, climate, pleasantness of work, prestige, etc.;
3. Where each behavior alternative, a , maps on a set of n possible consequences, S_a , we may replace the model by one in which each alternative maps on a single consequence, but each consequence has as its pay-off the n -dimensional vector whose components are the pay-offs of the elements of S_a .

Figure 2

Partial Ordering
of Pay-Offs



This representation exhibits a striking similarity among these three important cases where the traditional maximizing model breaks down for lack of a complete ordering of the pay-offs. The first case has never been satisfactorily treated—the theory of the n -person game is the most ambitious attempt to deal with it, and the so-called “weak welfare principles” of economic theory are attempts to avoid it. The second case is usually handled by superimposing a complete ordering on the points in the vector space (“indifference curves”). The third case has been handled by introducing probabilities as weights for summing the vector components, or by using principles like minimaxing satisfaction or regret.

An extension of the notion of a simplified pay-off function permits us to treat all three cases in much the same fashion. Suppose we regard a pay-off as *satisfactory* provided that $V_i \geq k_i$ for all i . Then a reasonable decision rule is the following:

- E.** Search for a subset S' in S such that $V(s)$ is satisfactory for all s in S' (i.e., $V(s) \geq k$).

Then search for an a in A such that S_a lies in S' .

Again existence and uniqueness of solutions are not guaranteed. Rule E is illustrated in Figure 2 for the case of a 2-component pay-off vector.

In the first of the three cases mentioned above, the satisfactory pay-off corresponds to what I have called a *viable* solution in “A Formal Theory of the Employment Relation” and “A Comparison of Organization Theories.” In the second case, the components of V define the *aspiration levels* with respect to sev-

eral components of pay-off. In the third case (in this case it is most plausible to assume that all the components of k are equal), k_i may be interpreted as the *minimum guaranteed pay-off*—also an aspiration level concept.

III. Existence and the Uniqueness of Solutions

Throughout our discussion we have admitted decision procedures that do not guarantee the existence or uniqueness of solutions. This was done in order to construct a model that parallels as nearly as possible the decision procedures that appear to be used by humans in complex decision-making settings. We now proceed to add supplementary rules to fill this gap.

A. OBTAINING A UNIQUE SOLUTION

In most global models of rational choice, all alternatives are evaluated before a choice is made. In actual human decision-making, alternatives are often examined sequentially. We may, or may not, know the mechanism that determines the order of procedure. When alternatives are examined sequentially, we may regard the first satisfactory alternative that is evaluated as such as the one actually selected.

If a chess player finds an alternative that leads to a forced mate for his opponent, he generally adopts this alternative without worrying about whether another alternative also leads to a forced mate. In this case we would find it very hard to predict which alternative would be chosen, for we have no theory that predicts the order in which alternatives will be examined. But in another case discussed above—the sale of a house—the environment presents the seller with alternatives in a definite sequence, and the selection of the *first* satisfactory alternative has precise meaning.

However, there are certain dynamic considerations, having a good psychological foundation, that we should introduce at this point. Let us consider, instead of a single static choice situation, a sequence of such situations. The *aspiration level*, which defines a satisfactory alternative, may change from point to point in this sequence of trials. A vague principle would be that as the individual, in his exploration of alternatives, finds it *easy* to discover satisfactory alternatives, his aspiration level rises; as he finds it *difficult* to discover satisfactory alternatives, his aspiration level falls. Perhaps it would be possible to express the ease or difficulty of exploration in terms of the cost of obtaining better information about the mapping of A on S , or the combinatorial magnitude of the task of refining this mapping. There are a number of ways in which this process could be defined formally.

Such changes in aspiration level would tend to bring about a “near-uniqueness” of the satisfactory solutions and would also tend to guarantee the existence

of satisfactory solutions. For the failure to discover a solution would depress the aspiration level and bring satisfactory solutions into existence.

B. EXISTENCE OF SOLUTIONS: FURTHER POSSIBILITIES

We have already discussed one mechanism by which the existence of solutions, in the long run, is assured. There is another way of representing the processes already described. Up to this point little use has been made of the distinction between A , the set of behavior alternatives, and $\overset{\circ}{A}$, the set of behavior alternatives that the organism considers. Suppose now that the latter is a proper subset of the former. Then, the failure to find a satisfactory alternative in $\overset{\circ}{A}$ may lead to a search for additional alternatives in A that can be adjoined to $\overset{\circ}{A}$.⁸ This procedure is simply an elaboration of the information-gathering process previously described. (We can regard the elements of A that are not in $\overset{\circ}{A}$ as elements that are initially mapped on the whole set, S .)

In one organism, dynamic adjustment over a sequence of choices may depend primarily upon adjustments of the aspiration level. In another organism, the adjustments may be primarily in the set $\overset{\circ}{A}$: if satisfactory alternatives are discovered easily, $\overset{\circ}{A}$ narrows; if it becomes difficult to find satisfactory alternatives, $\overset{\circ}{A}$ broadens. The more *persistent* the organism, the greater the role played by the adjustment of $\overset{\circ}{A}$, relative to the role played by the adjustment of the aspiration level. (It is possible, of course, and even probable, that there is an asymmetry between adjustments upward and downward.)

If the pay-off were measurable in money or utility terms, and if the cost of discovering alternatives were similarly measurable, we could replace the partial ordering of alternatives exhibited in Figure 2 by a complete ordering (an ordering in terms of a weighted sum of the pay-off and the cost of discovering alternatives). Then we could speak of the optimal degree of persistence in behavior—we could say that the more persistent organism was more rational than the other, or vice versa. But the central argument of the present paper is that the behaving organism does *not* in general know these costs, nor does it have a set of weights for comparing the components of a multiple pay-off. It is precisely because of these limitations on its knowledge and capabilities that the less global models of rationality described here are significant and useful. The question of how it is to behave “rationally,” given these limitations, is distinct from the question of how its capabilities could be increased to permit action that would be more “rational” judged from the mountain-top of a more complete model.⁹

The two viewpoints are not, of course, completely different, much less antithetical. We have already pointed out that the organism may possess a whole hierarchy of rational mechanisms—that, for example, the aspiration level itself may be subject to an adjustment process that is rational in some dynamic sense. Moreover, in many situations we may be interested in the precise question of whether one decision-making procedure is more rational than another, and to

answer this question we will usually have to construct a broader criterion of rationality that encompasses both procedures as approximations. Our whole point is that it is important to make explicit what level we are considering in such a hierarchy of models, and that for many purposes we are interested in models of “limited” rationality rather than models of relatively “global” rationality.

IV. Further Comments on Dynamics

The models thus far discussed are dynamic only in a very special sense: the aspiration level at time t depends upon the previous history of the system (previous aspiration levels and previous levels of attainment). Another kind of dynamic linkage might be very important. The pay-offs in a particular trial might depend not only on the alternatives chosen in that trial but also on the alternatives chosen in previous trials.

The most direct representation of this situation is to include, as components of a vector pay-off function, the pay-offs for the whole sequence of trials. But then optimization would require the selection, at the beginning of the sequence, of a strategy for the whole sequence (see the Appendix). Such a procedure would again rapidly complicate the problem beyond the computational capacity of the organism. A possible middle ground is to define for each trial a pay-off function with two components. One would be the “immediate” pay-off (consumption), the other, the “position” in which the organism is left for future trials (saving, liquidity).

Let us consider a chess game in which the players are paid off at the end of each ten moves in proportion to arbitrarily assigned values of their pieces left on the board (say, queen, 1; rook, 10; etc.). Then a player could adopt some kind of planning horizon and include in his estimated pay-off the “goodness” of his position at the planning horizon. A comparable notion in economics is that of the depreciated value of an asset at the planning horizon. To compute such a value precisely would require the player actually to carry his strategy beyond the horizon. If there is time-discounting of pay-offs, this has the advantage of reducing the importance of errors in estimating these depreciated values. (Time-discounting may sometimes be essential in order to assure convergence of the summed pay-offs.)

It is easy to conjure up other dynamic complications, which may be of considerable practical importance. Two more may be mentioned—without attempting to incorporate them formally. The consequences that the organism experiences may change the pay-off function—it doesn’t know how well it likes cheese until it has eaten cheese. Likewise, one method for refining the mapping of A on S may be to select a particular alternative and experience its consequences. In these cases, one of the elements of the pay-off associated with a particular alternative is the information that is gathered about the mapping or about the pay-off function.

V. Conclusion

The aim of this paper has been to construct definitions of “rational choice” that are modeled more closely upon the actual decision processes in the behavior of organisms than definitions heretofore proposed. We have outlined a fairly complete model for the static case, and have described one extension of this model into dynamics. As has been indicated in the last section, a great deal remains to be done before we can handle realistically a more completely dynamic system.

In the introduction, it was suggested that definitions of this kind might have normative as well as descriptive value. In particular, they may suggest approaches to rational choice in areas that appear to be far beyond the capacities of existing or prospective computing equipment. The comparison of the I.Q. of a computer with that of a human being is very difficult. If one were to factor the scores made by each on a comprehensive intelligence test, one would undoubtedly find that in those factors on which the one scored as a genius the other would appear a moron—and conversely. A survey of possible definitions of rationality might suggest directions for the design and use of computing equipment with reasonably good scores on some of the factors of intelligence in which present computers are moronic.

The broader aim, however, in constructing these definitions of “approximate” rationality is to provide some materials for the construction of a theory of the behavior of a human individual or of groups of individuals who are making decisions in an organization context. The apparent paradox to be faced is that the economic theory of the firm and the theory of administration attempt to deal with human behavior in situations in which that behavior is at least “intendedly” rational; while, at the same time, it can be shown that if we assume the global kinds of rationality of the classical theory the problems of internal structure of the firm or other organization largely disappear.¹⁰ The paradox vanishes, and the outlines of theory begin to emerge when we substitute for “economic man” or “administrative man” a choosing organism of limited knowledge and ability. This organism’s simplifications of the real world for purposes of choice introduce discrepancies between the simplified model and the reality; and these discrepancies, in turn, serve to explain many of the phenomena of organizational behavior.

VI. Appendix: Example of Rational Determination of an Acceptable Pay-off

In the body of this paper, the notion is introduced that rational adjustment may operate at various “levels.” That is, the organism may choose rationally within a given set of limits postulated by the model, but it may also undertake to set these limits rationally. The house-selling illustration of Section IIA provides an example of this.

We suppose that an individual is selling a house. Each day (or other unit of time) he sets an acceptance price: $d(k)$, say, for the k th day. If he receives one or more offers above this price on the day in question, he accepts the highest offer; if he does not receive an offer above $d(k)$, he retains the house until the next day, and sets a new acceptance price, $d(k+1)$.

Now, if he has certain information about the probability distribution of offers on each day, he can set the acceptance price so that it will be optimal in the sense that it will maximize the expected value, $V[d(k)]$, of the sales price.

To show this, we proceed as follows. Let $p_k(y)$ be the probability that y will be the highest price offered on the k th day. Then:

$$P_k(d) = \int_{d(k)}^{\infty} p_k(y) dy \quad (\text{A.1})$$

is the probability that the house will be sold on the k th day if it has not been sold earlier.

$$E_k(d) = \int_{d(k)}^{\infty} yp(y, k) dy \quad (\text{A.2})$$

will be the expected value received by the seller on the k th day if the house has not been sold earlier. Taking into account the probability that the house will be sold before the k th day,

$$E_k(d) = E_k(d) \prod_{j=1}^{k-1} (1 - P_j(d)) \quad (\text{A.3})$$

will be the unconditional expected value of the seller's receipts on the k th day; and

$$V\{d(k)\} = \sum_{k=1}^{\infty} E_k(d) \quad (\text{A.4})$$

will be the expected value of the sales price.

Now we wish to set $d(k)$, for each k , at the level that will maximize (A.4). The k components of the function $d(k)$ are independent. Differentiating V partially with respect to each component, we get:

$$\frac{\partial V}{\partial d(i)} = \sum_{k=1}^{\infty} \frac{\partial E_k(d)}{\partial d(i)} \quad (i = 1, \dots, n) \quad (\text{A.5})$$

But:

$$\frac{\partial E_i(d)}{\partial d(i)} = \frac{\partial E_i(d)}{\partial d(i)} \prod_{j=1}^{i-1} (1 - P_j(d)), \quad (\text{A.6})$$

and

$$\frac{\partial E_k(d)}{\partial d(i)} = E_k(d) \prod_{j=1}^{k-1} (1 - P_j(d)) \left(-\frac{\partial P_i(d)}{\partial d(i)} \right) \quad \text{for } i < k \quad (\text{A.7})$$

and

$$\frac{\partial E_k(d)}{\partial d(i)} = 0 \quad \text{for } i > k. \quad (\text{A.8})$$

Hence for a maximum:

$$\frac{\partial V}{\partial d(i)} = -d(i)p_i(d) \prod_{j=1}^{i-1} (1 - P_j(d)) + \sum_{k=i+1}^{\infty} E_k(d) \prod_{j=1}^{k-1} (1 - P_j(d)) p_i(d) = 0 \quad (\text{A.9})$$

Factoring out $p_i(d)$, we obtain, finally:

$$d(i) = \frac{\sum_{k=i+1}^{\infty} E_k(d) \prod_{j=1}^{k-1} (1 - P_j(d))}{\prod_{j=1}^{i-1} (1 - P_j(d))} = \sum_{k=i+1}^{\infty} E_k(d) \prod_{j=i+1}^{k-1} (1 - P_j(d)). \quad (\text{A.10})$$

For the answer to be meaningful, the infinite sum in (A.10) must converge. If we look at the definition (A.2) for $E_k(d)$ we see this would come about if the probability of offers shifts downward through time with sufficient rapidity. Such a shift might correspond to (a) expectations of falling prices, or (b) interpretation of y as the *present value* of the future price, discounted at a sufficiently high interest rate.

Alternatively, we can avoid the question of convergence by assuming a reservation price $a(n)$, for the n th day, which is low enough so that $P_n(d)$ is unity. We shall take this last alternative, but before proceeding, we wish to interpret the equation (A.10). Equation (A.10) says that the rational acceptance price on the i th day, $d(i)$, is equal to the expected value of the sales price if the house is not sold on the i th day and acceptance prices are set optimally for subsequent days. This can be seen by observing that the right-hand side of (A.10) is the same as the right-hand side of (A.4) but with the summation extending from $k = (i + 1)$ instead of from $(k = 1)$.¹¹

Hence, in the case where the summation is terminated at period n —that is, the house will be sold with certainty in period n if it has not been sold previously—we can compute the optimal $d(i)$ by working backward from the terminal period, and without the necessity of solving simultaneously the equations (A.10).

It is interesting to observe what additional information the seller needs in order to determine the rational acceptance price, over and above the information he needs once the acceptance price is set. He needs, in fact, virtually complete information as to the probability distribution of offers for all relevant subsequent time periods.

Now the seller who does not have this information, and who will be satisfied with a more bumbling kind of rationality, will make approximations to avoid using the information he doesn't have. First, he will probably limit the planning horizon by assuming a price at which he can certainly sell and will be willing to sell in the n th time period. Second, he will set his initial acceptance price quite high, watch the distribution of offers he receives, and gradually and approximately adjust his acceptance price downward or upward until he receives an offer he accepts—without ever making probability calculations. This, I submit, is the kind of rational adjustment that humans find “good enough” and are capable of exercising in a wide range of practical circumstances. ▼

-
- 1 Herbert Simon, *A Behavioral Model of Rational Choice*, 69 (1) Q. J. OF ECON. 99-118 (Feb.1955). Reprinted by permission from MIT Press Journals.
 - 2 The ideas embodied in this paper were initially developed in a series of discussions with Herbert Bohnert, Norman Dalkey, Gerald Thompson, and Robert Wolfson during the summer of 1952. These collaborators deserve a large share of the credit for whatever merit this approach to rational choice may possess. A first draft of this paper was prepared in my capacity as a consultant to the RAND Corporation. It has been developed further (including the Appendix) in work with the Cowles Commission for Research in Economics on “Decision Making Under Uncertainty,” under contract with the Office of Naval Research, and has been completed with the aid of a grant from the Ford Foundation.
 - 3 See Kenneth J. Arrow, *Alternative Approaches to the Theory of Choice in Risk-Taking Situations*, *ECONOMETRICA*, XIX, 404-37 (Oct. 1951).
 - 4 See the Appendix. It might be remarked here that the simple risk function, introduced by Wald to bring problems in statistical decision theory within the bounds of computability, is an example of a simple pay-off function as that term is defined here.
 - 5 This procedure also dispenses with the necessity of estimating explicitly the cost of obtaining additional information. For further discussion of this point see the comments on dynamics in the last section of this paper.
 - 6 I have estimated roughly the actual degree of simplification that might be realized in the middle game in chess by experimentation with two middle-game positions. A sequence of sixteen moves, eight by each player, might be expected to yield a total of about 10^{24} (one septillion) legally permissible variations. By following the general kind of program just described, it was possible to reduce the number of lines of play examined in each of these positions to fewer than 100 variations—a rather spectacular simplification of the choice problem.
 - 7 *ECONOMETRICA*, XIX (July 1951), 292-305 and *REVIEW OF ECONOMIC STUDIES*, XX (1952-53, No. 1), 40-49.
 - 8 I might mention that, in the spirit of crude empiricism, I have presented a number of students and friends with a problem involving a multiple pay-off—in which the pay-off depends violently upon a very contingent and uncertain event—and have found them extremely reluctant to restrict themselves to a set of behavior alternatives allowed by the problem. They were adverse to an alternative that promised very large profit or ruin, where the relevant probability could not be computed, and tried to invent new alternatives whose pay-offs were less sensitive to the contingent event. The problem in question is Modigliani's “hot-dog stand” problem described in *AMERICAN ECONOMIC REVIEW*, proceedings, XXXIX (1949), 201-8.

- 9 One might add: "or judged in terms of the survival value of its choice mechanism."
- 10 See HERBERT A. SIMON, *ADMINISTRATIVE BEHAVIOR*, 39-41, 80-84, 96-102, 240-44 (1947).
- 11 Equation (A.10) appears to have been arrived at independently by D.A. Darling and W.M. Kincaid. See their abstract, *An Inventory Problem*, in the *JOURNAL OF OPERATIONS RESEARCH SOCIETY OF AMERICA*, I, 80 (Feb. 1953).