An Engine, Not a Camera

How Financial Models Shape Markets

Donald MacKenzie

The MIT Press
Cambridge, Massachusetts
London, England

This book does not constitute financial advice, and it is sold with the understanding that neither the author nor the publisher is engaged in rendering investing, legal, accounting, or other professional service. If investment advice or other expert assistance is required, the services of a competent professional person should be sought.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

Set in Baskerville by SNP Best-set Typesetter Ltd., Hong Kong. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

MacKenzie, Donald A.
An engine, not a camera : how financial models shape markets / Donald MacKenzie.
p. cm. — (Inside technology)
Includes bibliographical references and index.
ISBN 0-262-13460-8
1. Capital market—Mathematical models. 2. Derivative securities—Mathematical models. 3. Financial crises—Mathematical models. 4. Financial crises—Case studies. I. Title. II. Series.

HG4523.M24 2006 332'.01'5195—dc22 2005052115

10 9 8 7 6 5 4 3 2 1

1

Performing Theory?

Chicago, late evening, October 19, 1987. Leo Melamed leaves a dinner meeting in the Metropolitan Club on the sixty-seventh floor of the Sears Tower. He walks along Wacker Drive to the twin skyscrapers of the Mercantile Exchange, where his office is on the nineteenth floor, high above the exchange's now-silent trading pits. His assistant greets him with a stack of pink message slips from those who have telephoned in his absence. As midnight approaches, "with sweating hands" he makes his first return call, to the Adolphus Hotel in Dallas. It is to Alan Greenspan, who two months earlier had been appointed to chair the Federal Reserve System's Board of Governors.[1]

Leo Melamed's life is a quintessential twentieth-century story. He was born in Bialystok, Poland. In 1939, at the age of seven, he watched, peeking through a crack in the shutters of his parents' home, as German troops entered the city. He witnessed the macabre ceremony in which Bialystok was handed over, under the temporary pact between Hitler and Stalin, to the Soviet Union. He and his family took the last train from Bialystok across the closing border into Lithuania. Almost certainly, they owed their lives to one of the good people of a bad time: Chiune Sugihara, who headed Imperial Japan's consulate in Kovno, Lithuania.

Against his government's instructions, Sugihara was issuing letters of transit to Lithuania's Jewish refugees—hundreds every day. One of Sugihara's visas took Melamed's family to Moscow, to Vladivostok, and to Kobe. The American embassy in Tokyo (Japan and the United States were not yet at war) provided them with a visa, and in 1941 they reached Chicago, where Melamed eventually became a "runner" and then a trader at the Chicago Mercantile Exchange.[2]

The "Merc" had been Chicago's junior exchange. The Board of Trade, with its glorious art deco skyscraper towering over LaSalle Street, dominated futures on grain, the Midwest's primary commodity. The Merc traded futures on humbler products—when Melamed joined it, eggs and onions. As Melamed's influence grew, he took the exchange in a new direction. Trading futures on currencies, on Treasury bills, on "Eurodollar" interest rates, and on stock indices, it was the first modern financial derivatives exchange. By the mid 1980s, it was central to global finance.

That October night in 1987, however, all Melamed had built—indeed much of the U.S. financial system—was close to ruin. During the day, America's financial markets had crashed. The Dow Jones industrial average had plummeted 22.6 percent. The Standard and Poor's (S&P) 500 index had lost about 20 percent. Futures on the S&P 500 were traded on the Mercantile Exchange, and they should have moved in tandem with the index. Instead, they had "disconnected," falling 29 percent (Jackwerth and Rubinstein 1996, p. 1611).

What Greenspan wanted to know from Melamed was whether the Mercantile Exchange would be able to open the following morning. Melamed was not able to promise that it would. Every evening, after a futures exchange such as the Merc closes, the process of clearing is undertaken. Those whose trading positions have lost money must transfer cash or collateral to the exchange's clearinghouse for deposit into the accounts of those whose positions have gained. After a normal day on the Merc in the late 1980s, $120 million would change hands. On the evening of October 19, however, those who had bought S&P futures contracts owed those who had sold such contracts twenty times that amount (Melamed and Tamarkin 1996, p. 359).

Across the United States, unknown numbers of securities-trading firms were close to failure, carrying heavy losses. Their banks, fearing that the firms would go bankrupt, were refusing to extend credit to see them through the crisis. That might leave those firms with no alternative other than "fire sales" of the stocks they owned, which would worsen the price falls that had generated the crisis. It was the classic phenomenon of a run on a bank as analyzed by the sociologist Robert K. Merton[3] (1948)—fears of bankruptcy were threatening to produce bankruptcy—but at stake was not an individual institution but the system itself.

For example, by the end of trading on Monday October 19, the New York Stock Exchange's "specialists," the firms that keep stock trading going by matching buy and sell orders and using their own money if there is an imbalance, had in aggregate exhausted two-thirds of their capital. One such firm was rescued only by an emergency takeover by Merrill Lynch, the nation's leading stockbroker, sealed with a handshake in the middle of that Monday night (Stewart and Hertzberg 1987, p. 1).

If clearing failed, the Mercantile Exchange could not open. That would fuel the spreading panic that threatened to engulf America's financial institutions in a cascade of bankruptcies. Melamed knew, that Monday night, just how important it was that clearing be completed. Frantic activity by Melamed and his colleagues throughout the night (including a 3 A.M. call to the home of the president of Morgan Stanley to tell him that his bank owed them $1 billion) achieved the transfer of $2.1 billion, but as morning approached $400 million was still owed to Continental Illinois Bank, which acted as the Merc's agent.

"We hadn't received all the pays," says Barry Lind, who had chaired the Mercantile Exchange's Clearing House Committee and who was called upon in 1987 to advise the Merc's board. "We were missing one huge pay." Some members of the board, which was meeting in emergency session, felt that the Merc should not open. Lind told them to think of the bigger picture, especially the Federal Reserve's efforts to shore up the financial system: "The Fed just spent all these billions of dollars that you are about to demolish. If we don't open, we may never open again. You will have ruined everything they did. Closing the Merc will not help. If you're broke, you're broke."[4]

Around 7 A.M., with 20 minutes to go before the scheduled opening of the Merc's currency futures, Melamed called Wilma Smelcer, the executive of the Continental Illinois Bank who oversaw its dealings with the exchange. This is how he recalls the conversation:

> "Wilma . . . You're not going to let a stinking couple of hundred million dollars cause the Merc to go down the tubes, are you?"
>
> "Leo, my hands are tied."
>
> "Please listen, Wilma; you have to take it upon yourself to guarantee the balance because if you don't, I've got to call Alan Greenspan, and we're going to cause the next depression."
>
> There was silence on the other end of the phone. . . . Suddenly, fate intervened. "Hold it a minute, Leo," she shouted into my earpiece, "Tom Theobald just walked in." Theobald was then the chairman of Continental Bank. A couple of minutes later, but what seemed to me like an eternity, Smelcer was back on the phone. "Leo, we're okay. Tom said to go ahead. You've got your money." I looked at the time, it was 7:17 A.M. We had three full minutes to spare. (Melamed and Tamarkin 1996, pp. 362–363)

The crisis was not over. By lunchtime on Tuesday, the New York Stock Exchange was on the brink of closing, as trading in even the most "blue chip" of corporations could not be begun or continued. But the NYSE, the Chicago Mercantile Exchange, and the U.S. financial system survived. Because the wider economic effects of the October 1987 crash were remarkably limited (it did not spark the prolonged depression Melamed and others feared), the threat it posed to the financial system has largely been forgotten by those who did not experience it firsthand.

The resolution of the crisis shows something of the little-understood network of personal interconnections that often underpins even the most global and apparently impersonal of markets. The Merc's salvation was, as we have seen, a verbal agreement among three people who knew and trusted each other. What eased Tuesday's panic was likewise often quite personal. Senior officials from the Federal Reserve telephoned top bankers and stockbrokers, pressuring them to keep extending credit and not to hold back from settling transactions with firms that might be about to fail. Those to whom they spoke generally did what was asked of them. Bankers telephoned their corporate clients to persuade them to announce programs to buy back stock (First Boston, for example, called some 200 clients on Tuesday morning), and enough of their clients responded to help halt the plunge in stock prices.[5]

The crisis of October 1987 is also the pivot of the twin stories told in this book. One story is of the changes in the financial markets in the United States since 1970, in particular the emergence of organized exchanges that trade not stocks but "derivatives" of stocks and of other financial assets. (The S&P 500 futures traded on Melamed's Mercantile Exchange, for example, are contracts that "derive" their value from the level of the index and thus permit what might be called "virtual ownership" of large blocks of stock.[6])

In 1970, the market in financial derivatives in the United States and elsewhere was very small by today's standards (there are no reliable figures for its total size), and to trade many of today's derivatives, such as the Merc's S&P 500 futures, would have been illegal. By 1987, derivatives played a central role in the U.S. financial system, which is why the fate of the Mercantile Exchange was so critical to that system. Derivatives markets were also beginning to emerge around the world.

By June 2004, derivatives contracts totaling $273 trillion (roughly $43,000 for every human being on earth) were outstanding worldwide.[7] The overall sum of such contracts exaggerates the economic significance of derivatives (for example, it is common for a derivatives contract to be entered into to "cancel out" an earlier contract, but both will appear in the overall figure), and the total must be deflated by a factor of about 100 to reach a realistic estimate of the aggregate market value of derivatives. Even after this correction, derivatives remain a major economic activity. The Bank for International Settlements estimated the total gross credit exposure[8] in respect to derivatives of the sixty or so largest participants in the over-the-counter (direct, institution-to-institution) market at the end of June 2004 as $1.48 trillion, roughly equivalent to the annual output of the French economy. If the dense web of interconnected derivatives contracts represented by that exposure figure were to unravel, as began to happen in the 1998 crisis surrounding the hedge fund Long-Term Capital Management, the global financial system could experience extensive paralysis.

This book's other story is the emergence of modern economic theories of financial markets. Finance was a mainstream subject of study in the business schools of U.S. universities, but until the 1960s it was treated largely descriptively. There was little or nothing in the way of sophisticated mathematical theory of financial markets. However, a distinctive academic specialty of "financial economics," which had begun to emerge in the 1950s, gathered pace in the 1960s and the 1970s. At its core were elegant mathematical models of markets.

To traditional finance scholars, the new finance theory could seem far too abstract. Nor was it universally welcomed in economics. Many economists did not see financial economics as central to their discipline, viewing it as specialized and relatively unimportant in almost the same way as the economics of ketchup, studied in isolation, would be trivial. ("Ketchup economics" was how the economist Lawrence Summers once memorably depicted how work on finance could appear to the discipline's mainstream.[9])

The academic base of financial economics was not in economics departments; it remained primarily in business schools. This often brought higher salaries,[10] but it also meant an institutional separation from the wider discipline and a culture that differed from it in some respects. Nevertheless, by the 1990s finance had moved from the margins of economics to become one of the discipline's central topics. Five of the finance theorists discussed in this book—Harry Markowitz, Merton Miller, William Sharpe, Robert C. Merton, and Myron Scholes—became Nobel laureates as a result of their work in finance theory, and other economists who won Nobel Prizes for their wider research also contributed to finance theory.

The central questions addressed by this book concern the relationship between its two stories: that of changing financial markets and that of the emergence of modern finance theory. The markets provided financial economists with their subject matter, with data against which to test their models, and with some of at least the more elementary concepts they employed. Part of the explanation of why financial economics grew in its perceived importance is the gradual recovery of the stock market's prestige—badly damaged by the Great Crash of 1929 and the malpractices it brought to light—and its growing centrality, along with other financial markets, to the U.S. and world economies. But how significant was the other direction of influence? What were the effects on financial markets of the emergence of an authoritative theory of those markets?

Consider, for example, one of the most important categories of financial derivative: options. (A "call option" is a contract that gives its holder the right but does not oblige the holder to buy a particular asset at a set price on or up to a given future date. A "put option" conveys the right to sell the asset at a set price.) The study of the prices of options is a central topic of financial economics, and the canonical work (Black and Scholes 1973; Merton 1973a) won Scholes and Merton their 1997 Nobel Prizes. (Their colleague Fischer Black had died in 1995, and the prize is never awarded posthumously.)

In 1973, the year of the publication of the landmark papers on option theory, the world's first modern options market opened: the Chicago Board Options Exchange, an offshoot of Melamed's rivals at the Board of Trade. How did the existence of a well-regarded theoretical model of options affect the fortunes of the Options Exchange and the pattern of prices in it? More generally, what consequences did the emergence of option theory have for financial markets?

Models and Their "Assumptions"

The question of option theory's practical consequences will be answered, at least tentatively, in the chapters that follow. However, before I turn to the effect of finance theory on markets I must say more about the nature of the models the theorists developed. "Models" are now a major topic of the history, philosophy, and sociology of science, but the term covers a wide range of phenomena, from physical analogies to complex sets of equations, running on supercomputers, that simulate the earth's climate.[11]

The models discussed in this book are verbal and mathematical representations of markets or of economic processes. These representations are deliberately simplified so that economic reasoning about those markets or processes can take a precise, mathematical form. (Appendix E contains a very simple example of such a model, although to keep that appendix accessible I have expressed the model numerically rather than algebraically.)

The models described in the chapters that follow are the outcomes of analytical thinking, of the manipulation of equations, and sometimes of geometric reasoning. They are underpinned by sophisticated economic thinking, and sometimes by advanced mathematics, but computationally they are not overwhelmingly complex. The Black-Scholes-Merton model of option pricing, for example, yields as its central result a differential equation (the "Black-Scholes equation"—equation 1 in appendix D) that has no immediately obvious solution but is nevertheless a version of the "heat" or "diffusion" equation, which is well known to physicists. After some tinkering, Black and Scholes found that in the case of options of the most basic kind the solution of their equation is a relatively simple mathematical expression (equation 2 in appendix D). The numerical values of the solution can be calculated by hand using standard mathematical tables.

The theoretical work discussed in this book was conducted primarily with pen or pencil and paper, with the computer in the background. The computer's presence is nevertheless important, as would be expected by readers of Philip Mirowski's (2002) account of the encounter between modern economics and the "cyborg sciences." Two major contributors to finance theory, Harry Markowitz and Jack Treynor, worked in operations research (a field whose interweaving with computing and whose influence on economics have been investigated by Mirowski), and the exigencies of computerization were important to William Sharpe's development of Markowitz's model.

Computers were needed to apply finance theory's models to trading. They also were needed to test the models against market data. As will be discussed in chapter 4, the results of those tests were by no means always positive, but as in analogous cases in the natural sciences (Harvey 1981) the very fact of finance theory's testability added to its credibility. It also helped the field to grow by creating roles in financial economics for those whose skills were primarily empirical rather than theoretical. Without computers, testing would have been very laborious if not impossible.

The "mathematicization" of the academic study of finance that began in the 1950s paralleled changes in the wider discipline of economics. Economics had developed in the eighteenth and nineteenth centuries predominantly as what the historian of economics Mary Morgan calls a "verbal tradition." Even as late as 1900, "there was relatively little mathematics, statistics, or modeling contained in any economic work" (Morgan 2003, p. 277). Although the use of mathematics and statistics increased in the first half of the twentieth century, economics remained pluralistic.[12]

However, from World War II on, "neoclassical" economics, which had been one approach among several in the interwar period, became increasingly dominant, especially in the United States and the United Kingdom. The "full-fledged neoclassical economics of the third quarter of the [twentieth] century" gave pride of place to "formal treatments of rational, or optimizing, economic agents joined together in an abstractly conceived free-market, general equilibrium[13] world" (Morgan 2003, p. 279). This approach's mathematical peak was for many years the sophisticated set-theoretical and topological reasoning that in the early 1950s allowed the economists Kenneth Arrow and Gerard Debreu to conclude that a competitive economy, with its myriad firms, consumers, and sectors, could find equilibrium.[14] In 1951, just over 2 percent of the pages of the flagship journal, the *American Economic Review*, contained an equation. In 1978, the percentage was 44 (Grubel and Boland 1986, p. 425).

The mathematicization of economics was accompanied, especially in the United States, by a phenomenon that is harder to measure but real nonetheless: the recovery of confidence, in the economics profession and in the surrounding culture, in markets. The Great Depression of the interwar years had shaken faith in the capacity of markets to avoid mass unemployment. In response, economists following in the footsteps of John Maynard Keynes emphasized the possibility of far-from-optimal market outcomes and the consequent need for government action to manage overall levels of demand. Their analyses were influential both within the economics profession and among policy makers in many countries.[15]

As Melamed's telephone call to Smelcer shows, even in 1987 the fear of a repetition of the interwar catastrophe was still alive. Gradually, however, disenchantment with Keynesian economics and with government intervention grew. The experience of the 1970s—when the tools of such intervention often seemed powerless in the face of escalating inflation combined with faltering growth—was a factor in the growing influence of free-market economists such as Milton Friedman of the University of Chicago, with his "monetarist" theory that the cause of inflation lay in over-expansion of the money supply.

Within economics, the rational-expectations approach became increasingly prominent. In this approach, economic actors are modeled as having expectations consistent with the economic processes posited by the model being developed: the actor "knows as much" as the economist does. From such a viewpoint, much government intervention will be undercut by actors anticipating its likely effects.[16]

No simple mechanical link can be drawn between the way economics as a whole was changing and the way financial markets were theorized. The unity of orthodox, neoclassical economics in the postwar United States is easy to overstate, as Mirowski and Hands (1998) have pointed out, and, as was noted above, even in the 1960s and the 1970s the financial markets did not seem to many economists to be a central topic for their discipline. The mainstream economists who did take finance seriously—notably Franco Modigliani, Paul Samuelson, and James Tobin—often had Keynesian sympathies, while Milton Friedman was among the economists who doubted that some of finance theory counted as proper economics (see chapter 2). Nevertheless, the mathematicization of scholarship on finance paralleled developments in the wider discipline of economics, and finance theorists largely shared their colleagues' renewed faith in free markets and in the reasoning capacities of economic agents. There is, for example, an affinity between rational-expectations economics and the "efficient-market" theory to be discussed in chapter 2.[17]

Like their "orthodox" colleagues in the wider profession, financial economists saw systematic knowledge about markets as flowing from precisely formulated models. As was noted above, finance theory's models are often computationally quite simple. The solutions they yield are often single equations, not large and elaborate sets of equations to be fitted painstakingly to huge amounts of data. To social scientists in disciplines other than economics, to many practitioners in and commentators on financial markets, and perhaps to some of the financial economists' colleagues in the wider discipline, this immediately raises the suspicion that finance theory is *too* simple in its models of markets.

The suspicion of over-simplicity can often be heightened when one examines the "assumptions" of finance theory's models—in other words, the market conditions they posit for the purposes of economic analysis. Typically, those assumptions involve matters such as the following: that stocks and other financial assets can be bought and sold at prevailing market prices without affecting those prices, that no commissions or other "transaction costs" are incurred in so doing, that stocks can be "sold short" (e.g., borrowed and sold, and later repurchased and returned) freely and without penalty, and that money can be borrowed and can be lent at the same "riskless" rate of interest. (The model in appendix E is an example of those assumptions.) Surely such assumptions are hopeless idealizations, markedly at odds with the empirical realities of markets?

For half a century, economists have had a canonical reply to the contention that their models are based on unrealistic assumptions: Milton Friedman's 1953 essay "The Methodology of Positive Economics," which was to become "the central document of modernism in economics" (McCloskey 1985, p. 9). Friedman was already prominent within the discipline by the 1950s, and in later decades his advocacy of free markets and of monetarism was to make him probably the living economist best known to the general public.

In his 1953 essay, Friedman distinguished "positive" economics (the study of "what is") from "normative" economics (the study of "what ought to be"). The goal of positive economics, he wrote, "is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances. Its performance is to be judged by the precision, scope, and conformity with experience of the predictions it yields. In short, positive economics is, or can be, an 'objective' science, in precisely the same sense as any of the physical sciences." (1953a, p. 4)

To assess theories by whether their assumptions were empirically accurate was, Friedman argued, fundamentally mistaken: "Truly important and significant hypotheses will be found to have 'assumptions' that are wildly inaccurate descriptive representations of reality. . . . A hypothesis is important if it 'explains' much by little . . . if it abstracts the common and crucial elements from the mass of complex and detailed circumstances . . . and permits valid predictions on the basis of them alone. To be important, therefore, a hypothesis must be descriptively false in its assumptions." (p. 14) The test of a theory was not whether its assumptions were "descriptively 'realistic,' for they never are, but . . . whether the theory works, which means whether it yields sufficiently accurate predictions" (p. 15).

To a reader versed in the philosophy of science, aspects of Friedman's position—especially his insistence that "factual evidence can never 'prove' a hypothesis; it can only fail to disprove it" (p. 9)—are immediately reminiscent of the writings of Karl Popper. Economic methodologists question, however, just how close Friedman's views are to Popper's, and indeed have found the former hard to classify philosophically.[18]

Popper and Friedman were founding members of the Mont Pèlerin Society, a meeting place of opponents of postwar statist collectivism set up in April 1947 by the free-market economist Friedrich von Hayek. (The society was named after the site of the society's ten-day inaugural meeting, a gathering that Friedman later said "marked the beginning of my involvement in the political process."[19]) Friedman himself certainly sees a similarity between his and Popper's stances. "My position is, essentially, the same as Popper's," he says, "though it was developed independently. . . . I met Popper in 1947, at the first meeting of the Mont Pèlerin Society, but I had already developed all of these ideas before then." (Friedman interview[20])

Ultimately, though, Friedman's "Methodology of Positive Economics" was oriented not to the philosophy of science but to economics,[21] and his stance provoked sharp debate within the profession. The best-known opponent of Friedman's position was Paul Samuelson, an economist at the Massachusetts Institute of Technology. With *Foundations of Economic Analysis* (1947) and other works, Samuelson played a big part in the mathematicization of economics in the postwar United States. He wrote the discipline's definitive postwar textbook (*Economics*, which sold some 4 million copies[22]), and in 1970 he was the third recipient of the Prize in Economic Sciences in Memory of Alfred Nobel. Samuelson ended his Nobel Prize lecture by quoting the economist H. J. Davenport: "There is no reason why theoretical economics should be a monopoly of the reactionaries." (Samuelson 1971, p. 287)

Samuelson seemed to share, at least in part, the suspicion of some of Friedman's critics that Friedman's methodological views were also political, a way of defending what Samuelson called "the perfectly competitive laissez faire model of economics." It was "fundamentally wrong," wrote Samuelson, to think "that unrealism in the sense of factual inaccuracy even to a tolerable degree of approximation is anything but a demerit for a theory or hypothesis. . . . Some inaccuracies are worse than others, but that is only to say that some sins against empirical science are worse than others, not that a sin is a merit. . . . The fact that nothing is perfectly accurate should not be an excuse to relax our standards of scrutiny of the empirical validity that the propositions of economics do or do not possess." (1963, pp. 233, 236)

Just as there is no unitary "scientific method," faithful following of which guarantees scientific advances,[23] there is not likely to be a productive, rule-like economic methodology. For example, Friedman noted that the "rules for using [a] model . . . cannot possibly be abstract and complete." How the "entities in [a] model" are to be connected to "observable phenomena . . . can be learned only by experience and exposure in the 'right' scientific atmosphere, not by rote" (1953a, p. 25). And on page 9 of the 1953 essay he inserted a crucial parenthetical phrase into the passage putting forward falsificationism, writing that a hypothesis should be "rejected if its predictions are contradicted ('frequently' or more often than predictions from an alternative hypothesis)"—a formulation that left room for the exercise of professional judgment.

By the standards of a strict falsificationism, for example, virtually all the models discussed in this book should have been discarded immediately on the grounds that some of their predictions were empirically false. Yet financial economists did not discard them, and they were right not to. For instance, the Capital Asset Pricing Model (discussed in chapter 2) led to the conclusion that all investors' portfolios of risky assets are identical in their relative composition. That was plainly not so, and it was known not to be so, but the model was still highly prized.

Friedman's methodological views were, therefore, not a precise prescription for how economics should be done. His view that economic theory was "an 'engine' to analyze [the world], not a photographic reproduction of it" (1953a, p. 35) was in a sense a truism: a theory that incorporates all detail, as if photographically, is clearly as much an impossibility as a map that reproduces exactly every aspect and feature of terrain and landscape. Nevertheless, the view that economic theory was an "engine" of inquiry, not an (infeasible) camera faithfully reproducing all empirical facts, was important to the developments discussed in this book.

When, in the 1950s and the 1960s, an older generation of more descriptively oriented scholars of finance encountered the work of the new finance theorists, their reaction was, as has already been noted, often a species of "the perennial criticism of 'orthodox' economic theory as 'unrealistic'" (Friedman 1953a, p. 30) that Friedman's essay was designed to rebut. Friedman made explicit a vital aspect of what, borrowing a term from Knorr Cetina (1999), we might call the "epistemic culture" of modern orthodox economics. In so doing, he gave finance theorists a defense against the most common criticism of them, despite his doubts as to whether some parts of finance theory were genuine contributions to economics.[24]

"Around here," the prominent finance theorist Merton Miller told me, "we just sort of take [Friedman's viewpoint] for granted. Of course you don't worry about the assumptions." (Miller interview) By "here" Miller meant the University of Chicago, but he could as easily have been describing much of finance theory. Attitudes to the verisimilitude of assumptions did differ, with Samuelson and (to a lesser extent) his student Robert C. Merton distancing themselves somewhat from the more Friedmanesque attitudes of some of their colleagues. However, that a model's assumptions were "unrealistic" did not generally count, in the epistemic culture of financial economics, as a valid argument against the model.

The Infrastructures of Markets

The "machineries of knowing" (Knorr Cetina 1999, p. 5) that make up finance theory's engines of inquiry are among this book's topics. More central to the book, however, is another issue. Financial economics, I argue, did more than analyze markets; it altered them. It was an "engine" in a sense not intended by Friedman: an active force transforming its environment, not a camera passively recording it.[25]

Economists themselves have had interesting things to say about how their subject affects its objects of study,[26] and there is a variety of philosophical, sociological, and anthropological work that bears on the topic.[27] However, the existing writing that best helps place this theme in a wider context is that of the economic sociologist and sociologist of science Michel Callon. Callon rightly refuses to confine economic sociology to the role economists often seem to expect it to take—as an effort to demonstrate irrational "social" elements intruding into market processes—and sees it instead as what might be called an "anthropology of calculation" which inquires into the processes that make calculative economic action and markets possible:

> . . . if calculations are to be performed and completed, the agents and goods involved in these calculations must be disentangled and framed. In short, a clear and precise boundary must be drawn between the relations which the agents will take into account and which will serve in their calculations and those which will be thrown out of the calculation. . . . (Callon 1998, p. 16)

Callon contrasts modern market transactions with the "entangled objects" described by ethnographers such as Thomas (1991). An object linked to specific places and to particular people by unseverable cultural and religious ties cannot be the subject of market transactions in the same way that, for example, today's consumer durables can. The contrast should not be overdrawn (Callon emphasizes the new entanglements without which markets could not function, and also the way in which market transactions "overflow" their frames[28]), but it helpfully focuses attention on the infrastructures of markets: the social, cultural, and technical conditions that make them possible.

Markets' infrastructures matter. Consider, for example, the market in "futures" on agricultural products such as grain, which is relevant to this book because it was from agricultural futures markets—the Chicago Mercantile Exchange and Board of Trade—that the first of today's financial derivatives exchanges emerged. A "future" is a standardized, exchange-traded contract in which one party undertakes to sell, and the other to buy, a set quantity of a given type of asset at a set price at a set future time. Futures markets did not originate in the United States. What seem to have been in effect rice futures were traded in eighteenth-century Osaka (Schaede 1989), and some European markets also predated those of the United States (Cronon 1991, p. 418). But futures trading developed in Chicago on an unprecedented scale in the second half of the nineteenth century, and it is the subject of a justly celebrated analysis by the historian William Cronon (1991).[29]

A futures market brings together "hedgers" (for example, producers or large consumers of the grain or other commodity being traded), who benefit from being certain of the price at which they will be able to sell or to buy the grain, and "speculators," who are prepared to take on risk in the hope of profiting from price fluctuations. However, successful futures trading requires more than the existence of economic actors who may benefit from it.

For futures trading to be possible, the underlying asset has to be standardized, and that involves a version of Callon's "disentanglement" and "framing." The grain to which a futures contract makes reference may not even have been harvested yet, so a buyer cannot pick out a representative sack, slit it open, and judge the quality of the grain by letting it run through his or her fingers. "Five thousand bushels of Chicago No. 2 white winter wheat" has to be definable, even if it does not yet physically exist.

As Cronon shows, the processes that made Chicago's trading in grain futures possible were based on the disentanglement of grain from its grower that took place when transport in railroad cars and storage in steam-powered grain elevators replaced transport and storage in sacks. Sacks kept grain and grower tied together, the sacks remaining the latter's property, identified as such by a bill of lading in each sack, until they reached the final purchaser. In contrast, grain from different growers was mixed irreversibly in the elevators' giant bins, and the trace of ownership was now a paper receipt, redeemable for an equivalent quantity of similar grain but not for the original physical substance.

The standardization of grain was both a technical and a social process. In Chicago the bushel, originally a unit of volume, became a unit of weight in order to permit measurement on scales on top of each elevator. A team of inspectors—employed first by the Chicago Board of Trade and then by the state of Illinois—checked that the scales were fair and made the inevitably contestable judgments that the contents of this bin were good enough to be classed as "No. 1 white winter wheat," which had to "be plump, well cleaned and free from other grains," while that bin contained only "No. 2," which was defined as "sound, but not clean enough for No. 1" (Cronon 1991, p. 118).

With grains thus turned into "homogeneous abstractions" (Cronon 1991, p. 132), disentangled at least partially from their heterogeneous physical reality, it was possible to enter into a contract to buy or to sell 5,000 bushels (the standard contract size) of, for example, "Chicago No. 2 white winter wheat" at a set price at a given future time. Such a contract had no link to any *particular* physical entity, and because its terms were standardized it was not connected permanently to those who had initially entered into it.[30]

If, for example, one of the parties to a futures contract wished to be free of the obligation it imposed, he or she did not have to negotiate with the original counterparty for a cancellation of the contract, but could simply enter into an equal-but-opposite futures contract with a third party. Although when the specified delivery month arrived a futures contract could in principle be settled by handing over elevator receipts, which could be exchanged for actual grain, in practice delivery was seldom demanded. Contracts were normally settled by payment of the difference between the price stated in the contract and the current market price of the corresponding grade of grain. A future was thus "an abstract claim on the golden stream flowing through [Chicago's] elevators" (Cronon 1991, p. 120).

The disentanglement of the abstract claim from grain's physical reality and the framing of the latter into standardized grades were never entirely complete. The standardization of grain depended on a "social" matter, the probity of the grain inspectors, and in nineteenth-century Chicago that was seldom entirely beyond question. The possibility of settlement by physical delivery and the role played by the current market price of grain in determining cash settlement sums kept the futures market and the "spot" (immediate delivery) market tied together.

However infrequently the physical delivery of grain was demanded, its possibility was essential to the legal feasibility of futures trading in the United States. If physical delivery was impossible, a futures contract could be settled only in cash, and that would have made it a wager in U.S. law. There was widespread hostility toward gambling, which was illegal in Illinois and in most other states. The consequent need for Chicago's futures exchanges to preserve the possibility of physical delivery—the chief criterion demarcating their activities from gambling—cast a long historical shadow. As chapter 6 will show, even in the 1970s this shaped the development of financial derivatives.

A further, particularly dramatic way in which futures trading was sometimes tied to the underlying physical substance was a "corner," in which a speculator or group of speculators purchased large amounts of grain futures and also sought to buy up most or all of the available physical grain. If a corner succeeded, those who had engineered it had at their mercy those who had sold futures short (that is, without owning corresponding amounts of grain). The success of a corner could depend on far-from-abstract matters, such as whether ice-free channels could be kept open in Duluth Harbor or in Thunder Bay long enough to allow sufficient grain to be shipped to Chicago to circumvent the corner. One such attempted corner, the failed "Leiter corner" of 1897–98, was the basis for Frank Norris's classic 1903 Chicago novel *The Pit*.[31]

Another aspect of the infrastructure of agricultural futures trading in the United States was a specific architectural feature of the physical space in which trading took place: the "pit" that gave Norris's novel its title. Overcrowding on the floor of the Board of Trade—which had 2,187 members by 1869 (Falloon 1998, p. 72)—led to the introduction of stepped "amphitheaters," traditionally octagonal in shape.

Despite the name, pits are generally raised above the floor of an exchange, not sunk into it. Standing on the steps of a pit, rather than crowded at one level, futures traders can more easily see each other, which is critical to facilitating Chicago's "open outcry" trading, in which deals are struck by voice or (when it gets too noisy, as it often does) by an elaborate system of hand signals and by eye contact. Where one stands in a pit is important both socially and economically: one's physical position can, quite literally, be worth fighting for, even though throwing a punch can bring a $25,000 fine from an exchange.[32]

The Performativity of Economics

The infrastructures of markets are thus diverse. As we have just seen, the infrastructure of grain futures trading included steam-powered elevators, grain inspectors who were hard to bribe, crowded pits, and contracts that reflected the need to keep futures trading separate from gambling. One important aspect of Callon's work is his insistence that economics itself is a part of the infrastructure of modern markets: ". . . economics, in the broad sense of the term, performs, shapes and formats the economy, rather than observing how it functions" (1998, p. 2).

By "economics, in the broad sense of the term" Callon means "all the activities, whether academic or not . . . aimed at understanding, analyzing and equipping markets" (2005, p. 9)—a definition that obviously goes well beyond the academic discipline. However, it is at least sometimes the case that economics in the narrower, academic sense "performs, shapes and formats the economy." Consider, for example, the "Chicago Boys," economists from the Universidad Católica de Chile trained by Milton Friedman and his University of Chicago colleagues between 1955 and 1964 as part of a Cold War U.S. program "to combat a perceived leftist bias in Chilean economics" (Valdés 1995; Fourcade-Gourinchas and Babb 2002, p. 547). Especially under the government of General Pinochet, the "Chicago Boys" did not simply analyze the Chilean economy; they sought to reconstruct it along the free-market, monetarist lines whose advantages they had been taught to appreciate.

The Chicago Boys are a well-known and politically controversial example, unusual in that it involves particularly direct access by economists to the levers of political power, but this example is a vivid manifestation of a general phenomenon. The academic discipline of economics does not always stand outside the economy, analyzing it as an external thing; sometimes it is an intrinsic part of economic processes. Let us call the claim that economics plays the latter role *the performativity of economics*.

The coiner of the term "performative" was the philosopher J. L. Austin. He admitted that it was "rather an ugly word," but it was one that he thought necessary to distinguish utterances that *do* something (performative utterances) from those that report on an already-existing state of affairs. If I say "I apologize," or "I name this ship the *Queen Elizabeth*," or "I bet you sixpence it will rain tomorrow," then "in saying what I do, I actually perform the action" (Austin 1970, p. 235).[33]

Many everyday utterances in financial markets are performative in Austin's sense. If someone offers to buy from me, or to sell to me, a particular asset for a particular price, and I say "done" or "agreed," then the deal *is* agreed—at least if I am in a market, such as the Chicago futures exchanges, in which a verbal agreement is treated as binding. But what might it mean for economics, or a particular subset of it such as financial economics, to be performative? Plainly, that is a far more complex matter than the analysis of specific, individual utterances.

At least three levels of the performativity of economics seem to me to be possible (figure 1.1).[34] The first, weakest level is what might be called "generic performativity." For an aspect of economics to be performative in this sense

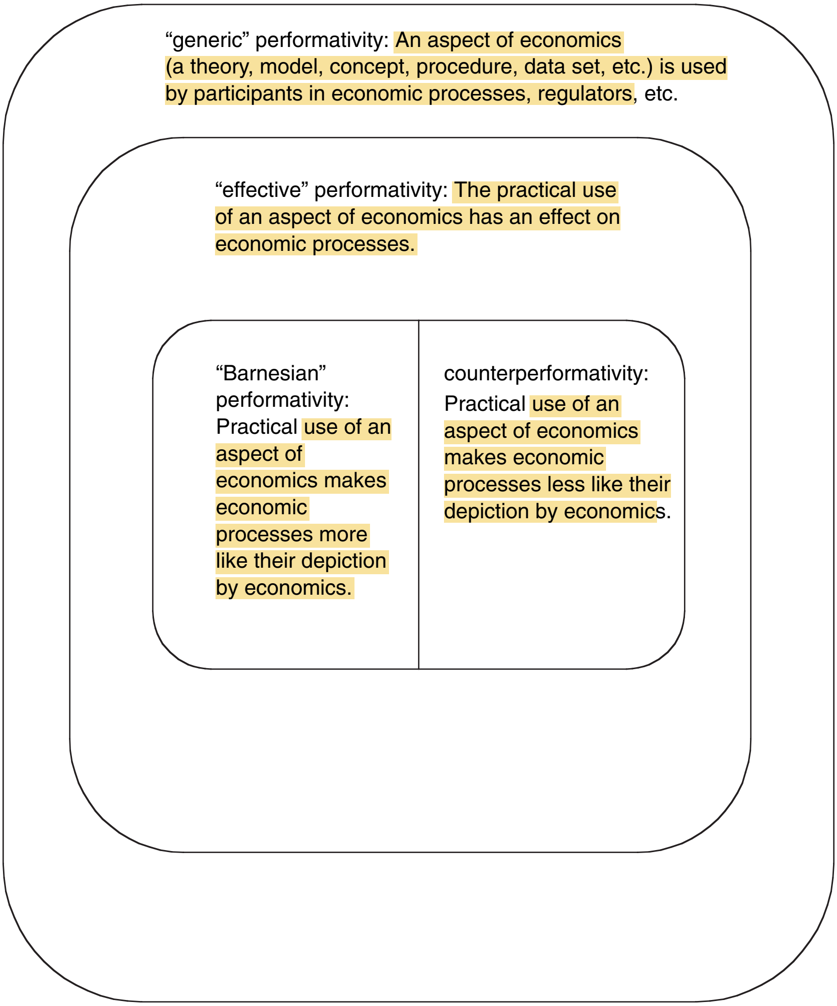

Figure 1.1
The performativity of economics: a possible classification. The depicted sizes of the subsets are arbitrary; I have not attempted to estimate the prevalence of the different forms of performativity.

means that it is used, not just by academic economists, but in the "real world": by market participants, policy makers, regulators, and so on. Instead of being external to economic processes, the aspect of economics in question is "performed" in the generic sense of being used in those processes. Whether this is so is, in principle, a straightforward empirical matter: one simply observes whether economics is drawn on in the processes in question. (In practice, of course, the available sources—historical or current—may not be sufficient to allow one to be certain how matters stand in this respect, and one must remember to look not just at what participants say and write but also at whether the processes in question involve procedures and material devices that incorporate economics.)

What is less straightforward conceptually, and more complicated empirically, is to determine what effect, if any, the use of economics has on the economic process in question. The presence of such an effect is what is required for a stronger meaning of "performativity": the subset of generic performativity that one might call "effective performativity." For the use of a theory, a model, a concept, a procedure, a data set, or some other aspect of economics to count as effective performativity, the use must *make a difference*. Perhaps it makes possible an economic process that would otherwise be impossible, or perhaps a process involving use of the aspect of economics in question differs in some significant way (has different features, different outcomes, and so on) from what would take place if economics was not used.

Except in the simplest cases, one cannot expect observation alone to reveal the effect of the use of an aspect of economics. One cannot assume, just because one can observe economics being used in an economic process, that the process is thereby altered significantly. It might be that the use of economics is epiphenomenal—an empty gloss on a process that would have had essentially the same outcomes without it, as Mirowski and Nik-Khah (2004) in effect suggest was the case for the celebrated use of "game theory" from economics in the auctions of the communications spectrum in the United States.

Ideally, one would like to be able directly to compare processes with and without use of the aspect of economics in question. Such comparisons, however, are seldom entirely straightforward: the relevant situations will typically differ not just in the extent of the usage of economics but in other respects too. There will thus often be an element of conjecture and an element of judgment in attributing differences in outcome to the use of economics rather than to some other factor.

Most intriguing of all the varieties of the performativity of economics depicted in figure 1.1 are the two innermost subsets. There the use of economics is not simply having effects on economic processes: those processes are being altered in ways that bear on their conformity to the aspect of economics in question. In the case of the use of an economic model, for example, one possibility is that economic processes or their outcomes are altered so that they better correspond to the model. Let me call this possibility "Barnesian performativity," because the sociologist Barry Barnes has emphasized (especially in a 1983 article and a 1988 book) the central role in social life of self-validating feedback loops. (In earlier work, I called this type of performativity "Austinian." That had the disadvantage of being read as invoking not sociology, which is what I wanted to invoke, but linguistic philosophy.)[35]

As Barnes notes, if an absolute monarch designates Robin Hood an "outlaw," then Robin *is* an outlaw. Someone is a "leader" if "followers" regard him or her as such. A metal disk, a piece of paper, or an electronic record is "money" if, collectively, we treat it as a medium of exchange and a store of value.[36]

The strong, Barnesian sense of "performativity," in which the use of a model (or some other aspect of economics) makes it "more true," raises the possibility of its converse: that the effect of the practical use of a theory or model may be to alter economic processes so that they conform less well to the theory or model. Let me call this possibility—which is not explicit in Callon's work—"counterperformativity."[37] An aspect of economics is being used in "real-world" processes, and the use is having effects, but among those effects is that economic processes are being altered in such a way that the empirical accuracy of the aspect of economics in question is undermined.

"Barnesian performativity" could be read as simply another term for Robert K. Merton's famous notion of the "self-fulfilling prophecy" (1948), and "counterperformativity" as another word for its less-well-known converse, the self-negating prophecy. I have three reasons for preferring the terminology I use here.

First, I want the terminology to reflect the way in which the strongest senses of "performativity" are subsets of a more general phenomenon: the incorporation of economics into the infrastructures of markets.

Second, the notion of "prophecy," whether self-fulfilling or self-negating, can suggest that we are dealing only with beliefs and world views. While beliefs about markets are clearly important, an aspect of economics that is incorporated only into beliefs "in the heads" of economic actors may have a precarious status. A form of incorporation that is in some senses deeper is incorporation into algorithms, procedures, routines, and material devices.[38] An economic model that is incorporated into these can have effects even if those who use them are skeptical of the model's virtues, unaware of its details, or even ignorant of its very existence.

Third, in Robert K. Merton's original article on "the self-fulfilling prophecy," as in much subsequent discussion, the notion carries the connotation of pathology: an incorrect belief, or at least an arbitrary one, is made true by the effects of its dissemination. It is emphatically not my intention to imply that in respect to finance theory. For example, to say of Black-Scholes-Merton option-pricing theory that it was "performative" in the Barnesian sense is not to make the crude claim that any arbitrary formula for option prices, if proposed by sufficiently authoritative people, could have "made itself true" by being adopted. Most such formulas could not do so, at least other than temporarily.

Even if a formula for option pricing had initially been adopted widely, it would soon have ceased to hold sway if it led those using it systematically to lose money, or if it gave rise to unconstrained opportunities for others to conduct arbitrage. (Arbitrage is trading that exploits price discrepancies to make riskless or low-risk profits.) Imagine, for example, that as a result of a mistake in their algebra Black and Scholes had produced a formula for the value of a call option that was half or double their actual formula (expression 2 in appendix D), that no one noticed, and that the formula was then used widely to price options. It would not have been a stable outcome: the sellers or buyers of options would have incurred systematic losses, and attractive arbitrage opportunities would have been created.

There was, furthermore, much more to the Black-Scholes-Merton model than an equation that could be solved to yield theoretical option prices. The model was an exemplar (in the sense of Kuhn 1970) of a general methodology for pricing a derivative: try to find a continuously adjusted portfolio of more basic assets that has the same payoffs as the derivative. (Such a portfolio is called a "replicating portfolio.") If one can do that, then one can argue that the price of the derivative must equal the cost of the replicating portfolio, for otherwise there is an arbitrage opportunity. Today it would be unusual to find the Black-Scholes-Merton model being used directly as a guide to trading options: in options exchanges, banks' trading rooms, and hedge funds, the model has been adapted and altered in many ways. However, the model's "replicating portfolio" methodology remains fundamental.

The methodology offers not just "theoretical" prices but also a clear and systematic account of the economic process determining those prices. This account altered how economists conceived of a broad range of issues: the pricing not just of derivatives but also of more "basic" securities, such as bonds, and even the analysis of decisions outside of the sphere of finance that can be seen as involving implicit options. It affected how market participants and regulators thought about options, and it still does so, even if the phase of the Barnesian performativity of the original Black-Scholes-Merton model has passed.

Detecting Barnesian Performativity

One way of detecting the Barnesian performativity of an aspect of economics such as a theory or a model is by comparing market conditions and patterns of prices before and after its widespread adoption. (By "market conditions" I mean matters such as the typical level of transaction costs or the feasibility and expense of short sales.) If those conditions or prices have changed toward greater conformity to the theory or model, that is evidence consistent with Barnesian performativity. It does not *prove* performativity, because the change could have taken place for reasons other than the effects of the use of the theory or model. Unfortunately, certainty in this respect tends to be elusive, but that is no reason to abandon the inquiry. All it means is that, as with "effective" performativity, we are dealing with a question of historical or social-science causation on which evidence can throw light but which it would be naive to expect to be resolved unambiguously.

Inquiring into Barnesian performativity thus involves more than an examination of the extent, the manner, and the general effects of the use of economics in economic practice. In investigating market conditions and prices and in judging whether they have moved toward (or away from) conformity to an aspect of economics, one is not just examining economics and those who develop and use it; inevitably one is also studying the "objects" that economics analyzes. That is something that the field to which much of my work has belonged—the sociology of scientific knowledge—has sometimes been reluctant to do.[39]

Certainly one should not underestimate the complexity of judging whether patterns of market prices, for example, have moved toward greater conformity with a model such as Black-Scholes-Merton. One way of formulating the question is to examine the extent to which the model's predictions are borne out. However, what a model predicts is often not straightforward. The Black-Scholes-Merton model, for example, yields an option price only after the characteristics of the option and the values of the parameters of the Black-Scholes equation have been set. One parameter, the volatility of the stock, is acknowledged not to be directly observable, so there is no unique theoretical price to compare with "actual" prices.

Furthermore, "actual" or "real-world" market prices are complex entities. As Koray Caliskan (2003, 2004) points out in a delightful ethnographic discussion of cotton trading, markets abound with prices and price quotations of many kinds. What gets reported as cotton's "world price," for instance, is a complicated construction involving not only averaging but also subjective adjustments.

Difficulties remain even if one restricts oneself to the prices at which transactions are actually concluded. For example, the most thorough early empirical tests of option-pricing models were conducted by the financial economist Mark Rubinstein (see chapter 6). He obtained the Chicago Board Options Exchange's own electronic records of transactions, so he did not have to rely on price quotations or on closing prices, but he still faced problems. For instance, it was common for options to trade at different prices when the price of the underlying stock did not alter at all. In a typical case, "while the stock price [of the Polaroid Corporation] apparently remained constant at $37\frac{1}{2}$ [$37.50], one July/40 call [option] contract was traded at $3\frac{1}{4}$ [$3.25], eight at $3\frac{3}{8}$ [$3.375], and, one at $3\frac{1}{2}$ [$3.50]. Will the true equilibrium option price please stand up?" (Rubinstein 1985, p. 465[40])

Thus, Rubinstein had to analyze not "raw" prices of options, but weighted averages. He also filtered out large numbers of price records that he regarded as problematic. For example, he excluded any record that referred to either the first or the last 1,000 seconds of the options exchange's trading day. He removed transactions close to the start of the day because they often reflected the "execution of limit orders[41] held over from the previous day." He eliminated those in the final minutes before the close of trading because prices then were influenced by "trades to influence market maker margin" (1985, p. 463)—in other words, the level of deposit that had to be maintained in order to be allowed to continue holding a position.

Transactions close to the start or the end of the day involved what Rubinstein called "artificial pricing" (p. 463). Filtering them out from the analysis was a perfectly sensible procedure (Rubinstein had been a trader on an options exchange and so had an insider's understanding of trading-floor behavior), but embedded in the exclusion of what were often the periods of most frantic trading activity was a view of the "natural" operations of markets.

The potentially problematic nature of "real-world" prices is only an example of the complexities of econometric testing: many of the points that historians and sociologists of science have made about scientific experiment can also be made about the testing of finance theory's models. As Callon's colleague Bruno Latour (among many others) has pointed out, detailed attention to the active, transformative processes by which scientific knowledge is constructed breaks down the canonical view in which there is a "world" entirely distinct from "language" and thus undermines standard notions of reference in which "words" have discrete, observable "things" to which they refer.[42]

Replication and the reproducibility of results are at least as problematic in econometrics as the sociologist Harry Collins has shown them to be in the natural sciences.[43] A later test will often contradict an earlier one—see the extensive lists of examples in Goldfarb's 1995 and 1997 papers. In that situation, there may be no a priori way of knowing whether the original test was at fault, whether the new one is incompetent, or whether the discrepancy is to be explained by historical and geographical variation or other differences in the economic processes being studied.

It is also the case that, as was noted above in respect to volatility, what a finance-theory model implies for a specific situation depends not on the model alone but also on auxiliary assumptions about that situation. What is being tested, therefore, is not the model in isolation but the model plus auxiliary assumptions, just as is always the situation in scientific experiment. (This is the "Duhem-Quine" thesis of the philosophy and sociology of science. See, for example, Barnes 1982, pp. 73–76.) An empirical result that apparently falsifies a model can therefore be blamed on a fault in one of the auxiliary assumptions.

For example, many efforts were made empirically to test two developments in finance theory: the Capital Asset Pricing Model and the efficient-market hypothesis. Normally it was not possible to disentangle these entirely so that only one was being tested at a time; typically the tests were of both the model and the hypothesis simultaneously. Tests of market efficiency usually involved examining whether investment strategies were available that systematically generated "excess" risk-adjusted returns. A criterion for what constitutes an "excess" return was thus needed, and in the early years of such testing the Capital Asset Pricing Model was usually invoked as the criterion.[44] When "anomalies" were found in the results of the tests, how to interpret them was therefore debatable: were they cases of market inefficiency, or evidence against the Capital Asset Pricing Model?

Conversely, central to the Capital Asset Pricing Model was what the model posited about the returns expected by investors on assets with different sensitivities to market fluctuations, but typically no attempt was made to measure these expected returns directly—for example, by surveying investors. (The results of any such survey would have been regarded as unreliable by most financial economists.) Instead, in empirical tests of the Capital Asset Pricing Model, more easily measurable after-the-fact realized returns were used as a proxy for expected returns—a substitution that rested on an efficient-market, rational-expectations view of the latter.[45]

Even something as basic as the "cleaning" of price data to remove errors in data entry can, in a sense, involve theory. The main original data source against which finance theory's models were tested was the tapes of monthly stock returns produced by the Center for Research in Security Prices at the University of Chicago. An already-known (and in one sense a theoretical) feature of stock-price changes was used as the basis for the computerized algorithm for detecting data-entry errors:

> Rather than coding and punching all prices twice and then resolving discrepancies manually, we found a better procedure. We know that the change in the price of a stock during one month is very nearly independent of its change during the next month. Therefore, if a price changes a large amount from one date to a second date, and by a similar amount in the opposite direction from the second date to a third, there is a reason to believe that at the second date the price was misrecorded. A "large change" was rather arbitrarily taken to mean a change in magnitude of more than 10 per cent of the previous price plus a dollar. (Lorie 1965, p. 7)

Because of the complexities of econometric testing, the extent of the "fit" between a theoretical model and patterns of prices cannot be determined by simple inspection. "There just isn't any easy way to test a theory," said Fischer Black (1982, p. 32). Knowledge of whether patterns of prices have moved toward greater conformity with a theory is the outcome of a difficult, and often a contested, process. It is therefore tempting to set the issue aside, and to abandon the strongest meanings of performativity (Barnesian performativity and counterperformativity). However, to do that would involve also abandoning a central question: Has finance theory helped to create the world it posited—for example, a world that has been altered to conform better to the theory's initially unrealistic assumptions?

Has the practical use of finance theory (for example, as a guide to trading, or in the design of the regulatory and other frameworks within which trading takes place) altered market processes toward greater conformity to theory? If the answer to that question is at least partially in the affirmative, we have identified a process shaping the financial markets—and via those markets perhaps even the wider economies and societies of high modernity—that has not received anything like sufficient attention. If, on the other hand, the practical use of finance theory sometimes undermines the market conditions, processes, and patterns of prices that are posited by the theory, we may have found a source of danger that it is easy to ignore or to underestimate if "reality" is conceived of as existing entirely independently of its theoretical depiction.

As the economist and economic policy maker Alan Blinder has pointed out, in many respects global economies have in recent decades moved closer to the standard way in which economists model them, with, for example, its assumption of "single-minded concentration on profit maximization." Blinder suspects that "economists . . . have bent reality (at least somewhat) to fit their models" (2000, pp. 16, 18). The anthropologist Daniel Miller likewise asserts that "economics has the authority to transform the world into its own image" (1998, p. 196).

Whether Blinder and Miller are right is a question this book seeks to answer, at least for one area of economics. The question requires us to examine the strongest level of performativity, despite the methodological difficulties it poses. The reader is warned, however, that there are complexities in the judgment of the correspondence of patterns of prices to models that are only touched on here. This is a study of finance theory and of its relations to markets, not a study of financial econometrics. I have done little more than distinguish those issues about which econometricians seem to agree (for example, the existence after 1987 of the "volatility skew") from those on which there is no clear consensus.

The Book's Goals

If academic pursuits are not to be narrow, they ought to seek to contribute to what Donald (now Deirdre) McCloskey called the conversations of humankind. One such set of conversations, a very old one,[46] is about markets. Those conversations are not always as free-flowing or as civilized as they should be. This is partly because of inequalities of wealth or power and the desire for outcomes economically beneficial to particular sets of participants, but it is also because those who come to those conversations often bring strong, deeply felt preconceptions. Some are convinced that markets are sources of human freedom and prosperity; others believe markets to be damaging generators of alienation, exploitation, and impoverishment. Currently, that divide tends to map onto a disciplinary one, with mainstream economists approving profoundly of markets and with sociologists and anthropologists frequently manifesting deep, albeit often unexplicated, reservations about them.[47]

This book plainly is not economics, although some of it is history of what eventually became one of the most important branches of modern economics. Nor is it economic sociology, at least as traditionally conceived, although it touches on some of that field's concerns. Instead, it is intended in the first instance as a contribution to "social studies of finance."[48] The term has a variety of possible meanings, but one way of describing the underlying enterprise is as drawing on, and developing, the intellectual resources of the social studies of science and technology in order to embark on a conversation about the *technicality* of financial markets. Economic sociology, for example, has been strong in its emphases on matters such as the embedding of markets in cultures, in politics, and in networks of personal interconnections.[49] It has traditionally been less concerned with the systematic forms of knowledge deployed in markets or with their technological infrastructures,[50] yet, if the social studies of science and the history and sociology of technology are right, those too are social matters, and consequential ones.

"We have taken science for realist painting," writes Bruno Latour, "imagining that it made an exact copy of the world. The sciences do something else entirely—paintings too, for that matter. Through successive stages they link us to an aligned, transformed, constructed world." (1999, pp. 78–79) If finance theory is one of Latour's sciences—and this book's conjecture is that it is—then simply to praise it is not to add much to humanity's conversations about markets, and simply to denounce it is to coarsen those conversations. To try to understand how finance theory has "aligned, transformed [and] constructed" its world—which is also everyone's world, the world of investment, savings, pensions, growth, development, wealth, and poverty—may, in contrast, contribute a little to conversations about markets.

Humanity's conversations about markets are not just intellectual; they bear on the question of the appropriate role for markets in our societies. Debates about that role sometimes remind me of debates about technology in the 1960s and the 1970s. Technology was then often taken as either to be adulated or to be condemned, and each of the apparent options frequently involved an implicit view of technological change as following an autonomous logic. The surrounding culture could choose to conform to that logic or to reject its products, but could not modify it fundamentally.

If the history and sociology of technology of the last 25 years have had a single dominant theme, it is that the view of technological change as following an autonomous logic is wrong, and the stark choice between conformity and refusal that it poses is an impoverished one. Technologies can develop in different ways according to circumstances, the design of technical systems can reflect a variety of priorities, and "users" frequently reshape technical systems in important ways. Ultimately, the development and the design of technologies are political matters.[51]

A nuanced and imaginative politics of technology is thus a better option than either uncritical acceptance or downright rejection of technical change. An equivalent approach to markets—one that is more nuanced and more specific than most current ways of thinking about them and of acting in relation to them—is badly needed. I do not claim to provide such an approach (that is a task beyond one book and one author), but my hope for this book is that it helps to begin a conversation with that aim in mind.

Sources

This book takes the form of a series of historical narratives of the development of finance theory and of its interaction with the modern financial markets. Although I touch on what I think would widely be agreed to be the theory's most salient achievements, I have not attempted a comprehensive account of its history. I am even more selective in my discussion of markets, focusing on developments that seem to me to be of particular relevance from the viewpoint of the issues, especially those to do with the performativity of economics, sketched in this chapter. Indeed, the core of the book—chapters 5, 6, and 7—is in a sense a single, extended case study of the development of option theory, of its impact on markets, and of the empirical history of option pricing.

Since relevant, accessible archival material for a book such as this is still sparse, the book's main unpublished source is a set of more than 60 oral-history interviews of the finance theorists and market participants listed in appendix H, and of a number of others who do not wish their names to be disclosed. In the case of the theorists, these interviews complement what can be gleaned from the published literature of their field, and the interviews with practitioners were crucial in helping me to disentangle complex matters such as the impact on markets of option theory or the celebrated debacle of Long-Term Capital Management. I was, however, also fortunate enough to be allowed access to finance theory's most important archive: the papers of Fischer Black, held in the Institute Archives at MIT.

Reasonably comprehensive interview coverage of the most influential finance theorists was possible.[52] Plainly, no such comprehensive coverage is possible in the case of the much larger and more heterogeneous body of people who have played important roles in the development of modern financial markets, even in a limited segment of those markets such as financial derivatives exchanges. My interviewing of market participants was therefore much more ad hoc, and was focused on episodes of specific interest like the emergence of modern derivatives trading in Chicago. These interviews were supplemented by the use of sources such as the trade press and by examination of econometric analyses. In particular, I had the good fortune that the analysis of the prices of options in the Chicago markets has become a locus classicus of modern financial econometrics.

Oral-history interviews have well-known disadvantages. In particular, interviewees' memories of events, especially of specific events long in the past, may be fallible, and they may wish a particular version of events to be accepted. In consequence, I have tried to "triangulate" as much I can, checking one interviewee's testimony against that of others and (where possible) against the published record or econometric analyses of the markets they were describing. In the case of Long-Term Capital Management, for example, I checked for any "exculpatory" bias in insiders' views of the fund's 1998 crisis by interviewing others who had been active in the same markets at the same time. The account of LTCM's crisis presented in chapter 8 was also checked for its consistency with price movements in the relevant markets in 1998 (see MacKenzie 2003b).

I have also had the advantage of having previous historical and sociological work on finance theory and financial markets to build on. Particularly worth singling out is Peter Bernstein's history of finance theory, *Capital Ideas* (1992).[53] Bernstein's emphases differ from mine; for example, he does not address what I call Barnesian performativity, and in regard to the theory's applications he is concerned more with stock-portfolio management than with derivatives markets. However, I owe a great debt to Bernstein, as will future historians of finance theory.

Effectively the only existing sociological analyses of the rise of modern finance theory are those by Richard Whitley (1986a,b). Although I disagree with him in some respects (for example, I think he understates the tension between finance theorists and practitioners), I have been influenced heavily, particularly in chapter 3, by his analysis of the role of changes in the business schools of American universities in creating a favorable environment for the development of the new financial economics.

Overview of the Book

Although finance theory is a mathematical domain, I have kept the book as non-mathematical as possible, banishing equations to endnotes or appendixes. Finance theory's technical terminology cannot be avoided entirely, but I have used it as sparingly as I can, explaining it in the chapters and in a glossary. (The glossary also contains explanations of relevant financial-market terms.) I hope the resultant account will be accessible to readers with no background in economics or in the financial markets, yet not too tediously simplistic for those with such backgrounds.

Chapter 2 describes the shift in the United States in the 1950s and the 1960s from descriptive scholarship in finance to the new analytical, mathematical, economics-based approach. The first of the three strands in my discussion is the work of Franco Modigliani and Merton Miller, whose "irrelevance" propositions were the most explicit challenge to the older approach. The second strand is Harry Markowitz's work on the selection of optimal investment portfolios and its development by William Sharpe into the Capital Asset Pricing Model, finance theory's canonical account of the way stock prices reflect a tradeoff between expected return and risk (in the sense of sensitivity to overall market fluctuations). The third strand is random-walk models of stock-price changes and the eventual culmination of those models in the efficient-market hypothesis. It is easy to imagine that by diligent study one can find patterns in stock-price changes that permit profitable prediction, but random-walk models denied the existence of such patterns. The efficient-market hypothesis generalized this denial into the assertion that prices in mature capital markets always, and effectively instantaneously, take into account all available price-relevant information, including not only the record of previous price changes but also economic information about corporations of the kind that stock analysts pore over. Since all available information is already incorporated into prices, Eugene Fama and other efficient-market theorists argued, it is not possible to make systematic, risk-adjusted, excess profits on the basis of it. Stock prices are moved by new information, but by virtue of being new such information is unpredictable and thus "random."

Chapter 3 broadens the discussion from the specific ideas of finance theory discussed in chapter 2. It discusses how the new finance scholarship developed into the distinct academic subfield of financial economics. (I use the term "financial economics" to include not only finance theory but also efforts to test theories and more general empirical and econometric work on finance.) The chapter also describes the ambivalent and frequently hostile reaction by market practitioners to finance theory. The theory could be drawn on to subject the performance of investment managers to a disconcerting mathematical gaze, and its central tenet—the efficient-market hypothesis—suggested that practitioners' beliefs about markets were often mistaken, that many of their activities were pointless, and that often their advice was of no real benefit to their clients.

Amidst the general hostility, however, there were pockets of practitioners who saw merits in finance theory. Indeed, some found in it ideas with which they could make money—for example, by calculating and selling values of the Capital Asset Pricing Model's most important parameter, beta, which indicates the extent to which the returns on a stock or some other financial asset are sensitive to fluctuations in the market as a whole.

The most significant early practical innovation to emerge from financial economics was the index fund. If, as financial economics suggested, managers' stock selections failed systematically to outperform broad market indices such as the S&P 500, then why not simply invest in the stocks that made up the index in such a way that the performance of one's portfolio would automatically track the level of the index? Such an index fund, its proponents suggested, would perform as well as the portfolios chosen by traditional managers, and it would not be hampered by the high fees those managers charged.

Index funds, first launched in the early 1970s, have become a major feature of modern stock markets. Rooted in financial economics, they can be seen as one way in which that field has been performed in the markets. There is even a Barnesian strand to this performance: the popularity of indexing has made a prediction of the Capital Asset Pricing Model that troubled Sharpe (the prediction that all investors would hold the same portfolio of risky assets, the market itself) less untrue.

Chapter 4 discusses the empirical testing of the strands of finance theory described in chapter 2 and begins to broaden the discussion of performativity. Given the difficulties of econometric testing discussed above, it is not surprising that it proving or disproving the empirical validity of finance theory's models turned out to be difficult. It was hard to construct empirically testable versions of the Modigliani-Miller propositions, and even the apparently directly testable Capital Asset Pricing Model was argued not to be testable at all. Central to the model was the "market portfolio" of all risky assets, but this, the financial economist Richard Roll argued, is not the same as the S&P 500 index or even the entire stock market; its true composition is unknown.

Tests of the efficient-market hypothesis by those who generally supported it led to the identification of "anomalies"—phenomena apparently at variance with the hypothesis—but the "failed" tests frequently led to practical action that had performative effects. The identification of anomalies gave rise to investment strategies to exploit them, and the pursuit of those strategies seems often to have reduced or eliminated the anomalies.

Chapter 4 also describes a path not taken by mainstream finance theory. In what became the standard model of changes in stocks' prices, the statistical distribution of changes in the natural logarithms of stock prices is the normal distribution—the canonical "bell-shaped" curve of statistical theory. This "log-normal" model is an example of what the mathematician and chaos theorist Benoit Mandelbrot calls "mild" randomness: the tails of the normal distribution, representing the probabilities of extreme events, are "thin." In the 1960s, Mandelbrot put forward a different model: one in which the tails are so "fat" that the standard statistical measure of a distribution's spread (the standard deviation or its square, the variance) is infinite.

Mandelbrot's model was of "wild" randomness: periods of limited price fluctuation can be interrupted unpredictably by huge changes. The model initially attracted considerable interest within financial economics (Eugene Fama, in whose work the efficient-market hypothesis crystallized, was the most prominent enthusiast for it), but, as chapter 4 describes, it also met fierce opposition because it undermined standard statistical procedures. In the words of one critic quoted in chapter 4, adopting Mandelbrot's model meant "consigning centuries of work to the ash pile."

Chapter 5 deals with how much options ought to cost, an apparently minor and esoteric problem in finance theory that nevertheless gave rise to a model that some see as "the biggest idea in economics of the century" (Fama interview). In the period in which much of option theory was developed, options were "specialized and relatively unimportant financial securities" (Merton 1973a, p. 141) and were stigmatized by being associated widely with gambling and with market manipulation. However, option pricing seemed a tantalizingly straightforward "normal science" problem, in the terminology of Kuhn (1970). With an established model (the log-normal model) of how stock prices fluctuate, it did not seem too difficult to work out how much an option on that stock should cost. Options—and a particular form of option called a "warrant"—also offered opportunities to perform arbitrage (that is, to make low-risk profits from price discrepancies), and that was another reason for interest in the problem.

Finding a satisfactory solution to the problem of option pricing turned out to be harder than it looked. In chapter 5 the development of the eventually successful solution by Fischer Black and Myron Scholes is contrasted with the work of Edward Thorp, a mathematician famous for showing how to beat the house at blackjack by "card counting" (that is, keeping a careful, systematic mental record of the cards that have been played). Black and Scholes were trying to solve a theoretical problem by applying the Capital Asset Pricing Model; Thorp was working on option pricing without use of the CAPM, and his chief goal was identifying arbitrage opportunities.

The work by Black and Scholes, published in 1973, unleashed a torrent of further theoretical innovation. As they suggested, many other securities that on the surface did not look like options nevertheless had option-like features and so could be valued following the same approach, and, as noted above, the approach was also extended to the analysis of decisions as well as of securities. Among other contributors to option theory were Robert C. Merton and his mentor Paul Samuelson. They believed that the original version of the Capital Asset Pricing Model rested on objectionable assumptions.

Merton developed an approach to option pricing that led to the same equation that Black and Scholes had derived but which rested on different foundations. Merton's derivation did not invoke the Capital Asset Pricing Model, although it too involved assumptions about markets that were markedly at odds with the actual conditions of the early 1970s. Other contributions to option theory quickly followed, and their level of mathematical sophistication rapidly grew. Merton had introduced the use of rigorous stochastic calculus, and by the end of the 1970s the problem of derivatives pricing was reformulated in terms of martingale theory, an advanced area of "pure mathematics." The mathematical repertoire of Wall Street's quantitative finance specialists (first called "rocket scientists," then "quants") was being assembled.[54]

Chapter 6 turns to the Chicago derivatives markets, the most important early site in which option theory was performed. It was above all in Chicago that the apparently quite unrealistic Black-Scholes-Merton model began to gain verisimilitude. (Black and Scholes took on board enough of Merton's derivation to justify the joint attachment all three names to the form eventually taken by the model.) The chapter traces how the Chicago financial derivatives exchanges emerged, how economics was deployed to provide the proposals for these exchanges with legitimacy in the face of suspicion that derivatives were dangerous wagers on price movements, and how the establishment of the new markets required collective action on the part of the memberships of the parent agricultural futures exchanges, the Chicago Mercantile Exchange and the Board of Trade. The chapter emphasizes the intensely bodily experience of trading in Chicago's apparently chaotic open-outcry pits, yet also notes how closely patterns of derivatives prices in those pits came to resemble those posited by the theory.

Was the theory's empirical success performative, and if so in what sense? Did "theory" and "reality" mesh because the former discovered preexisting patterns in the latter, or was "reality" transformed by the performance of theory? What made the Black-Scholes-Merton model, apparently an abstract, unrealistic professors' product, attractive to hard-bitten Chicago floor traders? When first formulated, the Black-Scholes equation was only an approximate fit to patterns of options prices. During the 1970s, however, the fit improved rapidly. Two processes seem to have been involved: market conditions began to change (albeit in many respects slowly) in ways that made the Black-Scholes-Merton model's assumptions more realistic; and, crucially, the model was employed in arbitrage, in particular in an arbitrage called "spreading," in which it was used to identify options that were cheap, or expensive, relative to each other.

Given the above discussion of econometric testing, it is worth remarking that a trader using spreading would have been exploiting—and thus reducing—precisely the discrepancies in options prices that were the focus of the most sophisticated econometric testing of the Black-Scholes-Merton model in this period. (As has been noted, this testing was conducted by the financial economist Mark Rubinstein.) It therefore seems plausible that the use of the model in spreading did more than add generally to its verisimilitude; spreading may have had a direct effect on specific features of price patterns examined in the model's econometric tests.

"Truth" *did* emerge—the fit between the Black-Scholes-Merton model and the Chicago option prices of 1976–1978 was good, by social-science standards, on Rubinstein's tests—but it inhered in the process as a whole; it was not simply a case of correspondence between the model and an unaltered external reality. Knowledge, according to Latour, "does not reside in the face-to-face confrontation of a mind with an object. . . . The word 'reference' designates the quality of the chain in its entirety. . . . Truth-value circulates." (1999, p. 69, emphases removed) The Black-Scholes-Merton model itself became a part of the chain by which its fit to "reality" was secured, or so chapter 6 conjectures.

"I have conceived of a society," writes Barnes, "as a distribution of self-referring knowledge substantially confirmed by the practice it sustains" (1988, p. 166). The Black-Scholes-Merton model informed practices such as spreading, and those practices in their turn helped to create patterns of prices of which the model was a good empirical description. In that sense, the performativity of the model was indeed Barnesian.

The already reasonably close fit between the Black-Scholes-Merton model and Chicago stock-option prices became even better for index options, once such options, and also futures on stock-market indices, were introduced in the early 1980s. However, perhaps the Black-Scholes-Merton model succeeded because it was simply the right way to price options, but market participants learned that only slowly, with their markets only gradually becoming efficient? If that were so, "Barnesian performativity" would be an empty gloss on a process that could better be described in simpler, more conventional terms.

In chapter 7, however, I draw on the econometric literature on option pricing to note that after the 1987 crash the fit between the Black-Scholes-Merton model and patterns of market prices deteriorated markedly. A "volatility skew" or "smile" at odds with Black-Scholes emerged, and it seems to be durable; it has not subsequently diminished or vanished. Option theory has left its permanent imprint on the options markets: the theory is embedded in how participants talk and in technical devices that are essential to their markets. But what is performed in patterns of prices in those markets is no longer classic option-pricing theory.

The volatility skew thus reveals the historicity of economics, at least of the particular form of economics examined here. The U.S. markets priced options one way before 1987 and have priced them differently since, and the change was driven by a historical event: the 1987 crash. Chapter 7 also inquires into the mechanisms of that crash, focusing on the possible role in it of "portfolio insurance," a technique for setting a "floor" below which the value of an investment portfolio will not fall. Since portfolio insurance was an application of option theory, this raises the issue of counterperformativity: perhaps the fit between option theory and reality was ended by an event in which one of its own applications was implicated? Unfortunately from the viewpoint of analytical neatness, however, it seems impossible to determine how large a role portfolio insurance played in exacerbating the crash.

Chapter 8 turns to an episode with echoes of 1987: the 1998 crisis surrounding the hedge fund Long-Term Capital Management (LTCM). Because the partners who ran the fund included the finance-theory Nobel laureates Robert C. Merton and Myron Scholes, its near-failure (it was re-capitalized by a group of the world's leading banks) has often been blamed on blind faith in finance theory's models and has been seen as suggesting fatal flaws in those models. Much of the commentary on LTCM has been dismissive, and some of it has included personal attacks on those who were involved. Indeed, the tone of much of the commentary is an example of the coarsening of the "conversations" about markets referred to above.

Merton, Scholes, and the other partners in LTCM were well aware of the status of finance theory's models as engines of inquiry rather than exact reproductions of markets, and the details of the models used by LTCM were often far less critical to its activities than is commonly imagined. The episode *is* interesting from the viewpoint of the relationship between models and "reality," but not by way of the banal observation that the former are imperfect approximations to the latter.

What is crucial is that LTCM conducted arbitrage, the central mechanism invoked by finance theory. To be sure, there are differences between the "arbitrage" that theory posits and arbitrage as market practice. However, some of what LTCM did was quite close to the paradigmatic arbitrages of finance theory. Aspects of its trading were similar to the arbitrage invoked in Modigliani and Miller's classic proof, and LTCM's option-market activities resembled the arbitrage that imposes Black-Scholes-Merton option pricing.

One way to pursue a better understanding of the relationship between financial models and "reality" is by means of empirical research on arbitrage as market practice, and the case of LTCM is of interest from that viewpoint. A social process—imitation—was at the heart of LTCM's crisis. The hedge fund and its predecessor group of bond-market arbitrageurs, led by LTCM's founder John Meriwether at the investment bank Salomon Brothers, were extremely successful. That success led others to begin similar trading, to devote more capital to it, or (in the case of mortgage-backed securities) even to adjust the models they were using in order to bring them into harmony with the features that they inferred the model being used by Meriwether's group must possess.

The eventual result was what chapter 8 calls a "superportfolio": a large, unstable structure of partially overlapping arbitrage positions. An event that was in itself less than cataclysmic—the Russian government's default on its ruble-denominated bonds on August 17, 1998—caused that superportfolio to begin to unravel. Arbitrageurs who suffered losses in Russia had to begin selling other assets in the superportfolio; in an increasingly illiquid market, those sales caused prices to move sharply against the holders of the superportfolio, forcing further sales; and so on. Finally, in September 1998, LTCM itself became the subject of a self-fulfilling prophecy of failure strikingly similar to the classic example of such a process given in 1948 by Robert K. Merton.

Chapter 9, the book's conclusion, discusses the model-building epistemic culture of finance theory, noting in particular the field's ambivalent attitude to the empirical adequacy of its models, an ambivalence that can be seen not only in what theorists say on the topic but also in the practical actions they take in markets. The chapter draws together the threads of the book's investigation of performativity, focusing especially on the extent to which finance theory brought into being that of which it spoke. A number of broader issues are then discussed, including "behavioral finance," which draws on work in psychology on biases in human decision making to contest orthodox finance's claims of market efficiency. This chapter pays particular attention to arbitrage, which is pivotal both in market practice and in the relations among orthodox finance, behavioral finance, and social studies of finance. The book ends by returning to the analogy between markets and technologies, and to the need for an informed politics of market design analogous to the politics of technology.