Michigan Law Review

Volume 101 | Issue 2

2002

The Fable of Entry: Bounded Rationality, Market Discipline, and Legal Policy

Avishalom Tor
Boston University School of Law

THE FABLE OF ENTRY: BOUNDED RATIONALITY, MARKET DISCIPLINE, AND LEGAL POLICY

*Avishalom Tor**

TABLE OF CONTENTS

* Lecturer, Boston University School of Law. Fellow, Olin Center for Law, Economics, and Business, Harvard Law School. Fellow, Program on Negotiation, Harvard Law School. LL.B., B.A. 1996, Hebrew University of Jerusalem; LL.M. 1998, S.J.D. candidate, Harvard. — Ed. This research was supported by fellowships from the Olin Center for Law, Economics, and Business, and the Program on Negotiation, Harvard Law School. I thank Jennifer Arlen, Max Bazerman, Assaf Hamdani, Louis Kaplow, Sarah Levine, Paul Slovic, and the editors for valuable comments. Jamie Byrnes provided helpful research assistance. Special thanks to Christine Jolls for providing extensive and most valuable comments on this Article.

INTRODUCTION

Legal scholars have recently advanced a behavioral approach to the law and economics school of thought in an attempt to improve its external validity and predictive power.[1] The hallmark of this new approach is the replacement of the perfectly rational actor with a "boundedly rational" decisionmaker who, apart from being affected by emotion and motivation, has only limited cognitive resources.[2] To function effectively in a complex world, boundedly rational individuals must rely on cognitive heuristics — simplifying mental shortcuts — that inevitably lead people to make some systematic decision errors; as a result, their behavior necessarily deviates from that predicted by rational actor models.[3]

---

1. Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998) (offering a broad vision of how law and economics could be improved by increasing its attention to insights about actual human behavior); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051 (2000) (examining the role of the rational actor in law and economics and suggesting replacing it with a behaviorally informed actor).

2. The term "bounded rationality" is used here broadly, encompassing the major findings of behavioral decision research over the last thirty years on the deviations of actual human behavior from rational models due to the limitations of human information processing abilities and the effects of motivation and emotion on human cognition. For instructive reviews of this enormous literature, see, for example, Colin Camerer, *Individual Decision Making*, *in* 1 THE HANDBOOK OF EXPERIMENTAL ECONOMICS 587 (John H. Kagel & Alvin E. Roth eds., 1995); Robyn M. Dawes, *Behavioral Decision Making and Judgment*, *in* THE HANDBOOK OF SOCIAL PSYCHOLOGY 497 (Daniel T. Gilbert et al. eds., 4th ed. 1998). The concept of bounded rationality was originally developed by Herbert A. Simon. *See* Herbert A. Simon, *A Behavioral Model of Rational Choice*, 69 Q.J. ECON. 99 (1956); Herbert A. Simon, *Rational Choice and the Structure of the Environment*, 63 PSYCHOL. REV. 129 (1958); *see also* HERBERT A. SIMON, REASON IN HUMAN AFFAIRS 17-23 (1983). In Simon's terminology, however, bounded rationality denoted only the *cognitive* limitations of the human mind. Contrast Jolls et al.'s depiction of bounded rationality as only one of three "bounds" differentiating actual human behavior from the rational actor (the other two being bounded self-interest and bounded willpower, but without reference to the broader effects of motivation and emotion on cognition). Jolls et al., *supra* note 1, at 1476-79.

3. *See, e.g.*, Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124, 1124 (1974), *reprinted in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 3 (Daniel Kahneman et al. eds., 1982) (stating, in an early formulation of the authors' highly influential "heuristics and biases" research paradigm, that: "[P]eople rely on a limited number of heuristic principles which reduce the complex task of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors."). Thus, Jolls et al. explain:

> Bounded rationality . . . refers to the obvious fact that human cognitive abilities are not infinite . . . . [P]eople sometimes respond rationally to their own cognitive limitations . . . . [b]ut even with these remedies, and in some cases because of these remedies, human behavior differs in systematic ways from that predicted by the standard economic model of unbounded rationality. Even when the use of mental shortcuts is rational, it can produce predictable mistakes.

Jolls et al., *supra* note 1, at 1477.

In response, advocates of traditional law and economics have argued that boundedly rational behavior is of little significance for the analysis of economic activities in market environments, most notably because competitive pressures will discipline such behavior. According to this view, the boundedly rational will underperform and consequently fail and exit the market.[4] Some cautious supporters of the behavioral approach have been quick to agree that bounded rationality is of limited importance for the analysis of market behaviors because of competitive discipline.[5] In fact, even the most influential proponents of behavioral law and economics have found it necessary to state, "law is a domain where behavioral analysis would appear to be particularly promising in light of the fact that *nonmarket behavior* is frequently involved."[6]

This Article heartily agrees that behavioral insights are highly applicable to those numerous areas in the law addressing nonmarket behaviors. Questioning the accepted wisdom on market discipline, however, it argues that advocates of the behavioral approach have conceded too much too quickly. It shows in the context of new entry into industry that, although intense competition eliminates many boundedly rational actors, competitive forces inevitably select some other such actors for success. Consequently, and because of their very large numbers, boundedly rational actors become overrepresented, *as a group*, among the ranks of successful entrants. In other words, this Article does not argue against the existence of market discipline; instead, it highlights how competitive forces unexpectedly facilitate bounded rationality in the market. The profound role of boundedly rational action in markets therefore renders its understanding supremely important for the legal regulation of economic phenomena. A study of the competition for profitability and survival among new entrants into industry thus highlights the unique contribution a behaviorally informed approach stands to make to legal and economic scholarship writ large, while shedding new light on the important topic of entry competition specifically.

4. *See, e.g.*, Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L. REV. 1551, 1570-71 (1998) (arguing that "selection effects" in markets largely discipline boundedly rational behavior and therefore render experimental findings on human decisionmaking largely irrelevant); *see also* RICHARD A. POSNER, FRONTIERS OF LEGAL THEORY 280 (2001) [hereinafter POSNER, FRONTIERS] (same).

5. *E.g.*, Thomas S. Ulen, *The Growing Pains of Behavioral Law and Economics*, 51 VAND. L. REV. 1747, 1748-49, 1758-60 (1998); *see also* Jennifer Arlen, Comment, *The Future of Behavioral Economic Analysis of Law*, 51 VAND. L. REV. 1765, 1782 (1998) (suggesting that behavioral findings from nonmarket settings may not necessarily generalize to market settings).

6. Jolls et al., *supra* note 1, at 1473 (emphasis added).

New entry into industry has been a major topic of analysis in both legal and nonlegal literatures.[7] This Article offers a behavioral analysis of entrant decisionmaking that explains many empirical findings about entry that appear highly puzzling and seemingly unrelated from a traditional economic perspective. Part I describes these phenomena, which include the prevalence of excess entry, the relative insensitivity of entrants to market predictors of success, and the inferior performance of startup entrants as compared to those who enter by diversification. This Article reveals how these phenomena all result from the operation of the various processes of entrant overconfidence.[8]

On the most basic level, the processes of overconfidence — most notably, optimistic and desirability-related biases — explain the puzzling empirical phenomenon of excess entry, wherein the high rate of entry appears economically inexplicable in the face of its low expected value.[9] This conclusion is the main theme of an important recent experimental study in the nonlegal literature of the relationship between overoptimism and excess entry.[10]

Notwithstanding the findings of this leading study, the literature on entry, both legal and nonlegal, has thus far failed to examine whether, beyond revealing the basic fact of excess entry, the empirical data corroborate or contradict the claim of entrant overconfidence. By engaging in a closer analysis of both the economics and the psychology of entry, this Article breaks new ground in two important ways. First, it links additional puzzling empirical findings — such as entrants' relative insensitivity to entry barriers and other market predictors of success and, most importantly, the lower survival rates and inferior average performance of startups as compared to diversifying entrants — to

---

7. *E.g.*, 1 PHILLIP E. AREEDA & HERBERT J. HOVENKAMP, ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION ¶ 112b (2d ed. 2000) (discussing the important role entry plays in numerous antitrust doctrines); 2A PHILLIP E. AREEDA & HERBERT J. HOVENKAMP, ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION ¶ 420b (2d ed. 2002); Richard J. Gilbert, *The Role of Potential Competition in Industrial Organization*, 3 J. ECON. PERSP. 107 (1989) (surveying the main roles of entry in economic theory).

8. For the sake of clarity, this Article distinguishes between overconfidence — a characterization of certain behaviors of entrants or other decisionmakers — and the psychological processes that generate manifestations of overconfidence. The latter are referred to either globally, as "the processes of overconfidence," or individually, by the term or terms assigned to them in the behavioral literature.

9. *See* Giovanni Dosi & Dan Lovallo, *Rational Entrepreneurs or Optimistic Martyrs? Some Considerations on Technological Regimes, Corporate Entries, and the Evolutionary Role of Decision Biases*, *in* TECHNOLOGICAL INNOVATION: OVERSIGHTS AND FORESIGHTS 41 (Raghu Garud et al. eds., 1997) [hereinafter Dosi & Lovallo, *Rational Entrepreneurs*] (reporting findings of excess entry in experimental games and suggesting, more generally, that various decision biases may play a role in generating excess entry); *see also infra* Section I.B.1.

10. *See* Colin Camerer & Dan Lovallo, *Overconfidence and Excess Entry: An Experimental Approach*, 89 AM. ECON. REV. 306 (1999). For further discussion of this study, see *infra* text accompanying notes 101-103.

the same psychological processes of entrant overconfidence that explain the basic fact of excess entry. Second, it provides a unified framework for a new understanding of the various aspects of entry decisionmaking, its economic consequences, and its implications for legal policy.

Part II shows how, from a behavioral perspective, deciding whether to attempt new entry is, first and foremost, an *investment decision with significant personal stakes that must be made under conditions of extreme uncertainty.* In these circumstances, entrants are likely to be overconfident in their prospects, thereby making excessive entry attempts. Moreover, the behavioral analysis of entry shows that entrants' insensitivity to pale, if potentially important, background market, whose impact on their fate is statistical and indirect, is a likely by-product of many of the processes of entrant overconfidence. A close analysis of these processes also exposes important and hitherto unnoticed behavioral differences between startups and diversifying entrants, differences that closely follow the economic differences between these two entrant types. The effects of two important psychological variables determining the extent of entrant overconfidence — the intensity of preferences and the ambiguity of the decision environment — are likely to cause startup entrants to exhibit a greater bias than diversifying ones. This novel finding explains the perplexing inferior average performance of startups.

Part III reveals that the same competitive pressures that weed out many overconfident entrants also generate a post-entry landscape in which the more extremely biased startups are over-represented, at the expense of some of their less-biased, ex ante better-qualified, competitors. After exploring why the private action of financiers fails to curb entrant overconfidence significantly, this Part considers the desirability and possibility of government intervention.

In Part IV, a careful study of the social costs and benefits of boundedly rational entry and the possibility of implementing an effective policy on this matter reveals, however, that the governmental regulation of entry is probably undesirable. An effort to regulate boundedly rational entry is impractical because of the difficulty of identifying negative expected value entry, determining when exactly to limit it at the margin, and finding an effective means of actually doing so. This Article finds, moreover, that although overconfident entrants generate social and private losses, they also fulfill an important economic function, serving as the voluntary cannon fodder of innovation. These entrants, even when they fail and more so when they succeed, facilitate economic growth and expand the range of consumer choice. Equally important, they make a significant long-term contribution to the disciplining of incumbent firms.

This analysis also has important implications for antitrust law. The Article suggests that the role assigned to entry barriers in judicial and

regulatory determinations of market power should be modified, because even when such barriers create only a limited impediment to entry they may still protect incumbents by making post-entry survival less likely. Relying on the insights of the behavioral analysis of entry, the Article reevaluates legal doctrine in two important areas of antitrust law, suggesting modifications in the rules concerning predatory pricing by dominant firms and horizontal mergers between competitors.

Finally, Part V highlights the profound impact of bounded rationality on market behavior and market outcomes outside the domain of antitrust law. The Article concludes by discussing the nature of and providing specific examples for the unique contribution a psychologically informed approach to legal scholarship stands to make to the legal analysis of market and nonmarket behaviors alike.

I. THE PUZZLES OF ENTRY

Current economic theories of entry commonly assume that entrants are rational profit maximizers. However, the evidence from the field of industrial organization regarding entry, exit, entrant market penetration, and related economic variables appears to contradict the assumption of entrant rationality. This Part examines the apparent tension between the empirical findings and the economic theory of entry, showing it to span a number of related phenomena, including the prevalence of excess entry, the relative insensitivity of entrants to market predictors of future profitability, and the inferior average performance of startup entrants. After exploring these puzzling findings, the Part concludes by showing that attempts to explain why entry may still be rational are not very compelling and fall short of accounting for the full magnitude and range of the various puzzles, although they may provide a partial account for some of the empirical evidence.

A. *The Assumption of Entrant Rationality: Empirical Findings vs. Economic Theory*

The dynamics of competition among new entrants into industry[11] are an important subject of economic analysis: the existence of actual,

11. This Article defines entry as the construction of a new plant by a firm that was not previously manufacturing in the particular industry. While other definitions exist (e.g., plant-level entry that counts as entrants new plants constructed by firms already manufacturing in the industry, or firm-level entry that counts as entrants firms who enter by acquisition of incumbents or their existing plants without adding capacity to the industry), the one used here is both common in the literature and highlights the relationship between the number of new decisionmakers in the industry and production capacity, which is foundational to the economic theories of entry examined in this Article. *See, e.g.*, JOHN R. BALDWIN, THE DYNAMICS OF INDUSTRIAL COMPETITION: A NORTH AMERICAN PERSPECTIVE 11-12 (1995) (explaining the importance of both firm-level and plant-level information); Timothy

and even potential, entry exerts competitive pressure on extant firms in an industry.[12] In antitrust law, for instance, the idea "that unnecessary restrictions on new business entry are socially harmful" is a basic proposition.[13] Examples of this important role include how the law infers that firms enjoy market power from observations of market concentration only when barriers to entry are high, and how the conduct of firms that creates an unjustified barrier to entry may be deemed illegal.[14]

Importantly, current theories of entry assume that entrants are *rational decisionmakers who will attempt entry only if it is profit maximizing*.[15] The assumption of profit maximization is therefore commonly relied upon in the many antitrust doctrines founded upon economic theories of entry.[16] Profit maximization is accomplished by maximizing the net present value ("NPV") of investments.[17] If entrants were to follow the economic theory of investment decisionmaking, they would attempt entry only if its expected returns — where the values of outcomes are multiplied by their respective probabilities — were to exceed the necessary investment to enter the industry after

---

Dunne et al., *Patterns of Entry and Exit in U.S. Manufacturing Industries*, 19 RAND J. ECON. 495, 500-01 (1988) [hereinafter Dunne et al., *Patterns*] (a leading, longitudinal empirical study of entry and exit using a similar definition).

12. This effect was observed a century ago by JOHN BATES CLARK, THE CONTROL OF TRUSTS: AN ARGUMENT IN FAVOR OF CURBING THE POWER OF MONOPOLY BY A NATURAL METHOD 13 (1901); *see also* GEORGE J. STIGLER, THE ORGANIZATION OF INDUSTRY 18 (1968) (a more modern version of this argument).

13. *See, e.g.*, 1 AREEDA & HOVENKAMP, *supra* note 7, ¶112b, at 120.

14. 2A AREEDA & HOVENKAMP, *supra* note 7, ¶420b, at 60.

15. Specifically, the "structuralist" school of industrial organization pictured a world in which the rate of entry is a function of its profitability. JOE S. BAIN, BARRIERS TO NEW COMPETITION 4-5, 11-19 (1956). Profit maximization also underlies the classical theory of "limit pricing" — the notion that established firms may deter entry by cutting prices before entry occurs and thereby signal that post-entry prices will not be sufficiently profitable to justify rational entry. Franco Modigliani, *New Developments on the Oligopoly Front*, 66 J. POL. ECON. 215 (1958). Similarly, one of the most influential theories regarding the effect of potential entry on the market — Baumol's theory of contestable markets — is founded on the notion that entrants will enter if and only if entry is profitable. WILLIAM J. BAUMOL ET AL., CONTESTABLE MARKETS AND THE THEORY OF INDUSTRY STRUCTURE (1982); William J. Baumol, *Contestable Markets: An Uprising in the Theory of Industry Structure*, 72 AM. ECON. REV. 1 (1982). Last, like their neoclassical counterparts, modern game theoretic models of entry decisionmaking used by scholars of the New Industrial Organization also assume profit maximization. *See, e.g.*, JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION 34-35 (1988).

16. 1 AREEDA & HOVENKAMP, *supra* note 7, ¶113, at 137 (stating that "[a]s a general proposition business firms are (or must be assumed to be) profit-maximizers"); 2A AREEDA & HOVENKAMP, *supra* note 7, ¶422a, at 71 (stating, when discussing the likelihood of entry, that " 'likely' generally means 'profitable,' for entry *will not occur in the absence of expected profits, after taking all costs and risks into account*") (emphasis added).

17. RICHARD A. BREALEY & STEWART C. MYERS, PRINCIPLES OF CORPORATE FINANCE 11-28, 85-108 (5th ed. 1996).

being discounted for risk.[18] In other words, rational entrants would not embark on negative expected value ventures.[19]

Although entrants are assumed to maximize the expected value of entry,[20] the large body of empirical data on entry reveals a set of puzzling phenomena that appear to contradict this assumption of entrant rationality.

1. *The First Puzzle: The Prevalence of Negative Net Present Value Entry*

The empirical evidence strongly suggests that negative net present value entry is commonplace. First, while entry is pervasive, amounting on average to about 50% of all existing firms every five years across all domestic manufacturing industries,[21] entrants also exhibit strikingly high mortality rates. Within ten years, only about 20% of any entrant cohort still operates. Attrition, moreover, begins right from the start, with more than 25% of new entrants exiting within two years, over 60% disappearing within five years.[22] In fact, high-volume exit accom-

18. *See, e.g., id.* at 12-17, 179-83, 204-29.

19. *See, e.g.*, ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 26 (3d ed. 2000) (noting that in decisions involving monetary outcomes economists assume decision-makers are risk neutral or, at times, risk averse); A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS 51 (1983) (same).

20. The terms "expected value" and "net present value" are used interchangeably.

21. *E.g.*, Dunne et al., *Patterns*, *supra* note 11, at 497 & n.4 (the average rate of gross entry in the United States during the period 1963-1982 is greater than 10% per year, amounting to more than 25,000 annual new entrants). International comparisons, especially from other industrialized countries, report high rates of gross entry as well. *See, e.g.*, P.A. Geroski, *Domestic and Foreign Entry in the United Kingdom*, *in* ENTRY AND MARKET CONTESTABILITY: AN INTERNATIONAL COMPARISON 63, 64, 76 (P.A. Geroski & Joachim Schwalbach eds., 1991) [hereinafter Geroski, *Domestic and Foreign Entry*] (United Kingdom data); Joachim Schwalbach, *Entry, Exit, Concentration, and Market Contestability*, *in* ENTRY AND MARKET CONTESTABILITY, *supra*, at 121, 121-22 (German data). *But see* BALDWIN, *supra* note 11, at 17, 401-02 (reporting somewhat smaller figures in a study disregarding those small firms that together account for 40%-54% of all manufacturing establishments).

22. David B. Audretsch & Talat Mahmood, *The Post-Entry Performance of New Firms*, *in* MARKET EVOLUTION: COMPETITION AND COOPERATION 245, 250, tbl.1 (Arjen van Witteloostuijn ed., 1995) [hereinafter Audretsch & Mahmood, *Post-Entry Performance*] (analysis of data in table); Dunne et al., *Patterns*, *supra* note 11, at 509, tbl.8; Geroski, *Domestic and Foreign Entry*, *supra* note 21, at 79 (reporting even more striking figures from the U.K. between 1974-1982: "Roughly 12.4 percent of entrants survived no longer than 6 months, 27.3 percent no longer than a year, 55 percent no longer than 2 years, and roughly 85 percent no longer than 4 years. *Only 0.1 percent of the cohort of 1974 entrants were still operating in 1982.*") (emphasis added); *see also* Arnold C. Cooper et al., *Entrepreneurs' Perceived Chances for Success*, 3 J. BUS. VENTURING 97, 99 (1988) [hereinafter Cooper et al., *Entrepreneurs' Perceived Chances*] (citing earlier studies reporting extremely high failure rates).

panies the high volume of entry in most industries, such that the two phenomena are strongly correlated,[23] and result in little net[24] entry.[25]

Second, the limited success of entrants is even more apparent from market penetration figures.[26] When measured by either output or employment, the share of new entrants in the industry is even smaller than their numbers suggest; it is nearly negligible in the short term — rarely amounting to more than a few decimal points and often negative.[27] This minimal penetration reveals, importantly, that the success of those entrants who survive and grow does not compensate sufficiently for the strong effect of their peers' extremely high attrition rate, a conclusion that studies following specific cohorts of entrants longitudinally corroborate.[28] These findings of high mortality and low penetration also suggest that most entrants simply displace preceding ones rather than diminish the market share of incumbents,[29] leading

23. *See, e.g.*, Dunne et al., *Patterns*, *supra* note 11, at 507-08, tbls.7-8; *see also* Geroski, *Domestic and Foreign Entry*, *supra* note 21, at 77 (U.K. data); Schwalbach, *supra* note 21, at 123 (German data).

24. Net entry is calculated by deducting the number of firms exiting the industry in a particular year from the number of firms entering it. References below will generally be to *gross* entry, however, which reflects actual entry rates. *See, e.g.*, BALDWIN, *supra* note 11, at 12 (noting that net entry measures expansion rather than entry, underestimating the amount of entry by the amount of exit). *But see* Jose Mata, *Sunk Costs and the Dynamics of Entry in Portuguese Manufacturing*, *in* MARKET EVOLUTION, *supra* note 22, at 270 (discussing the limits of gross entry measures).

25. *E.g.*, Dunne et al., *Patterns*, *supra* note 11, at 503, tbl.2 (exit rates averaging 95% of the entry rates); *id.* at 506, tbl.5 (the similarity appears at all levels: the particular industry, the industrial sector, and all manufacturing industries together). These figures are representative of net entry findings in numerous other studies. *See, e.g.*, P.A. Geroski, *What Do We Know About Entry?*, 13 INT'L J. INDUS. ORG. 421, 423 (1995) [hereinafter Geroski, *What Do We Know*] (concluding, in a recent review of the empirical findings on entry, that "[e]ntry and exit rates are highly positively correlated, and net entry rates and penetration are modest fractions of gross entry rates and penetration").

26. "Entry is common. Large numbers of firms enter most markets in most years, *but entry rates are far higher than market penetration rates*." Geroski, *What Do We Know*, *supra* note 25, at 422 (emphasis added).

27. Dunne et al., *Patterns*, *supra* note 11, at 504 tbl.3; *id.* at 505, tbl.4 (analysis of data in tables yielding a net negative market share penetration of 0.1%); *see also* BALDWIN, *supra* note 11, at 16 tbl.2-2 (Canadian data showing a negative penetration of 0.3% employment, with the best of the twelve years reported showing a positive net penetration of 0.3%); John Cable & Joachim Schwalbach, *International Comparisons of Entry and Exit*, *in* ENTRY AND MARKET CONTESTABILITY, *supra* note 21, at 256, 260 tbl.14.2 (a review of eight, mostly international studies, showing similar findings); Geroski, *Domestic and Foreign Entry*, *supra* note 21, at 76 (finding an almost "negligible" net market share penetration in the UK).

28. Dunne et al. show how the already limited market share of entrants further shrinks with time when all entrants are included. Dunne et al., *Patterns*, *supra* note 11, at 509, tbl.8. *But see* BALDWIN, *supra* note 11, at 21-23 & n.14 (asserting his Canadian data show an *increase* in entrant cohorts' value-added share (indexed on entrants' initial share), but accomplishing this feat only by cumulating the data of successive cohorts and using only about the larger half of the entrant population. Even then, moreover, he admits that "if employment were used rather than value-added, the results would show a decline in share after several years . . .").

29. *See, e.g.*, Cable & Schwalbach, *supra* note 27, at 266.

scholars to term post-entry competition as "turbulence," "churning," or a "revolving door" phenomenon.[30]

Scholars familiar with these striking findings could not fail to observe how entry seems excessive given poor rates of success.[31] With high attrition and many costly years to profitability, entrants would have to expect much higher returns to success than they appear to enjoy, on average, to be making only rational entry attempts.[32]

2. *The Second Puzzle: Entrants' Insensitivity to Predictors of Future Profitability*

In addition to the basic puzzle of excess entry, other puzzling findings on the effects of economic variables that should facilitate or inhibit entry remain largely unexplained by the literature. Specifically, rational entry should be proportional to its anticipated profitability. The presence of factors that rationally indicate lower profits or diminish the probability of success — such as entry barriers, strategic deterrence, or the anticipated intensity of competition — should inhibit entry. The data, however, portray a picture of entrant behavior that clearly differs from these expectations.[33]

The behavior of entrants is not completely irrational: they are attracted to markets exhibiting growth, a factor associated with future

---

30. BALDWIN, *supra* note 11, at 359 ("churning"); Laurie Beth Evans & John J. Siegfried, *Entry and Exit in United States Manufacturing Industries from 1977 to 1982*, *in* EMPIRICAL STUDIES IN INDUSTRIAL ORGANIZATION: ESSAYS IN HONOR OF LEONARD W. WEISS 253, 254 (1992) ("revolving door"); P.A. Geroski, *Some Data-Driven Reflections on the Entry Process*, *in* ENTRY AND MARKET CONTESTABILITY, *supra* note 21, at 282, 295 [hereinafter Geroski, *Data-Driven Reflections*] ("turbulence").

31. *See, e.g.*, BALDWIN, *supra* note 11, at 359; Geroski, *Data-Driven Reflections*, *supra* note 30, at 282, 295.

32. As pointed out by Camerer and Lovallo, however, "even if cumulative industry profits are actually negative at some point in time, it is possible that positive returns will roll in later .... So it is hard to imagine how to establish conclusively that *expected* industry returns were negative." Camerer & Lovallo, *supra* note 10, at 307. In fact, Schumpeter already pointed to a similar ambiguity when discussing the possibility that entry may bring "negative return[s] to entrepreneurs ... as a group," explaining, "[w]hether this actually is so in any particular case is, of course, extremely difficult to establish." Joseph A. Schumpeter, *The Creative Response in Economic History*, 7 J. ECON. HIST. 149, 156 & n.14 (1947).

33. While the evidence cited below supports the claim that entrants are not rational decisionmakers, a finding that the rate of entry does vary with those factors rationally related to entrants' objective prospects would only suggest that entrants' judgments are *coherent*; it would not show, however, that the "baseline" rate of entry is itself rational. Entrants could respond to differences in the attractiveness of entry, while still being strongly biased overall. *Cf.* Dan Ariely et al., "Coherent Arbitrariness": Stable Demand Curves Without Stable Preferences 3-4 (Nov. 2001) (unpublished manuscript, on file with author) (providing evidence that consumers' absolute valuations of pain are arbitrary, even while their relative valuations of stimuli are internally coherent, and arguing that such data cast doubt on whether data showing that *changes* in circumstances cause theoretically predicted changes in valuation can support the claim that individuals' *fundamental* valuations are rational).

survival,[34] although they are also attracted to industries with a more volatile rate of growth, a factor suggesting higher risk.[35] It is unclear, moreover, whether industries with higher past profitability or profitability growth — both of which are associated with future profitability — attract more entry.[36] In any case, entrants seem slow in reacting to high profits,[37] and to the extent entrants attempt entry more frequently in high profit industries, they seem to ignore the anticipated increase in competition from other similarly attracted entrants that they will likely face upon entry.[38]

In addition to showing insensitivity to the expected intensity of competition in high-profit industries, entrants also disregard many entry barriers.[39] They are not significantly deterred from industries where capital intensity and scale economics play an important role,[40]

34. *See, e.g.*, BALDWIN, *supra* note 11, at 366-72 (finding industry growth associated with an increase in gross entry but not gross exit rates); Timothy Dunne & Mark J. Roberts, *Variation in Producer Turnover Across U.S. Manufacturing Industries*, *in* ENTRY AND MARKET CONTESTABILITY, *supra* note 21, at 187, 191, 193-94 [hereinafter Dunne & Roberts, *Variation*] (finding growth associated with increased entry rates and decreased exit rates). *But see* Geroski, *Data-Driven Reflections*, *supra* note 30, at 81 (reporting variations in industry growth rates have little effect on entry flows).

35. *See* BALDWIN, *supra* note 11, at 366-72 (finding the increase in entry rate in industries with greater variability robust, without clear evidence of increased survival rates).

36. Evans & Siegfried, *supra* note 30, at 265-66 (finding that all entrants are not more likely to enter industries with higher profitability, and that historical growth is unrelated to startup entry and citing various studies); *see also* BALDWIN, *supra* note 11, at 368-78 (finding, using a variety of estimation procedures, that neither past profits nor profitability growth is significantly associated with entry rates). *But see* Dunne & Roberts, *supra* note 34, at 194 (finding that price-cost margins are positively correlated with entry rates); Geroski, *Data-Driven Reflections*, *supra* note 30, at 81 (finding effect of expected profits on gross entry in U.K. study).

37. *See* Geroski, *What Do We Know*, *supra* note 25, at 427.

38. Because industries with higher rates of entry also show higher rates of exit, they exhibit little *net* entry (i.e., entrants into these industries are not significantly more successful than those entering other industries). *See supra* text accompanying notes 22-28; *see also* BALDWIN, *supra* note 11, at 349, 352-53 (reporting greater turnover in high-profit industries); Dunne & Roberts, *supra* note 34, at 191 (finding that price-cost margins are *negatively correlated* with entry shares).

39. In summarizing the current state of the findings on entry Geroski concludes: "Although there is a very large cross-section variation in entry, differences in entry between industries do not persist for very long. In fact, *most of the total variation in entry across industries and over time is 'within' industry variation rather than 'between' industry variation.*" Geroski, *What Do We Know*, *supra* note 25, at 423 (emphasis added). Hence, while "[e]conometric estimates of the height of entry barriers suggest that they are high," *id.* at 429, "*[e]ntry rates are hard to explain using conventional measures of profitability and entry barriers*," *id.* at 430 (emphasis added).

40. *E.g.*, Audretsch & Mahmood, *Post-Entry Performance*, *supra* note 22, at 245 & n.2 ("One of the more striking empirical results to emerge . . . was that the entry of new firms into an industry is apparently not substantially deterred in industries where capital intensity and scale economies play an important role."); Evans & Siegfried, *supra* note 30, at 268-69 (finding little evidence for any effect of these barriers).

unless those industries are concentrated,[41] even though these barriers diminish their prospects for success.[42] The evidence on sunk cost effects, although limited, similarly suggests that they do little to deter most entrants.[43] Last, instead of deterring entry, the intensity of investment in R&D is positively correlated with entry,[44] and the intensity of investment in advertising that is often considered an impediment to entry seems to have non-significant effects.[45]

3. *The Third Puzzle: An Inferior Average Performance of Startup Entrants*

Significantly, the empirical findings on entry also reveal that startup entrants — new firms entering industry by new plant construction[46] — attempt entry more frequently than diversifying entrants[47] — those already existing firms entering a new industry by new plant construction. Startups also exhibit higher failure rates[48] and an *inferior average performance* as compared to diversifying entrants[49].

---

41. Hence, although Dunne & Roberts, *supra* note 34, at 200, find capital intensity negatively associated with entry, BALDWIN, *supra* note 11, at 378, examines further a similar pattern in his data and finds it is in fact limited to concentrated industries.

42. *See, e.g.*, Audretsch & Mahmood, *Post-Entry Performance*, *supra* note 22, at 252; Ioannis N. Kessides, *Entry and Market Contestability: The Evidence from the United States*, *in* ENTRY AND MARKET CONTESTABILITY, *supra* note 21, at 23, 41, 44.

43. *See, e.g.*, Kessides, *supra* note 42, at 41; Jose Mata, *Sunk Costs and Entry by Small and Large Plants*, *in* ENTRY AND MARKET CONTESTABILITY: AN INTERNATIONAL COMPARISON, *supra* note 21, at 49, 52-58 (finding that sunk costs deter only large entrants, those with over 250 employees).

44. BALDWIN, *supra* note 11, at 368-73 & tbl.14.2; Audretsch & Mahmood, *Post-Entry Performance*, *supra* note 22, at 252; Evans & Siegfried, *supra* note 30, at 265-66.

45. BALDWIN, *supra* note 11, at 368-73; Evans & Siegfried, *supra* note 30, at 265-66 & tbl.3.

46. This is a typical definition of startup entry. *See, e.g.*, Dunne et al., *Patterns*, *supra* note 11, at 501; *see also* BALDWIN, *supra* note 11, at 10-15 (discussing different definition and categories of entry). Note, however, that this definition includes *all* new firms in the startup category, regardless of whether they are started by individuals or are subsidiaries of existing firms. Consequently, some startup entrants resemble diversifying firms more than they resemble the quintessential startup. This ambiguity only suggests, however, that the empirical puzzles discussed in this Article would be even more striking if the data were to allow a sharper distinction between prototypical and atypical startups and diversifying entrants respectively.

47. Dunne et al., *Patterns*, *supra* note 11, at 504 & tbl.3 (analysis of data shows that startups comprise 87% of all entry by new plant creation in the United States, while diversifying firms make up only the remaining 13%).

48. Thus, Geroski's summary states, "[d]e novo entry is more common but less successful than entry by diversification." Geroski, *What Do We Know*, *supra* note 25, at 424; *see also* BALDWIN, *supra* note 11, at 11-15; Dunne et al., *Patterns*, *supra* note 11, at 501, 513, tbl.11 (providing cumulative exit rates for both startups and diversifying entrants). Note that startups are compared to diversifying firms who enter by new plant construction rather than to all diversifying firms, including those who enter by changing their product mix in an existing facility. While the latter entrants comprise a significant portion of diversifying entry, they only partly resemble the prototypical diversifying firm — they do not construct new produc-

Taken alone, the higher failure rates of startups could simply indicate that these ventures tend to be riskier but also more profitable, with startups exhibiting an average performance at least resembling that of diversifying entrants.[50] The data show, however, that startups do not perform as well as diversifying entrants. The performance of startups — as indicated by their relative size, growth, productivity, profitability, and mortality — strongly suggests that the higher average risk they bear is not coupled with an appropriately high return, as NPV maximization demands. In fact, these entrants' performance is significantly worse than the already unimpressive performance of their diversifying competitors. Hence, while the overall findings on entry have generated a consensus that entry rates are excessive, the data on startups show them to deviate from the rationality standard of entry to an even greater degree, establishing the third puzzle of entry.

Startups enter at a small scale[51] and fail to reach the average size in the industry even after fifteen years.[52] Diversifying entrants, on the other hand, begin operating at levels comparable to industry averages,[53] and grow at spectacular rates.[54] Smaller ventures are typically of a suboptimal size, which prevents them from enjoying the benefits of

---

tion facilities, and they enter under the same management of the existing facility. Moreover, most empirical studies do not consider them entrants and the data concerning their performance are therefore very limited. *See also infra* note 50.

49. *See infra* notes 50-56 and accompanying text.

50. Startups could face higher risk because their limited resources force them to begin operating at a smaller scale (assuming that capital lending markets are imperfect; otherwise startups should have been able to obtain the financing necessary to enter at optimal size). Ventures taking the startup route may also be biased towards higher-risk products that entrepreneurs find difficult to sell to incumbents or that existing firms do not want to produce via diversification, under the same corporate ownership, because of the associated exposure to increased liability. *Cf. infra* Section III.C (analyzing the relationship between risk, innovation, and startup entry). Last, diversifying entrants who want to reduce their risk exposure and "test the waters" may choose to enter without constructing new facilities, changing the mix of products they produce in existing facilities, an option that is unavailable to startups. Thus, the rate of entry of diversifying product-mix entrants is between 2.3 to 6.2 times greater than that of diversifying firms entering by new plant creation, although still lower than the rate of startup entry. Dunne et al., *Patterns*, *supra* note 11, at 504, tbl.3. These patterns of entry and exit show a greater resemblance to those exhibited by startups than to those of diversifying firms entering by new plant formation, lending support to the "testing the waters" hypothesis. *Id.* at 504 (product-mix entrants are only slightly larger than startups); *id.* at 513, tbl.11 (the exit rates of product-mix entrants are as high as those of startups, sometimes even higher).

51. Thus, Dunne et al. report that startup entrants are on average 28.4% as large as incumbent producers, while diversifying entrants are as much as 87.1% of the size of the latter. Dunne et al., *Patterns*, *supra* note 11, at 504.

52. *Id.* at 512, tbl.10.

53. *See, e.g.*, Timothy Dunne et al., *The Growth and Failure of U.S. Manufacturing Plants*, 104 Q. J. ECON. 671, 676 (tbl.1), 689 (1989) [hereinafter Dunne et al., *Growth*].

54. Dunne et al., *Patterns*, *supra* note 11, at 512.

scale economics[55] and results in productivity rates that are lower than industry-wide averages.[56]

Furthermore, when looking at the longitudinal performance of new entrant cohorts we find that the market share of startup entrants *as a group* declines with the passage of time, indicating that the increase in the share of successful startups does not compensate for the decline in share resulting from those who fail. The opposite obtains for diversifying entrants, however, where the impressive growth of surviving entrants more than offsets the loss of market share due to the failure of others, resulting in an overall gradual increase in their market share.[57]

Interestingly, when comparing the performance of *successful plants alone*, startups exhibit higher growth and fare better than diversifying firms, in relative terms.[58] These findings indicate that while the average fate of startups is worse, a successful startup may be better off than a successful diversifying entrant. This pattern suggests, in turn, that the higher risk associated with startups may be coupled with somewhat higher returns to success.[59] These higher returns are nevertheless insufficient to compensate the average startup for the significantly increased risk of failure it is must bear upon entry.

---

55. *See, e.g.*, David B. Audretsch & Michael Fritsch, *Creative Destruction: Turbulence and Economic Growth in Germany, in* BEHAVIORAL NORMS, TECHNOLOGICAL PROGRESS, AND ECONOMIC DYNAMICS: STUDIES IN SCHUMPETERIAN ECONOMICS 137, 139-40 & n.2 (Ernst Helmstadter & Mark Perlman eds., 1996) (citing numerous studies indicating that small entrants enter at suboptimal scale that often forces them to exit unless they can expand); Audretsch & Mahmood, *Post-Entry Performance*, *supra* note 22, at 245-46 (arguing that smaller entrants are of sub-optimal scale).

56. The evidence suggests that all entrants have lower productivity than incumbents because of their smaller size. *See, e.g.*, BALDWIN, *supra* note 11, at 209-10, 217-18.

57. Dunne et al., *Growth*, *supra* note 53, at 672-73, 689-93. Diversifying firms also show a higher growth rate even after controlling for the systematic size differences between these two entrant types. *See id.* at 686-93; Dunne et al., *Patterns*, *supra* note 11, at 509-13.

58. Dunne et al., *Growth*, *supra* note 53, at 672-73, 689-93. Thus, Dunne et al. note:

> When the failure probability is integrated into the analysis, however, the relationship between plant growth and size is negative for plants owned by single-plant firms but positive for plants owned by multiplant firms .... *For single-plant firms the decline of the growth rate of successful plants with size overwhelms the reduction in the failure rate.* Expected growth rates then decline with size. *For plants owned by multiplant firms, the decline in the failure rate with size is the dominant effect*, and expected growth increases with size.

Id. at 672-73 (emphases added).

59. *See* Leo A. Weiss, *Start-up Businesses: A Comparison of Performances*, 23 SLOAN MGMT. REV. 37 (1981) (comparing the performance of *successful startups* started by individual entrepreneurs with those started by large firms, finding the former to grow faster and achieve higher profitability sooner, and speculating that the possibly higher risk associated with individually started ventures is reflected in their higher returns).

B. *Can the Rationality Assumption Be Salvaged?*

The rationality assumption — that entrants will make only positive, risk-adjusted net present value entry attempts — is difficult to reconcile with the empirical data. The first and second puzzles — entry that is both excessive and insensitive to market predictors of success — challenge the applicability of the rationality assumption to all entrants, while the third puzzle — the systematically inferior average performance of startups — suggests that startup entrants exhibit behaviors that are even more difficult to reconcile with the rationality assumption.

The following section examines two categories of responses to the puzzles of entry that attempt to salvage the rationality assumption: those suggesting that entrants may still be attempting to make positive net present value entry, and those modifying the rationality assumption to allow "rational" entry without positive NPV.

1. *Maintaining the Rationality Assumption: Windows of Opportunity, Limited Information, and the Fruits of Learning*

a. Entrants Exploit Short-Term "Windows of Opportunity." Theoretically, entrants may recognize "windows of opportunity" in the market that allow "hit and run" entry — a quick, low-cost entry that yields short-term profits and is followed by exit once the opportunity for profits has dissipated.[60] If this were the case, it should not be surprising to find that many entrants exit soon after they enter. The empirical data, however, cast doubt on the likelihood of low-cost, highly profitable short-term entry. First, in most manufacturing industries new plant creation requires significant and partially irretrievable investments of labor and capital,[61] while hit-and-run entry should occur only where there is little risk of capital loss.[62] Second, the findings on

60. *See, e.g.*, Camerer & Lovallo, *supra* note 10, at 307; Geroski, *Data Driven Reflections*, *supra* note 30, at 282-83.

61. Thus, the economic literature distinguishes between fixed costs (which should not create a barrier to entry even while they give rise to scale economics) and "sunk costs." *See* PAUL A. GEROSKI ET AL., BARRIERS TO ENTRY AND STRATEGIC COMPETITION 26-37 (1990); William J. Baumol & Robert D. Willig, *Fixed Costs, Sunk Costs, Entry Barriers, and Sustainability of Monopoly*, 96 Q. J. ECON. 405 (1981). The measurement of the proportion of investment that must be sunk is difficult and has thus not been attempted often. One study examining various industries concluded that the range of sunk costs ranged from a low of 23% in petroleum refining to a high of 59% in electric motors. Richard J. Gilbert, *Preemptive Competition*, *in* NEW DEVELOPMENTS IN THE ANALYSIS OF MARKET STRUCTURE 90, 108, tbl.3.1 (Joseph E. Stiglitz & G. Frank Mathewson eds., 1986).

62. *See, e.g.*, GEROSKI ET AL., *supra* note 61, at 4 (reviewing the relevant arguments in the literature).

entrant profitability indicate that high profitability does not occur in the short-term but, instead, quite the opposite.[63]

b. Entrants Make Rational Decisions Based on Limited Information. Another possible argument is that entrants are guided by net present value considerations, but possess limited information.[64] They cannot predict with certainty the expected costs of and revenues from entry, which depend on numerous future events;[65] they may even be unable to determine accurately in advance their ability to manage and develop the new enterprise.[66] If this were the entrants' predicament, many entrants would misestimate their net present values and attempt entry when they should not.

While the limits of information available to entrants could explain the empirical findings, they could do so only given extreme assumptions about entry.[67] Entrants with incomplete information should still make accurate predictions *on average* in the absence of a bias in the available sources of information, although they would be somewhat less likely to succeed as a group than would be the case if perfect information were available. Entrants with incomplete information should be expected to make inaccurate NPV calculations at times. Some will wrongly conclude that entry is worthwhile; others will make the opposite mistake. These mistakes will affect the rate of entry, but only when they are large enough to cause a reversal of the sign of the perceived NPV of their prospective ventures — making a negative NPV seem positive or vice versa.

---

63. *See* E. RALPH BIGGADIKE, CORPORATE DIVERSIFICATION: ENTRY, STRATEGY, AND PERFORMANCE 191 (1979) (a study of diversifying ventures of Fortune 200 firms showing that these entrants required, on average, more than eight years since entry to "break even"); Weiss, *supra* note 59 (a study finding that even the most successful surviving startups take about four years to begin making profits).

64. Thus, GEROSKI ET AL., *supra* note 61, at 36, state that "[t]he threat of excessive entry is particularly severe when entrants act with incomplete information about the decisions of potential competitors," without explaining why incomplete information would be expected to lead to a skewed distribution of errors in favor of more frequent entry. The few economic models of entry decisionmaking with incomplete information generate diverse conclusions depending on their particular assumptions. *See, e.g.*, Roger Sherman & Thomas D. Willett, *Potential Entrants Discourage Entry*, 75 J. POL. ECON. 400, 402-403 & n.9 (1967) (concluding, in an early model of simultaneous entry, that limited information about other entrants' decision rules can discourage entry, while the effects under imperfect information are indeterminate).

65. *Cf.* MICHAEL T. HANNAN & JOHN FREEMAN, ORGANIZATIONAL ECOLOGY (1989) (using "learning" to refer not to an active adaptation process but, instead, simply to removing uncertainty by finding out whether the entering firm is sufficiently able to meet market demand efficiently vis-à-vis its rivals):

66. *Cf.* Boyan Jovanovic, *Selection and the Evolution of Industry*, 50 ECONOMETRICA 649 (1982) (an early, influential, learning model in which entrants face random costs that differ among firms, with the assumption that new entrants do not know their ability to manage the new startup but rather learn it from their actual post-entry performance).

67. Thus, the analysis below highlights only the implausibility of the information-based explanation as the sole cause of observed entry patterns, as limited information could explain any observation with appropriate assumptions.

Furthermore, the overall effect of these errors on entrants' *success rates* should be even smaller than the number of mistaken entry attempts alone suggests. All of the mistaken entrants reversing the NPV sign from positive to negative will simply avoid entry. At the same time, not all of those mistaking negative NPV entry as positive will fail. Depending on their objective success probability, these latter entrants will sometimes succeed, albeit at rates lower than those exhibited by their truly positive NPV competitors. Moreover, rationally mistaken entrants would not be likely to make frequent entry attempts that have a very low probability of success; instead, the attempts of such entrants are likely to represent ventures whose NPV are only moderately negative.

Such mistaken entry could partially explain increased entry rates, especially given the relatively large number of potential entrants with negative NPVs as compared to their few counterparts with positive NPV ventures.[68] This rational mistaken entry could also explain increased rates of failure; after all, the more negative NPV entrants that attempt entry, the greater the proportion of actual entrants that will eventually fail. This type of entry, however, would be highly unlikely to reduce survival rates to 40% or less within five years, as observed in data.[69]

Given reasonable assumptions, the argument based on limited information explains only a small amount of negative NPV entry, not the large proportion of failed entry observed in the data.[70] Moreover, the argument does not explain entrant insensitivity to market predictors — a form of information that rational actors would find relatively cheap to obtain and incorporate into the decision-making process — as the second puzzle reveals. Last, the information argument may suggest that rational startups have less information than diversifying entrants and therefore make more frequent errors and fail at higher rates.[71] The analysis of the first puzzle, however, has shown that ra-

68. This would occur, for example, under the reasonable assumption that potential ventures are not normally distributed around a zero NPV, but are rather mostly negative NPV ventures, with only a few bearing positive NPVs.

69. For rationally mistaken entry alone to bring average survival rates down to the observed levels, rational entrants would have to enter when their objective probability of success is even lower than industry averages since the few entrants who are not mistaken will be attempting higher probability entry. Ventures with a very low success probability, however, must expect extremely high returns that the masses of new entrants could not be rationally anticipated to obtain.

70. In reality, the proportion of failed entrants would have probably been even smaller than that suggested by the above analysis to the extent the entrants' probability of success were not fixed but rather normally distributed and determined by the rate of entry. If this were the case, the replacement of some more-effective entrants by less-effective ones would result in an overall increase in the probability of success of other entrants rather than an increased failure rate.

71. Although one could also argue that many of the objective predictors of success — economic and demographic alike — can be estimated by professional analysts at a cost that

tional mistakes alone do not explain extremely high failure rates. Rationally erring startups, moreover, would exhibit inferior performance not only on average, but upon success as well: These mistaken entrants would be riskier ventures with average returns that are unlikely to generate the superior performance in case of success that the data suggest.

c. Entrants Engage in a Long-Term Learning Process. The third possible defense of the rationality assumption is that entrants engage in a learning process:[72] They may knowingly attempt negative NPV entry because the information and experience they obtain from the entry attempt, even if it fails, increase their probability and magnitude of success in future ventures. According to this view, entrants will attempt entry as long as the NPV of all their entry attempts, taken together, is positive.

This learning-by-entering explanation seems compatible with the image of an entrepreneurial individual who repeatedly starts new ventures, unfazed by early failures, until striking gold.[73] Whether this prototypical entrepreneur is representative of new entrants is anybody's guess; statistics on repeated entry are not available. Learning, however, is very difficult in real-world settings where feedback is ambiguous and not always available.[74] The limited evidence on learning by entrants suggests that no learning occurs in the early years of the venture, and only some learning occurs as the venture progresses.[75]

For learning alone to account for the striking failure statistics in the data (as opposed to contributing to the success probability of experienced entrants) — as the case with the limited information argument. First, for entrants to plan repeated entry rationally they must be able to estimate the NPV of *all* of their expected attempts. Such a calculation, while theoretically possible, is highly implausible. It also contradicts the fundamental intuition behind the learning argument, that good estimates of complex future events are hard to come by. The

---

is unlikely to be large in comparison with the financial benefits and costs of entry. In addition, as limited as any objective analysis may be, a startup with very limited information could reduce the rate of error significantly by obtaining and properly analyzing these data.

72. *See, e.g.*, Geroski, *Data Driven Reflections*, *supra* note 30, at 283-84; *see also* Jovanovic, *supra* note 66 (modeling entry as a learning experiment).

73. *See* JOSEPH A. SCHUMPETER, THE THEORY OF ECONOMIC DEVELOPMENT: AN INQUIRY INTO PROFITS, CAPITAL, CREDIT, INTEREST, AND THE BUSINESS CYCLE 91-94 (Redvers Opie trans., 1934) (Oxford Univ. Press 1961) (describing the motives and character of the entrepreneurial "type").

74. *See infra* note 273 (discussing the difficulty of learning to overcome decision biases in real-world environments).

75. *See* Andrew H. Van de Ven & David N. Grazman, *Technological Innovation, Learning, and Leadership*, *in* TECHNOLOGICAL INNOVATION: OVERSIGHTS AND FORESIGHTS 279-80 (Raghu Garud et al. eds., 1997) (reviewing recent empirical findings to this effect).

possibility that this calculation could actually be made appears even more remote given the type of information necessary to determine the NPV of potential future ventures. Paradoxically, entrants would have to forecast future events based on information they do not have, when the need for repeated entry itself results from a similar, albeit less complicated, absence of information.[76]

Second, even if entrants had all the relevant information but simply needed to improve their ability to perform an entry task, they would be unlikely to attempt a significantly negative NPV entry. For entry to be rational, it is not enough for the aggregate of all attempts to have positive NPV. Rather, one must deduct the expected sunk costs, including time value, of all previous negative NPV entry attempts. Consequently, the positive NPV attempt has to be highly successful — as only few ventures are.[77]

While the learning argument may explain some excess entry, it resolves less adequately the other puzzles of entry, much like the previous arguments attempting to reconcile the empirical findings with the rationality assumption.[78]

2. *Modifying the Rationality Assumption: Maximizing Expected Utility with Negative Net Present Values*

The arguments attempting to reconcile the data with the rationality assumption do not adequately account for the various puzzles of entry under reasonable assumptions. It is therefore appropriate to examine whether modifying the definition of what rational entry consists of could save the assumed rationality of entrants. According to this line of reasoning, entrants may be acting rationally while attempting negative NPV entry if they derive other, non-monetary benefits from entry. To the extent the utility entrants derive from these benefits is greater than their disutility from the expected economic losses of entry, their behavior is still "rational."[79] Entrants may wish to fulfill a dream, enjoy running their own business, or even derive utility directly from the

76. Entrants with unlimited resources could keep attempting entry until successful, but that would not make their overall investment rational, of course.

77. Failed entry is an expensive teaching tool. It provides only a limited range of information about the market, and even less information that would be particularly relevant to future ventures unless they were to repeat closely the original attempt.

78. Another possible argument for the rationality of diversifying entrants is that they maximize the overall profits of their firms, using entry to punish or threaten rivals, protecting their profits in those industries where they are well established incumbents. We do not discuss this argument in detail, however, since not only does it apply to diversifying entry alone but it also explains — at best — but a small proportion of such entry.

79. *Cf.* Posner, *supra* note 4, at 1553-55 (claiming that the evidence presented by Jolls et al. of bounded rationality can be explained by various ad hoc hypotheses regarding individuals' utility functions).

entry attempt, regardless of its consequences — a sort of gambling in a business context.[80] A stronger version of this argument would claim that entrants *must* be maximizing their subjective expected utility. This view would assert that the studies of entry all misinterpret the data, or otherwise fail to identify what is it that entrants in fact maximize.[81]

These and similar suggestions may accommodate puzzling data, but at the cost of abandoning the NPV standard[82] together with its underlying justification as "the managerial compass" in financial decisionmaking.[83] Once individual tastes and values enter the calculation, managers lose their NPV-based legitimacy vis-à-vis investors, who do not share the managers' personal preferences.

Moreover, the arbitrary addition of "value" sources reduces the rationality assumption to a tautology.[84] It protects the rationality assumption from empirical criticism by eliminating its predictive power

---

80. Using anecdotal observation and some empirical studies, the entrepreneurship literature often states as obvious the notion that entrants, while motivated by profit seeking, also seek to fulfill other preferences. *See, e.g.*, Thomas M. Begley & David P. Boyd, *Psychological Characteristics Associated with Performance in Entrepreneurial Firms and Smaller Businesses*, 2 J. BUS. VENTURES 79, 82 (1987) (citing studies); James W. Carland et al., *Differentiating Entrepreneurs from Small Business Owners: A Conceptualization*, 9 ACAD. MGMT. REV. 354, 356 & tbl.1 (1984). Nevertheless, one wonders whether unsuccessful entry that ends with quick failure is even utility maximizing in the minimal sense.

81. Compare Posner's telling defense of the traditional economic analysis of law in his important introductory book:

> The reader who lacks previous acquaintance with economics may be troubled by what appear to be the severely unrealistic assumptions that underlie economic theory. The basic assumption, that human behavior is rational, seems contradicted by the experiences and observations of everyday life. The contradiction is less acute when one understands that . . . [r]ationality means little more to the economist than a disposition to choose . . . an apt means to *whatever ends the chooser happens to have.*

RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW § 1.3, at 17 (5th ed. 1998) (emphasis added); *see also* Posner, *supra* note 4, at 1553-55 (using a similar approach to explain evidence of bounded rationality). *See generally* Korobkin & Ulen, *supra* note 1, at 1060-66 (citing sources explaining that according to this type of definitional approach, all behavior is necessarily rational and necessarily "worth it" to the actor).

82. *See, e.g.*, BREALEY & MYERS, *supra* note 17, at 180 (NPVs do not include the enjoyment of risks or other non-pecuniary values).

83. *Id.* at 24 ("The remarkable thing is that managers of firms can be given one simple instruction: Maximize net present value."). NPV is the basis on which managers maximize the firm's, and thereby its shareholders' wealth. *Id.*; *see also* FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 23 (1991) (stating that management decisions should aim to maximize shareholder wealth). It justifies the separation of ownership and control, "the fundamental condition for the successful operation of a capitalist economy." BREALEY & MYERS, *supra* note 17, at 24.

84. *See* MARK BLAUG, THE METHODOLOGY OF ECONOMICS 230-32 (2d ed. 1992) (citing examples of similar tautological claims in modern economic thought); *see also* DANIEL M. HAUSMAN, THE INEXACT AND SEPARATE SCIENCE OF ECONOMICS 205-26, 247-80 (1992) (providing an overview of the limits of economic methodology vis-à-vis rationality assumptions).

and converting it to an irrefutable axiom, capable of accommodating any data.[85]

II. A BEHAVIORAL ANALYSIS OF ENTRY DECISIONMAKING

Unlike traditional theories of entry, a behavioral approach does not expect entrants to make decisions under uncertainty according to norms of strict rationality. Instead, it accepts that entrants are boundedly rational: they may weigh the pros and cons of entry, but their ability to do so rationally is impeded by the limitations of human cognition and affected by motivation and emotion. To utilize their limited mental resources effectively, entrants employ simplifying decision heuristics that enable them to function reasonably well in a complex environment, but also lead them at times to systematic, predictable errors.[86] More specifically, the psychological literature reveals a number of cognitive processes that lead entrants, like other individuals making judgments with significant personal stakes under uncertainty, to be overconfident about the prospects of their ventures and insensitive to background statistical information.[87]

85. *Cf.* Christine Jolls et al., *Theories and Tropes: A Reply to Posner and Kelman*, 50 STAN. L. REV. 1593, 1600 (1998) (responding to Posner's argument by stating that "[i]t is difficult to see what conclusions should be drawn from the fact that evolution can be shown to produce a behavior *and* the absence of that behavior"); *see also* ANATOL RAPOPORT, DECISION THEORY AND DECISION BEHAVIOR 8 (2d rev. ed. 1998) ("[I]t would be futile to infer the decision-maker's values entirely *ad hoc* . . . . For this sort of inference would make any hypothesis and hence any descriptive theory of decision *unfalsifiable, hence worthless.*") (second emphasis added).

86. *See supra* note 2 (various sources citing evidence for and discussing the systematic nature of judgmental errors); *see also* Jeffery J. Rachlinski, *The New Law and Psychology: A Reply to Critics, Skeptics, and Cautious Supporters*, 85 CORNELL L. REV. 739, 750-52 (2000) (describing the general psychological parameters of the behaviorally informed actor). For some general, though not comprehensive, reviews of legal articles applying a behavioral approach, see Donald C. Langevoort, *Behavioral Theories of Judgment and Decision Making in Legal Scholarship: A Literature Review*, 51 VAND. L. REV 1499 (1998); Cass R. Sunstein, *Behavioral Law and Economics: A Progress Report*, 1 AM. L. & ECON. REV. 115 (1999).

87. Entrants are likely to use many of those cognitive heuristics that impact individuals' judgments under uncertainty even when these judgments bear no significant personal consequences for the decisionmaker. The analysis below nevertheless focuses on those powerful, robust, and systematic biases that generate overconfidence when decisionmakers must make self-interested judgments. Other cognitive phenomena are mentioned only to the extent they are likely to have any systematic impact on entrants' propensity to enter. *See, e.g.*, *infra* notes 136-140 and the accompanying text. In doing so, this Article follows the practical approach suggested by Jolls et al., *supra* note 1, at 1481 (noting, when describing the main application of a behavioral law and economics, that they "do not emphasize behavioral patterns that depart from standard economic assumptions but fail to point in systematic directions" but instead "focus [on] robust, empirically documented phenomena that have *reasonably precise implications* for legal issues") (emphasis added).

A. *Solving the First Puzzle: The Processes of Entrants' Overconfidence*

From a behavioral perspective, the most significant characteristics of the judgments that entrants must make when deciding whether to enter are two. First, entrants must make their decision under a veil of *uncertainty*. They are unable to predict the fate of their venture with certainty, and they cannot know with certainty all the relevant information about the market they consider entering nor can they process it fully. In fact, entrants cannot even know with certainty all their personal characteristics pertinent to the entry task or even what exactly these abilities, skills, and experiences might be. Second, the consequences of the decisions that potential entrants must make are incredibly significant for them. These consequences implicate their self-perception and self-esteem as effective, successful business decision-makers and managers. The decisions they make also have important positive and negative financial and reputational implications for them.

Ironically, a wealth of psychological data show that in circumstances of this kind, where rational judgment would arguably be most beneficial to the decisionmaker, people tend to exhibit a significant bias.[88] As the analysis below shows, this bias results from a number of psychological processes that affect entrants' judgments of both the *probability* and *value* of their prospective ventures.

Some scholars, recognizing the puzzle of excess entry, have suggested that entrants' behavior may primarily reflect an optimistic bias, in addition to other psychological phenomena.[89] An important recent set of experiments provides direct evidence of overoptimism in a market entry game.[90] This Part both reviews these findings and develops a systematic framework showing how various psychological processes are highly likely to generate entrants' overconfidence — making this biased behavior common and robust even in those real-world contexts that have not been examined experimentally.

---

88. The numerous, mostly egocentric (self-related) biases in perception, evaluation, and prediction that generate entrant overconfidence result from a combination of motivational and cognitive factors. Dale W. Griffin & Carol A. Varey, *Towards a Consensus on Overconfidence*, 65 ORGANIZATIONAL BEHAV. & HUM. DECISION. PROC. 227, 228 (1996). More generally, the identification of an important "egocentric" component of these phenomena does not exclude the role of "cold" cognitive processes in generating the biases, but only emphasizes the role of self-implication as the thread unifying them. *See* ZIVA KUNDA, SOCIAL COGNITION: MAKING SENSE OF PEOPLE 211-46 (1999) (reviewing some findings); Michael Ross & Fiore Sicoly, *Egocentric Biases in Availability and Attribution*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 179 (Daniel Kahneman et al. eds., 1982) (reviewing the early literature on egocentric biases in attribution).

89. *See, e.g.*, Geroski, *Data Driven Reflections*, *supra* note 30, at 282, 284 (suggesting that "low survival rates may reflect either systematic errors made by entrants in forecasting post-entry returns, or a positively risk-taking attitude" without explaining why entrants make mistakes or which mistakes they actually make).

90. Camerer & Lovallo, *supra* note 10.

The comprehensive framework developed in this Part will be further applied to the yet unaddressed and unresolved second and third puzzles of entry. Consequently, this Part describes the processes of overconfidence in detail, laying important foundations for the remainder of the behavioral analysis of entry decisionmaking.

1. *Optimistic Bias*

The evidence on optimistic bias indicates that entrants are likely to have inflated views of both their absolute and comparative ability, with a resulting bias in the perception of their probability of success (and the concomitant risk associated with their ventures). To the extent that the value (e.g., net profits) of successful entry depends on the entrant's business acumen, moreover, overconfident entrants will also expect their value to be higher than objectively warranted.

Individuals have a strong tendency to exhibit optimistic bias. They overestimate their positive traits, abilities, skills, and likelihood of experiencing positive events, while they underestimate their vulnerability to certain risks.[91] Overoptimism is especially pronounced in *comparative* contexts, in which people judge themselves in relation to others.[92]

---

91. *See, e.g.*, David Dunning et al., *Self-Serving Prototypes of Social Categories*, 61 J. PERSONALITY & SOC. PSYCHOL. 957 (1991) (showing how people judge positive traits to be overwhelmingly more characteristic of themselves than negative attributes, and define personal attributes in idiosyncratic ways that emphasize their perceived strengths); Shelley E. Taylor & Jonathon D. Brown, *Illusion and Well-Being: A Social Psychological Perspective on Mental Health*, 103 PSYCHOL. BULL. 193 (1988) [hereinafter Taylor & Brown, *Illusion and Well-Being*]; Shelley E. Taylor & Jonathon D. Brown, *Positive Illusions and Well-Being Revisited: Separating Fact from Fiction*, 116 PSYCHOL. BULL. 21, 22-23 (1994) [hereinafter Taylor & Brown, *Separating Fact from Fiction*] (reviewing and discussing findings on individuals' mildly distorted positive perceptions); Neil D. Weinstein, *Unrealistic Optimism About Future Life Events*, 39 J. PERSONALITY & SOC. PSYCHOL. 806 (1980) [hereinafter Weinstein, *Unrealistic Optimism*]; Neil D. Weinstein & William M. Klein, *Unrealistic Optimism: Present and Future*, 15 J. SOC. & CLINICAL PSYCHOL. 1, 1-6 (1996) [hereinafter Weinstein & Klein, *Present and Future*] (discussing findings showing that people think themselves invulnerable to certain risks). For some pertinent recent findings, see Valerie A. Clarke et al., *Unrealistic Optimism and the Health Belief Model*, 23 J. BEHAV. MED. 367, 372-74 (2000) (healthy subjects exhibit bias in *absolute* (non-comparative) judgments of the expected risk and severity of cancer, as well as the potential benefits and costs they will experience in using screening methods); Marsha T. Gabriel et al., *Narcissistic Illusions in Self-Evaluations of Intelligence and Attractiveness*, 62 J. PERSONALITY 143 (1994); Janet Metcalfe, *Cognitive Optimism: Self-Deception or Memory-Based Processing Heuristics?*, 2 PERSONALITY & SOC. PSYCHOL. REV. 100 (1998) (resembling positive illusions).

92. *See, e.g.*, David Dunning et al., *Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Appraisals of Ability*, 57 J. PERSONALITY & SOC. PSYCHOL 1082 (1989) [hereinafter Dunning et al., *Ambiguity and Self-Evaluation*] (academic skills, leadership ability, marriage prospects, and health); Gabriel et al., *supra* note 91 (intelligence and attractiveness); Ola Svenson, *Are We All Less Risky and More Skillful Than Our Fellow Drivers?*, 47 ACTA PSYCHOLOGIA 143 (1981); Taylor & Brown, *Illusion and Well-Being*, *supra* note 91, at 195-96; Taylor & Brown, *Separating Fact from Fiction*, *supra* note 91, at 22-23 (stating that people choose dimensions of comparison in which they excel, and select worse-off comparison targets that guarantee a favorable comparison); Weinstein &

Over the last twenty years a wealth of empirical econometric data evidencing boundedly rational behavior in financial markets has accumulated,[93] with recent studies documenting the effects of optimistic bias on the performance of investors and markets alike.[94] Optimistic bias is common in investment decisionmaking, as exemplified by a recent study examining portfolio allocation decisions of eighty MBA students from the Kellogg Graduate School of Management at Northwestern University in a computerized investing simulation.[95] The participants in this study were not only business-oriented, but had an average investment experience of almost five years.[96]

The study created a simulated market based on real performance data of the nine largest mutual funds as of 1985, plus an S&P 500 index fund. The game covered ten years, which corresponded to the period 1985-1994, although the subjects were not aware of that. Participants had $100,000 to invest and the opportunity to review their investments every six months, resulting in a twenty-period game. They could invest their money in any combination of the ten funds or leave it in the "bank." Participants also had full information about the performance of the funds and the market for each period.[97]

As expected, study participants consistently overestimated their predicted performance relative to the market. Even more strikingly, they overestimated the past performance of their portfolios — believing at the end of the game that, on average, they matched the per-

---

Klein, *Present and Future*, *supra* note 91, at 4 n.2 (mentioning a list of more the 200 studies showing unrealistic optimism in expectations regarding positive and negative life events the author had accumulated by 1996).

93. *See* ADVANCES IN BEHAVIORAL FINANCE (Richard H. Thaler ed., 1993) (a classic collection of articles on behavioral finance); ANDREI SHLEIFER, INEFFICIENT MARKETS: AN INTRODUCTION TO BEHAVIORAL FINANCE (2000) (a readable and comprehensive review of the behavioral finance literature).

94. *See, e.g.*, Terrance Odean, *Volume, Volatility, Price, and Profit When All Traders Are Above Average*, 53 J. FIN. 1887 (1998) (testing a model of overconfidence and surveying the literature). Other empirical studies show that stock markets overreact, that both individual and professional investors cannot "beat the market" in the long run, that investors trade too much, SHLEIFER, *supra* note 93 (reviewing and discussing many of these studies), and that this excessive trading has a negative effect on their returns. Brad M. Barber & Terrance Odean, *Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment*, 116 Q.J. ECON. 261 (2001) (finding that men underperform women, who are already overconfident, because the former trade even more frequently); Brad M. Barber & Terrance Odean, *Trading is Hazardous to Your Health: The Common Stock Investment Performance of Individual Investors*, 55 J. FIN. 773 (2000) (finding in a sample of 66,465 households a high portfolio turnover and a negative correlation between turnover and profitability).

95. Don A. Moore et al., *Positive Illusions and Forecasting Errors in Mutual Fund Investment Decisions*, 79 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 95 (1999).

96. They also reported average investments of over $20,000, and rated their investment knowledge as close to amateur, on a novice-amateur-expert scale; their ages ranged from 23 to 45 years old. *Id.* at 100-01,103.

97. Measured by the S&P 500 index. *Id.* at 103.

formance of the market, while in fact their performance was significantly below the market. This finding was also accompanied by the participants' overestimation of the consistency of their portfolios' performance over the twenty periods of the game. It was obtained despite the unusually clear feedback the subjects received at the end of every game period. Last, the subjects switched their investments too frequently, a behavior that correlated with (and possibly caused) poorer investment performance.[98] These and similar findings[99] are not surprising to those familiar with the investment behavior of both amateurs and professionals, who report how analysts commonly overestimate their investment selection ability and consequently overestimate the expected returns on their investments.[100]

The optimistic bias pervasive in business environments has also been directly tested in simulated entry decisionmaking. Experimental economists have recently shown that entrants' overestimates of their comparative skill generate excessive entry.[101] Camerer and Lovallo created an experimental game, where participants chose simultaneously, without communicating with each other, whether to enter the market. Participants were told in advance the "capacity" of the market — the number of participants who can make a positive profit from entering. The larger the number of participants who decide to enter the experimental market, the smaller the average returns to the entrants. Participants were also informed that not all entrants will receive the same returns. Instead, their returns depended on how they ranked in comparison with their counterparts, with the ranking determined based on either a random drawing or their performance in a skill or trivia task. Under these circumstances, most participants realized that the average profit of entrants would be negative.[102] They erroneously

98. *Id.* at 104-10

99. *See, e.g.*, William N. Goetzmann & Nadav Peles, *Cognitive Dissonance and Mutual Fund Investors*, 20 J. FIN. RES. 145 (1997) (also finding that mutual fund investors exhibit overly optimistic perceptions of past mutual fund performance). *See generally* MAX H. BAZERMAN, JUDGMENT IN MANAGERIAL DECISION MAKING 96-108 (5th ed. 2002) (discussing a variety of common investment mistakes including overconfidence and optimism).

100. *See, e.g.*, Arnold S. Wood, *Behavioral Risk: Anecdotes and Disturbing Evidence*, J. INVESTING, Spring 1997, at 8; Arnold S. Wood, *Fatal Attractions for Money Managers*, FIN. ANALYSTS J., May-June 1989, at 3.

101. Camerer & Lovallo, *supra* note 10, at 310-12.

102. In fact, earlier studies have shown that when the market is simple, and skill and self-selection into competition play no role, entrants are a little biased but after a sufficient number of repeated trials converge close to the rational rate of entry. *See* Daniel Kahneman, *Experimental Economics: A Psychological Perspective*, *in* BOUNDED RATIONAL BEHAVIOR IN EXPERIMENTAL GAMES AND MARKETS 11 (Reinhard Tietz et al. eds., 1988) (introducing the entry game paradigm and reporting that experiment participants enter roughly at market capacity in a simple game where all participants know the market's capacity, where there are obvious profits to those who decide not to enter, and where market outcomes depend *only on the number* of entrants — not on their skill or performance); Amnon Rapoport, *Individual Strategies in a Market-Entry Game*, 4 GROUP DECISION & NEGOTIATION 117 (1995)

expected, however, that their own profits would be positive. This bias was especially strong where only those participants who self-selected to participate (though not necessarily enter) in a skill-based entry game competed among themselves. These participants exhibited significant overconfidence in their comparative skill, and failed to anticipate the greater intensity of competition they were bound to face in the self-selection condition, where only participants who considered themselves skilled enough to win positive returns chose to participate.[103]

2. *The Desirability Bias and Related Phenomena*

While overoptimism affects entrants' perceptions and expectations regarding their future performance by inflating their self-perception, desirability related phenomena bias a far broader range of judgments. These biases occur when people predict future events that implicate their self-perception, emotions, or interests. They cause decisionmakers to align their expectations with preferences for particular outcomes of events even when these events *do not* concern their own actions specifically, as when predicting market-wide trends or the behavior of other actors. Such desirability related biases affect predictions of factors that determine a future venture's *value* as well as factors that bias *probability* judgments.

a. The Desirability Bias. Numerous studies show that people exhibit biased predictions of external events that are not under their control, but whose outcomes nevertheless implicate their self-perception or are otherwise important to the predictors who hold clear preferences regarding them.[104] When manifesting such biased predictions regarding desirable outcomes, entrants may not only overestimate the profitability of successful entry, but also underestimate the investments and the time necessary for the venture to become viable. This

(replicating the design using larger stakes with similar findings). The subjects' ability in Camerer and Lovallo's study to forecast average returns with reasonable accuracy provides only a limited consolation, however; the clear and simple nature of the game constrained the subjects' wishful thinking dramatically more than most real-world situations do.

103. Camerer & Lovallo, *supra* note 10, at 310-16 (describing this phenomenon as "reference group neglect"); *see also* Edward J. Zajac & Max H. Bazerman, *Blind Spots in Industry and Competitor Analysis: Implications of Interfirm (Mis)Perceptions for Strategic Decisions*, 16 ACAD. MGMT. REV. 37 (1991) (reviewing studies showing the scope of competitive blind spots and suggesting these play a role in new business failures).

104. *See, e.g.*, Elisha Babad & Yosi Katz, *Wishful Thinking — Against All Odds*, 21 J. APPLIED SOC. PSYCHOL. 1921 (1991) [hereinafter Babad & Katz, *Against All Odds*]; David V. Budescu & Meira Bruderman, *The Relationship Between the Illusion of Control and the Desirability Bias*, 8 J. BEHAV. DECISION MAKING 109 (1995); Roy M. Poses & Michele Anthony, *Availability, Wishful Thinking, and Physicians Diagnostic Judgments for Patients with Suspected Bacteremia*, 11 MED. DECISION MAKING 159 (1991) (reporting a "value bias"). For a clear, early study, see Douglas McGregor, *The Major Determinants of the Prediction of Social Events*, 33 J. ABNORMAL PSYCHOL. 179 (1938).

may happen if they underestimate the full scope of the required investments for ventures of the type they contemplate, the various future events in the industry or economy at large that may negatively affect the prospects of the entry attempt, or the time they — like other entrants — will need to become profitable. In addition, entrants may also underestimate their likely losses upon failure.[105]

The desirability bias has been measured directly, with many experiments showing that estimates of the likelihood of future events are correlated with people's desirability ratings of these events.[106] The effects of this bias even occur when people have no pre-existing stake in the subject of their judgment and cannot affect the outcome of the predicted event.[107]

The evidence shows that even professional investment managers exhibit the desirability bias. In two related studies, American and Taiwanese investment managers were requested to make predictions of whether certain economic events or trends of importance to them and their clients (e.g., trends of merger activity, effects of high debt levels, and relevant regulation) would occur or change a year from the prediction date. Both studies revealed systematic positive correlations between the managers' predictions and their ratings of these events' desirability.[108]

While even ad hoc designations create desirability and affect predictions, the bias is especially strong and pervasive where people have preexisting, vested interests in the outcomes of a predicted event, such as when voters predict election outcomes[109] and sport fans predict the

105. This latter misperception differs from an inability to recognize the full scope of necessary investment, reflecting an underestimation of the proportion of the investment that cannot be recouped upon failure (i.e., the "sunk costs").

106. *See* Budescu & Bruderman, *supra* note 104; George Wright & Peter Ayton, *Subjective Confidence in Forecasts: A Response to Fischhoff and McGregor*, 5 J. FORECASTING 117 (1986) (reviewing these findings).

107. For example, participants in a recent study who where designated either "plaintiffs" or "defendants," exhibited systematically biased expectations of the decision an objective judge would arrive at in a tort case, each group in accordance with its designation, although their roles were merely ad hoc designations and they had no opportunity to address the judge. George Loewenstein et al., *Self-Serving Assessments of Fairness and Pretrial Bargaining*, 22 J. LEGAL STUD. 135, 151, tbl.2 (1993); *see also* Linda Babcock et al., *Biased Judgments of Fairness in Bargaining*, 85 AM. ECON. REV. 1337 (1995). *But see* Maya Bar-Hillel & David Budescu, *The Elusive Wishful Thinking Effect*, 1 THINKING & REASONING 71, 72-74 (1995) (not always finding a desirability bias for emotionally neutral outcomes that are only endowed with value ad hoc in the experiment).

108. Robert A. Olsen, *Desirability Bias Among Professional Investment Managers: Some Evidence from Experts*, 10 J. BEHAV. DECISION MAKING 65, 66-70 (1997).

109. McGregor, *supra* note 104, at 181-83, 191, tbl.5. For a review of election studies and a large-scale analysis of the preference-expectation link in U.S. presidential elections, see Donald Granberg & Edward Brent, *When Prophecy Bends: The Preference-Expectation Link in U.S. Presidential Elections, 1952-1980*, 45 J. PERSONALITY & SOC. PSYCHOL. 477, 477-79, tbl.1 (1983).

outcomes of matches.[110] As noted above, entry is also characterized by the significant positive and negative consequences it harbors for entrants, who are therefore likely to exhibit a strong desirability bias as well.

b. The Affect Heuristic. In addition to exhibiting the desirability bias, people often use an "affect heuristic," making judgments based on affective "tags" they associate with the subjects of their judgment.[111] They frequently simplify judgmental processes by consulting readily available affective impressions, using affective shortcuts very much like they utilize cognitive heuristics in decisionmaking.[112] For example, when assessing the risks and benefits of different activities, people rely on the affect they associate with them; consequently, judgments of the risks and benefits of hazards are often *negatively* correlated (e.g., people deem those activities they consider beneficial as low-risk), although risks and benefits tend to be *positively* correlated.[113]

Thus, the affect heuristic contributes to individuals' tendency to align expectations with preferences, especially when preferences are strong and the predicted events carry significant affective importance to the predictor. This judgmental process provides an important complement to our understanding of the processes of entrant overconfidence: while the basic desirability bias has been recorded mainly in the realm of probabilistic or binary (i.e. yes or no) predictions, the evidence on the affect heuristic reveals its impact on judgments of *benefits* as well as those of *risks*. The affect heuristic therefore explains how desirability can impact entrants' estimates of the value of a potential venture and their judgments of the risk associated with it. The evidence on the positive correlation between individuals' predictions of benefits and risks further suggests that entrants will not only overestimate the value of entry and underestimate its risks, but that these biased perceptions will further support and reinforce one another.

c. The Planning Fallacy. Findings on the "planning fallacy" further document how desirability leads people to underestimate the *time* and *costs* required for completing projects, especially when these projects

110. *See, e.g.*, Elisha Babad, *Wishful Thinking and Objectivity Among Sports Fans*, 2 SOC. BEHAV. 231 (1987) [hereinafter Babad, *Objectivity Among Sports Fans*]; Babad & Katz, *Against All Odds*, *supra* note 104, at 1923-24, 1929-32 (finding that high incentives for accuracy do not eliminate fans' bias).

111. Melissa L. Finucane et al., *The Affect Heuristic in Judgments of Risks and Benefits*, 13 J. BEHAV. DECISION MAKING 1 (2000). Special thanks to Paul Slovic for directing my attention to the relevance of the affect heuristic research for the present analysis.

112. *See, e.g.*, *id.* at 3.

113. *See id.* at 3-4 (reviewing studies showing this effect and discussing the objective relationship between risk and benefit).

are complex and protracted,[114] as in the case of entry. In addition to countless real-world examples of this fallacy,[115] a series of recent studies highlights the prevalence of optimistic predictions of task completion times.[116] Participants in these studies exhibited the planning fallacy when predicting the completion of a significant and somewhat novel academic task,[117] in familiar everyday activities,[118] and in the familiar (though infrequent) annual task of completing tax forms.[119]

---

114. *See* Roger Buehler et al., *Exploring the "Planning Fallacy": Why People Underestimate Their Task Completion Times*, 67 J. PERSONALITY & SOC. PSYCHOL. 366 (1994) [hereinafter Buehler et al., *The Planning Fallacy*]; Roger Buehler et al., *The Role of Motivated Reasoning in Optimistic Time Predictions*, 23 PERSONALITY& SOC. PSYCHOL. BULL. 238 (1997) [hereinafter Buehler et al., *Optimistic Time Predictions*]; Ian R. Newby-Clark et al., *People Focus on Optimistic Scenarios and Disregard Pessimistic Scenarios While Predicting Task Completion Times*, 6 J. EXPERIMENTAL PSYCHOL.: APPLIED. 171 (2000). As Kahneman and Tversky note:

> The context of planning provides many examples in which the distribution of outcomes in past experience is ignored. Scientists and writers. for example, are notoriously prone to underestimate the time required to complete a project, even when they have considerable experience of past failures to live up to planned schedules. A similar bias has been documented in engineers' estimates of the completion time for repairs of power stations. Although this planning fallacy is sometimes attributable to motivational factors such as wishful thinking, it frequently occurs even when underestimation of duration or cost is actually penalized.

Daniel Kahneman & Amos Tversky, *Intuitive Prediction: Biases and Corrective Procedures* in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 414, 415 (Daniel Kahneman et al. eds., 1982) [hereinafter Kahneman & Tversky, *Intuitive Prediction*] (citation omitted).

115. *See* Buehler et al., *The Planning Fallacy*, *supra* note 114, at 366 (reviewing such evidence). Other striking and oft-cited examples appear in Rand's large-scale Pioneer Plants Study, conducted for the U.S. Department of Energy. The study reviewed "what was known about the problem of poor cost estimation and performance for advanced technologies in particular and very large projects in general" and concluded: "Severe underestimation of capital costs in the norm for *all* advanced technologies; the underestimation for energy process technologies mirrored that seen in major weapon systems acquisition, very large advanced construction projects, and major public works activities." EDWARD W. MERROW ET AL., RAND, UNDERSTANDING COST GROWTH AND PERFORMANCE SHORTFALLS IN PIONEER PROCESS PLANTS 2 (1981). In their own data, based on a large sample of plants, the authors of this study found, moreover, that the earliest estimates made for the plants "averaged *less than one-half* of actual costs. And many of the early estimates . . . reflected little more than *one-third* of what the plants actually cost to design and construct . . ." *Id.* at 38.

116. For a brief review of some earlier studies documenting the planning fallacy in laboratory settings, see Buehler et al., *The Planning Fallacy*, *supra* note 114, at 367.

117. Only 10.8% of the subjects finished their tasks by the time they indicated it would take them if "everything went as well as it possibly could," and fewer than half of them finished by the time they indicated the task would take them if "everything went as poorly is it possibly could." *Id.* at 369 (Study 1).

118. These tasks included school assignments and various tasks around the home. *Id.* at 370-71 (Study 2), 372 (Study 3), 374-75 (Study 4); *see also* Newby-Clark et al., *supra* note 114, at 174, 175, 177, and 178 (reporting Experiments 1-4 showing, inter alia, a main effect for bias in final predictions).

119. Buehler et al., *Optimistic Time Predictions*, *supra* note 114, at 241.

3. *The Illusion of Control*

Individuals often deem themselves able to control chance occurrences and risky eventualities, especially when final outcomes depend on a mixture of skill and chance, exhibiting the illusion of control.[120] This illusion stems from people's inability to distinguish between "skill" and "chance" situations and their more general desire to believe they can control the world around them. It leads to inflated expectations of personal success in tasks whose outcomes depend, in part or in whole, upon chance factors.[121]

Studies show that decisionmakers often behave in chance situations as if skilled behavior could influence outcomes; consequently, the same factors that people associate with success in skill-dependent tasks affect their expectations regarding the outcomes of pure chance events, such as lotteries and gambles.[122] They think they are more likely to win, for example, when they can choose among options; when they are more familiar or skilled with either the stimulus (e.g., the particular lottery ticket) or the necessary response; when they are actively (versus passively) involved in the task; and, importantly, when competition is present.[123]

In addition to its direct effect on judgments, the illusion of control also facilitates[124] both optimistic bias[125] and the desirability bias,[126] fur-

120. *See* Ellen J. Langer, *The Illusion of Control*, 32 J. PERSONALITY & SOC. PSYCHOL. 311 (1975) (citing earlier studies); *see also* Budescu & Bruderman, *supra* note 104, at 110 (citing additional studies).

121. Budescu & Bruderman, *supra* note 104, at 109-10; Langer, *supra* note 120, at 313.

122. While the illusion of control operates even in pure chance environments, it is clearly more pernicious in real-world settings, where outcomes depend on a mix of skill and chance factors, and where the respective contribution of skill and chance to final outcomes cannot be determined with certainty. As Langer noted, "there is an element of chance in every skill situation and an element of skill in almost every chance situation." Langer, *supra* note 120, at 324.

123. *Id.* at 315 (Experiment 1, competition); *id.* at 316-17 (Experiment 2, choice); *id.* at 318 (Experiment 3, stimulus familiarity); *id.* at 319-20 (Experiment 4, response familiarity), *id.* at 320-22 (Experiments 5-6, type of involvement). Thus, for example, participants wagered more money in competitive situations when they perceive their opponent as less skilled, even when the outcome of the bet was completely determined by chance, *id.* at 313-15, and placed a higher value on lottery tickets they choose for themselves than those chosen for them. *Id.* at 315-17; *see also* Budescu & Bruderman, *supra* note 104, at 109-10, 114-15 (illusion of control results in experiment 1).

124. Some studies even purport to show that optimistic bias is completely dependent on perceptions of control. *See, e.g.*, Frank P. McKenna, *It Won't Happen to Me: Unrealistic Optimism or Illusion of Control?*, 84 BRIT. J. PSYCHOL. 39 (1993) (reporting studies showing optimistic overconfidence regarding risk of accidents for drivers but not for passengers, who realize that their driving skill would have no impact on the outcome of accidents). Many studies, however, provide evidence that overoptimism exists regardless of perceived control. *See, e.g.*, Weinstein, *Unrealistic Optimism*, *supra* note 91; Weinstein & Klein, *Present and Future*, *supra* note 91.

125. *See* Weinstein, *Unrealistic Optimism*, *supra* note 91, at 808:

ther increasing individuals' willingness to engage in risk-taking behaviors under the erroneous belief that that they can control the risk.[127] This effect appears to be especially pervasive in business contexts: managers do not consider themselves risk takers but rather "risk controllers," sophisticated actors who only take calculated, controlled risks; they think that "managerial risk taking is an endeavor in which a manager can use his *judgment*, exert *control*, and utilize *skills*."[128]

Entrants, like other business decisionmakers, are especially prone to exhibit the illusion of control, often perceiving the risks associated with their new ventures as significant but largely controllable.[129] The

---

The greater the perceived controllability of a negative event, the greater the tendency for people to believe that their own chances are less than average; the greater the perceived controllability of a positive event, the greater the tendency for people to believe that their own chances are greater than average.

Thus, typical expressions of optimistic bias appear in areas such as health, crime, and accidents, where the outcome is often dependent to some extent on the people's actions or psychological attributes, *see, e.g, supra* notes 91-92 (citing studies), while in other studies participants show a strong illusion of control as one component of their optimistic bias, even regarding events that are objectively completely out of their control. *See, e.g.*, Shelley E. Taylor et al., *Beliefs About Control, and Adjustment to Breast Cancer*, 46 J. PERSONALITY & SOC. PSYCHOL. 489 (1984) (finding breast cancer patients asserting they have a high degree of personal control over the cancer though there is little scientific evidence of such control, with many exhibiting overoptimistic assessments of their survival likelihood despite medical records to the contrary); *see also* Peter Harris, *Sufficient Grounds for Optimism?: The Relationship Between Perceived Controllability and Optimistic Bias*, 15 J. SOC. & CLINICAL PSYCHOL. 9, 24-25 (1996) (reviewing the findings on the relationship between the two phenomena, finding strong evidence for its existence in predictions of positive events and weaker evidence for its existence in predictions of negative ones).

126. Much like optimistic bias, the desirability bias has been shown to exist independently of control perceptions, although the two phenomena often coexist. *See* Budescu & Bruderman, *supra* note 104 (reviewing previous studies and providing further evidence on this relationship).

127. This is true even for behavior with life and death consequences. One study reports, for example, that only few of those gay men who engage in high-risk sexual behavior admit it, justifying their beliefs and attendant behavior by some objectively ineffective precautions they take. Laurie J. Bauman & Karolynn Siegel, *Misperception Among Gay Men of the Risk for AIDS Associated with Their Sexual Behavior*, 17 J. APPLIED SOC. PSYCHOL. 329 (1987).

128. ZUR SHAPIRA, RISK TAKING: A MANAGERIAL PERSPECTIVE 46-49 (1995) (managers' emphasis on control and skills). Reporting the findings of this in-depth study of managerial perspectives on risk using a sample of over 700 managers, Shapira notes:

> The managers .... believed that risk was manageable. Seventy-three percent of the respondents saw risk as controllable. As a result, they made a sharp distinction between gambling (where the odds are exogenously determined and uncontrollable) and risk taking (where skill or information can reduce the uncertainty). The situations they faced seemed to them to involve risk taking but not gambling.

Id. at 73 (also citing earlier studies reporting managers' perceptions of their ability to control risks).

129. Unsurprisingly, an experimental study that presented MBA students with a detailed case involving a possible venture found participants to exhibit a significant illusion of control that was further found to affect their decision to start a venture both directly and by decreasing the risk they perceived the venture to entail. Mark Simon et al., *Cognitive Biases, Risk Perception, and Venture Formation: How Individuals Decide to Start Companies*, 15 J. BUS. VENTURING 113 (1999).

factors that facilitate the illusion of control are typically found in entry decisionmaking, where entrants choose their ventures and strategies, often consider themselves familiar with both the task and the necessary behaviors (regardless of the accuracy of such perceptions), are most actively involved in the venture, and make judgments in highly competitive settings.

The illusion of control enhances the already significant effects of overoptimism and the desirability biases in causing entrants to underestimate the risks associated with entry. Often, this illusion will not only bias entrants' *probability* estimates, but also indirectly inflate their predictions of the venture's *value*, which they will not reduce enough to account for the true risk associated with their ventures.

B. *Solving the Second Puzzle: The Side Effects of Entrants' Overoptimism*

The analysis of the processes of overconfidence resolved the first puzzle of entry, showing the various biases likely to make many entrants overoptimistic in their judgments of the attractiveness of entry. These phenomena also shed light on the puzzling insensitivity of entrants to many of those market characteristics that affect their prospects.

Significantly, this framework explains why entrants are somewhat better in responding to the past profitability of industry and even to growth trends, while otherwise appearing less rational. Profitability and growth are frequently the focus of entrants' attention, the very factors that lead many of them to embark on new ventures. Although entrants will not analyze these factors perfectly, given their bounded rationality, entrants will nevertheless focus much attention on them. Entrants are likely to ignore other background variables, however, especially if these variables do not affect entrants' ability to embark on a new venture but "only" the venture's prospects. The analysis of such variables, if done at all, would therefore be more likely to fall prey to the processes of overconfidence, leading entrants to exhibit a relative insensitivity to their presence.

1. *Optimistic Bias*

The entry games of Camerer and Lovallo have shown that optimistic bias inflates experimental entrants' perceptions of their prospects, also leading them to ignore the anticipated intensity of competition.[130] The behavioral findings on optimistic bias do not explain, however,

130. Recall that the participants in these studies showed the greatest degree of excess entry when competing with others who similarly self-selected to enter skill-based competition. Camerer & Lovallo, *supra* note 10, at 311-12.

why entrants are largely insensitive to those market factors that affect their prospects but have only little to do with their knowledge or skill.

2. *Desirability Biases*

Desirability biases, unlike optimistic bias, generate overconfidence in predictions of events that entrants clearly cannot influence. An entrant predicting, for example, the direction and extent of future profitability growth in a particular industry will tend to align expectations with preferences, arriving at overly positive estimates of future trends.[131] Such biased projections may not be wholly unrelated to the objective evidence, but are nevertheless likely to be quite insensitive to variations in the data.

Some of the factors affecting entrants' prospects do not require entrants to make any significant projections, however. Evidence on entry barriers, for example, need only be rationally incorporated into determinations of a venture's attractiveness. When present in a particular industry, barriers such as capital intensity or sunk costs do not disappear rapidly. Entrants suffering from the desirability bias alone would thus be less likely to exhibit the insensitivity, revealed by the evidence of the second puzzle, to such market characteristics whose presence and impact on the prospects of entry involve little uncertainty.[132]

The bias of these entrants is better explained, however, by other desirability-related phenomena. First, recent studies of the affect heuristic have shown that when the consequences of an event carry a sharp affective meaning, variations in the probability of these consequences do not impact people's judgments. In the extreme cases of winning a large lottery or becoming ill with cancer, for example, individuals' images and feelings change little whether the probability is one in ten million or one in ten thousand, although the objective difference between these two eventualities is thousand-fold. Because these images and feelings are responsible for the impact of the affect heuristic on judgments, biased evaluations of different probabilistic outcomes vary far less than rational decisionmaking requires.[133] One study even found that when the potential outcome of a gamble is emotionally powerful, judgments of its attractiveness or unattractiveness

131. *See, e.g.*, Olsen, *supra* note 108.

132. Strictly speaking, of course, any judgment of entry involves "projections" in the minimal sense of anticipating the continued existence of market characteristics or estimating the impact of these factors on the particular venture.

133. George Lowenstein et al., Risk as Feelings (unpublished paper), *cited in* Slovic et al., *supra* note 111, at 20; *cf.* Cass R. Sunstein et al., *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)*, 107 YALE L.J. 2071, 2074 (1998) (finding that while moral judgments are widely shared, people have a great deal of difficulty in translating such judgments onto a scale of dollars).

had only a very limited impact on changes in probability as dramatic as the change from 0.99 to 0.01![134]

Second, while the planning fallacy results from a combination of cognitive and motivational factors,[135] one of the most important of these factors, the "inside view," is especially likely to make entrants insensitive to background market predictors. The inside view, also known as the "internal approach" to prediction, is very common in managerial and business decisionmaking.[136] This approach produces underestimation of time to completion and the risk of failure, because it does not adequately represent the combinatorial effect of all those things that can go wrong, and directs the decisionmakers' attention instead to their positive plans regarding the implementation of the task.[137]

When using an internal approach to prediction, people focus "on the constituents of the specific problem rather than on the distribution of outcomes in similar cases."[138] The inside view contributes to entrants' insensitivity, in addition to increasing their overconfidence, by leading them to focus on their plans and disregard distributional data.[139] Because those market characteristics affecting entrants' prospects are inevitably "distributional data" that provide general information affecting all potential entrants, the inside view leads entrants to be largely insensitive to this important statistical base-rate information.[140]

---

134. Yuval Rottenstriech & Chris K. Hsee, *Money, Kisses, Electric Shocks: On the Affective Psychology of Probability Weighting* (unpublished paper), *cited in* INTUITIVE JUDGMENT: HEURISTICS AND BIASES 20 (Tom Gilovich et al. eds., forthcoming) (on file with author). These findings are especially striking because the opposite effect — an underweighting of outcomes and an excessive reliance on probabilistic information — usually obtains when outcomes are affectively vague. *See, e.g.*, Slovic et al., *supra* note 111, at 17-19 (discussing the findings on the latter effect).

135. *See* Buehler et al., *Optimistic Time Predictions*, *supra* note 114, at 239-40, 245-46; Kahneman & Tversky, *Intuitive Prediction*, *supra* note 114.

136. Daniel Kahneman & Dan Lovallo, *Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking*, 39 MGMT. SCI. 17, 25-26 (1993); Kahneman & Tvesky, *Intuitive Prediction*, *supra* note 114, at 414.

137. *See* Maya Bar-Hillel, *On the Subjective Probability of Compound Events*, 9 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 396 (1973) (finding that people consistently exaggerate the probability of the conjunction of a series of likely events); Kahneman & Tversky, *Intuitive Prediction*, *supra* note 114, at 415 (also adding that because the combinatorial consideration is not adequately represented in people's intuitions, "[a]ttempts to combat this error by adding a slippage factor are rarely adequate, since the adjusted value tends to remain too close to the initial value that acts as an anchor") (citation omitted).

138. Kahneman & Tversky, *Intuitive Prediction*, *supra* note 114, at 415.

139. *See* Kahneman & Lovallo, *supra* note 136, at 26; Kahneman & Tversky, *Intuitive Prediction*, *supra* note 114, at 415.

140. The impact of the inside view is likely to be further reinforced by the general tendency of decisionmakers to overweight vivid anecdotal evidence and underweight pallid base rates. *See, e.g.*, Amos Tversky & Daniel Kahneman, *Evidential Impact of Base Rates*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 153 (Daniel Kahneman et al.

3. *The Illusion of Control*

The illusion of control diminishes managers' sensitivity to probability estimates, leading them to give more weight to the anticipated value of the "downside" risk of a venture than to the probability that this risk would materialize.[141] Of course, these managers are not completely insensitive to risk. When they find a venture too risky or when they think that they cannot manage the risks associated with the venture effectively, they are likely to refrain from embarking on it.[142] When a given alternative promises good enough returns but poses an unacceptable danger, however, they will seek additional information or a different perspective on the problem in an attempt to revise their probability estimates and make the venture acceptable.[143]

The illusion of control therefore contributes to entrants' relative insensitivity to objective predictors of success. Any barrier or impediment reflected in market statistics is considered a mere obstacle to be conquered. Once conquered — in the entrant's mind — the barrier is irrelevant for the prospects of the future venture and entrants will therefore be prone to disregard it when judging the venture's attractiveness.[144]

4. *Underestimating the Importance of Indirect Effects in Competitive Settings*

Another phenomenon that is especially likely to lead entrants to insensitivity toward market predictors is the tendency to underestimate indirect effects. Many decision errors people make in competitive settings result from this fundamental cognitive characteristic of human decisionmaking.[145] Indirect effects are those factors other than

---

eds., 1982) (providing a brief summary and further references to studies showing various sources of this tendency, including the representativeness and availability heuristics).

141. SHAPIRA, *supra* note 128, at 43-53; *see also* James G. March & Zur Shapira, *Managerial Perspectives on Risk and Risk-Taking*, 33 MGMT. SCI. 1404, 1411-12 (1987) (describing and discussing reasons for managerial insensitivity to probability in risk estimates).

142. *Cf.* SHAPIRA, *supra* note 128, at 74-78.

143. In doing so, the managers exhibited not only the illusion of control, but also desirability biases, adjusting their estimates to fit their preferences and aspirations. These managers also showed additional optimistic bias, thinking that *they* could perform better than expected even after making upward revisions of their estimates. *Id.*

144. *Cf.* Leslie E. Palich & D. Ray Bagby, *Using Cognitive Theory to Explain Entrepreneurial Risk-Taking: Challenging Conventional Wisdom*, 10 J. BUS. VENTURING 425 (1995) (finding, in an experimental study comparing entrepreneurs to non-entrepreneurs, that the former consistently frame ambiguous business prospects more positively, viewing a fictitious retailing company described in various scenarios as having more strengths and opportunities, fewer weaknesses and threats, and a greater future potential).

145. *See* Avishalom Tor & Max H. Bazerman, *Understanding Indirect Effects in Competitive Environments: Explaining Decision Errors in the Monty Hall Game, the Acquiring a*

the decision-maker's own behavior that affect the consequences of the decision-maker's decisions.[146] These indirect effects include background variables, such as market conditions and the decisions that other market participants make, both of which statistically determine the outcomes of the game but do not impact the actor's ability to take any particular course of action initially.

Until recently, studies have mainly explored how market actors' failure to consider the decisions made by others in the market generate common decision errors.[147] A recent comprehensive study shows that common failures in three seemingly unrelated decision problems all result from the difficulty people experience in properly addressing indirect effects in decisionmaking.[148] Strikingly, in each of the tested problems, individuals frequently made the wrong decision, although the problems required no complex reasoning.[149]

In one famous problem, decisionmakers in a takeover game systematically made negative expected value offers for a target firm.[150] The problem describes the Target and requires the decisionmaker to determine whether another company, the Acquirer, should make a tender offer for the target's shares; it then states that the Target is worth 50% more under the Acquirer than under its present management, and that the value of the Target ranges between $0-$100 per share, all values being equally likely, with the ultimate value of the Target depending on the yet unknown outcomes of an oil drill. The problem also explains, however, that the Acquirer must make its offer now, without knowing the ultimate value of the Target, while the Target — who wishes to be acquired by the Acquirer but will only agree to a profitable offer — will make the decision of whether to accept the offer only after it finds out the outcomes of its oil drill.[151]

Company Problem, and Multi-party Ultimatums (May 2002) (manuscript under review, on file with author).

146. *Id.* (manuscript at 3).

147. *See, e.g.*, John S. Carroll et al., *Negotiator Cognitions: A Descriptive Approach to Negotiators' Understanding of their Opponents*, 41 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 352 (1988) (using protocol analysis to show how participants' failure in a well-known decision problem is related to their tendency to ignore the cognitions of others); *see also* Tor & Bazerman, *supra* note 145 (manuscript at 5) (listing various studies whose findings can be explained, at least in part, by the indirect effects framework).

148. Most studies documenting such decision failures do not use the indirect effects framework that was only recently developed by Tor & Bazerman, *supra* note 145.

149. *Id.* (manuscript at 5-11) (detailing the three problems tested in this study, some of them in more than one version).

150. This problem was originally developed by William F. Samuelson & Max H. Bazerman, *The Winner's Curse in Bilateral Negotiations*, *in* 3 RESEARCH IN EXPERIMENTAL ECONOMICS 105 (Vernon L. Smith ed., 1985).

151. *See* Tor & Bazerman, *supra* note 145 (manuscript at 7-8) (providing the full text of the problem in Appendix A).

Most subjects make offers in the range between $50-$75, the current expected values of the Target under present management and the Acquirer, respectively.[152] The correct answer, however, is making no offer at all, because all offers have a negative expected value. The reason for this is simple: if the Target accepts the offer, knowing the outcomes of the oil drill, the true value under current management must be less than the offer. Since all values are equally likely, the expected value of an accepted offer is ½ of the offer price (e.g., the expected value of a $60 offer is $30). At the same time, the Target is worth 50% more under the Acquirer, resulting in an expected value of exactly ¾ of any given offer (i.e., $45 for a $60 offer). Hence, at any offer price, the expected value of acquiring the Target is negative for the Acquirer.[153]

Using the process-tracing method of protocol analysis,[154] the study found that participants fail to recognize the simple trap in the problem. They reason correctly that the Target is always worth more under the Acquirer but fail to recognize the informational asymmetry between the parties, wherein the Target makes its decision to accept only *after* the outcomes of the oil drill are known but the Acquirer must make its offer *before* these outcomes are known. Moreover, even those few subjects who recognize the asymmetry usually fail to draw the normative conclusion — that they should make no offer at all.[155]

Numerous studies show the robustness of the decision errors that result from the difficulty of incorporating even simple indirect effects into competitive decisions. Replications of the Acquiring a Company problem, for example, produced similar results with sophisticated participants including accounting firm partners, CEOs, investment bankers, and many other skilled decisionmakers.[156] Moreover, participants who were paid based on performance and given multiple trials to foster learning continued to make similar errors.[157]

---

152. *See id.* (manuscript at 8) (citing various studies).

153. *See also id.* (manuscript at 8-9).

154. Protocol analysis allows experimenters to test participants' conscious thought processes by coding their verbalizations while solving experimental problems based on a pre-existing coding scheme, and then conducting statistical analyses of these codes. For a good introduction to and description of this method, see K. ANDERS ERICSSON & HERBERT A. SIMON, PROTOCOL ANALYSIS: VERBAL REPORTS AS DATA (rev. ed. 1993); Ola Svenson, *Eliciting and Analyzing Verbal Protocols in Process Studies of Judgment and Decision making*, *in* PROCESS AND STRUCTURE IN HUMAN DECISION MAKING 65 (Henry Montgomery & Ola Svenson eds., 1989).

155. Tor & Bazerman, *supra* note 145 (manuscript at 18-27).

156. *See* BAZERMAN, *supra* note 99, at 145-49.

157. *See generally* Sheryl B. Ball et al., *An Evaluation of Learning in the Bilateral Winner's Curse*, 48 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 1 (1991); Brit Grosskopf & Yoela Bereby-Meyer, Learning to Avoid the Winner's Curse (manuscript in preparation, Harvard Business School) (on file with author).

The common difficulty decisionmakers have in integrating indirect effects therefore provides a cognitive explanation for the overconfidence of entrants, who inevitably give greater weight to their plans than to other important factors such as the behavior of potential competitors.[158] Even more significantly, entrants' difficulty in incorporating indirect effects into their judgments also explains why they appear insensitive to many market characteristics objectively found to affect their ultimate survival prospects.

C. *Solving the Third Puzzle: All Entry is Not the Same*

The behavioral framework has offered a solution to the first two puzzles of entry, suggesting that both phenomena may result from the processes of overconfidence; it also provides, however, a novel solution to the third puzzle of entry. Although the psychological literature has yet to direct significant attention to the variables that facilitate and inhibit overconfidence, a close study of these variables suggests that startups are likely to exhibit a greater bias than their diversifying competitors. The enhanced bias of startups explains both their extremely high rate of entry and their otherwise puzzling inferior performance. In providing this new solution to the third puzzle, the behavioral approach also reveals how the various perplexing findings on entry all reflect closely related aspects of entrant overconfidence.

The most important variables moderating entrant overconfidence are the intensity of the entrant's preferences (i.e., how desirable he finds the potential venture) and the ambiguity surrounding those judgments entrants must make when determining the attractiveness of entry.

1. *The Intensity of Preferences*

a. The Moderating Effects of Preference Intensity. One would expect individuals with a greater stake in the subject of their predictions to exhibit a greater degree of optimistic and desirability biases. There is some evidence that overoptimism is more pronounced for events associated with greater affect for the predictors,[159] and the magnitude of various desirability biases is clearly and strongly related to their degree of desirability.[160]

158. In fact, Tor & Bazerman, *supra* note 145 (manuscript at 5), already suggest this phenomenon may explain, at least in part, the findings of Camerer & Lovallo, *supra* note 10.

159. *See, e.g.*, Weinstein, *Unrealistic Optimism*, *supra* note 91, at 807, 810-13 (demonstrating that strength of preference affects extent of bias for positive events).

160. As McGregor already noted in his early work that predates the modern findings on the desirability bias:

> The influence of [preferences] on prediction will undoubtedly vary with the importance for the predictor of the predicted event. If his own welfare, or pride, or ethical ideals are in-

Predictions of natural occurrences, including world events,[161] election outcomes,[162] and the outcomes of sport matches[163] show, unsurprisingly, that the desirability bias depends on the strength of the predictors' preference for a particular outcome. The effects of strong preexisting preferences are practically unchanged, moreover, when decisionmakers are given monetary incentives for correct predictions.[164] In fact, a strong bias remains even when predictions are made *solely* for the purpose of profit-making that is totally dependent on the accuracy of these predictions.[165]

The intensity of preference also affects the degree of bias in predictions of otherwise neutral events, such as the drawing of a particular card from a deck. Many early studies have shown that one can enhance desirability effects by providing increasing levels of rewards for successful prediction.[166] Recent studies similarly show that incentives

---

volved, we may expect the intensity of his wishes concerning the outcome of the situation to be greater than when the situation and its outcome are relatively remote from his personal life.

McGregor, *supra* note 104, at 189-90 (citation omitted).

161. *See, e.g.*, McGregor, *supra* note 104, at 183 tbl.1 (predicting if and how likely the King of England is to announce his marriage before a certain date in 1937).

162. *See* Elisha Babad, *Wishful Thinking Among Voters: Motivational and Cognitive Influences*, 9 INT'L. J. PUB. OPINION RES. 105, 120 (1997) [hereinafter Babad, *Voters*] (finding a strong desirability bias in voters' predictions of mayoral elections, with the extent of bias being a linear function of the intensity of preference); Granberg & Brent, *supra* note 109, at 480-81 (reporting that the correlation between preference and expectation increases somewhat with political interest, concern about the outcomes of the election, and political party affiliation); *id.* at 482 (reporting additional data on how the tendency to predict one's candidate as the winner increases as a function of how favorable one's attitude is toward the candidate); McGregor, *supra* note 104, at 191 tbl.5 (showing a monotonic increase across five levels of preference among voters' predictions).

163. *See* Babad & Katz, *Against All Odds*, *supra* note 104, at 1929-32 (showing that fans' bias is a function of their level of team support).

164. *E.g.*, Babad, *Voters*, *supra* note 162, at 116 (finding that monetary rewards for accurate predictions have a significant effect, but of a very small magnitude as compared to the remaining bias).

165. Thus, one study of sports fans' desirability bias examined whether fans exhibited the bias when making anonymous predictions of game outcomes in pre-game betting stations. These fans were filling betting forms, betting on the outcomes of fourteen games that were to take place the following day, obviously aiming to make profits. While accurate predictions are always important for winning bets, moreover, this particular form of betting provided little reward for bettors predicting correctly thirteen of fourteen game outcomes, but guaranteed enormous prizes for those making fourteen correct predictions. Babad & Katz, *Against All Odds*, *supra* note 104, at 1924-26. Nevertheless, bettors still exhibited a strong bias in favor of "their" teams when betting, and even those who did not consider themselves "fans" still showed a significant bias in favor of their "preferred" teams, albeit weaker than the bias exhibited by "true" fans. *Id.* at 1931-32.

166. *See, e.g.*, Francis W. Irwin, *Stated Expectations as Functions of Probability and Desirability of Outcomes*, 21 J. PERSON. 329 (1953); Francis W. Irwin & Cynthia N. Graae, *Tests of the Discontinuity Hypothesis on the Effects of Independent Outcome Values upon Bets*, 76 J. EXPERIMENTAL PSYCHOL. 444 (1968); Francis W. Irwin & Joan Gay Snodgrass, *Effects of Independent and Dependent Outcome Values upon Bets*, 71 J. EXPERIMENTAL PSYCHOL. 282 (1966); *see also* Rose W. Marks, *The Effect of Probability, Desirability, and "Privilege" on*

for early task completion *exacerbate* the desirability-related planning fallacy, making time-to-completion predictions more overoptimistic.[167]

Experiments studying the affect heuristic reinforce the picture emerging from research on other desirability-related biases, finding the extent of this bias positively correlated with the strength of the positive or negative affect associated with the evaluated activity.[168]

b. The Enhanced Preference Intensity of Startup Entrants. The behavioral evidence confirms that those entrants for whom the potential venture is more important — financially, emotionally, or otherwise — will also tend to be more biased, exhibiting more optimistic overconfidence and, especially, stronger desirability biases. Startup entrants, however, will often harbor a greater intensity of preference for embarking on a particular entry venture than diversifying entrants. Startup entrepreneurs tend to be more involved in their ventures. They commonly have greater emotional and personal interest in the specific idea of the particular venture, often seeking to fulfill a dream or introduce an innovation to the market.[169] They may consult others,[170] but will ultimately be the ones developing the new idea into a startup venture and managing it, at least in its early stages. In addition, startup decisionmakers are also likely to invest more of their own private resources in the particular venture than decisionmakers at diversifying entrants.[171] As a result, both the failure of the particular venture and its success will often be of greater personal economic consequence for the former.[172]

---

the Stated Expectations of Children, 19 J. PERSON. 332 (1951) (showing value bias in children's estimates of the probability of drawing a picture card).

167. Stephanie J. Byram, *Cognitive and Motivational Factors Influencing Time Prediction*, 3 J. EXPERIMENTAL PSYCHOL.: APP. 216, 233 (Experiment 5) (1997) (reporting that financial incentives for speed of performance *exacerbate* the bias, with participants making shorter time predictions but performing that task no more quickly); *see also* Buehler et al., *Optimistic Time Predictions*, *supra* note 114, at 241 (finding that hopes of income tax refund increase bias in time predictions of tax forms completion); *id.* at 243 (finding that rewards for speed increase bias, although the provision of incentives for accuracy alone decreases it).

168. *See, e.g.*, Ali Siddiq Alhakami & Paul Slovic, *A Psychological Study of the Inverse Relationship Between Perceived Risk and Perceived Benefit*, 14 RISK ANALYSIS 1085, 1094-95 (1994); *see also* Finucane et al., *supra* note 111, at 9-13 (study 2: manipulating affect by providing risk and benefit information).

169. *See, e.g.*, Joshua Ronen, *Some Insights into the Entrepreneurial Process*, *in* ENTREPRENEURSHIP 137, 140-41 (Joshua Ronen ed., 1983) (discussing the objectives of entrepreneurs).

170. Although, strikingly enough, Kaish & Gilad report that those entrepreneurs who perceived themselves as successful have showed a tendency to *minimize* their use of information from any external source. Stanley Kaish & Benjamin Gilad, *Characteristics of Opportunities Search of Entrepreneurs Versus Executives: Sources, Interests, General Altertness*, 6 J. BUS. VENTURING 45, 49-50 (1991).

171. *See, e.g.*, Cooper et al., *Entrepreneurs' Perceived Chances*, *supra* note 22, at 102.

172. Also, diversifying entrants, as the empirical evidence in Part I shows, are typically larger firms. These tend to be managed professionally, by decisionmakers other than the firms' owners.

For diversifying managers, on the other hand, any new venture may be only one of several ventures and will likely be only one of their many managerial responsibilities in the diversifying firm; ultimately, therefore, they will not find most ventures as uniquely desirable as do startup entrants.[173] Moreover, when considering the prospects of a potential venture, diversifying managers will typically examine the characteristics of the specific venture and then seek the appropriate management for it; they will look within or outside the diversifying firm, but will almost never expect to run the future venture themselves. Consequently, they will not have to consider their own managerial qualities when judging the venture's prospects, thereby decreasing the potential effects of optimistic bias.[174]

While diversifying managers still face significant organizational and reputational consequences in the case of failure[175] and therefore have much at stake as well, the literature suggests that such factors make corporate decisionmakers more timid in their actual choices. Even if they maintain bold forecasts,[176] in this case, the managers' timidity will decrease the effects of their overconfident predictions, further distinguishing them from startup entrants.[177]

---

173. Managers in diversifying firms often examine or even embark on a "portfolio" of ventures, consequently having less at stake in any given one of these ventures. *Cf.* CHRISTOPHER FREEMAN, THE ECONOMICS OF INDUSTRIAL INNOVATION 156 (2d ed. 1982) (suggesting that the high degree of uncertainty associated with innovation is likely to be acceptable for large firms who can afford to use a "portfolio" approach, offsetting a very few uncertain investments against a large number of low risk projects); Richard E. Caves & Michael E. Porter, *Barriers to Exit*, *in* ESSAYS ON INDUSTRIAL ORGANIZATION IN HONOR OF JOE S. BAIN 43 (Robert T. Masson & P. David Qualls eds., 1976) (arguing that the top management of diversifying firms is well situated to make dispassionate decisions regarding the necessity of exit).

174. *See, e.g.*, Nicholas Epley & David Dunning, *Feeling "Holier than Thou": Are Self-Serving Assessments Produced by Errors in Self- or Social Prediction*, 79 J. PERSON. & SOC. PSYCHOL. 861, 871-72 (2000) (finding that participants were more likely to overestimate the likelihood they will act in generous or selfless ways, whereas their predictions of others' behavior were considerably more accurate, and presenting evidence that this divergence arises, in part, because people are unwilling to consult population base rates when predicting their own behavior but use this diagnostic information more readily when predicting the behavior of others). Of course, the venture decisions of diversifying managers' implicate their self-perception to some degree, since they are significant managerial actions. If the venture ultimately fails, the manager responsible for attempting it will be associated with the failure inasmuch as either his decision to enter or his choice of management for the venture are proven wrong.

175. There is, for example, some evidence that mature managers (i.e., older and more senior in the organization), and those belonging to larger organizations tend to be more averse to risks. *See* Kenneth R. MacCrimmon & Donald A. Wehrung, *Characteristics of Risk Taking Executives*, 36 MGMT. SCI. 422, 425-32 (1990) (studying more than 500 top-level executives testing the relationship between personal, financial, and professional characteristics and measures of risk taking).

176. *See, e.g.*, Kahneman & Lovallo, *supra* note 136.

177. *See, e.g.*, Kahneman & Lovallo, *supra* note 136. Managers may also be more averse to taking the risk of a new venture because the negative consequences they must face in case of failure are far greater than those awaiting them in case of missing a profitable opportu-

Last, because diversifying managers are situated within established organizations, their concern about organizational consequences would combine with the structured environment in which they operate, leading them to engage in a more formal decisionmaking process.[178] And while such a process would be prone to bias as well, it could direct greater attention to relevant economic information than is the case for startups, which operate in a more unstructured environment.[179]

It therefore appears that startups are highly likely to exhibit a more extreme bias, overestimating the value of entry and the likelihood of success and underestimating the risks, time, and costs associated with entry more than their diversifying competitors.[180]

2. *Judgmental Ambiguity*

Much like the intensity of preference, the amount and type of information available to entrants and the resulting level of ambiguity they face when making their judgments regarding the attractiveness of entry is another variable that moderates the processes of overconfidence.

a. Differentiating Ambiguity from Lack of Information. If all relevant information were known with certainty, there would be no bias; entrants would simply *know* the fate of their ventures and would only enter upon knowing they will succeed. In reality, there is no certainty as to the future of the venture and all predictions and judgments are

---

nity, which is also much harder to identify. *See, e.g.*, Raghu Garud et al., *Technological Choices and the Inevitability of Errors*, *in* TECHNOLOGICAL INNOVATION, *supra* note 9, at 20, 23-27, 32-34 (showing also how these two decision errors relate to and affect one another).

178. *See, e.g.*, Ken G. Smith et al., *Decision Making Behavior in Smaller Entrepreneurial and Larger Professionally Managed Firms*, 3 J. BUS. VENTUR. 223 (1988) (reviewing studies suggesting managers are more likely to follow rational, formal decision processes than entrepreneurs, corroborating these conclusions with a field study, and finding a correlation between decision process and organizational performance).

179. Thus, one of the frequently suggested methods in the managerial and decisionmaking literatures for battling the effects of biases is to engage in a more formal and structured judgmental process, resembling more of an "outside" rather than an "inside" view. *See, e.g.*, BAZERMAN, *supra* note 99, at 157-64 (suggesting a number of methods to debias managerial judgment); Kahneman & Tversky, *Intuitive Prediction*, *supra* note 114, at 417-21 (describing a corrective procedure for prediction); *see also* Epley & Dunning, *supra* note 174, at 872 (finding that people predict others' behavior better than they predict their own behavior in part because they use diagnostic information more readily when predicting the behavior of others).

180. There is also some evidence that decisionmakers interacting with others or making a decision together often provide more accurate responses. *E.g.*, Carl Martin Allwood & Par Anders Granhag, *Realism in Confidence Judgments as a Function of Working in Dyads or Alone*, 66 ORG. BEHAV. & HUM. DEC. PROC. 277, 277-79, 286-88 (1996) (reporting better performance of pairs as compared to individual subjects). If this were true, diversifying managers would probably enjoy an additional advantage over startups.

probabilistic. The degree of uncertainty, however, will vary among ventures and entrants. It is therefore plausible that those entrants whose prospects are more uncertain will also be more biased.[181]

Importantly, however, an increase in the *amount* of information alone may not decrease entrants' bias, as the psychological literature clearly shows that additional information rarely changes people's pre-existing attitudes.[182] To the extent entrants' preferences shape their judgments regarding the attractiveness entry, a mere change in the amount of information may not exert a significant impact on their judgments. Nonetheless, the findings we discuss below suggest that when the decision environment is more ambiguous — a function of the *quality and type* of one's knowledge as well the amount of information — entrants are likely to be more biased. The presence of uncertainty provides opportunity for bias while the ambiguity of the decision environment "legitimates" the operation of egocentric, self-serving perceptions. A less ambiguous environment, to the contrary, limits the manifestation of these biases.[183]

Two recent studies examining the effects of desirability on voters' predictions of election outcomes provide a clear example of the distinction between the respective roles of information and ambiguity in the processes of overconfidence. In one study, voters who were better informed about the outcomes of previous elections exhibited a degree of bias similar to that exhibited by uninformed voters. These voters'

181. *See, e.g.*, Jovanovic, *supra* note 66 (offering a model in which learning from post-entry performance reduces uncertainty and improves judgment).

182. *See, e.g.*, Loewenstein et al., *supra* note 107, at 145-55 (showing how the interpretation of neutral information can be biased even by an ad-hoc designation of participants to one side in a hypothetical litigation). *See generally* KUNDA, *supra* note 88, at 111-160 (discussing general evidence on biased hypothesis testing). Studies show, for example, that the supply of additional information increases overconfidence in the accuracy of one's judgment because it tends to increase judgmental confidence much more than it increases accuracy. *See, e.g.*, Stuart Oskamp, *Overconfidence in Case-Study Judgments*, 29 J. CONSULTING PSYCHOL. 261 (1965), *reprinted in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 287 (Daniel Kahneman et al. eds., 1982) (a classic study of this effect in clinical judgments). *But see* Dane K. Peterson & Gordon F. Pitz, *Effects of Amount of Information on Predictions of Uncertain Quantities*, 61 ACTA PSYCHOLOGICA 229 (1986) (finding that the provision of relevant information decreases overconfidence in the accuracy of judgment and suggesting that this result does not conflict with findings like those of Oskamp who provided participants with redundant information).

183. *See* Granberg & Brent, *supra* note 109, at 479 (concluding, after providing evidence on the role of ambiguity in the effect of voter preferences on election outcome expectations that "a certain degree of ambiguity may be necessary for the effect to occur, and a lack of ambiguity may diminish the effect . . ."); McGregor, *supra* note 104, at 193-195 (discussing the effects that the amount of information and its nature have on predictions). More generally, Kunda notes when commenting generally on the processes of motivated reasoning:

> Motivation can color our judgments, but we are not at liberty to conclude whatever we want to conclude simply because we want to. Even when we are motivated to arrive at a particular conclusion, we are also motivated to be rational and to construct a justification for our desired conclusion that would persuade a dispassionate observer.

KUNDA, *supra* note 88, at 224.

bias was moderated by their intensity of preference alone.[184] In another study, however, voters who were more familiar with the results of recent election polls were less biased than voters who were unfamiliar with or had mistaken recall of these poll outcomes.[185]

The different kinds of information possessed by the voters explain the seeming contradiction between these two studies. In the first study, informed voters had information about past election outcomes that, although relevant, still left an enormous degree of ambiguity as to the outcomes of the present elections. Under these circumstances, the moderating role of preference intensity dominated and the incremental informational advantage of some voters had no discernible effect on their performance. The second study, on the other hand, examined the effects of highly relevant information. Given accurate knowledge of recent poll outcomes, voters faced a far less ambiguous prediction task. They could still be highly biased, as shown by the study, because many uncertain factors may cause election outcomes to differ from those reflected in recent polls. The lower-ambiguity setting of the latter voters, however, constrained their bias due to ambiguity more than the knowledge possessed by the first-study voters, causing informed voters to exhibit a significantly lower bias than did their uninformed counterparts.[186]

b. The Moderating Effects of Judgmental Ambiguity. The experimental literature provides strong support for the existence of an intuitive relationship between bias and ambiguity. In the case of optimistic overconfidence, a number of studies show that predictions of ambiguous tasks and target events increase the predictors' bias, presumably because ambiguity allows individuals to rely more on the information that puts them in a more favorable light. For example, subjects exhibit more overoptimistic absolute and comparative judgments of themselves when they can choose on which dimension to focus their judgments.[187] They also show a greater bias when making judgments on matters where only limited feedback is available to them being, for example, more likely to overestimate the morality of their behavior than the intelligence it reflects,[188] more likely to overestimate their

184. This similarity persisted, moreover, regardless of whether voters were better informed due to their pre-existing knowledge or because the experimenter provided them with partial or full base-rate information about previous elections' outcomes. Elisha Babad, *Can Accurate Knowledge Reduce Wishful Thinking in Voters' Predictions of Election Outcomes*, 129 J. PSYCHOL. 285 (1995).

185. Although informed voters were highly biased as well, the magnitude of their improvement due to better information was very small as compared to their overall bias. Babad, *Voters*, *supra* note 162.

186. *Id.* (discussing the differences between these studies).

187. *See* Dunning et al., *Ambiguity and Self-Evaluation*, *supra* note 92.

188. Scott T. Allison et al., *On Being Better But Not Smarter than Others: The Muhammad Ali Effect*, 7 SOC. COGNITION 275 (1989) (especially studies 2 and 3).

physical attractiveness than their intelligence,[189] and more likely to deem themselves sophisticated than tall.[190]

These findings are further corroborated by a study of negotiation and conflict resolution showing that the ambiguity of complex situations increases self-serving perceptions of fairness and negatively affects negotiators' bargaining behavior.[191] Interestingly, in this study an increase in information was associated with increased, rather than diminished, ambiguity, since subjects had no background knowledge of the case they were bargaining over.[192] Participants were told the experiment involved negotiation and conflict resolution, and were randomly assigned to the roles of representatives negotiating a labor agreement between a teachers' union and a Board of Education.[193] The results showed, as expected, that those teachers' representatives who received a greater amount of neutral background information, which was interpreted by objective decisionmakers as equally supportive of both parties' positions, exhibited exacerbated egocentric perceptions of what would constitute a fair salary for the teachers.[194]

In addition to the evidence of the impact of ambiguity in overoptimism, recent findings suggest that highly unskilled actors exhibit the most extreme overconfidence. These actors' incompetence saddles them with a double burden: Not only do they perform poorly because they lack relevant skill and knowledge, but they also find it difficult to identify their errors. When their skill and knowledge are enhanced, however, their overconfidence is reduced because their performance

189. Gabriel et al., *supra* note 91, at 149 (finding a medium correlation between participants' self-rated intelligence and their performance on objective measures and no correlation at all between self-rated and other measures of attractiveness, for both men and women).

190. *See, e.g.*, Dunning et al., *Ambiguity and Self-Evaluation*, *supra* note 92 (reporting a series of studies showing that subjects show a greater bias when making self-judgments of more ambiguous traits that are reflected in a wider range of potential behaviors); *see also* Judith Weiner Regan et al., *Do People Have Inflated Views of Their Own Ability?*, 31 J. PERSONALITY & SOC. PSYCHOL. 295, 295 (1975) ("Self-esteem needs must be aroused in a situation in which distortion of one's ability level is possible. This requires that the ability dimension be important enough to arouse self-esteem needs and that actors be *unsure of how skilled they are* on this dimension.") (emphasis added).

191. Leigh Thompson & George Loewenstein, *Egocentric Interpretations of Fairness and Interpersonal Conflict*, 51 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 176, 184-96 (1992) (especially Experiment 2).

192. This is in striking contrast to real-world judgments of entrants, which are always made under conditions of extreme uncertainty, and where more and better relevant information would be likely to decrease the effects of egocentric biases on prediction to the extent it decreased ambiguity. The authors of the study, however, seem to suggest that increases in the amount (though not the quality) of information may be generally associated with a greater degree of egocentric bias. *Id.* at 182.

193. *Id.* at 184-86.

194. *Id.* at 188-89 (noting, however, that the effect was asymmetric, with Board representatives with more background information showing a level of bias not significantly different from that exhibited by those with less background information).

improves while they also develop a more realistic appraisal of their limited ability.[195]

The clarity of either the task or the relevant information is also an important moderator of the effects of desirability biases.[196] Surveys find voters more biased in favor of their preferred candidate, for example, when the objective winning odds (as measured by the overall proportion of surveyed voters' support) are better[197] and less biased when they have more accurate knowledge of relevant information.[198] Similarly, sport fans exhibit less biased predictions in favor of "their" teams when making predictions at half-time as compared to pre-game predictions.[199] Last, experimental studies have also shown participants to make more biased outcome predictions when the probability of the target event is about equally likely to occur or not to occur,[200] or when making judgments on the basis of more ambiguous stimulus phrases (e.g., "a good chance") as compared to less ambiguous ones (e.g., "improbable").[201]

c. The Enhanced Ambiguity Facing Startup Entrants. Not all types of entrants possess similar information or face the same level of ambiguity when deciding whether to embark upon entry. Diversifying entrants who already have relevant managerial and business experience face a less ambiguous task, because they know their relevant strengths and weaknesses with greater certainty.[202] Moreover, since these entrants are already established in related industries, much of their specific market knowledge, regarding factors affecting the costs of production and marketing, potential customers and suppliers, and similar factors affecting their prospects, will be relevant to the decision on the

195. Justin Kruger & David Dunning, *Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments*, 77 J. PERSONALITY & SOC. PSYCHOL. 1121, 1130-31 (1999).

196. *See, e.g.*, McGregor, *supra* note 104, at 183-89.

197. *See, e.g.*, Granberg & Brent, *supra* note 109, at 478-79 ("[T]he tendency to predict a preferred outcome is stronger in the high-ambiguity years.").

198. *See, e.g., id.* at 482 tbl.2; Babad, *Voters*, *supra* note 162, at 118-19.

199. Babad & Katz, *Against All Odds*, *supra* note 104, at 1923 (citing an earlier study by the first author) (although a significant desirability bias remained at half-time for trailing team).

200. Irwin, *supra* note 166, at 331-33.

201. Brent L. Cohen & Thomas S. Wallsten, *The Effect of Constant Outcome Value on Judgments and Decision Making Given Linguistic Probabilities*, 5 J. BEHAV. DECISION MAKING 53, 61-62, 65-66 (1992); *see also* Dan Zakay, *The Relationship Between the Probability Assessor and the Outcomes of an Event as a Determiner of Subjective Probability*, 53 ACTA PSYCHOLOGIA 271, 277-78 (1983) (finding a greater bias in the prediction of ambiguous target events).

202. This is not to say, of course, that managers are not overconfident in their ability and skill, as the findings discussed *supra* Section II.A have shown. Also, some startup entrants will have better knowledge and experience than others. *See supra* Section I.C.1.c (addressing learning from repeated entry).

new venture.[203] In addition, they will often have better access to background statistical data and better means to analyze this data more effectively.[204]

Startup entrants, in contrast, will generally have less experience and knowledge. Some may have experience in the industry but often lack managerial expertise. Those startup entrants who have never started a new venture will be even less aware of the factors likely to affect its long-run performance. They will not only have less specific knowledge of performance, but fewer resources to obtain such information and analyze it effectively. Startups will therefore face a more ambiguous decision task, and will tend to exhibit more overconfident predictions regarding the future of their contemplated ventures.[205]

A survey of 2,994 new small business owners from a variety of industries[206] supports these conclusions. On average, these entrants exhibited a great degree of optimistic bias. Almost all of the entrepreneurs in this survey (95%) perceived their odds of success as being greater than five out of ten, 81% perceived odds greater than seven out of ten, and a full 33% thought their chances were a "dead certain" ten out of ten.[207] Moreover, when making comparative judgments, only 5% thought their chances poorer than those of similar businesses, about 27% thought their chances as good as others, and as many as 68% of the respondents perceived their odds of success as better than those of similar entrants.[208] These inflated perceptions did not change with time since entry, with entrants who have been in business sixteen months or more exhibiting patterns similar to those who had just started their businesses.[209] In fact, perceptions of the entrants in this study were also remarkably unrelated to various demographic factors that affect one's probability of success. Those who were poorly prepared for the entry task were just as optimistic as their better-prepared

203. *See* BIGGADIKE, *supra* note 63, at 90-115 (discussing and analyzing the benefits diversifying entrants derive from the "relatedness" of their new venture to the experience of the parent company).

204. Better and cheaper access to information and its processing may provide a rational justification for the search for and use of more information, and the opposite is true when information is costly to obtain and process. This would matter, however, only when the marginal costs of information acquisition and processing are of the same order as the benefits of more accurate judgments, which would rarely be the case for entry with its high attendant costs and benefits.

205. *See* Kruger & Dunning, *supra* note 195 (showing how low skill promotes biased estimates).

206. Notably, most of these new entrants were not in manufacturing industries, *see* Cooper et al., *Entrepreneurs' Perceived Chances*, *supra* note 22, at 102, although there is no reason to expect this to impact their overoptimistic perception.

207. *Id.* at 103.

208. *Id.* at 104.

209. *Id.* at 105-06.

counterparts,[210] presumably because of their strongly inflated self-perceptions.[211]

Similarly, the small entrants participating in the study also showed strong desirability effects, revealing an extreme bias when judging the odds of "any business like theirs." With estimates far exceeding those warranted, 78% of the surveyed entrants estimated their odds of success at five of ten or better, 39% estimated them at seven of ten or better, and 16% thought these odds nine of ten or better![212]

The conclusion that startup entrants exhibit a greater degree of boundedly rational decisionmaking is further buttressed by the anecdotal evidence on patterns of information gathering and processing by startup entrants.[213] The few studies to survey these entrepreneurs have generally revealed patterns incompatible with rational decisionmaking. Specifically, one study found that those who were inexperienced searched more intensely for information, but preferred personal sources, such as family and friends, to professional ones.[214] When planning a new venture in business domains where they had less relevant experience, these entrepreneurs tended to search less, rather than more, for information.[215] Another study has similarly reported that while entrepreneurs were generally more alert to new opportunities, they directed their information search to less formal sources. Importantly, they were less concerned with economic and market factors and

210. *Id.* at 105.

211. These survey findings strongly echo the experimental data of Kruger & Dunning, *supra* note 195, at 1121, who showed how low skill not only decreases performance but also deprives decisionmakers of the necessary meta-cognitive tools to observe their low level of skill.

212. Cooper et al., *Entrepreneurs' Perceived Chances*, *supra* note 22, at 103.

213. While the few existing studies of decisionmaking and information acquisition by entrepreneurs do not make a formal distinction between the two entrant types we examine, they tend to equate "entrepreneurs" with younger and often smaller startup entrants. *See* Lowell W. Busenitz & Jay B. Barney, *Differences Between Entrepreneurs and Managers in Large Organizations: Biases and Heuristics in Strategic Decision-Making*, 12 J. BUS. VENTURING 9, 17 (1997) (comparing the decisionmaking of entrepreneurs to that of managers in large organizations, using a sample of 124 founders of firms who started their ventures, on average, 1.7 years before to participating in the study); Arnold C. Cooper et al., *Entrepreneurial Information Search*, 10 J. BUS. VENTURING 107, 111 (1995) [hereinafter Cooper et al., *Entrepreneurial Information Search*] (using a sample of 1,176 founders whose new ventures were, on average, 11 months old at the time of the survey); Kaish & Gilad, *supra* note 170, at 50 (surveying 51 founders of New Jersey companies appearing on the mailing list of the state's Small Business Administration Agency, without reporting time since venture inception); Ken G. Smith et al., *Decision Making Behavior in Smaller Entrepreneurial and Larger Professionally Managed Firms*, 3 J. BUS. VENTURING 223, 227 (1988) (using size as measured by number of employees as proxy for "entrepreneurial" firms, with the fifteen such firms in the sample having between two to fifty employees).

214. Cooper et al., *Entrepreneurial Information Search*, *supra* note 213, at 114-15.

215. *Id.* at 115.

more concerned with factors relating to the implementation of their plans, in contrast to managers in a large financial conglomerate.[216]

In sum, the two most important moderators of entrants' overconfidence — the *intensity* of entrants' preference for the particular venture and the level of *ambiguity* they face when making their judgments — vary systematically with entrant type, making startups prone to exhibit an enhanced bias. This novel conclusion solves the third puzzle of entry, suggesting that startups enter more frequently because of their extreme overconfidence, which contributes to their inferior average performance that would not have been exhibited by non-biased entrants.[217] Finally, this analysis shows how the behavioral framework provides a coherent and effective explanation for the various empirical findings on entry.

III. THE CONSEQUENCES OF ENTRANT-BOUNDED RATIONALITY

The psychology of entrant overconfidence provides a clear explanation for the prevalence of boundedly rational, negative expected value entry and the relative insensitivity of entrants, particularly startups, to many of the economic factors affecting the future success of their ventures. While providing answers to the puzzles of entry, these conclusions also confirm the widespread presence of boundedly rational entry. The following sections examine how the presence of significant boundedly rational entry alters the dynamics of entry competition.

A. *How the Bounded Rationality of Entrants Transforms the Competitive Landscape*

Entrants who overestimate their prospects are more likely to fail than entrants who make accurate average estimates, but their presence also decreases other entrants' probability of success and changes the composition of the final cohort of successful entrants. Boundedly rational entrants diminish the prospects of all entrants because their high-volume attempts substantially increase the intensity of competition. Their influence on market outcomes for all entrants, however, goes well beyond an across-the-board detraction from the net present value of entry.

---

216. Kaish & Gilad, *supra* note 170, at 55-56. This latter finding suggests that these startups applied an "inside view" to a greater degree than the other managers in the sample, resembling the greater willingness of decisionmakers to use base-rates in predictions concerning others, as observed by Kruger & Dunning, *supra* note 195.

217. Thus, rational startups might have been expected to enter at high rates given the very large absolute number of potential startups, but they would not have exhibited an inferior average performance. *See supra* Section I.A.3 (discussing the third puzzle of entry).

First, overconfident startups that "luck out" displace some diversifying entrants with better ex ante prospects, effectuating a process of *probabilistic replacement*.[218] Probabilistic replacement is the inevitable consequence of competition under uncertainty: some competitors with a low ex ante probability of success will always succeed, while some having better ex ante prospects will ultimately fail.[219] The impact of this replacement is more pronounced in the competition because of the highly disparate base-rates (i.e., the number of potential entrants) of startup versus diversifying entrants and the extreme overconfidence of startup entrants.[220]

Because startup entrants exhibit a strong bias regarding the prospects of their potential ventures, the proportion of startups attempting entry is larger than it would have been if their entry predictions were less biased. Their absolute numbers end up being very large, since there is an almost unlimited number of potential negative expected value ventures that could appear subjectively attractive to some startup entrant. Diversifying firms, on the other hand, exhibit the opposite pattern: not only are they far fewer in number, being limited to those already established firms that may have an interest in the particular industry, but they are also less biased than startups. The proportion of startups among successful entrants is thus far greater than what it would have been if ex post outcomes were to mirror the respective ex ante probabilities of success of the two entrant types. In other words, while startup entrants are individually *less* likely to sur-

218. The major role of luck in economic selection processes has been persuasively discussed in the influential work of Armen A. Alchian, *Uncertainty, Evolution, and Economic Theory*, 58 J. POL. ECON. 211 (1950). Building on this work, scholars have shown through simulation that luck may be a more parsimonious explanation for differences in the profitability of firms than many other factors known to be associated with profitability. *See, e.g.*, Richard B. Mancke, *Causes of Interfirm Profitability Differences: A New Interpretation of the Evidence*, 88 Q.J. ECON. 181 (1974). This claim has met with some disagreement, although contrary arguments have been found hard to prove. For a brief review, see Jay B. Barney, *On Flipping Coins and Making Technology Choices: Luck as an Explanation of Technological Foresight and Oversight*, *in* TECHNOLOGICAL INNOVATION: OVERSIGHTS AND FORESIGHTS 13, 16-17 (Raghu Garud et al. eds., 1997).

219. *See* Richard R. Nelson & Sidney G. Winter, *Forces Generating and Limiting Concentration Under Schumpeterian Competition*, 9 BELL J. ECON. 524, 524-25 (1978) (introducing an influential evolutionary model by describing competitive outcomes saying that "[i]ndeed, a situation that is regarded as "highly competitive" is typically one in which *luck is the principal factor* that finally distinguishes winners from near-winners . . . .") (emphasis added).

220. Note that absent competition between startups and diversifying firms, probabilistic replacement would have occurred only within each of the two types. In reality, however, startup entrants and diversifying firms do compete with one another, even if just because a particular market can accommodate only so many entrants at any particular point in time. To the extent that the two groups of entrants compete for market share, therefore, some overconfident startups will succeed despite their low ex ante probability of success, while some diversifying firms with much better prospects will fail.

vive on average, they are likely, as a group, to replace some of their diversifying competitors who appeared better-qualified ex ante.[221]

The effects of probabilistic replacement on the composition of the successful entrant cohorts are further exacerbated when some diversifying entrants engage in a conscious process which may be termed *negative selection*, whereby they choose to avoid entry altogether. Because of diversifying entrants' ability to process the relevant information for entry more rationally, some of these entrants are likely to take into account the reduction in the attractiveness of their ventures due to the increased intensity of competition from extremely overconfident startups. Consequently, at the margin, some of these entrants might refrain from entering altogether.

The initially puzzling statistics on entry — revealing systematically different survival rates for the two entrant types which are altogether lower than what would be expected under rational entry — may therefore reflect not only the effects of disparate base rates and differential overconfidence, but also the additional impact of probabilistic replacement and negative selection that further lower diversifying entrants' entry and survival rates.[222] In this way, the operation of competitive pressures on a background of entrant bounded rationality draws a post-entry landscape with an increased proportion of ex ante more biased, less qualified entrants, in contradiction to the conventional view of competition as the "survival of the fittest."[223]

---

221. *Cf.* J. Bradford De Long et al., *The Survival of Noise Traders in Financial Markets*, 64 J. BUS. 1 (1991) (developing important early model). This model shows how noise traders (i.e., investors with incorrect expectations about return variances), as long as they do not affect market prices, can earn higher expected returns than rational investors with similar risk aversion due to the relationship between risk and return. *Id.* at 3. Even more importantly, under these assumptions, when "survival" is defined as an investor's group share of total wealth almost surely not approaching zero and "dominance" is defined as a higher-than-half long run probability of an investor's group increasing its share of total wealth:

> Noise traders as a group might survive and come to dominate rational investors in wealth even when on average a rational investor dominates any noise trader of a fixed type in wealth .... But the wealth of noise traders as a group relative to that of rational investors as a group need not tend toward zero, for the downward drift imparted by idiosyncratic risk does not affect noise traders' collective wealth. *If idiosyncratic risk is large, each individual noise trader with high probability fails to survive in the market, but noise traders as a whole can nevertheless survive.* Evolution may leave an ever-shrinking army of ever-richer fools who collectively dominate the market.

Id. at 3-4 (emphasis added).

222. Probabilistic replacement and negative selection create a selection bias not only in favor of startups as compared to diversifying entrants, but also in favor of the more overconfident entrants within each of these entrant types. For simplicity's sake, however, our analysis will focus on the main between-type effects and on the prevalence of negative expected value entry.

223. One scholar has recently used the extreme metaphor of a large-scale coin-flipping contest to highlight that the ex post survival of firms "cannot be taken as definitive proof that some firms are somehow 'better,' 'more efficient,' or 'more able' " than their failed competitors. *See* Barney, *supra* note 218, at 15. The study describes a hypothetical competition between many thousands of technology management experts, who train hard in ad-

B. *The Limited Efficacy of Financier Gate Keeping*

One obvious question raised by the prevalence of boundedly rational entry is *how* overconfident entrants are able to obtain the necessary resources for rationally unjustified entry attempts. The empirical findings of entry make it clear that financiers whose support is often necessary for entrants to embark on new ventures do not prevent effectively boundedly rational entry from taking place.[224] The seeming ease with which boundedly rational entrants attempt entry is surprising, however, since private financial actors pursuing their own interests should be loath to supply them with funds. A brief analysis of this phenomenon will not only help understand why financiers are not as effective gate-keepers as might have been expected, but also provide a first glimpse at the difficulties involved in any attempt to limit boundedly rational entry.

In contrast to diversifying entrants who can rely on internal resources, startups rarely have sufficient resources to embark on an entry venture. Unless their venture is small, they will often have to seek outside financing beyond family and friends.[225] One would therefore expect these entrants to face the close scrutiny of financiers, resulting in few startups making negative NPV attempts. In fact, if financiers were rational actors and the private optimality of financiers and entrants were fully aligned, we would expect to find no boundedly rational entry by startups. If this were the case, the *average* performance of startup entrants would have been far superior to that exhibited by diversifying firms relying on their internal resources.[226]

---

vance. At the end of a number of rounds only 156 contestants remained, and many of their failed competitors sought their advice and secrets so they would be better prepared for future competition. Ultimately, after 18 rounds, the tale finds one successful winner, whose success from that moment on was insured by a combination of the value of the coins won from other contestants and the high value others were willing to pay for participation in his coin-flipping seminars. *Id.* at 13-15. A similar metaphor is already attributed by Alchian, *supra* note 218, at 214-15, to the French mathematician Borel to show that pure chance processes may generate long-lived firms of disparate sizes.

224. A full-fledged analysis of the interaction between financiers' judgments and entrant bounded rationality is outside the scope of the present analysis, which only highlights some major aspects of this interaction.

225. These sources include professional venture capital funds, private investors, and even banks. *See, e.g.*, Paul A. Gompers, *Resource Allocation, Incentives and Control: The Importance of Venture Capital in Financing Entrepreneurial Firms*, *in* ENTREPRENEURSHIP, SMALL AND MEDIUM-SIZED ENTERPRISES AND THE MACROECONOMY 206, 207-11 (Zoltan J. Acs et al. eds., 1999) (describing the main sources of funding available to entrepreneurs, with an emphasis on the role of venture capitalists); William E. Wetzel, Jr., *Angels and Informal Risk Capital*, 24 SLOAN MGMT. REV. 23 (1983) (discussing the presence and role of an intermediate source of financing for startups).

226. As noted *supra* Section I.A.3, in the discussion of the third puzzle of entry, rational startups could enter more frequently, but their average performance would then have to be better than that of diversifying entrants.

Even given the inevitable bounded rationality of financiers, we would still expect them to be no more biased than diversifying entrants, and possibly even less. The low intensity of preference financiers have for any particular venture and the reduced ambiguity in the decision environment they face given their continuous exposure to market information both suggest they would exhibit a weaker bias. These financiers should therefore be able to reduce the flood of overconfident startup entry attempts dramatically.

A comparison of the decision environment facing startup financiers and startup entrants explains, however, why the former do not significantly curb the boundedly rational behavior of the latter. First, financiers and entrants face differing risk exposures.[227] Some financiers are quite diversified[228] and may therefore rationally support some riskier ventures to complement low-risk, low-return investments.[229] Moreover, the financiers' risk exposure is determined, at least in principle, by their financing arrangement with the entrants rather than by the entrants' prospects directly. Consequently, a negative expected value venture might still promise the financier a positive net present value. This could be done by limiting expected losses via security interests in property and other guarantees and enhancing expected gains by allocating to the financier a proportionally large portion of the returns to success.

Second, financiers may have a limited ability to distinguish overconfident entrants from others. This problem would not hinder many financiers' ability to remain in business, because of diversification and deep pockets, as well as because some overconfident entrants will do well and provide them with large returns. These successes will further cement financiers' erroneous conviction in their ability to identify attractive startups even when this is not the case.[230]

Third, the more an entrant has succeeded in exciting the venture capitalist and interesting him in the details of the planned venture, the more he has converted the latter from an objective observer to an in-

227. *See* MARK CASSON, THE ENTREPRENEUR: AN ECONOMIC THEORY 210-13 (1982) (describing the problem faced by the entrepreneur, broadly defined, in obtaining financing on the basis of private, non-consensus information, and the ways in which financiers can reduce their risk exposure when lending to entrepreneurs).

228. *See* Gompers, *supra* note 225, at 212-15, tbl.8.1 (summarizing selected data on the U.S. venture capital industry from 1980-1994).

229. Although financiers should still never finance an entry attempt that is negative NPV for them.

230. The high level of uncertainty associated with venture capitalists' business may also lead to a selection bias, where overconfident startups circulate among financiers until, at times, one happening to find their proposal more attractive exhibits a winner's curse and funds them. *See* RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE (1994) (discussing evidence of the tendency of winning bidders to overpay, since the average rather than the highest bid provides the best value estimate).

sider. Inasmuch as this conversion has occurred, financiers will be likely to exhibit a degree of bias, losing some of their relative advantage with respect to variables that impact overconfidence.[231]

It is possible, however, that financiers do perform an important, if not fully effective, gate-keeping role that is already reflected in the statistics of entry. Even though they assist overconfident entrants in their attempts, they reject some negative expected value ventures that consequently never take place.[232] This is especially true of professional venture capital funds. These funds, despite their high status and resources, finance only an exceedingly small proportion of the multitude of new business ventures started annually.[233] And, when they do provide financing, venture capitalists provide close supervision and advice to the entrants, thereby arguably increasing those ventures' likelihood of survival.[234]

Some overconfident entrants who are unable to secure outside funding, however, are still likely to attempt entry, albeit with personal or other non-professional sources of funding and at a smaller scale. The evidence on the relationship between entrant size and survival lends some support to this hypothesis,[235] suggesting that the smaller

---

231. Once they become more interested in a particular venture, venture capitalists will show some of the effects of preference on judgment, mainly because of the desirability-related affect heuristic and inside view. *See supra* Section II.C.1. To the extent they intend to support the venture by managerial guidance in addition to funding, moreover, financiers inevitably increase the impact of their own preferences on their judgments (e.g. "I would love to be involved in this type of venture."). *See supra* Section II.C.2.

232. *See* Wetzel, *supra* note 225, at 28 (describing how "angels" reject many funding proposals).

233. *See* Gompers, *supra* note 225, at 212-15 tbl.8.1 (reporting that venture capitalists never funded more than 1,729 companies a year in the period between 1980-1994). Moreover, even those entrants that succeed in secure VC financing, typically do so when their ventures are more advanced and have an established track record. Rapahel Amit et al., *Venture Capital Financing of Entrepreneurship: Theory, Empirical Evidence and a Research Agenda*, *in* THE BLACKWELL HANDBOOK OF ENTREPRENEURSHIP 259 (Donald L. Sexton & Hans Landstrom eds., 2000) (reporting findings based on Canadian data).

234. There is evidence that venture capital support extends beyond mere financing and is correlated with higher survival rates (although venture capitalists may merely detect some superior ex ante entrants). *See, e.g.*, Jeffry A. Timmons & William D. Bygrave, *Venture Capital's Role in Financing Innovation for Economic Growth*, 1 J. BUS. VENTURING 161 (1986); *see also* Harry J. Sapienza, *When Do Venture Capitalists Add Value?*, 7 J. BUS. VENTURING 9 (1992) (finding that entrepreneurs and venture capitalists share this perception of the VC role in promoting entry success). This may also be the case, to some extent, with private financing. *See* Wetzel, *supra* note 225, at 27 (describing how "angels" are typically active investors who provide more than just capital). *See generally* Sophie Manigart & Harry Sapienza, *Venture Capital and Growth*, *in* THE BLACKWELL HANDBOOK OF ENTREPRENEURSHIP 248-52 (Donald L. Sexton & Hans Landstrom eds., 2000) (reviewing evidence on VCs' selection of ventures for funding and how they promote the latter's growth by supplying monitoring, advising, and assistance).

235. *See, e.g.*, Dunne et al., *Growth*, *supra* note 53, at 676 (raw statistics in tbl.1); *id.* at 686-89 (larger entrant size associated with increased survival rates); *see also supra* notes 55-56 and accompanying text (discussing the effects of small entrant size).

and typically suboptimal scale of many startups is not only a *cause* of their greater mortality, but also a *symptom* of their inability to obtain external financing[236] or their more extreme underestimates of the resources necessary for their ventures, both results of their enhanced bias.

C. *The Consequences of Boundedly Rational Innovative Entry*

NEV entry exerts a far-reaching effect on market outcomes in an unregulated environment that even those private economic actors best situated to do so fail to prevent. Before examining the natural question of whether the government should regulate entry, this Part highlights the close association between the introduction of innovation, broadly construed,[237] and boundedly rational entry.[238]

1. *Bounded Rationality and Innovative Entry*

Startups' inferior performance results from both psychological and economic factors: the variables affecting overconfidence lead these entrants to exhibit an enhanced bias, while the economics of entry simultaneously channel higher-risk ventures to the startup route and make startup more risky than diversifying entry. These factors also suggest, however, that the introduction of innovation is related to a higher frequency of negative expected value entry attempts, and that innovative

236. This is especially true for companies that lack substantial tangible assets and have a large degree of uncertainty about their future, who are unlikely to receive significant bank loans given that they face many years of negative earnings and are unable to make interest payments or meet principal repayments. Consequently, such firms face severe financial constraints and are largely dependent on the availability of venture capital funding.

237. *See* Giovanni Dosi, *The Nature of the Innovative Process*, *in* TECHNICAL CHANGE AND ECONOMIC THEORY 221, 222 (Giovanni Dosi et al. eds., 1988) ("In an essential sense, innovation concerns the search for, and the discovery, experimentation, development, imitation, and adoption of new products, new production processes and new organisational setups."); Schumpeter, *supra* note 32, at 151 ("Seen in this light, the entrepreneur and his function are not difficult to conceptualize: the defining characteristic is simply *the doing of new things or the doing of things that are already being done in a new way* (innovation).") (emphasis added).

238. *See* Kevin Bryant, *Promoting Innovation: An Overview of the Application of Evolutionary Economics and Systems Approaches to Policy Issues*, *in* FRONTIERS OF EVOLUTIONARY ECONOMICS: COMPETITION, SELF-ORGANIZATION AND INNOVATION POLICY 361, 371 (John Foster & J. Stanley Metcalfe eds., 2001) [hereinafter Bryant, *Promoting Innovation*] (concluding the empirical studies establish that various "market imperfections" — including imperfect knowledge, bounded rationality and the inclusion of non-financial consideration in decisionmaking, "*are universal — and are necessary to drive change*"); Giovanni Dosi & Yuri Kaniovski, *The Method of Generalized Urn Schemes in the Analysis of Technological and Economic Dynamics*, *in* THE ECONOMICS OF GROWTH AND TECHNICAL CHANGE: TECHNOLOGIES, NATIONS, AGENTS 261, 280 (Gerald Silverberg & Luc Soete eds., 1994) (using a new modeling paradigm to show how " 'market imperfections' and 'informational imperfections' often tend to foster technological variety").

entry is more closely associated with startup rather than diversifying entry.[239]

Innovative entry is likely to involve more risk than non-innovative entry because the introduction of new products and technologies is associated with a greater degree of uncertainty. An increased uncertainty as to development time, costs and success, market acceptance, future competition and other factors affecting the venture's ultimate fate also implies an increased likelihood of extreme outcomes.[240] These outcomes may be positive, as in the case of early success or unexpected demand, but often they are negative as a result of the many unanticipated obstacles that entrants must overcome to achieve profitability.[241]

Because of its greater risk, innovative entry is more likely to be undertaken by startup entrants.[242] New innovators may prefer to sell their innovations to incumbent firms rather than take the risk of inde-

239. *See* Roy Rothwell & Mark Dodgson, *Innovation and Size of Firm*, *in* THE HANDBOOK OF INDUSTRIAL INNOVATION 310 (Mark Dodgson & Roy Rothwell eds., 1994) (reviewing evidence showing that small and very large firms — but not medium-size firms — enjoy innovative advantages, which vary for small firms depending on the costs of entry); Zoltan J. Acs et al., *Productivity Growth and Firm Size Distribution*, *in* ENTREPRENEURSHIP, SMALL AND MEDIUM-SIZED ENTERPRISES AND THE MACROECONOMY 367, 369-71, 392-93 (Zoltan J. Acs et al. eds., 1999) (providing a brief review of some of arguments for and against the relative innovative advantages of small versus large firms and, ceteris paribus, of startup versus diversifying entrants); *see also* FREDRIC M. SCHERER, *Corporate Size, Diversification, and Innovative Activity*, *in* INNOVATION AND GROWTH 222, 237 (1984) (concluding from the analyses of various data sources that large corporations invest greater relative resources in R&D, but "contribute[] fewer significant innovations, contest-winning technical advances, and invention patents . . . than smaller enterprises").

240. *See* Dosi, *supra* note 237, at 222 (stating "innovation involves a fundamental element of *uncertainty*"); *see also* Kenneth J. Arrow, The Rand Corporation, *Economic Welfare and the Allocation of Resources for Invention*, *in* NAT'L BUREAU OF ECON. RESEARCH, THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS 609 (1962) (arguing that more frequent innovation may be associated with greater uncertainty, not only as to the development of the product but also as to consumer demand); FREEMAN, *supra* note 173, at 148-68 (discussing the relationship between uncertainty and innovation and ways to deal with this risk).

241. *See, e.g.*, Giovanni Dosi, *Finance, Innovation and Industrial Change*, 13 J. ECON. BEHAV. & ORG. 299 (1990) (arguing that the search for innovation that extends beyond the existing competencies of incumbents might be heavily biased towards mistaken attempts); Giovanni Dosi, *Sources, Procedures, and Microeconomic Effects of Innovation*, 26 J. ECON. LIT. 1120 (1988) (suggesting that the increased uncertainty associated with innovation decreases the probability of survival of the innovating business).

242. *Cf.* Janet E. L. Bercovitz et al., *Firm Capabilities and Managerial Decision Making: A Theory of Innovation Biases*, *in* TECHNOLOGICAL INNOVATION, *supra* note 9, at 233 (arguing that large incumbents tend to introduce less radical innovations than do small firms because of the various individual and organizational biases exerted on the former); *see also* FREEMAN, *supra* note 173, at 135-37 (suggesting that the evidence on the relationship between firm size and innovation is not fully conclusive, but that small firms tend to introduce more innovations, while larger firms are instrumental in developing many innovations and bringing these innovations to the market).

pendent entry,[243] but incumbents' willingness to acquire innovations will be negatively correlated with their originality. The more innovative the new invention, the more incumbents' benefits and costs from the adoption of the innovation,[244] as well their objective opinion of its attractiveness will diverge from those of inventors.[245]

In addition, since risk is positively associated with originality, interested incumbents will be more likely to introduce more original innovations via startup rather than under their own name and full liability. Innovative entrants who are unable (or unwilling) to convince incumbents of the attractiveness of their innovation are likely to embark on even riskier ventures however.[246] Such innovators face the choice of either giving up or attempting independent new entry. Those determined ones who proceed to enter by themselves in the face of

---

243. *See, e.g.*, David B. Audretsch & Zoltan J. Acs, *Entrepreneurial Activity, Innovation, and Macroeconomic Fluctuations*, in INNOVATION IN TECHNOLOGY, INDUSTRIES, AND INSTITUTIONS: STUDIES IN SCHUMPETERIAN PERSPECTIVES 173 (Yuichi Shionoya & Mark Perlman eds., 1994). When discussing the role of new entry in the introduction of innovation, Audretsch & Acs explain:

> In principle, the inventor of the idea should be able to sell that idea to an incumbent enterprise, in the form of a higher wage or royalty payment that roughly equals the expected net value of the idea. As long as there are even the slightest economies of scale associated with any aspect of the enterprise, such as production, marketing, distribution, advertising, or hiring, it should be more economical for the incumbent firm to integrate the innovation into its organization than for the inventor to start a new firm.

Id.

244. Thus, Audretsch & Acs, following Schumpeter as well as more recent scholarship, suggest that "the more radical an innovation is — that is, the degree to which the competence of a firm is destroyed by that innovation — the more costly it will be for the firm to pursue that innovation." *Id.* at 174.

245. Some suggest that because of the "subjectivity of knowledge" involved in estimating the benefits of innovation, "a differential in the expected net value of a potential innovation between the innovator and incumbent firm is likely to emerge. As this gap gets large enough, the inventor will weight the costs of starting his own firm against the net benefits accruing from such a new start-up." *Id.* at 174. Of course, to the extent incumbents are convinced of the innovation's attractiveness, they will also be embarking on higher risk projects than they normally do. *See* FREEMAN, *supra* note 173, at 157 (counting among those likely to embark on high risk innovative projects such firms that unwittingly accept very high risk because they are convinced by the entrepreneurs and share their overoptimism).

246. This will typically happen when the divergence between the innovators' valuation of their innovations and that of incumbents is great. In these cases, innovators may be unwilling to sell for value they deem unreasonable and incumbents may even be unwilling to make any offer for innovations they consider too risky or of a very low value. *Cf.* DAVID B. AUDRETSCH, INNOVATION AND INDUSTRY EVOLUTION (1995) (proposing a model in which an individual agent within an organization, who possesses new knowledge that may or may not have positive economic value because of uncertainty, may decide to exit and start a new venture due to asymmetries in knowledge and valuation between the individual and the organization); *see also* CASSON, *supra* note 227, at 201-09 (describing some problems faced by the entrepreneur who believes he has privileged commercial information in exploiting it either by himself or by approaching others).

adversity will consequently be embarking more often on higher-risk, negative expected value ventures.[247]

The economic factors associated with innovative entry increase the likelihood of bias in those entrants' assessments of their ventures' prospects.[248] Specifically, innovative entrants face a greater degree of ambiguity when making their entry decisions than do less- or non-innovative entrants. Hence, highly innovative entrants will be prone to exhibit more extreme estimates of their potential ventures' value and probability of success.[249]

Non-innovative entrants, in contrast, face less risk and uncertainty.[250] They may therefore be less biased, making fewer negative expected value attempts. It thus appears that while some NEV entry results from the inflated assessments of non-innovative ventures, both economic and psychological factors direct highly biased innovative entrants to attempt negative NPV entry frequently, with startups being the likely venue.

2. *The Consequences of Innovative Entry*

The association between overconfidence and innovative entry suggests boundedly rational entrants facilitate innovation and its attendant benefits,[251] from the expansion of consumer choice, through technological "spillovers" and an increased rate of growth, to the increase of competitive pressures upon incumbents.

When overconfident entrants attempt entry more frequently than rationality dictates, they increase the range of the possible outcomes

247. In certain cases, the innovator may be unable to convince incumbents of the truly positive NPV of the venture or prefer, for solid economic reasons, to attempt new entry. Under these uncommon circumstances, the innovative new entry will be rational. *Cf.* Arrow, *supra* note 240 (suggesting that given some simple assumptions small firms are likely to generate a large proportion of innovative research and only some production, while large firms engage more often in the mass production on the basis of ideas generated by small firms).

248. *Cf.* Paul J. H. Schoemaker & M. Laurentius Marais, *Technological Innovation and Firm Inertia*, *in* ORGANIZATION AND STRATEGY IN THE EVOLUTION OF ENTERPRISE 179, 189-93 (Giovanni Dosi & Franco Malerba eds., 1996) (discussing the likely of effects of various cognitive biases on investment decisions).

249. See *supra* Section II.C.2 for an analysis of how ambiguity increases the magnitude of entrant overconfidence. Also, innovative entrants may have a greater personal stake in their proposed ventures; in this case, they will also be prone to exhibit more overconfidence. *Supra* Section II.C.1; *see also* FREEMAN, *supra* note 173, at 150-55 (citing numerous economic studies showing evidence of a strong optimistic bias when discussing the highly limited reliability of project estimation techniques for innovations).

250. *See* FREEMAN, *supra* note 173, at 151 ("[I]t must never be forgotten that estimates can only be really accurate if uncertainty is reduced, and uncertainty can only be significantly reduced either by further research or by making a project less innovative.").

251. *Cf.* Harvey Leibenstein, *Entrepreneurship and Development*, 58 AM. ECON. REV. 72, 82 (1968) ("[A] lower profit investment that releases entrepreneurial energies and capacities may be more fruitful in the long run than a higher profit investment.").

of entry selection, bringing to the market many innovative products, services, and methods of operation that otherwise would not have been introduced.[252] While many high-risk innovative ventures fail, those that survive expose the economy to a significant amount of innovation. And, even those who fail, infuse other market participants with new ideas and information that may later provide a basis for successful ventures by other entrants or incumbents.[253] Failed entrants may also facilitate and shape consumer demand for innovative products, by making consumers — and not just producers — aware of their possibility of production.[254] Together, these effects of boundedly rational, innovative entry stimulate economic growth.[255]

---

252. *See, e.g.*, Martin Carree & Roy Thurik, *Industrial Structure and Economic Growth*, *in* INNOVATION, INDUSTRY EVOLUTION, AND EMPLOYMENT 86, 88 (David B. Audretsch & A. Roy Thurik eds., 1999) (suggesting that "small businesses may contribute to higher growth because of their contribution to the selection process due to their variety"); *id.* at 106 (citing among some additional benefits of small firms the ability to "satisfy a fragmented and differentiated demand," in a study showing, however, that the relative advantages and disadvantages of small firms vary depending on industry characteristics); Geroski, *What Do We Know*, *supra* note 25, at 436-37 (suggesting that while innovations are often supply driven, potential consumers must get acquainted with new products before determining how they value their various characteristics, adding that the role of entry in introducing a variety of products may be more important in the earlier stages of the development of new markets).

253. *See, e.g.*, William J. Baumol, *Innovation and Creative Destruction*, *in* CREATIVE DESTRUCTION: BUSINESS SURVIVAL STRATEGIES IN THE GLOBAL INTERNET ECONOMY 21, 23-26 (Lee W. McKnight et al eds., 2001) (arguing the positive externalities from "spillovers" of innovation are important and of a larger magnitude than commonly recognized); Henry Capron & Michele Cincera, *Exploring the Spillover Impact on Productivity of World-Wide Manufacturing Firms*, *in* THE ECONOMICS AND ECONOMETRICS OF INNOVATION 543 (David Encaoua et al. eds., 2000) (discussing the definitions of "spillovers" in the literature and reporting some evidence of their presence to varying degrees in different economies).

Ironically, a highly innovative environment may simultaneously increase the likelihood of the successful development of innovation and decrease the likelihood that the innovation will become a viable and marketable product. *See, e.g.*, David B. Audretsch, *Entrepreneurship and Economic Restructuring: An Evolutionary View*, *in* ENTREPRENEURSHIP, SMALL AND MEDIUM-SIZED ENTERPRISES AND THE MACROECONOMY 79, 84-85 (Zoltan J. Acs et al. eds., 1999) (describing a recent study exploring the relationship between the persistently asymmetric firm size distribution in industry, whereby small firms dominate, and the fact that entry is not substantially deterred in industries where scale economics and innovative activity play an important role).

254. *See, e.g.*, Dosi & Lovallo, *Rational Entrepreneurs*, *supra* note 9, at 57-58 (suggesting that both the success and the failure of entrants fulfills an important role in industry learning, inter alia, by contributing to increased collective knowledge, in which case "they represent a sort of externality for the whole system"); *cf.* FREEMAN, *supra* note 173, at 201 (stating that the direction of present research determines "the range of real choice available to consumers" when arguing for governmental support of R&D activity). Schumpeter stated the early view of the relationship between innovation and consumer choice:

> Yet innovations in the economic system do not as a rule take place in such a way that first new wants arise spontaneously in consumers and then the productive apparatus swings round through their pressure. We do not deny the presence of this nexus. It is, however, the producer who as a rule initiates economic change, and consumers are educated by him if necessary; they are, as it were, taught to want new things, or things which differ in some respect or other from those which they have been in habit of using.

SCHUMPETER, *supra* note 73, at 65.

Successful innovative NEV entrants also provide an important competitive check on the behavior of incumbents.[256] Thus, the empirical findings on entry suggest that incumbent firms should have little concern with most new entrants, except possibly some of the largest diversifying ones, at the time of entry. Nevertheless, some successful entrants who prosper and grow eventually pose a competitive threat to incumbents. As time goes by, those remaining entrants become incumbents themselves, in part by replacing older and less competitive predecessors. As today's incumbents, yesterday's successful entrants exert significant competitive pressure on extant incumbents, including dominant firms, requiring them to become more efficient and competitive in order to maintain market share and profitability.

Successful innovative entrants are likely to pose a greater threat to incumbents than do other entrants, because their successful innovations differ to a greater degree from the products and technologies used by incumbents. The success of such entrants indicates that consumers want new products and technologies, and that incumbents might become obsolete if they fail to provide them. At a minimum, threatened incumbents will be more likely to seek ways to increase the efficiency of their production and the competitiveness of their product prices, to the benefit of consumers and society alike. Successful innovative entrants may therefore exert a proportionately greater competitive pressure upon other incumbents and dominant firms than the

255. Thus, Bryant recently summarized the present state of the evidence by stating: "There is a general observation at the macro level that long-run economic growth depends on innovation . . . ." Bryant, *Promoting Innovation*, *supra* note 238, at 371. Innovation leads to growth by fostering a greater menu of options for market selection. *See, e.g.*, Uwe Cantner & Horst Hanusch, *Heterogeneity and Evolutionary Change: Empirical Conception, Findings and Unresolved Issues*, *in* FRONTIERS OF EVOLUTIONARY ECONOMICS, *supra* note 238, at 228, 229-34 (suggesting that technological heterogeneity reflects innovation and is responsible for technological change); Pier Paolo Saviotti, *Variety, Economic and Technological Development*, *in* INNOVATION IN TECHNOLOGY, *supra* note 243, at 27, 46 ("[I]nnovations . . . lead to qualitative change in the composition of the economic system, and this qualitative change is reflected in a growing variety" — a quantitative criterion the author develops to denote distinguishable products and economic actors.). *See generally* Chris Freeman, *Innovation and Growth*, *in* THE HANDBOOK OF INDUSTRIAL INNOVATION, *supra* note 239, at 78 (reviewing the relationship between innovation and growth in economic theory); Paul D. Reynolds, *Creative Destruction: Source or Symptom of Economic Growth?*, *in* ENTREPRENEURSHIP, SMALL AND MEDIUM-SIZED ENTERPRISES AND THE MACROECONOMY 97 (Zoltan J. Acs et al. eds., 1999) (presenting research and describing additional empirical evidence showing that the volume of entry and its attendant turbulence have a major role in facilitating economic growth).

256. *See, e.g.*, J. STANLEY METCALFE, EVOLUTIONARY ECONOMICS AND CREATIVE DESTRUCTION 115 (1998) ("Effective competition depends on diversity in behaviour and over time this can only be maintained by the continual introduction of new and better products and new and better methods of production." Therefore, "it is the line between innovation and competition which has proved to be the mainspring of economic growth.").

competitive threat other, non-innovative successful entrants are capable of posing.[257]

IV. EVALUATING ENTRY AFRESH

Despite competitive pressures, and in some large measure because of them, boundedly rational entrants are overrepresented among the ranks of ex post successful entrants. Because of the diverging incentives and risks facing financiers and entrants, the latter do not always prevent the former effectively from entering when they should not. Consequently, some overconfident entrants introduce into the market negative NPV innovations that would never have been introduced if all entrants were rational. Building on these conclusions, this Part reevaluates the social costs and benefits of entry to determine the desirability and possibility of using governmental regulation to curb this pervasive boundedly rational activity.

A. *The Social Costs and Benefits of Boundedly Rational Entry*

The private benefits and costs of entry vary from one entrant to another: most biased entrants suffer overall losses, but a few become successful because of their boundedly rational behavior. Regardless of entrants' varying individual fates, as a group negative expected value entrants necessarily suffer losses, as the inevitable consequence of aggregate negative expected value entry.

The private costs of negative expected value entry thus translate into deadweight loss to society from the unrecoverable resources these entrants waste on their failed attempts.[258] This does not mean, however, that all entrants together generate social losses. The direct social costs of entry are a function of the relative proportion of negative expected value entrants to all other entrants: if the total costs incurred by the former group are greater than the total profits earned by the latter, than the phenomena of entry generates direct losses to society. If, however, the opposite situation obtains — that is, if the net costs of

257. Thus, while many incumbents will be less likely to be leading innovators, they may very well employ a strategy of "defensive" innovation, attempting incremental improvements in response to, or in anticipation of, more radical innovations by new entrants. *See* FREEMAN, *supra* note 173, at 176-83 (noting also how incumbents will often not engage in pure imitation, but instead seek to improve and modify new innovations). Moreover, as Geroski notes when counting among the "stylized facts" about entry that "[h]igh rates of entry are often associated with high rates of innovation and increases in efficiency," such facts "do not imply the entrants are always, or even often, the major source of innovation in markets. *Many case studies show that entry stimulates incumbents to introduce new products and processes* which they had been holding back." Geroski, *What Do We Know*, *supra* note 25, at 431 (emphasis added).

258. *See, e.g.*, Carree & Thurik, *supra* note 252, at 88 (suggesting that the "lower level of stability, [of small businesses] inherent in the selection process, leads to welfare losses").

negative expected value entry are smaller than the net gains of other entry — then entry may still be beneficial to society.

Negative expected value entrants, however, generate social costs beyond the resources they waste directly. They generate significant negative externalities, diminishing the success prospects of less biased entrants and interfering with the efficient allocation of resources through market competition.[259] Those entrants that had superior ex ante prospects but either failed after initially attempting entry because of probabilistic replacement[260] or refrained from entry altogether[261] instead must direct their resources to less beneficial uses, further reducing social welfare.[262]

Entrant overconfidence also interferes with the market mechanism of resource allocation.[263] Under ideal circumstances, market trading directs resources to their most valuable use given their original distribution in society.[264] Consumers can only purchase, however, from among those products actually available in the marketplace. Those products that would have been offered by rational entrants who either refrained from entry altogether or were replaced by overconfident competitors will inevitably be eliminated from the set of products on which consumer choice and selection operate. Consumers will also prefer some negative expected value products, which would not have been offered if entrants were strictly rational, over some other com-

259. This negative externality is in addition to the negative externality generated by any successful entrant — rational or not — and especially innovative ones, of diminishing the prospects of other entrants and incumbents alike. *See* Baumol, *supra* note 253, at 32 (providing a brief explanation of this effect). Schumpeter already suggested that innovative entry also spells losses to incumbents by decreasing their market share and profitability:

> It is similarly clear that entrepreneurial gain is not a net accretion to the returns of the industrial sector in which it occurs. The impact of the new product or method spells losses to the "old" firms. The competition of the man with a significantly lower cost curve is, in fact, the really effective competition that in the end revolutionizes the industry.

Schumpeter, *supra* note 32, at 156.

260. *See supra* Section II.A.

261. *See supra* Section II.A.

262. The conclusion that the success of some overconfident entrants causes the diversion of some rational entrants' resources to less productive uses becomes obvious when the exact definition of rational entry is substituted for the simplified definition used throughout this Article. Only such entry that yields a risk adjusted NPV greater than opportunity costs of entry is truly rational, and the opportunity costs of entry are the next most profitable use into which the entrant could put the resources spent on entry. Whenever rational entry is not made, therefore, the resources which would have been most productive if directed towards entry are instead channeled to other goals that are less beneficial by definition.

263. *See, e.g.*, JOSEPH E. STIGLITZ, ECONOMICS OF THE PUBLIC SECTOR 63-71 (2d ed. 1988) (an introductory exposition of the fundamental theorems of welfare economics).

264. *Id.*

peting products, substituting some negative NPV products or product-mixes including such products for those products not offered.[265]

While entrants' bounded rationality generates social costs, it is also a source of important positive external effects. The various social benefits brought about by an increased rate of innovation — from the expansion of consumer choice, through technological "spillovers" and an increased rate of growth, to the increase of competitive pressures upon incumbents — all appear to be associated with boundedly rational entry. Hence, if entrants were strictly rational decisionmakers, economic growth and development would slow down and competitive discipline would become even less effective than currently.

The ultimate social balance of negative expected value entry thus depends, to a significant degree, on the relative proportion of innovative to non-innovative entrants within this large group. Apparently, therefore, innovative negative expected value entrants serve as the voluntary cannon-fodder of our economy — embarking on ventures that are personally unprofitable on average while bestowing important positive externalities on society as a whole.[266]

B. *Regulating Negative Expected Value Entry?*

The behavioral analysis of entrant decisionmaking reveals the myriad negative and positive effects generated by the boundedly rational entrant behavior. One must therefore not rush to brand the boundedly rational behavior of entrants a new form of market failure, justifying the benevolent intervention of the regulator.[267] Any intervention in the decisionmaking process of entrants in an attempt to reduce the social costs of negative expected value entry may also limit its beneficial impact.

---

265. *See, e.g.*, WALTER NICHOLSON, MICROECONOMIC THEORY: BASIC PRINCIPLES AND EXTENSIONS 164-79 (7th ed. 1998) (providing a basic exposition of substitution); *cf.* STIGLITZ, *supra* note 263, at 78 (discussing the lack of complementary markets as market failure).

266. *See, e.g.*, Dosi, *supra* note 241 (suggesting that the close association between innovation and error implies that the search for new innovation is likely to lead to disappointing economic failures, on average, for the individual actors who embark upon it, who nevertheless serve as important agents of economic change); Dosi & Lovallo, *Rational Entrepreneurs*, *supra* note 9, at 57 (suggesting that, when entry is viewed as part of a learning process, a number of individual and organizational biases, but "especially overconfidence, inside view thinking, and illusions of control — *are essential to sustain exploration, even when the latter is not individually rewarding*") (footnote omitted); G. Silverberg et al., *Innovation, Diversity and Diffusion: A Self-Organisation Model*, 98 ECON. J. 1032 (1988) (developing a model of the diffusion of new technology where, under certain assumptions regarding the learning process, unequivocally superior innovations sometimes diffuse only if there are overoptimistic entrepreneurs — "Schumpeterian sacrificial lambs" — who pay the price of initial exploration, bestowing a positive externality upon the industry).

267. *See* STIGLITZ, *supra* note 263, at 71-80 (arguing that market failures justify governmental intervention in markets).

In principle, an effective detection and prevention of some negative net present value entry at the margin could be beneficial. The problem is, however, that there is no easy means of quantifying with any certainty the benefits and costs of entry. Measuring the direct deadweight losses of negative expected value entry is the least difficult, since — at least in principle — one could calculate the number of failed entrants per industry and the average costs they have sunk in their ventures. The calculation of net losses would be more complex, though, for those entrants who obtain some profits before exiting the market.

It is even less clear, moreover, how to measure the externalities of entry. Any quantification of the externalities of negative NPV entry requires a measurement of the reduction this entry causes in other entrants' success prospects. This requires finding the number of less biased potential entrants who would have entered but refrained from entry, and the alternative uses into which those potential entrants have put the resources they did not invest in entry. Similarly, it is not clear how to measure the social loss from negative expected value entry's effect on allocative efficiency — how does one compare the local optimum generated by competition among actual entrants to the global optimum brought about when the dynamics of competition operate on rational entry alone?[268]

An examination of the quantities that must be measured for any quantification of the positive externalities of negative expected value entry reveals another impenetrable maze. For example, how does one quantify the incremental benefits of additional competitive discipline? And the estimation of the benefits of innovation is even more elusive. Clearly, these benefits go well beyond the net profits of those few successful negative NPV entrants. But how does one measure the contribution of negative expected value entry — both successful and failed — to the introduction of new ideas into the market and the development of consumer demand for innovation?

As if these measurement difficulties were not enough, any attempt to regulate entry would face a number of practical impediments. First, it would be difficult and costly for a government agency, for example, to determine ex ante which ventures have a negative expected value. Second, any active regulation of entry is likely to harm disproportionately innovative entrants because expert regulators would not only be less optimistic regarding the prospect of entry than individual entrants, but would also lack access to the "soft" data that motivates entrants, especially innovative ones, to attempt entry. Third, in addition to raising additional impediments to entry, regulation — to the extent it would slow the process of entry — may decrease the attractiveness of

268. *Cf.* Francis M. Bator, *The Simple Analytics of Welfare Maximization*, 47 AM. ECON. REV. 22 (1957) (stating the classic argument showing how multiple optima arise).

certain ventures whose success depends on the timing of their entry into the market. Finally, government agents also suffer numerous shortcomings in their decisionmaking process: their incentive structure is not aligned with society's interests; they are often "captured" or affected by incumbents and dominant market participants; and they are not subject to effective pressure or discipline.[269]

Beyond the numerous difficulties involved in any governmental regulation of entry, most of the various potential forms of intervention are also likely to be ineffective in accomplishing the goal of reducing undesirable NEV entry.[270] An attempt to debias entrants — training them to overcome their boundedly rational predictions and judgments — would be the best method of intervention. Nonetheless, behavioral findings suggest that such an approach is doomed to fail in the case of negative expected value entry. Numerous studies show that, even under circumstances far more conducive to learning, decisionmakers find it difficult to overcome the processes of overconfidence.[271] Additionally, even partially successful debiasing interventions require training, and subjects typically fail to generalize their learning from one setting to another.[272] A minimal requirement for effective debiasing is the

269. *See, e.g.*, STIGLITZ, *supra* note 263, at 198-202 (describing the institutional and personal limitations of governmental action); *see also* Robert C. Clark, *Contracts, Elites, and Traditions in the Making of Corporate Law*, 89 COLUM. L. REV. 1703, 1719-20 (1989) (arguing that governments lack proper incentives for efficient action); Jolls et al., *supra* note 1, at 1543-45 (discussing the limited effectiveness of "Behavioral Bureaucrats").

270. The most radical intervention in the process of entry — short of a complete ban of new entry — is the erection of entry barriers. The second puzzle of entry has shown, however, that most impediments to entry have little effect on its rate; raising entry barriers, moreover, would increase the costs of all entry, contribute to the social losses from failed entry attempts and — to the extent it would reduce entry — provide incumbents with a highly visible, socially undesirable protection from competitive discipline.

271. *See, e.g.*, Byram, *supra* note 167 (finding that various debiasing manipulations fail to diminish the effects of the planning fallacy on time predictions); Neil D. Weinstein & William M. Klein, *Resistance of Personal Risk Perceptions to Debiasing Interventions*, 14 HEALTH PSYCHOL. 132 (reporting how a variety of debiasing methods were found to have little impact on optimistic overconfidence, while conditions using the opposite manipulations increased the bias); *see also* Lisa E. Bolton, The Effects of Nonanalytic and Analytic Thinking in New Product Forecasting (2000) (unpublished Ph.D. Dissertation, University of Florida) (on file with author) (providing an extensive set of product forecasting showing the extreme difficulty of debiasing decisionmakers after they made an initial intuitive business forecast). In general, the psychological evidence shows the difficulty of debiasing decisionmakers even under those conditions most conducive to learning, such as a controlled environment and clear feedback. *See, e.g.*, Baruch Fischhoff, *Debiasing*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 422 (Daniel Kahneman et al. eds., 1982) (reviewing studies and discussing the difficulty of debiasing).

272. *See, e.g.*, Frank P. McKenna & Ian P. Albery, *Does Unrealistic Optimism Change Following a Negative Experience?*, 31 J. APP. SOC. PSYCHOL. 1146 (2001) (using a severe threat manipulation that highlighted vividly the possible consequences of optimistic risk taking and finding that while this strong manipulation improved estimates in the particular domain, the effect was domain specific, failing to affect risk estimates in other domains). *See generally* Fischhoff, *supra* note 271, at 422-23 (describing how decisionmakers who are debiased in one context fail to transfer their learning to judgments in other settings).

availability of clear feedback on performance, which is nonexistent before entry occurs.[273]

The provision of easily understandable relevant statistical information, which negative expected value entrants may fail to obtain, would be a weaker form of intervention. This Article's analysis of entrant bias does suggest that better information may help somewhat in reducing the extent of negative expected value entry. The efficacy of this intervention might nevertheless be limited by entrants' tendency to discount pallid statistical information regarding the market at large that seems irrelevant to the fate of their unique venture.[274]

C. *Antitrust Law and Boundedly Rational Entry*

The balance of boundedly rational entry's social costs and benefits appears uncertain, and its regulation of questionable desirability and little practicality. But if intervention is neither clearly necessary nor effectively possible, the law should take the pervasive presence of overconfident entry as given and examine whether its doctrines require modification. A primary area in which this should be done is antitrust law, whose doctrines frequently rely on analyses of entry. This Part therefore examines briefly some potential implications of the behavioral analysis of entry for antitrust law.

1. *General Conclusions: Entry and Market Power in Antitrust*

Among the findings of the behavioral economic analysis of entry, the following are the most significant for a critical evaluation of the role of entry in antitrust law:

1) Entry is not exceptionally difficult; however, post-entry success and survival are unlikely for most entrants.[275]
2) Most startups, and small entrants generally, pose no short-term competitive threat to incumbents.[276]

273. *See, e.g.*, Hillel J. Einhorn, *Learning from Experience and Suboptimal Rules in Decision Making*, *in* COGNITIVE PROCESSES IN CHOICE AND DECISION BEHAVIOR 1 (Thomas S. Wallsten ed., 1980) (emphasizing the importance of unambiguous feedback for learning); Richard E. Nisbett et al., *Improving Inductive Inference*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 445, 445-46 (Daniel Kahneman et al. eds., 1982) (noting that decisionmakers need to know that an error has occurred, how it has occurred, and how to improve the decision process).

274. *See supra* note 148; *see also* Weinstein & Klein, *supra* note 271, at 132 (citing various studies that found the provision of information to be largely ineffective in modifying risk perceptions and even less effective in producing changes in behavior).

275. *See supra* Section I.A.1; *e.g.*, GEROSKI ET AL., *supra* note 61, at 3 ("[E]ntry appears to be easy but post-entry market penetration and, indeed, survival is not.").

276. *See supra* Section I.A.3.

3) Some large diversifying entrants may pose a competitive threat to incumbents even in the short-term, although the prospects of these entrants for survival and prosperity are far from certain.[277]
4) In the long run, the few successful boundedly rational innovative entrants are an important source of competitive pressure on incumbents.[278]
5) Those market characteristics known as "entry barriers" do little to deter entry (except possibly when industry is also concentrated).[279]
6) Such barriers nevertheless affect the post-entry performance of entrants, and significantly decrease their survival. Thus, in the presence of bounded rationality, entry barriers become *survival barriers* — economic factors that often, instead of inhibiting entry, make post-entry survival more difficult and less likely.[280]

These findings suggest that the fundamental hostility of antitrust law to unnecessary restrictions on new business entry[281] is well-founded, given entry's important pro-competitive benefits. In fact, the analysis here shows that survival barriers in the market should be viewed with additional concern. Much like intra-industry "mobility barriers," which prevent entrants from moving from one industry group to another,[282] survival barriers create impediments to new entrants' survival and growth that are especially pernicious because of entrants' relative insensitivity to their presence.

Additionally, the presence of significant benefits from boundedly rational, small, innovative entry also indicates that the long-discarded "populist" goals of antitrust law[283] do not contradict the now-accepted

277. *See supra* Section I.A.3.

278. *See supra* Section III.C.

279. *See supra* Section I.A.2.

280. *Cf.* Geroski, *What Do We Know*, *supra* note 25, at 436 (suggesting, after discussing the puzzling co-existence of high entry barriers and high rates of entry that "[i]f . . . barriers to entry are thought of as an obstacle which prevents new firms from surviving long in a market, then the data present less of a puzzle").

281. *See, e.g.*, 1 AREEDA & HOVENKAMP, *supra* note 7, ¶112a.

282. R. E. Caves and M. E. Porter, *From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition*, 91 Q. J. ECON. 241 (1977).

283. Such goals include the dispersion of economic and political power and the protection of small competitors from larger and more powerful rivals, the latter of which was especially significant in the legislative history of the Sherman Act. *See, e.g.*, 21 CONG. REC. 2569 (Mar. 24, 1890) (Senator Sherman's statement when arguing the merits of the proposed Act regarding the negative consequences of "combinations" with "a purpose to prevent competition, so that if a humble man starts a business in opposition to them, solitary and alone, in Ohio or anywhere else, they will crowd him down"); *id.* at 3147 (Apr. 8, 1890) (Senator Geroge's argument that by "the use of this organized force of wealth and money the small

goal of promoting economic efficiency as much as commonly thought. The conclusions from this Article's analysis regarding the important role of small entrants in both facilitating the introduction of innovation and increasing competitive pressures in the market suggest that early populist views may have reflected, to a degree, a well-founded concern with maintaining the contribution of small, boundedly rational, entry to society. Nonetheless, while this partial reconciliation of economic and "populist" objectives is useful in highlighting their common ground, as a practical matter, the analysis here confirms that an economic approach, albeit modified, still provides the most coherent framework for the interpretation and application of the antitrust laws.[284]

At the same time, however, these conclusions indicate that the law should be wary of relying on findings of low barriers alone to guarantee competitive pressure on incumbents in the short-run because most new entrants detract little from incumbents' market power.[285] When low barriers are accompanied with larger scale, actual or potential, diversifying entry, on the other hand, incumbents are more likely to face competitive pressure even in the short term. Additionally, the effects of barriers are especially pernicious in concentrated markets.[286] The empirical evidence shows that when markets are concentrated, en-

---

men engaged in competition with [the trusts] are crushed out, and that is the great evil at which all this legislation ought to be directed"); *see also* United States v. Trans-Missouri Freight Ass'n, 166 U.S. 290, 323-24 (1897) (emphasizing the harmful consequences of driving out small and independent dealers). The emphasis on noneconomic concerns continued through the years of the Warren Court. *See, e.g.*, United States v. Von's Grocery Co., 384 U.S. 270, 274 (1966) ("From this country's beginning there has been an abiding and widespread fear of the evils which flow from monopoly . . . . On the basis of this fear, Congress in 1890 . . . passed the Sherman Act in an attempt to prevent further concentration and to preserve competition among a large number of sellers."); Brown Shoe Co. v. United States, 370 U.S. 294, 344 (1962) ("[W]e cannot fail to recognize Congress' desire to promote competition through the protection of viable, small, locally owned businesses. Congress appreciated that occasional higher costs and prices might result from the maintenance of fragmented industries and markets. It resolved these competing considerations in favor of decentralization.").

For a discussion of the various conflicting goals attributed to the antitrust law both historically and at the present, see 1 AREEDA & HOVENKAMP, *supra* note 7, ¶¶ 101-12.

284. Antitrust commentators have long pointed out that such noneconomic goals fail to provide proper guidance to the courts in its implementation of the antitrust laws. *See, e.g.*, 1 AREEDA & HOVENKAMP, *supra* note 7, ¶ 103; (arguing that the traditional economic approach to antitrust law, even if imperfect, is still far more coherent than alternative approaches); ROBERT BORK, THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF 6-11 (2d. ed. 1993) (discussing the conflicting goals of the antitrust laws and arguing for the alignment of antitrust policy with efficiency considerations); RICHARD A. POSNER, ANTITRUST LAW, at vii-x (2d. ed. 2001).

285. This conclusion is also supported by the findings on limited incumbent reaction to entry generally, and the rarity of price-related entry-deterring strategies on the part of incumbents specifically. *See infra* note 300; *cf.* Geroski, *What Do We Know*, *supra* note 25, at 437 (arguing in favor of antitrust law's emphasis on entry barriers but suggesting that the pro-competitive effects of entry can be easily exaggerated, especially in the short run).

286. *See supra* note 41.

trants not only face high mortality rates, but are also significantly less likely to attempt entry and challenge incumbent firms. The doctrines of antitrust law that look to entry barriers as part of market power determinations should therefore be reexamined.

Theoretically, all determinations of whether an individual firm or a number of firms together enjoy market power — that is, the power to profitably charge supra-competitive prices for their products — implicate entry: When entry is sufficient to drive prices back to competitive levels, firms do not enjoy market power and, therefore, pose no significant threat to competition and need not concern the law.[287] In practice, however, the law treats impediments to entry differently depending on the nature of the potentially anti-competitive behavior at stake. It declares certain behaviors illegal per-se — such as horizontal price fixing — based on the understanding that these practices are so likely to be harmful in the short-term that a detailed inquiry into their particular circumstances is unnecessary. Accordingly, an analysis of entry is less important in those instances.[288] Other practices, however, require the court to inquire further into the circumstances of the case and thus raise the question of the effectiveness of entry. The following sections examine briefly two of the settings in which such inquiries

287. *See* Ball Mem'l Hosp., Inc. v. Mut. Hosp. Ins., 784 F.2d 1325, 1335 (7th Cir. 1986) (rejecting the argument that market share indicates market power even in the absence of entry barriers, stating that "the lower the barriers to entry, and the shorter the lags of new entry, the less power existing firms have"); Will v. Comprehensive Accounting Corp., 776 F.2d 665, 672 n.3 (7th Cir. 1985) ("Unless barriers to entry prevent rivals from entering the market at the same cost of production, even a very large market share does not establish market power."); United States v. Waste Mgmt. Inc., 743 F.2d 976, 982 (2d Cir. 1984) ("[A] market definition artificially restricted to existing firms competing at one moment may yield market share statistics that are not an accurate proxy for market power when substantial potential competition able to respond quickly to price increases exists."); *see also* 2A AREEDA & HOVENKAMP, *supra* note 7, ¶ 420b & n.10 ("Entry conditions are therefore relevant to assessing the market power required by most antitrust doctrines," because "[m]arket power bears on the anticompetitive potential of challenged conduct. Indeed, challenged conduct may be deemed unreasonable or exclusionary, and therefore illegal, precisely because it creates a barrier to new competition without sufficient offsetting justification.").

288. *See, e.g.*, Nat'l Collegiate Athletic Ass'n v. Bd. of Regents, 468 U.S. 85, 100, 109-10 (1984) (stating that "[h]orizontal price fixing and output limitation are ordinarily condemned as a matter of law under an 'illegal per se' approach because of the probability that these practices are anticompetitive is so high" and that "as a matter of law, the absence of proof of market power does not justify a naked restriction on price or output"); Nat'l Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679, 692 (1978) (horizontal price-fixing agreements are "agreements whose nature and necessary effect are so plainly anticompetitive that no elaborate study of the industry is needed to establish their illegality — they are 'illegal per se' "); *see also* 12 AREEDA & HOVENKAMP, *supra* note 7, ¶¶ 2010-12 (introducing the basics of horizontal cartels and price fixing); *id.* ¶ 2011c (explaining why market power as traditionally understood is not necessary for this offense). *See generally* 7 AREEDA & HOVENKAMP, *supra* note 7, ¶¶ 1500-11 (providing an introduction to the issues surrounding the per-se rule and its main alternative — the rule of reason).

into entry are important — predatory pricing by dominant firms and horizontal mergers.[289]

2. *Predatory Pricing: Entry and Recoupment*

Entry plays an important role in the legal analysis of predatory pricing, the practice of selling at non-remunerative prices to drive out, exclude, or discipline rivals.[290] Because predatory pricing requires the predatory firm to make significant investments by selling at unprofitable prices, it is only deemed illegal if the predator has the opportunity to recoup its losses.[291] That is, the predator must enjoy a sufficiently long period in which it is able to sell at sufficiently high prices following predation.[292] According to the case law, for recoupment to be pos-

---

289. Entry and barriers to entry play an important role in numerous other antitrust doctrines, including vertical and conglomerate mergers and certain restraints of trade, such as tying arrangements. The present analysis is therefore far from exhaustive, only intended to outline some of the antitrust implications of the behavioral analysis of entry decisionmaking.

290. Predatory pricing violates both the Sherman Act, 15 U.S.C. § 2 (2000) (the offense of monopolization) and the Robinson-Patman Act, 15 U.S.C. § 13 (2000). *See generally* 3 AREEDA & HOVENKAMP, *supra* note 7, ¶ 723.

291. Thus, the Supreme Court had declared that when pursuing predatory pricing allegations:

> Recoupment is the ultimate object of an unlawful predatory pricing scheme; it is the means by which a predator profits from predation. Without it, predatory pricing produces lower aggregate prices in the market, and consumer welfare is enhanced. Although unsuccessful predatory pricing may encourage some inefficient substitution toward the product being sold at less than its cost, unsuccessful predation is in general a boon to consumers . . . . The plaintiff must demonstrate that there is a likelihood that the predatory scheme alleged would cause a rise in prices above a competitive level that would be sufficient to compensate for the amounts expended on the predation, including the time value of the money invested in it.

Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 224-25 (1993) (rejecting claims of competitive injury in a price discrimination suit, which the Court determines to be of the same general character as the injury inflicted by predatory pricing schemes actionable under § 2 of the Sherman Act).

Areeda & Hovenkamp introduce their discussion of recoupment by stating:

> No rational firm would bear the losses, difficulties, and possible legal troubles of trying to exclude or discipline rivals by predatory pricing unless the firm is reasonably confident of a payoff that exceeds the investment, taking into account all relevant risks, including the risk of antitrust litigation. This payoff is generally referred to as "recoupment" of the predation investment.

3 AREEDA & HOVENKAMP, *supra* note 7, ¶ 726a.

292. Importantly, the possibility of recoupment is not the only precondition for predatory pricing beyond the related requirement of market power discussed *infra*. The pricing employed by the alleged predator must also be below some measure of costs. *Brooke*, 509 U.S. at 222 ("[A] plaintiff seeking to establish competitive injury resulting from a rival's low prices must prove that the prices complained of are below an appropriate measure of its rival's costs.") (footnote omitted). For a discussion of the relationship between the different requirements for proving predatory pricing, see 3 AREEDA & HOVENKAMP, *supra* note 7, ¶ 725b.

Note, however, that the Court failed to articulate what is the appropriate measure of costs below which prices may be deemed predatory, *Brooke*, 509 U.S. at 222 n.1 (noting that "[b]ecause the parties in this case agree that the relevant measure of cost is average variable

sible, the alleged predator must enjoy, inter alia, the protection of high barriers to entry; in the absence of such barriers, the predator will be unable to sell at supracompetitive prices for a sufficiently long period after driving out or disciplining its victims, since new firms will simply enter the market.[293] Thus, the Supreme Court, in *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, declared that predatory pricing allegations can be rejected summarily when entry is easy.[294]

While the traditional approach to entry suggests that the presence of actual entry and the lack of high entry barriers guarantee that predatory pricing will be unprofitable, the present analysis suggests a somewhat different conclusion. Specifically, since a high rate of *overconfident* entry may be accompanied by very limited market penetration, the mere evidence of entry should not be sufficient to reject predatory pricing claims out of hand. Instead, the courts should focus on the *success of entrants in penetrating the market* as a better indicator of the short-term competitive threat such entrants pose for allegedly predatory incumbents. Such historical success in penetration can be based on the record of performance in the relevant market, which should not be more difficult to obtain than other evidence of market conditions or past performance that parties are often required to present in antitrust cases.[295]

Courts tend to dismiss predatory pricing allegations quickly, because they often believe that recoupment is impossible. Although the courts' analysis of entry is often summary, some courts nevertheless appear to consider evidence of actual penetration. Thus, in *Brooke*, the Court noted that an evidentiary finding of easy new entry would bar a reasonable jury from concluding that an alleged predatory

---

cost, however, we again decline to resolve the conflict among the lower courts over the appropriate measure of cost"), and a debate currently rages over the question of this definition, *see, e.g.*, Aaron S. Edlin, *Stopping Above-Cost Predatory Pricing*, 111 YALE L.J. 941 (2002) (arguing that above cost pricing can be predatory); Einer Elhauge, *Why Above-Cost Price Cuts to Drive Out Entrants Do Not Signal Predation or Even Market Power — and the Implications for Defining Costs* (manuscript at 9-12, 17-37) (forthcoming YALE L.J., on file with author) (reviewing the state of the law and offering a new definition of variable costs as the appropriate costs for determining the possibility of predatory pricing).

293. *Brooke*, 509 U.S. at 226. Importantly, where the alleged predator merely lowered its prices in response to a competitive threat, the antitrust laws do not condemn the predator's behavior because it benefits consumers and is "the very essence of competition." *Id.* (quoting Cargill Inc. v. Monfort of Colorado Inc., 479 U.S. 104, 122 n.17 (1986) (quoting Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574, 594 (1986))).

294. *Id.* at 225-26.

295. *See, e.g.*, 2A AREEDA & HOVENKAMP, *supra* note 7, ¶ 420b (discussing the requirements and burdens of parties to an antitrust suit involved in proving entry conditions); *see also infra* notes 296-298 and accompanying text (discussing the requirement of showing likely "effective" (and not merely "any") entry that would have a price-disciplining effect in horizontal merger cases).

scheme would likely result in sustained supracompetitive pricing.[296] When addressing the allegations at hand, however, the Court did not simply apply its earlier clear neoclassical statement, which equates easy entry with effective price competition. Instead of limiting the relevant portion of its inquiry to the evidence of mere *entry* following the alleged predation, the Court relied on evidence of *expansion* in the relevant segment of the market to reject the possibility of recoupment. Noting the rapid expansion in the relevant segment — the outcome of successful penetration as well as the continued growth of the alleged predator's sales in this segment — the Court explained that recoupment was unlikely. In the face of evidence of expansion, where the alleged predatory pricing did not cause a restriction in output, the supracompetitive pricing that is crucial for recoupment could not have occurred.[297]

Much like the *Brooke* decision, which formally followed the neoclassical entry criterion even while relying, to a degree, on evidence of penetration and growth, the First Circuit has also recognized the importance of successful penetration in rejecting predatory pricing allegations. In *R.W. Int'l Corp. v. Welch Food Inc.*, the circuit affirmed the dismissal of such allegations upon a motion for summary judgment in the district court.[298] In doing so, the Welch court explicitly referred to evidence of new product penetration at the time of the alleged predatory pricing, declaring that such evidence reveals that the alleged predatory pricing caused no injury to competition.[299]

Despite the fact that it appears courts sometimes consider market penetration when analyzing entry, they often fail to consider factors that would make predatory pricing claims seem more plausible. These factors are primarily the lack of entrant penetration into the market, the absence of large-scale entry, and the absence of innovative entry. In essence, the presence of small entrants without additional evidence of actual or potential significant market penetration should not be taken as sufficient evidence that recoupment by the predator is unlikely, as is frequently the case now.

The conclusion of this Article's entry decisionmaking analysis suggests that courts examining the ease of entry as part of recoupment

296. *Brooke*, 509 U.S. at 226 (enumerating easy entry among a number of market circumstances that would bar a reasonable jury from finding a likelihood of recoupment).

297. *Id.* at 233-34. To be precise, the Court still entertained the possibility that supracompetitive pricing might have occurred even given rapid expansion if the evidently fast growth was still slowed down by the alleged predatory pricing. This possibility was rejected by the Court, however, as lacking concrete evidence. *Id.* at 234.

298. 13 F.3d 478, 488 (1st Cir. 1994).

299. *Id.* ("Where . . . a new product is able to *deeply penetrate* the market during the challenged price-cutting period, it is evident that competition is unharmed and 'summary disposition of the case is appropriate.' ") (quoting *Brooke*, 509 U.S. at 226) (emphasis added).

analysis should also be attentive to the entry of relatively large-scale, often diversifying, firms. By posing a more significant short-term competitive check on the incumbents, these entrants are likely to make recoupment more difficult. They are also very likely to be the firms against which the alleged predator is directing its practices.[300]

The long term threat posed by innovative entrants provides a further reason for the courts to submit predatory pricing allegations to closer scrutiny: incumbents may be concerned not only with short-term diversifying entrants but also with the long-term threat posed to them by those few successful innovative entrants.[301] Thus, when taking into account the long-term benefits predatory incumbents can obtain from holding back the introduction and dissemination of significant innovations that fundamentally threaten their position, the possibility of recoupment appears greater than the traditional approach seems to recognize. Hence, while the quantification of such benefits to the predator may be highly uncertain, their presence should at least make courts more careful in rejecting predatory pricing allegations out of hand on the basis of the supposed impossibility of recoupment, especially when the alleged victim of the predation is an innovative entrant.

3. *Horizontal Mergers: Entry in the Merger Guidelines*

Traditionally, the antitrust laws have been concerned with horizontal mergers — mergers between competitor firms — because of the potential harm to competition between the merged firm and its remaining rivals in the market.[302] At the extreme, a merger may create a

300. The empirical data on incumbents' responses to entry are mixed, suggesting that they tend to ignore entry on most occasions, but may engage in various entry deterring strategies — most notably by increasing advertising and typically not by lowering prices — in specific cases. *See* Geroski, *What Do We Know*, *supra* note 25, at 431-34 (reviewing the various findings on incumbents' reactions and noting some of their limitations). However, predatory pricing cases typically revolve around allegations made by a sizable but smaller competitor against one or a number of its larger counterparts. *See, e.g.*, *Brooke*, 509 U.S. at 212-14 (plaintiff increased its shrinking share from two to five percent of the highly concentrated national cigarette market by pioneering the development of a new "generic" segment in the market by the time the alleged predation occurred, but this had happened only four years after the introduction of the new segment during which this segment grew from a fraction of a percent of the market — a share typical of new entrants — to over four percent of the total market); Cargill Inc. v. Monfort of Colo. Inc., 479 U.S. 104, 106-07 n.2 (1986) (plaintiff, who was challenging a merger between two of its larger competitors in an oligopolistic market inter alia on the grounds that the merger would be followed by predatory pricing, was the fifth-largest firm in the market with a five to six percent market share).

301. *Cf. Brooke*, 509 U.S. at 214-15 (the response of the alleged predator to the increasingly successful innovation of a new market segment by the plaintiff).

302. *See, e.g.*, E. THOMAS SULLIVAN & HERBERT HOVENKAMP, ANTITRUST LAW, POLICY AND PROCEDURE: CASES, MATERIALS, PROBLEMS 835-36 (3d ed. 1994) (summarizing the potential effects of horizontal mergers). *See generally* 1992 Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, § 2 (Apr. 2, 1992) (promulgated by the U.S. Department of

dominant firm or a monopoly, and even in less extreme cases, it may give the merged firm sufficient market power to raise prices to supracompetitive levels for a significant period of time. Unlike price-fixing between competitors, however, horizontal mergers also have the potential to create procompetitive efficiencies. The newly merged firm may enjoy, for example, new economies of scale or better integration, resulting in increased efficiency and lower prices, which benefit consumers.[303] Consequently, the law does not presume that every horizontal merger has anticompetitive effects. Instead, it requires a case-by-case determination of whether the newly-merged firm will obtain sufficient power over price to cause competitive harm.[304]

Unsurprisingly, the presence of potential entry is one of the major factors considered by the courts and the regulatory agencies in determining whether a merged firm would have the power to raise prices. If the post-merger firm were to raise prices to supracompetitive levels and firms would enter as a result of the higher profit potential, the merger would not pose a significant competitive threat.[305] The revised 1992 Horizontal Merger Guidelines[306] ("Guidelines") — used by the Department of Justice Antitrust Division and the Federal Trade Commission in reviewing mergers and deciding whether to challenge them — rely on potential entry analysis both for defining "the market" in which the power of the merged firm is examined and for determining that firm's market power.[307]

Justice and Federal Trade Commission) (reviewing the various potential adverse competitive effects of horizontal mergers).

303. *See generally* Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, § 4 (1992).

304. Thus, the statutory test for the legality of a merger under section 7 of the Clayton Act, 15 U.S.C. § 18, is whether the effect of the merger "may be substantially to lessen competition." Brown Shoe Co. v. United States, 370 U.S. 294, 324 (1962) (citing and analyzing the test under Section 7 in the context of a vertical merger); *see also* United States v. Philadelphia Nat'l Bank, 374 U.S. 321, 355 (1963) (applying the test to a challenged horizontal merger); Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, § 0.1 (1992) ("The unifying theme of the Guidelines is that mergers should not be permitted to create or enhance market power or to facilitate its exercise.").

305. *See, e.g.*, United States v. Waste Mgmt. Inc., 743 F.2d 976, 981-84 (2d. Cir. 1984) (reversing the district court's judgment on the grounds that a merger resulting in a large market share did not substantially lessen competition where entry was easy); *see also* Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, § 3.0 (1992) ("A merger is not likely to create or enhance market power or to facilitate its exercise, if entry into the market is so easy that market participants, after the merger, either collectively or unilaterally could not profitably maintain a price increase above premerger levels.").

306. Horizontal Merger Guidelines, 57 Fed. Reg. 41,552 (1992).

307. Another aspect of this Article's conclusions that has significance for the analysis of horizontal mergers is the impact of concentration on entry. Unsurprisingly, both the courts and the Guidelines agree that mergers are likely to create or enhance market power or facilitate its exercise only when it significantly increases concentration. *See, e.g.*, United States v. Baker Hughes Inc., 908 F.2d 981, 982 (D.C. Cir. 1990) ("The basic outline of a section 7 horizontal acquisition case is familiar. By showing that a transaction will lead to undue concentration in the market for a particular product in a particular geographic area, the gov-

The Guidelines include in the definition of potential entry such sellers who would enter the market given a price increase within a certain time period. For market definition purposes, the period is one year;[308] for price effects purposes, the period is two years.[309] In other words, only sellers that could enter at competitive prices within a year are "in the market" where the merged firm's power is examined. When determining post-merger market power, however, the Guidelines state that as long as sufficient entry would likely occur within two years from the date the merged firm were to begin charging supra-competitive prices, the merger will not be banned.[310] The Guidelines consider as potential entrants only "committed entry" — "new competition that requires expenditure of significant sunk costs of entry and exit."[311] The committed entrant definition encompasses all new entry by new plant creation — by diversifying entrants and start-ups alike — without making any distinction between the two entrant types.

This Article's findings suggest, however, that diversifying entrants, which tend to attempt entry at a much larger scale than startups, and larger-scale entrants more generally, are those most likely to provide a short-term competitive check on the newly-merged firm. Nevertheless, the Guidelines' language is flexible enough to accommodate these findings.

Specifically, the first component of the Guidelines' three-part analysis of the effectiveness of entry in counteracting anticompetitive

---

ernment establishes a presumption that the transaction will substantially lessen competition." (footnote omitted) (citing United States v. Citizens & S. Nat'l Bank, 422 U.S. 86, 120-22 (1975)); *Philadelphia Nat'l Bank*, 374 U.S. at 363; Horizontal Merger Guidelines 57 Fed. Reg. 41,552, § 1.0 (1992).

Even in mergers occurring in concentrated industries, however, one court has required defendants to provide only limited proof that a merger creating a dominant firm is unlikely to substantially lessen competition, relying, inter alia, on the traditional belief in the disciplinary power of potential entry. *Baker Hughes*, 908 F.2d at 983-988 (citing numerous cases to argue that when barriers to entry are low, potential entry serves as a check on the use of market power). The findings here suggest, however, that in concentrated industries the short-term disciplinary potential of new entry is especially limited, even more than it is in other industries. The *Baker Hughes* court was therefore probably mistaken in relying on potential entry to discipline the post-merger firm from exercising its power.

308. Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, § 1.32, § 1.321 (1992).

309. *Id.* at § 3.2.

310. The framework of this three-part test, although not typically applied as suggested below, is largely compatible with the conclusions of this Article regarding the competitive impact and limitations of new entry in the post-merger market.

311. Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, § 3.0 (1992). This definition differentiates these potential new entrants from those "uncommitted entrants" that are able to respond to a profit opportunity in the relevant market without making significant sunk investments. *Id.* at § 1.32. Such entrants are largely comparable to the diversifying product-mix entrants discussed briefly above. *See supra* Section I.A and notes 47 & 50. The Guidelines take into account their competitive impact when defining the relevant market.

effects is "timeliness."[312] To be timely enough, the potential new entry must not only occur within two years from initial planning; it must also be expected to "achieve a significant impact on price in the relevant market" within this time frame.[313] This component of the timeliness requirement — the demand that the anticipated new entry will have significant impact on price — can accommodate the distinction between different types of potential entrants. If the conditions and history of the market suggest that larger-scale, probably diversifying, entry is likely to occur following the merger, then a significant impact on price is more likely to follow. If, however, the evidence suggests that only smaller-scale, mostly startup, entry is likely to occur, then a significant timely impact on price would not be expected.[314]

The second component of the analysis is that entry must be "likely."[315] As one might expect, the Guidelines state that entry is likely only "if it would be profitable at premerger prices."[316] This standard might appear to understate the likelihood of new entry, at first, given the empirical evidence showing the pervasiveness of boundedly rational entry. In fact, when determining the likelihood of profitable entry, the Guidelines specifically explain that all the costs of entry must be taken into account, "including an appropriate rate of return on invested capital *given that entry could fail and sunk costs, if any, will be lost.*"[317]

However, while the Guidelines may understate the likelihood of entry per se, their definition is largely compatible with the evidence on likely penetration — the relevant source of competitive impact on post-merger prices. The definition does just that by declaring that entry is unlikely if new entrants would not have the opportunity to obtain minimum viable scale,[318] which the empirical findings examined in this Article suggest will often be the case with many smaller entrants.

---

312. Horizontal Merger Guidelines, 57 Fed. Reg. 41,552, §§ 3.0, 3.2 (1992).

313. *Id.* at § 3.2.

314. *See e.g.*, Rebel Oil Co. v. Atl. Richfield Co., 51 F.3d. 1421, 1440-41 (9th Cir. 1995) ("The fact that entry has occurred does not necessarily preclude the existence of 'significant' entry barriers. If the output or capacity of the new entrant is insufficient to take significant business away from the predator, they are unlikely to represent a challenge to the predator's market power."); Ohau Gas Serv., Inc. v. Pac. Res., Inc., 838 F.2d 360, 366-67 (9th Cir. 1988), *cert. denied*, 488 U.S. 870 (1998) (finding that evidence of entry that did not pose a competitive threat to incumbent — including small entry — did not preclude a jury finding that defendant monopolized the market); *In re* Time Warner, Inc., 123 F.T.C. 171 (1997), 1997 FTC LEXIS 13, *56 (the majority of the commissioners finding in a statement accompanying a consent order that evidence of small entry does not suggest a likely competitive threat to the large merging firms within the time frame set by the guidelines, due to the difficulty of achieving significant penetration).

315. *Id.* at § 3.3.

316. *Id.*

317. *Id.* (emphasis added).

318. *Id.*

Therefore, especially in those markets where economies of scale are important for the survival of new ventures — but not sufficiently taken into account by many prospective entrants — the requirement of likely rational entry minimum viable scale prospects aligns the Guidelines' definition with the more competitively significant prospect of "increased likelihood of penetration."[319]

Moreover, the factors enumerated by the Guidelines as potential sources of sales' opportunities available to entrants include, inter alia, "the ability of the latter to capture a share of reasonably expected growth in market demand" and "divert sales from incumbents."[320] By taking into account such factors, the Guidelines leave room for considering the special impact successful innovative entrants may have in markets that are characterized by a high rate of innovation.[321] In such markets, startup entrants may often contribute to growth and sometimes divert sales from incumbents, increasing the likelihood that some new entrants will obtain the requisite minimum viable scale, and, therefore, that new entry will impact post-merger prices.[322]

Finally, the Guidelines emphasize not only the timeliness and the likelihood of entry but also the sufficiency of its "magnitude, character and scope."[323] However, this requirement — which could have been used to guarantee that entry will be deemed easy only when sufficiently large-scale or significant innovative entry is likely to occur — is constructed narrowly by the Guidelines. In fact, the sufficiency requirement appears almost automatically fulfilled whenever entry is likely.[324]

---

319. *Cf.*, *e.g.*, FTC v. Cardinal Health, Inc., 12 F. Supp. 2d 34, 54-58 (D.D.C. 1998) (finding, when examining whether merging defendants have rebutted evidence of market power given high post-merger concentration, that where larger-scale entry was unlikely to occur within the time frame set by the guideline, the existing potential for expansion by small "fringe" firms was not sufficient to achieve the required impact on price).

320. *Id.*

321. Note, however, that to the extent an innovative entrant who can already be identified is "uncommitted" to entry, it would be taken into account at the market definition stage rather than in determining the likely impact of potential entry on post-merger prices, as explained *supra* note 311 and accompanying text.

322. Thus, the Guidelines explain that "[e]ntrants' anticipated share of growth in demand depends [inter alia] . . . on the relative appeal, acceptability, and reputation of incumbents' and entrants' products to the new demand." *Id.* at § 3.3 n.33.

323. *Id.* at §§ 3.0, 3.4.

324. The only two exceptions noted by the Guidelines are those occasions where incumbents control important assets that limit entrants' opportunities and on some mergers between producers of differentiated products. *Id.* at § 3.4.

V. BOUNDED RATIONALITY IN MARKETS: LESSONS FOR LEGAL POLICY

A. *The Prevalence of Bounded Rationality in Legally Relevant Market Settings*

The behavioral analysis of entry demonstrates why the law should often take into account the bounded rationality of legal actors in market and nonmarket settings. It shows how a legal policy that disregards either the effects of barriers on overconfident entrants or the long-term benefits of small-scale, innovative, negative expected value entry, stands a significant risk of decreasing instead of increasing social welfare.

More generally, some of the characteristics identified in the area of entry as facilitating overconfidence and preventing effective learning from taking place are typical of other market environments as well:

1) Motivation. Market participants must often make judgments about their future performance and the outcomes of events that are important for them; these judgments are therefore highly likely to reflect optimistic and desirability biases.[325]

2) Decision Ambiguity. Before making their decisions, market decisionmakers generally have some information but also face a significant degree of ambiguity. This ambiguity provides room for the preferences of these actors to bias their judgments.[326]

3) Noisy Feedback. Market participants find it difficult to discover their errors. Even when they recognize that their judgments caused negative outcomes, they rarely are able to associate these outcomes with specific judgmental mistakes.[327]

4) No Effective Arbitrage. Profitable opportunities for arbitrage in the face of significant uncertainty are rare. Consequently, arbitrageurs do not effectively limit the operation of bounded rationality in many markets.[328]

5) No Simple Intervention. Contrary to the suggestions of traditional law and economics advocates,[329] debiasing market decisionmakers is typically not a realistic op-

325. *See supra* Section II.C.1.

326. *See supra* Section II.C.2.

327. *See infra* note 333.

328. *See infra* note 334.

329. *E.g.*, Posner, *supra* note 4, at 1575.

> tion. Other interventions, such as the provision of information, may prove helpful in certain cases. Most often, however, these interventions cannot overcome the fundamental, pervasive forces that generate decision biases.[330]

Taken together, the prevalence of these characteristics in legally relevant market settings suggests that boundedly rational behavior may be far more common in markets than traditional analyses acknowledge. Legal scholars would therefore be well advised to examine the impact of behavioral forces in shaping market outcomes when determining the appropriate role of the law in regulating economic behavior.

B. *The Limits of Arguments Relying on Markets to Eliminate Bounded Rationality*

Advocates of traditional law and economics have claimed that the law need not take into account bounded rationality when examining the behavior of legal actors in market settings. The main argument marshaled for this position is that competitive forces will discipline boundedly rational actors. According to this theory, wherever competitive selection operates, boundedly rational actors will make more errors, underperform and exit.[331]

The fable of entry has shown the limits of competitive selection. In the following paragraphs, this Article explains why other arguments for the corrective effects of markets are of little relevance to the analysis of entry and many other market settings of legal interest.

First, traditional economists argue that because in markets decisionmakers pay a price for their mistakes they learn and correct their errors.[332] The learning argument assumes, however, that decisionmak-

---

330. *See supra* Section IV.B.

331. *See supra* note 4 and accompanying text. The arguments promoted by law and economics scholars mirror the traditional responses of economists to assertions that behavioral findings cast doubt on the applicability of rational actor models, most notably traceable to Milton Friedman. *See* MILTON FRIEDMAN, *The Methodology of Positive Economics*, *in* ESSAYS IN POSITIVE ECONOMICS 3, 21-22 (1953). Hogarth & Reder summarize this position stating:

> The economics paradigm focuses on actions taken in competitive circumstances. The underlying assumption is that through competition the action of individual agents are subject to feedback that forces them either to become effective or to withdraw from such actions . . . . Economists have little interest in modeling agents who do not behave according to rational principles since they believe that these agents will not survive in the market.

Robin M. Hogarth & Melvin W. Reder, *Introduction: Perspectives from Economics and Psychology*, *in* RATIONAL CHOICE: THE CONTRAST BETWEEN ECONOMICS AND PSYCHOLOGY 1, 6 (Robin M. Hogarth & Melvin W. Reder eds., 1986) [hereinafter RATIONAL CHOICE] (note omitted).

332. That is, before market discipline eliminates them.

ers are able to identify their mistakes, associate them with the costs they incur and proceed to correct them. These assumptions are rarely met either in the case of entry or in other legally significant real-world settings.[333]

Second, some theorists suggest that where markets are available, rational actors will identify, exploit, and consequently erode the profit opportunities resulting from the errors of boundedly rational decisionmakers.[334] The arbitrage argument assumes, however, the presence of a sufficiently large group of arbitrageurs who can both identify the opportunity and bear the risk involved with selling to or buying from the boundedly rational actors.[335] It also assumes the ready availability of substitutes for the products overpriced or underpriced by boundedly rational actors.[336] Since the conditions necessary for effective arbitrage rarely exist even in the most advanced financial markets, this argument fails to apply to other market settings.[337] For example, rational arbitrageurs would be hard pressed to find an easy way to benefit from negative expected value entry attempts of overconfident entrants.

Last, economists have asserted that individual behavior is of no import in market settings as long as markets perform in the aggregate

---

333. *See supra* note 273 and accompanying text; *supra* Section IV.B (discussing the difficulty of learning in real-world settings). As Amos Tversky and Daniel Kahneman explain:

> Effective learning takes place only under certain conditions: it requires accurate and immediate feedback about the relation between the situational conditions and the appropriate response. The necessary feedback is often lacking for decisions made by managers, entrepreneurs, and politicians because (i) outcomes are commonly delayed and not easily attributable to a particular action; (ii) variability in the environment degrades the reliability of the feedback, especially where outcomes of low probability are involved; (iii) there is often no information about what the outcome would have been if another decision had been taken; and (iv) most important decisions are unique and therefore provide little opportunity for learning.

Amer Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, *in* RATIONAL CHOICE, *supra* note 331, at 67, 90 (citation omitted).

334. *See* SHLEIFER, *supra* note 93, at 1, 3-4, 13-16 (2001) (providing a highly readable introduction to the argument from arbitrage and related evidence from financial markets).

335. This risk is especially large in a world fraught with uncertainty where rational actors sometimes make mistakes as well, the boundedly rational are probabilistically selected for success on some occasions, and profits materialize only in the long run.

336. SHLEIFER, *supra* note 93.

337. The case for arbitrage is obviously even more tenuous in noneconomic legal settings:

> Consider the proposition that a potential criminal will commit some crime if the expected gains from the crime exceed its expected costs. Suppose a criminal mistakenly thinks that the expected gains outweigh the expected costs, when in fact the opposite is true. First notice that no arbitrage will be possible in this situation. If someone is unfortunate enough to commit a crime with a negative expected value, then there is no way for anyone else to profit directly from his behavior. Outside of financial markets (and not always there), those who engage in low-payoff activities lose utility but do not create profit opportunities for others. Nor do they typically disappear from the market.

Jolls et al., *supra* note 1, at 1486.

"as if" individuals were strictly rational.[338] According to this view, there is reason to believe that rationality assumptions approximate actual market performance because the various erroneous decisions made by specific individuals will cancel out. For example, some decisionmakers will overestimate the risks associated with a certain investment while others will underestimate it. As a result, the former will demand an excessive return on this investment while the latter will only require a lower-than-rational return. On average, however, the return demanded for the investment in the market will resemble the return that would have been demanded if the various decisionmakers were demanding the "rational" rate of return individually.[339]

The problem with the "as if" argument is that individual errors will only cancel in the aggregate if two conditions are fulfilled: the errors must be randomly distributed and their mean must resemble the "rational" judgment.[340] These conditions, however, do not hold where bounded rationality causes *systematic, predictable errors* in human judgment.[341] The evidence of the significant and consistent entrant overconfidence and bias therefore undermines this argument, as the operation of cognitive heuristics and biases in other legal settings often does.

Thus, the arguments commonly used to support and explain the impact of markets are, at best, of limited relevance in legal settings such as the entry context examined here. Clearly, they do not justify a presumption that legal analysis can avoid examining the role of boundedly rational behavior and its consequences in market settings more generally.

C. *Some Broader Lessons for Legal Analyses of Market Behavior*

This Article's novel framework provides a clear example of the unique advantages of a behaviorally-informed approach over other uninformed alternatives. It can also be generalized to many other market settings of legal interest.

1. *When Should Legal Analyses of Market Behavior Take Bounded Rationality into Account?*

The findings of the behavioral approach to entry stem from a better understanding of the psychological processes of human judgment and decisionmaking. These processes not only make all entrants prone

338. *See, e.g.*, FRIEDMAN, *supra* note 331.

339. *E.g.*, POSNER, FRONTIERS, *supra* note 4, at 261; Posner, *supra* note 4, at 1556.

340. *E.g.*, Posner, *supra* note 4, at 1556.

341. *See supra* note 3 and the accompanying text; *see also* Jolls et al., *supra* note 1, at 1476-78 (emphasizing that patterns of bounded rationality are systematic and predictable).

to exhibit overconfidence and other biases, but also make startup entrants more biased than diversifying ones. Because of the large number of highly biased potential entrants, competitive forces generate an overrepresentation of highly biased entrants among the ranks of successful competitors, even while they eliminate many similarly biased ones from the market.

These basic characteristics of the dynamics of competition among entrants are present in other legally relevant market settings as well. In fact, any market interaction involving a sufficiently high proportion of biased legal actors will likely lead to the success and survival of some biased actors, significantly affecting market outcomes. Importantly, the relevant bias need not be overconfidence — although this is a pervasive phenomenon[342] — and the proportion of highly biased entrants need not be as extreme as it is in the case of entry. The degree of impact on the private and social outcomes of the interaction between these decisionmakers will depend, however, on the extremity of both the bias and the ratio of more- to less-biased legal participants, as well as the presence or absence of those market forces examined above[343] that may correct or eliminate boundedly rational behavior.

The area of securities regulation, for example, is a case in point. In this area, the law has long recognized the need for regulation of various market behaviors relating to both the original sale of securities and their trading by market participants.[344] Many of these participants, most notably individual investors but investment professionals as well, exhibit significant overconfidence; moreover, the bounded rationality of these investors not only generates private losses for them but also affects the overall performance of securities markets.[345] It is therefore not surprising to find the securities laws attempting to limit the impact of bounded rationality and facilitate rational investing in the market.[346]

342. *Cf.* Rachlinski, *supra* note 86, at 760-63 (arguing that because overconfidence is so pervasive, it is likely to exert a persistent effect on the behavior of contracting parties making decisions on liquidated damages than will other, less robust psychological factors).

343. *See supra* Sections V.A & V.B.

344. *See* 1 LOUIS LOSS & JOEL SELIGMAN, SECURITIES REGULATION §§ 1-A — 1-G (3d rev. ed. 2001).

345. *See supra* notes 194-199 and accompanying text.

346. The present regime does not necessarily appear optimal when examined from a behavioral perspective, however. For some behaviorally informed analyses of securities regulation, see Donald C. Langevoort, *Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (and Cause Other Social Harms)*, 146 U. PA. L. REV. 101 (1997); Donald C. Langevoort, *Selling Hope, Selling Risk: Some Lessons for Law from Behavioral Economics About Stockbrokers and Sophisticated Customers*, 84 CAL. L. REV. 627 (1996).

2. *How Should Legal Analyses of Market Behavior Take Bounded Rationality into Account?*

The behavioral analysis of entry suggests that the law must be sensitive not only to the bounded rationality of actors generally, but also to the presence of behaviorally different types of actors in the market. It is therefore important to develop an understanding of the variables that determine both whether decisionmakers will be biased and how biased they will be.[347]

So, in other words, the fable of entry teaches us that all legally-relevant economic actors are not the same. This important conclusion resonates with numerous legal doctrines that provide different rules for different types of actors. In tort law, for example, the standard of behavior that amounts to "due care" on the part of an allegedly negligent tortfeasor varies with the defendant's level of expertise.[348] The present analysis shows, however, that different legal actors may vary not only in obvious personal or economic characteristics, such as their information or experience, but also in the degree and kind of bounded rationality they are likely to manifest. It will therefore be beneficial for legal scholarship to examine whether and to what degree the law already addresses or should address such behavioral differences.

A related important lesson from the analysis in this Article is that the legal analyst should strive to develop an accurate understanding of those variables that determine whether and how biased different actors are likely to be. In the case of entry, an understanding of the role motivation and ambiguity play in shaping entrant bias has exposed systematic differences between entrants. These and similar behavioral factors also affect decisionmaking in numerous other legal settings.

For instance, contract law scholars have long attempted to explain the limitations imposed by the law on the bargain principle. According to this principle, contracting parties should be bound to the terms of their agreement. Many legal doctrines nevertheless limit freedom of contract and trump the bargain principle by defining circumstances in which parties are not bound to their original agreement.[349] The behavioral approach suggests, however, that certain limitations of the bargain principle, such as the special scrutiny with which the courts re-

347. This is not to say that one could necessarily quantify with accuracy the degree of bias of individual actors. *See supra* Section IV.B. A general understanding of the factors that create and facilitate decision errors is still helpful for the analysis of legal rules and doctrines, however, as the fable of entry shows.

348. *See* RESTATEMENT (SECOND) OF TORTS § 283 (1965) (the general standard of conduct only requires one to behave as would a reasonable man under the circumstances); *id.* at § 299A (stating that professionals are "required to exercise the skill and knowledge normally possessed by members of that profession or trade").

349. *See generally* Melvin Aron Eisenberg, *The Bargain Principle and Its Limits*, 95 HARV. L. REV. 741 (1982).

view liquidated damages, may reflect the law's recognition of the bounded rationality of contracting parties.[350]

The case of form contracts is another area of contract law that may have developed rules that limit the bargain principle based on behavioral differences between different types of legal actors. Both the Uniform Commercial Code and common law doctrine provide special rules that govern the validity of pre-printed terms,[351] tending to treat repeat market players and non-repeat market players differently.

More generally, informed analyses would build on the major strength of the behavioral approach as compared to traditional ones — its awareness of the effects of context. Because of its sensitivity to the effects of context on human behavior, a behaviorally informed analysis can often avoid the pitfall of over-generalization that is common in legal theory.[352] Instead of assuming decisionmakers always conform to a fixed set of norms, a behavioral approach can often identify *how* decisionmakers will behave, taking into account the environment in which they will make their decisions.[353]

A behaviorally informed scholarship can therefore provide important benefits beyond law and economics. Regardless of the normative theory the legal analyst applies to the question at hand, a scientific empirically-based understanding of human judgment and decisionmaking is bound to provide better predictions, and consequently more effective prescriptions, for legal policy.[354]

Finally, this Article has shown the importance of the behavioral approach for the legal analysis of competitive behaviors — where

350. This possibility has been explored recently by a number of legal scholars. *See* Melvin Aron Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211 (1995) (suggesting that various limitations of the bargain principle, including liquidated damages, are better explained by a behavioral approach than by other traditional explanations); Rachlinski, *supra* note 86 (providing a behavioral analysis of liquidated damages). *But see* Robert A. Hillman, *The Limits of Behavioral Decision Theory in Legal Analysis: The Case of Liquidated Damages*, 85 CORNELL L. REV. 717 (2000) (arguing that behavioral findings provide no clear directive as to the optimal legal regime for liquidated damages).

351. These rules include the provisions of U.C.C. § 2-207(1) and the common law doctrine of unfair surprise. *See* Eisenberg, *supra* note 350, at 245-48.

352. Some legal scholars argue that awareness of context effects is a potential disadvantage rather than an advantage of the behavioral approach. *See, e.g.*, Hillman, *supra* note 350, at 730-31. However, since human behavior does vary systematically depending on its context, whether or not the theorist assumes such variations away, remaining uninformed about robust and systematic effects of context on decisionmaking is not a sensible strategy of dealing with the complexity of human behavior.

353. *See, e.g.*, Rachlinski, *supra* note 86, at 743-44 ("Studying the effect of context has long been a part of psychology in general and [behavioral decision theory] in particular. It is a core principle of psychological research that understanding a phenomenon requires understanding when the phenomenon will occur and when it will not.").

354. *See, e.g.*, Rachlinski, *supra* note 86, at 743 (noting the scientific, empirical foundation of the behavioral approach).

many scholars have deemed it irrelevant. It follows that in other legal settings, where even the limited impact of market forces is absent and decisionmakers have greater freedom to engage in boundedly rational action, behavioral insights are often indispensable.[355]

CONCLUSION

Scholars have noted that the rate of entry seems excessive, given the data on entrant mortality and profitability, and recent experimental findings have linked excess entry to optimistic bias. The present analysis goes beyond the existing literature by developing a systematic behavioral framework for understanding the processes that generate robust and common overconfident entry. This novel framework also explains a set of additional puzzling empirical findings on entry, revealing that entrants' insensitivity to market predictors of their future profitability is a likely byproduct of the processes of overconfidence and showing that the inferior average performance of startup entrants stems from their greater degree of overconfidence as compared to their diversifying competitors.

In this way, the behavioral approach provides a unified, coherent solution to the puzzles of entry, showing them all to stem from related aspects of entrant bounded rationality. Furthermore, an analysis of the interaction between the differential economics of startup and diversifying entrants and their different degrees of overconfidence have painted the competition among entrants in a new light. It showed how the enhanced bias of startups and their large numbers lead to their overrepresentation as a group among the ranks of successful entrants, even while it diminished their average individual prospects.

Initially, the market's selection and promotion of many boundedly rational actors even while disciplining most others suggested that some form of entry regulation may be advisable. A comparison of the social costs and benefits of negative expected value entry found, however, that while overconfident entrants generate negative externalities, they also bring about significant social benefits, most importantly when they serve as a venue for the introduction of innovative ideas and products into the market and an important source of additional competitive pressure on incumbent firms in the longer term. Upon further analysis it also became apparent that the regulation of entry would not only be mostly undesirable, but also costly, impractical, and largely ineffective.

After dismissing the regulatory option, the analysis concluded by outlining the implications of the behavioral approach for antitrust law, taking the presence of boundedly rational, overconfident, entry as given. The framework developed here supported the traditional hos-

355. *See, e.g.*, Jolls et al., *supra* note 1, at 1473.

tility of the law to unnecessary restrictions on new entry and suggested that the legal view of the relationship between market share, market power, and entry barriers incumbents may require modification to account for the boundedly rational behavior of entrants.

Last, this study of the competition for profitability and survival among new entrants into industry also illustrated how a detailed understanding of the psychology of boundedly rational actors is not only helpful for clarifying empirical findings, but also crucial for developing effective, realistic legal policy towards complex economic phenomena. In view of the profound impact boundedly rational actors can exert on the market in an intensely competitive environment — on the occasions where market forces do not suffice to discipline them fully — it seems all the more important to examine those psychological processes affecting human judgment and decisionmaking in other legal contexts, where the lack of intense competitive pressures might allow an even more far-reaching expression of bounded rationality.