

S
cc 10

THE
ALCHEMY
OF
FINANCE

READING THE MIND OF THE MARKET

GEORGE SOROS

NEW PREFACE • FOREWORD BY PAUL TUDOR JONES II



John Wiley & Sons, Inc.

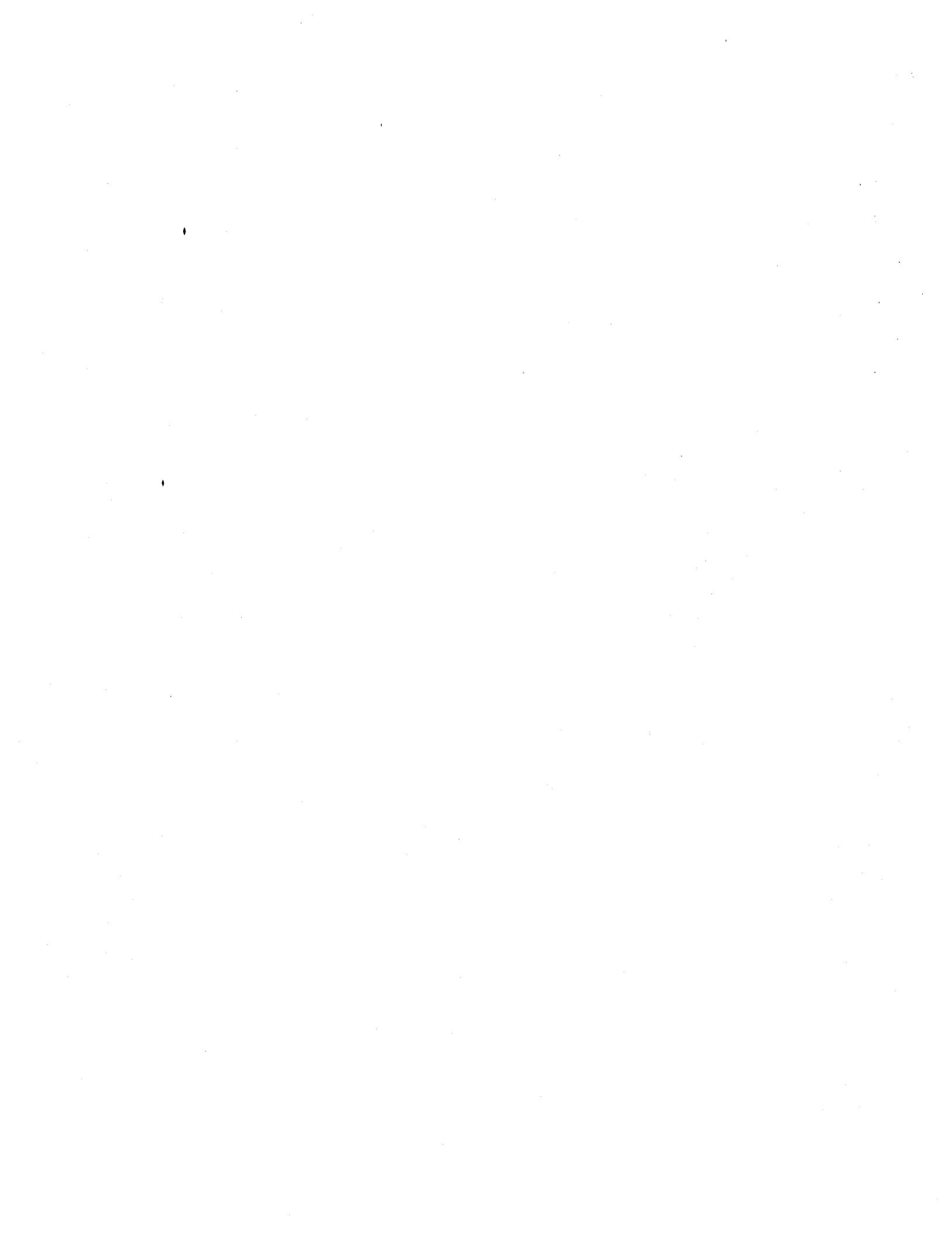
New York • Chichester • Brisbane • Toronto • Singapore

TO SUSAN,
without whom this book would have been ready much sooner

ACKNOWLEDGMENTS

A number of people have read all or part of the manuscript at various stages of its development. They are too numerous to be listed, but I want to thank them all for their help and criticism. Byron Wien, in particular, has gone beyond the call of duty in reading and commenting on, the manuscript at three different stages of development. Special thanks are due to Antonio Foglia, who generated the graphics that illustrate the real-time experiment. Larry Chiarello supplied the figures.

I also want to thank the team that contributed to the performance of Quantum Fund during the experiment: Bill Ehrman, Gary Gladstein, Tom Larkin, Robert Miller, Steven Okin, Joe Orofino, Stephen Plant, Allan Raphael, and Anne Stires.



CONTENTS

Foreword	1
Preface	4
Introduction	11

Part I THEORY

1. <i>The Theory of Reflexivity</i>	27
Anti-equilibrium	27
The Problem of Imperfect Understanding	31
The Problem of the Social Sciences	34
The Participants' Bias	40
The Concept of Reflexivity	41
Reflexivity versus Equilibrium	43
2. <i>Reflexivity in the Stock Market</i>	46
3. <i>Reflexivity in the Currency Market</i>	69
4. <i>The Credit and Regulatory Cycle</i>	81

Part II HISTORICAL PERSPECTIVE

5. <i>The International Debt Problem</i>	91
6. <i>The Collective System of Lending</i>	103
7. <i>Reagan's Imperial Circle</i>	110
8. <i>Evolution of the Banking System</i>	118
9. <i>The "Oligopolarization" of America</i>	128

Appendix**Part III
THE REAL-TIME EXPERIMENT**

10. <i>The Starting Point: August 1985</i>	141
11. <i>Phase 1: August 1985–December 1985</i>	148
12. <i>Control Period: January 1986–July 1986</i>	196
13. <i>Phase 2: July 1986–November 1986</i>	236
14. <i>The Conclusion: November 1986</i>	288

**Part IV
EVALUATION**

15. <i>The Scope for Financial Alchemy: An Evaluation of the Experiment</i>	299
16. <i>The Quandary of the Social Sciences</i>	309

**Part V
PRESCRIPTION**

17. <i>Free Markets Versus Regulation</i>	317
18. <i>Toward an International Central Bank</i>	324
Exchange Rates 328	
International Debt 329	
Oil 333	
An International Currency 338	
19. <i>The Paradox of Systemic Reform</i>	341
20. <i>The Crash of '87</i>	345

Epilogue 361

Notes 366

Appendix 368

FOREWORD

Four hundred seventy-three million to one. Those are the odds against George Soros compiling the investment record he did as the manager of the Quantum Fund from 1968 through 1993. His investing record is the most unimpeachable refutation of the random walk hypothesis ever!

As a trader coming of age in the latter half of the frenetic 1970s and the 1980s, *The Alchemy of Finance* was somewhat of a revolutionary book. Remember, this was the period when trend following and indexation were the vogue in investing. It was a time when technical analysis (the study of price movement as a forecasting tool) reached its zenith. Traders of my generation armed themselves with charts and computer-generated graphics that predicted future price direction. We sat day after day in front of screens, mesmerized by blinking lights and everchanging numbers in a deafening cacophony of information overload. With the possible exception of Elliott Wave Theory, an intellectual framework for understanding the course of social, political, and economic events was noticeably forgotten in favor of just making sure that one was part of the ever-quickenning process.

The Alchemy of Finance was a shot out of the dark for me. It let me take a giant step forward by first taking a step backwards, clarifying events that appeared so complex and so overwhelming. During an era when so much money was made in larger than life events, from the Hunt brothers' squeeze of the silver market in 1979 to KKR's takeover of RJR Nabisco in 1989, Mr. Soros's theory of reflexivity is the first modern, nontechnical effort to describe

and forecast the dynamic interplay between the participants in the process. That is the brilliance of this book. It describes the dynamics of the path between points of extreme valuation and equilibrium in the marketplace. This is particularly important for the average investor. How many times have we been correctly long near the bottom or short near the top of a major market move? But our staying power with these positions has been weak (as well as our returns) because of a lack of understanding of the path of big price moves. Soros gives us critical insight into that path and thus more confidence in our investments. This constitutes 70% of any successful investing campaign.

When I enter the inevitable losing streak that befalls every investor, I pick up *The Alchemy* and revisit Mr. Soros's campaigns. Studying how he coped with adversity provides an excellent tutorial for breaking the string of negative behaviors that occasionally besets any investor. Winning is infectious. And this book is replete with examples of trading behaviors all would want to emulate. Importantly, Mr. Soros's intellect gives him the confidence and strength of his own convictions to stay with his positions even during trying periods. In that sense, *The Alchemy* joins Edwin Lefèvre's *Reminiscences of a Stock Operator* as a timeless instructional guide of the marketplace. And as such, Soros should beware! In the World War II movie *Patton*, my favorite scene is when U.S. General George S. Patton has just spent weeks studying the writing of his Germany adversary Field Marshall Erwin Rommel and is crushing him in an epic tank battle in Tunisia. Patton, sensing victory as he peers onto the battle field from his command post, growls, "Rommel, you magnificent bastard. I read your book!" Enough said.

The Alchemy is also an excellent economic and political history of recent times. From unknowingly providing a blueprint as to how the savings and loan fiasco in the United States would be resolved six years in advance (page 124) to predicting the stock market crash of 1987 two years in advance (page 181), Soros reveals himself as the great market visionary of our time.

History will probably remember Mr. Soros as the speculator who tilted against the Bank of England in 1992 (and freed the English people from recession). His billion dollar score is simply too compelling a story for scribes to overlook. Mr. Soros himself would probably like to be remembered as a great economist or even scientist. But I am going to remember him for something even more important, for which he does not receive the credit he deserves. He is

someone who genuinely cares about the state of the human condition and tries to better it. His myriad and monumental philanthropical efforts will qualify him as one of history's great benefactors. Even today at age 62, he pursues the activities of his six foundations with the vigor and work ethic of a young turk on the way up the financial ladder, working 18-hour days around the globe on behalf of his causes. He does not just write checks, which any wealthy person can do. He is a hands-on workaholic who materially impacts the quality of the lives of people less fortunate than he. Now this, this is a sign of greatness.

Paul Tudor Jones II

PREFACE

Seven eventful years have passed since *The Alchemy of Finance* was first published. My investment fund, the Quantum Fund, has continued to flourish: Shareholders have enjoyed average annual gains of 35.8% in the last seven years in spite of a setback in the crash of 1987. Quantum Fund has also spawned a number of offspring, some of which are doing even better than the goose that is laying the golden eggs. Starting in 1989, we decided to distribute a portion of our earnings to shareholders, either in cash or in shares of the newly created funds. As a result, we now manage seven funds with combined equity of over \$10 billion.

I have become progressively less active in the management of the funds. I was fortunate in meeting Stanley Druckenmiller through *The Alchemy of Finance*. He was managing another fund at the time, and he sought me out because he was intrigued by my book. We started talking and, eventually, he joined my firm. At the beginning, he found it difficult to work with me. Although I gave him a great deal of authority, he was inhibited by my presence and felt that he was not doing as well as he had before joining my firm. Fortunately, I was becoming increasingly involved in the revolutionary process that led to the collapse of communism. I was establishing a network of foundations throughout the communist world and it involved travelling in places where communications were rather poor. In the summer of 1989, I told Stan that he must take full charge of running the Fund. Since then we have had no difficulties.

I became the coach, and he became the competitor. Our performance improved and we embarked on a period of sustained growth. In each of the last three years, we chalked up gains in excess of

50%. Although we have had two similar periods of prosperity previously, this qualifies as an exceptional performance in view of our outlandish size. Druckenmiller is not only a good fund manager, he is also a good partner. Under his leadership, we have been able to enlarge and improve our management team so that it now has a depth which it never had before. So it happened that I found the reward for my philanthropic activities in the prosperity of my business. That prosperity enabled me to expand the foundation network at a breakneck speed.

My participation in the collapse of communism is a different story which has to be told in a different place. In fact, I have already written two books on the subject, *Opening the Soviet System* in 1990, and *Underwriting Democracy* in 1991. The point that needs to be made here is that I was guided by exactly the same philosophy in my philanthropic activities in Eastern Europe as in the financial markets. As the reader will learn, I treat developments in financial markets as a historical process. That makes my theory eminently applicable to a historical process such as the collapse of communism. I did apply my theory and on the whole it enabled me to anticipate events better than most people. As I discovered, there is a great deal of similarity between a boom-bust process in the financial markets and the rise and fall of the Soviet system.

It is ironic that I became famous, not because of my activities in Eastern Europe, but because of the profit we made on sterling when Britain left the Exchange Rate Mechanism on September 16, 1992. I became an instant celebrity, first in Britain, then in the rest of the world. When it became known that the Quantum group of funds had bought a large block of Newmont Mines, the price of gold soared. Although I expressed no opinion on gold, all kinds of opinions were attributed to me. I made some attempts to rebut them, but to no avail. Although I had not sought guru status, I could not ignore it when it was thrust upon me. In fact, I welcomed it because I thought that it would be useful in having my voice heard on political issues. But that was not as simple as it seemed. When I said that the Bundesbank's high interest rate policy was becoming counterproductive, the markets responded by pushing down the German Mark. But when I inveighed against European policy on Bosnia, I was either ignored or told to stick to the field of my expertise. I fared particularly poorly in France, where I refrained from speculating against the franc because I did not want to be responsible for the collapse of what remained of the European Exchange Rate

Mechanism, but I was blamed for it anyhow. The French government resented my advice even more than it would have resented my speculative activities. It goes to show that speculators ought to speculate and keep their mouths shut.

My notoriety as a financial guru has created a tremendous demand for *The Alchemy of Finance*, hence this new edition. I must confess that my thinking has evolved a great deal since I wrote this book, but I have been concerned mainly with historical processes, not with financial ones. I cannot summarize my ideas in this preface—I need to write another book. I intend to do so as soon as time permits, but there is one important theoretical point I need to make in order to bring this book in line with my current thinking.

In *The Alchemy of Finance*, I put forward the theory of reflexivity as if it were relevant at all times. That is true in the sense that the two-way feedback mechanism that is the hallmark of reflexivity can come into play at any time, but it is not true in the sense that it is at play at all times. In fact, in most situations it is so feeble that it can be safely ignored. We may distinguish between near-equilibrium conditions where certain corrective mechanisms prevent perceptions and reality from drifting too far apart, and far-from-equilibrium conditions where a reflexive double-feedback mechanism is at work and there is no tendency for perceptions and reality to come close together without a significant change in the prevailing conditions, a change of regime. In the first case, classical economic theory applies and the divergence between perceptions and reality can be ignored as mere noise. In the second case, the theory of equilibrium becomes irrelevant and we are confronted with a one-directional historical process where changes in both perceptions and reality are irreversible. It is important to distinguish between these two different states of affairs because what is normal in one is abnormal in the other.

The idea of a distinction between near-equilibrium and far-from-equilibrium conditions is present in *The Alchemy of Finance*. At the end of Chapter 1, I distinguish between humdrum and historical change but I underestimate the importance of the distinction. I call it “tautological.” I now consider this a mistake. The tautology arises only because I do not probe deeply enough and cover up with a tautology what is a fundamental difference in the structure of events.

In most phenomena investigated by scientific method, one set of conditions follows another irrespective of what anybody thinks about them. The phenomena studied by social sciences, which

include the financial markets, have thinking participants and this complicates matters. As I have tried to show, the participants' views are inherently biased. Instead of a direct line leading from one set of conditions to the next one, there is a constant criss-crossing between the objective, observable conditions and the participant's observations and vice versa: participants base their decisions not on objective conditions but on their interpretation of those conditions. This is an important point and it has far-reaching consequences. It introduces an element of indeterminacy which renders the subject matter less amenable to the kind of generalizations, predictions, and explanations that have given natural science its reputation. Exactly because it is so disruptive, the social sciences in general and economic theory in particular have done their best to eliminate or to ignore the element of indeterminacy. I have taken issue with that endeavor and tried to develop an alternative approach which takes the participants' bias as its starting point.

In retrospect, I may have overstated my case. There are many situations that can be fruitfully studied by taking the participants' bias as given and ignoring the element of indeterminacy which it may generate. It is only in certain respects and in certain special circumstances that the indeterminacy becomes significant. It comes into play when expectations about the future have a bearing on present behavior—which is the case in financial markets. But even there, some mechanism must be triggered for the participants' bias to affect not only market prices but the so-called fundamentals which are supposed to determine market prices. Apparently I have failed to make this point sufficiently clear. The message of my book is usually summed up by saying that the participants' value judgments are always biased and the prevailing bias affects market prices. If that is all I had to say it would be hardly worth writing a book about it. My point is that there are occasions when the bias affects not only market prices but also the so-called fundamentals. This is when reflexivity becomes important. It does not happen all the time but when it does, market prices follow a different pattern. They also play a different role: they do not merely reflect the so-called fundamentals; they themselves become one of the fundamentals which shape the evolution of prices. This recursive relationship renders the evolution of prices indeterminate and the so-called equilibrium price irrelevant.

Nobody would deny that individual participants operate with biased views; but the prevailing wisdom holds that the participants'

bias can be dismissed as temporary aberrations, so-called random walks. That is the point on which I disagree. I now believe this point can be more effectively made by drawing a distinction between near-equilibrium and far-from-equilibrium conditions than by proposing a general theory of history based on the constant cross-crossing between perceptions and reality as I have done in *The Alchemy of Finance*. That does not mean that there is anything wrong with the general theory; it means only that the concept of reflexivity becomes more significant if it is reserved for those cases where the double feedback mechanism is actually at work.

The Alchemy of Finance is devoted to the study of such cases. The most obvious example is equity leveraging where a temporary overvaluation of shares is converted into per-share earnings through the issue of shares at inflated prices. In most of the cases discussed, the participants' bias involves an actual error in their thinking. For instance, in the late 1970s international bankers lent too much money to developing countries because they failed to recognize that the so-called debt ratios they used to measure the creditworthiness of the borrowing countries were reflexive in the sense that they were affected by their own lending activity. But it is not necessary for the bias to involve an actual error. As I show in Chapter 3, a freely fluctuating exchange rate system is inherently unstable because of the influence of trend-following speculation, yet speculators follow the correct strategy by following the trend.

Judging by the public reaction—which consists mainly of comments by journalists who read the book superficially or not at all—I have not been successful in demonstrating the significance of reflexivity. Only the first part of my argument—that the prevailing bias affects market prices—seems to have registered. The second part—that the prevailing bias can in certain circumstances also affect the so-called fundamentals and changes in market prices cause changes in market prices—seems to have gone unnoticed.

The fault is at least partially mine. Since reflexivity changes the structure of events, I have tried to put forward a reflexive structure as the universally valid way of looking at the evolution of market prices—a kind of general theory à la Keynes in which the absence of reflexivity constitutes a special case. It would have been better to present reflexivity as the special case because what endows reflexivity with significance is the fact that it operates intermittently.

Once the significance of reflexivity has sunk in and the inadequacy of the prevailing wisdom has been recognized, the time would have been ripe for proposing a general theory of reflexivity.

I have my excuses. I did not observe reflexivity in financial markets but developed reflexivity as an abstract philosophical concept before I entered the financial markets. In other words, I failed as a philosophical speculator before I succeeded as a financial one. Apparently, my failure as a philosopher carried over into my book because I did not make the concept of reflexivity—which can be observed and converted into profit—as clear as it could be. When one discovers something new, one has an understandable inclination to exaggerate its importance. This is what I did with reflexivity. By proposing a general theory of reflexivity, I may have gone too far too soon. I claimed that economic theory is false and social science is a false metaphor. These are exaggerated claims. Since far-from-equilibrium conditions arise only intermittently, economic theory is only intermittently false. And the dividing line between natural and social science is not quite as hard and fast as I made it appear when I wrote the book. These qualifications render reflexivity more rather than less significant.

Once the concept of reflexivity is established, the range of its applicability seems to widen. It is possible to treat the evolution of prices in all financial markets taken together as a reflexive, historical process. I have done so in *The Alchemy of Finance* when I analyzed Reagan's "Imperial Circle," and I have found other examples since the book was published, such as the German Imperial Circle after the fall of the Berlin Wall. (See appendix: "The Prospect of European Disintegration.") But there is a danger in pushing the concept of reflexivity too far, as I have learned at my own expense. There are long fallow periods when the movements in financial markets do not seem to follow a reflexive tune but rather resemble the random walks mandated by the efficient market theory. In these circumstances, it is better to do nothing than to pursue a reflexive hypothesis.

Treating reflexivity as an intermittent phenomenon rather than as a universally valid condition opens up fertile fields for investigation. For instance, the question poses itself: How can near- and far-from-equilibrium conditions be distinguished from each other? What is the criterion of demarcation? I have done a lot of thinking on that question and I have the beginnings of an answer. Whether

I can formulate it properly remains to be seen in my next book. It revolves around a question of values and it is relevant for society in general, not only for financial markets. My next book, if it is ever written, will be a theory of history, not a theory of finance. I am providing an example of how the boom-bust pattern of financial markets can be applied to larger historical processes in the appendix where I reproduce a lecture I delivered on September 29, 1993, entitled "Prospect for European Disintegration."

INTRODUCTION

In a very real sense, this book is my life's work. It touches on many of my most abiding interests and it brings together the two main strands in my intellectual development: one abstract and one practical.

The abstract came first. Ever since I became conscious of my existence I have had a passionate interest in understanding it, and I regarded my own understanding as the central problem that needed to be understood. To understand oneself—gnote auctor; nosce te ipsum—is an impossible task. To achieve anything resembling knowledge we must be able to draw a distinction between subject and object; yet in this case the two are the same. What one thinks is part of what one thinks about; therefore, one's thinking lacks an independent point of reference by which it can be judged—it lacks objectivity.

As an undergraduate I studied economics, but I found economic theory highly unsatisfactory because it failed to come to grips with this problem; indeed, it went through great contortions to avoid it. Economics seeks to be a science. Science is supposed to be objective and it is difficult to be scientific when the subject matter, the participant in the economic process, lacks objectivity.

I was greatly influenced at the time by Karl Popper's ideas on scientific method. I accepted most of his views, with one major exception. He argued in favor of what he called "unity of method"¹—that is, the methods and criteria that apply to the study of natural phenomena also apply to the study of social events. I felt that there was a fundamental difference between the

two: the events studied by the social sciences have thinking participants; natural phenomena do not. The participants' thinking creates problems that have no counterpart in natural science. The closest analogy is in quantum physics, where scientific observation gives rise to Heisenberg's uncertainty principle; but in social events it is the participants' thinking that is responsible for the element of uncertainty, not the outside observer.

Natural science studies events that consist of a sequence of facts. When events have thinking participants, the subject matter is no longer confined to facts but also includes the participants' perceptions. The chain of causation does not lead directly from fact to fact but from fact to perception and from perception to fact. This would not create any insuperable difficulties if there were some kind of correspondence or equivalence between facts and perceptions. Unfortunately, that is impossible because the participants' perceptions do not relate to facts, but to a situation that is contingent on their own perceptions and therefore cannot be treated as a fact.

Economic theory tries to sidestep the issue by introducing the assumption of rational behavior. People are assumed to act by choosing the best of the available alternatives, but somehow the distinction between perceived alternatives and facts is assumed away. The result is a theoretical construction of great elegance that resembles natural science but does not resemble reality. It relates to an ideal world in which participants act on the basis of perfect knowledge and it produces a theoretical equilibrium in which the allocation of resources is at an optimum. It has little relevance to the real world in which people act on the basis of imperfect understanding and equilibrium is beyond reach.

The relationship between the participants' understanding and the situation in which they participate continued to preoccupy me long after I left college. My first priority was to try and make a living but in my spare time I wrote a philosophical treatise on the subject with the catchy title "The Burden of Consciousness." Unfortunately, the title was the best part of it. By the time I finished, I disagreed with my own presentation. I spent three years revising it. One day I reread what I had written the day before and I could not make head or tail of it. It made me realize that I had reached a dead end, and I decided to give it up. That was when the practical streak in me began to dominate my intellectual development.

If I had to sum up my practical skills, I would use one word: survival. When I was an adolescent, the Second World War gave

me a lesson that I have never forgotten. I was fortunate enough to have a father who was highly skilled in the art of survival, having lived through the Russian revolution as an escaped prisoner of war. Under his tutelage the Second World War served as an advanced course at a tender age. As the reader shall see, the investment vehicle I created a quarter of a century later drew heavily on skills I learned as an adolescent.

After leaving college I had a number of false starts and finally became an international arbitrage trader in stocks, first in London and then in New York. When the European Common Market was formed in 1957, American investors became interested in European shares. I became a security analyst advising American institutions on their European investments and for a brief period I ruled as a one-eyed king among the blind. My glory came to an abrupt end when President Kennedy introduced a so-called interest equalization tax which effectively stopped purchases of foreign securities. I decided to put my money-making activities on the back burner and spent three years, from 1963 to 1966, revising "The Burden of Consciousness."

When I finally decided to return to the land of the living I started a model portfolio that became a hedge fund (a mutual fund that employs leverage and uses various techniques of hedging) in 1969. I have been in charge of that fund ever since, although I delegated much of the responsibility to others between September 1981 and September 1984. The fund has grown from about \$4 million at inception to nearly \$2 billion and most of the growth has been internally generated. Original investors have seen the value of their shares multiply 300-fold. No investment fund has ever produced comparable results.

In the first ten years of my business career I had not much use for anything I had learned in college and there was an almost total separation between my practical activities and my theoretical interests. Selling and trading in securities was a game I played without putting my true self on the line.

All this changed when I became a fund manager. I was putting my money where my mouth was and I could not afford to dissociate myself from my investment decisions. I had to use all my intellectual resources and I discovered, to my great surprise and gratification, that my abstract ideas came in very handy. It would be an exaggeration to say that they accounted for my success; but there can be no doubt that they gave me an edge.

I developed my own peculiar approach to investing, which was

at loggerheads with the prevailing wisdom. The generally accepted view is that markets are always right—that is, market prices tend to discount future developments accurately even when it is unclear what those developments are. I start with the opposite point of view. I believe that market prices are always wrong in the sense that they present a biased view of the future. But distortion works in both directions: not only do market participants operate with a bias, but their bias can also influence the course of events. This may create the impression that markets anticipate future developments accurately, but in fact it is not present expectations that correspond to future events but future events that are shaped by present expectations. The participants' perceptions are inherently flawed, and there is a two-way connection between flawed perceptions and the actual course of events, which results in a lack of correspondence between the two. I call this two-way connection "reflexivity."

In the course of my investment activities, I discovered that financial markets operate on a principle that is somehow akin to scientific method. Making an investment decision is like formulating a scientific hypothesis and submitting it to a practical test. The main difference is that the hypothesis that underlies an investment decision is intended to make money and not to establish a universally valid generalization. Both activities involve significant risk, and success brings a corresponding reward—monetary in one case and scientific in the other. Taking this view, it is possible to see financial markets as a laboratory for testing hypotheses, albeit not strictly scientific ones. The truth is, successful investing is a kind of alchemy.

Most market participants do not view markets in this light. That means that they do not know what hypotheses are being tested; it also means that most of the hypotheses that are submitted to market testing are quite banal. Usually they amount to nothing more than the assertion that a particular stock is going to outperform the market averages.

I had a certain advantage over other investors because at least I had an idea about the way financial markets operate. I would be lying, however, if I claimed that I could always formulate worthwhile hypotheses on the basis of my theoretical framework. Sometimes there were no reflexive processes to be found; sometimes I failed to find them; and, what was the most painful of all, sometimes I got them wrong. One way or another, I often invested

without a worthwhile hypothesis and my activities were not very different from a random walk. But I was attuned to reflexive processes in financial markets and my major successes came from exploiting the opportunities they presented.

My approach to the market was not as abstract as it sounds. It took an intensely personal, emotional form: testing was closely associated with pain and success with relief. When I asserted that "markets are always biased" I was giving expression to a deeply felt attitude: I had a very low regard for the sagacity of professional investors and the more influential their position the less I considered them capable of making the right decisions. My partner and I took a malicious pleasure in making money by selling short stocks that were institutional favorites. But we differed in our attitudes to our own activities. He regarded only the other participants' views as flawed, while I thought that we had as good a chance of being wrong as anyone else. The assumption of inherently flawed perceptions suited my self-critical attitude.

Operating a hedge fund utilized my training in survival to the fullest. Using leverage can produce superior results when the going is good, but it can wipe you out when events fail to conform to your expectations. One of the hardest things to judge is what level of risk is safe. There are no universally valid yardsticks: each situation needs to be judged on its own merit. In the final analysis you must rely on your instincts for survival. Thus my engagement in running a hedge fund brought together both my abstract interests and my practical skills.

I did not play the financial markets according to a particular set of rules; I was always more interested in understanding the changes that occur in the rules of the game. I started with hypotheses relating to individual companies; with the passage of time my interests veered increasingly toward macroeconomic themes. This was due partly to the growth of the fund and partly to the growing instability of the macroeconomic environment. For instance, exchange rates were fixed until 1973; subsequently, they became a fertile field for speculation.

For the past four or five years I have had a growing sense of impending financial disaster. I felt that a long-lasting expansionary cycle was becoming increasingly unsound and unsustainable and we were getting ready for a bust. That was one of the reasons I distanced myself from the active management of the fund in 1981 and reduced its overall level of exposure. My interest shifted

from my own survival to the survival of the system. I made a study of the international debt problem and published some papers on the subject. I used the same theoretical framework as in my investment activities and my analysis was not without merit. Unfortunately, the more complex the system, the greater the room for error. I made some mistakes in my analysis that detracted from the accuracy of my predictions; they also had a detrimental effect on my investment results until I revised my views in the course of writing this book.

The more successful I had been in applying my ideas in financial markets, the keener I became to express them in theoretical form. I continued to cherish the fantasy that the concept of reflexivity constitutes a major contribution to our understanding of the world in which we live. I believed that the participants' bias is the key to an understanding of all historical processes that have thinking participants, just as genetic mutation is the key to biological evolution. But a satisfactory formulation of the theory of reflexivity continued to elude me. I always ran into trouble when I tried to define what I meant by the imperfect understanding of the participants. To speak accurately of a distortion we must know what the situation would be if it were not distorted by the participants' perceptions. Unfortunately that does not seem possible because the participants' thinking is an integral part of the situation they have to think about. It is not surprising that the concept of reflexivity should present extreme difficulties; if it were an easier concept to work with, economists and other social scientists would not have gone to such lengths to banish it from their subject matter.

This book is a final attempt to explore the implications of reflexivity. I have tried to circumvent the difficulties I encountered in the past by approaching the subject from the opposite direction. Instead of getting bogged down in abstract theory, I am going to draw on my experimental, practical findings to the greatest possible extent. I cannot avoid an abstract discussion altogether, but I have confined it to a single chapter. In exploring the practical implications, I start with the simplest cases and gradually lead up to more complex ones. This approach happens to coincide with the historical order in which I encountered reflexive developments in practice: first the stock market, then the currency market, then the international debt problem, and finally what may be called the credit cycle.

The stock market provides some pure examples of a boom and bust pattern; freely floating currency rates allow me to explore well-formed wave patterns. The boom and bust in international bank lending is part of a more complex, historical process of credit expansion and eventual credit contraction. It has given rise to the configuration that I have dubbed "Reagan's Imperial Circle." The configuration prevailed from the international debt crisis of 1982 until the first half of 1985 but it was inherently unstable. How the instability will be resolved is one of the main questions considered in this book.

The experimental approach has borne unexpected results. I made two major discoveries in the course of writing: one is a reflexive connection between credit and collateral; the other is a reflexive relationship between the regulators and the economy they regulate.

It has long been assumed that monetary values are a passive reflection of the state of affairs in the real world. Classical economics focused on the real world and neglected the problems connected with money and credit; even Keynes couched his general theory in real terms. Monetarists sought to stand the relationship on its head: they argue that it is possible to control inflation by controlling the growth of the money supply.

In my opinion, all these views are based on a fundamental misconception. Money values do not simply mirror the state of affairs in the real world; valuation is a positive act that makes an impact on the course of events. Monetary and real phenomena are connected in a reflexive fashion; that is, they influence each other mutually. The reflexive relationship manifests itself most clearly in the use and abuse of credit.

Loans are based on the lender's estimation of the borrower's ability to service his debt. The valuation of the collateral is supposed to be independent of the act of lending; but in actual fact the act of lending can affect the value of the collateral. This is true of the individual case and of the economy as a whole. Credit expansion stimulates the economy and enhances collateral values; the repayment or contraction of credit has a depressing influence both on the economy and on the valuation of the collateral. The connection between credit and economic activity is anything but constant—for instance, credit for building a new factory has quite a different effect from credit for a leveraged buyout. This makes it difficult to quantify the connection between credit and

economic activity. Yet it is a mistake to ignore it. The monetarist school has done so, with disastrous consequences.

The reflexive interaction between the act of lending and collateral values has led me to postulate a pattern in which a period of gradual, slowly accelerating credit expansion is followed by a short period of credit contraction—the classic sequence of boom and bust. The bust is compressed in time because the attempt to liquidate loans causes a sudden implosion of collateral values.

Economic history has been punctuated by booms and busts. Nevertheless, the concept of a credit cycle is too simplistic to explain the course of events. For one thing, the connection between credit and economic activity is too tenuous and variable to yield a regular pattern. For another, the sequence of events is greatly complicated by the influence of economic policy. Periodic busts have been so devastating that strenuous efforts have been made to prevent them. These efforts have led to the evolution of central banking and of other mechanisms for controlling credit and regulating economic activity.

To understand the role of the regulators it must be realized that they are also participants: their understanding is inherently imperfect and their actions have unintended consequences. The relationship between the regulators and the economy is reflexive and it also exhibits cyclical characteristics in the sense that it tends to swing from one extreme to another.

What is the connection between the regulatory cycle and the credit cycle? At this point, my views become very tentative. I believe that the two cycles broadly overlap in time, that the minimum of regulations tends to coincide with the maximum of credit expansion and vice versa. But within this chronological coincidence there is constant interaction between the two cycles that influences the shape and duration of both. The interaction between the two cycles yields a unique path that cannot be fitted into any regular or repetitive pattern.

I have tried to apply this complex and tentative framework to an interpretation of recent economic and financial history. Needless to say, a great many factors come into play; but my focus is on the twin cycles in credit and regulation. The main topics I deal with are the transformation of banking from a highly regulated to a less regulated industry, the boom and bust in international lending, mergermania, and international capital movements.

Until 1982, the story is a fairly straightforward case of boom

and bust, but after 1982 the situation gets very complicated. If events had been allowed to take their course, the uncontrolled credit expansion of the 1970s would have come to an unhappy end; but exactly because the consequences would have been so disastrous, the financial authorities came to the rescue and successfully avoided a bust. Since then, we have been passing through uncharted waters. The great boom exhausted itself some time ago but its life span has been extended by artificial means in order to avoid a great bust.

I try to trace the unique path that events have taken: the preservation of the accumulated burden of bad international debt through the formation of what I call the "Collective" system of lending and the emergence of the United States government as the "borrower of last resort." Both of these are unprecedented developments. They gave rise to this strange constellation that I have called the Imperial Circle: a benign circle at the center and a vicious circle at the periphery of a worldwide system based on a strong dollar, a strong U.S. economy, a growing budget deficit, a growing trade deficit, and high real interest rates. The Imperial Circle held the international economic and financial system together but it was inherently unstable because the strong dollar and high real interest rates were bound to outweigh the stimulating effect of the budget deficit and weaken the U.S. economy. The Imperial Circle could not last indefinitely. What would happen next?

To answer that question, I conducted an experiment from August 1985 onward. In effect, I kept a diary in which I recorded the thoughts that went into my investment decisions on a real-time basis. Since I considered the future of the Imperial Circle of paramount importance, the experiment served as a test of my ability to predict the future course of events, using the framework developed in the book. The experiment was a roaring success in financial terms—my fund never did better. It also produced a surprising result: I came out of the experiment with quite different expectations about the future.

I started with the presumption that the benign circle was in danger of reversing itself: a weak dollar and a weak economy would combine to keep interest rates higher than they ought to be, and without any scope for further monetary or fiscal stimulus the decline of both the economy and the dollar would become irreversible. But the situation was once again saved by the inter-

vention of the monetary authorities. By changing from a system of freely floating exchange rates to a "dirty float," the decline of the dollar was cushioned, and, with the help of lower interest rates and booming financial markets, the economy was prevented from slipping into recession. We entered a new phase which I describe, with only a modicum of irony, as the "Golden Age of Capitalism."

It can be seen that this book seeks to accomplish a number of things at the same time. It propounds not just one general theory—the theory of reflexivity—but also another specific theory, that of a credit-cum-regulatory cycle. The latter idea is so tentative that it hardly qualifies as a theory. Yet I try to apply it not only to explain contemporary history but also to predict the outcome, and in the course of doing so I provide a practical demonstration of the fundamental difference between explaining and predicting reflexive phenomena. I also try to draw some general conclusions from the analysis. The most important ones are, first, that it is credit that matters, not money (in other words, monetarism is a false ideology), and, second, that the concept of a general equilibrium has no relevance to the real world (in other words, classical economics is an exercise in futility). Financial markets are inherently unstable; that leads to a third conclusion that is better stated as a question than an assertion: what policy measures are needed to reestablish the stability of our economic system?

The book would be easier to read if it tried to make just one point at a time. Unfortunately, that is not possible, because the various points are interconnected. If the theory of reflexivity were widely known, I could have taken it for granted and concentrated on an exploration of the credit and regulatory cycles. Similarly, if the fact that financial markets are inherently unstable were generally recognized, I need not have spent so much time on demonstrating that the concept of equilibrium as used in economic theory is totally unrealistic. As it is, I had to try to make several points more or less simultaneously.

To make matters worse, the book does not qualify as a finished product. When I started writing it, I thought I had a theory of reflexivity to present and my difficulties related only to its presentation. As I tried to apply the theory to various situations, I discovered that I do not actually have a well-formed theory. The idea that the participants' biases play an important causal role in historical events is both valid and interesting, but it is too general to

qualify as a theory that can help to explain and predict the course of events. The boom/bust pattern I have established applies to some developments but not to others. To try to fit every initially self-reinforcing and eventually self-defeating development into its mold can give rise to serious distortions. I feel like the early astronomers who tried to describe the elliptical paths of the planets in terms of circles and semicircles; the only difference is that the path of reflexive events is irregular to start with.

My fantasy was to present a general theory of reflexivity that would explain the great bust of the 1980s in the same way that Keynes's *General Theory of Employment, Interest and Money* explained the Great Depression of the 1930s. As it turns out, we do not have a great bust and I do not have a general theory. What I have is an approach that can help to illuminate the present precarious state of our financial system. It cannot explain and predict the course of events in the manner to which we have become accustomed during our long love affair with natural science for the simple reason that reflexive processes cannot be explained and predicted in that manner. A different approach is needed and this book is an attempt to develop one. It is best regarded as part of a process of discovery rather than the finished product.

All this makes for a difficult, dense book, although I can promise the reader that nothing in the rest of the book is quite as dense as this introduction. I explore a complex subject. I bring a complex mind to bear on it. I can argue in my defense that the complexity of my thinking mirrors the complexity of the financial markets rather well, as demonstrated by the financial outcome of the real-time experiment. There is, therefore, at least a *prima facie* case for giving me a hearing. I shall try not to abuse the privilege.

It may be helpful if I sketch out the structure of the book. Part I propounds the theory. The first chapter deals with the concept of reflexivity in general terms and explores the difficulties in understanding reflexive phenomena. In particular, it argues that the symmetry between explanation and prediction that characterizes the laws of natural science is not attainable. The next three chapters apply the theory to the financial markets: Chapter 2 to the stock market and Chapter 3 to the currency market, with Chapter 4 outlining a credit and regulatory cycle.

Part II seeks to explain contemporary economic and financial history using the hypotheses outlined in Chapter 4. The history

is, of necessity, selective, concentrating on those developments that are relevant to the concept of a credit and regulatory cycle. My main topics are banking, international lending, and merger-mania.

Part III consists of a real-time experiment which is both test and prediction at the same time. As a test it does not qualify as a scientific one by the standards of natural science; but it may serve as an example of how theories about reflexive developments can be tested.

Part IV evaluates the results of the real-time experiment. Chapter 15 explores the scope for what I provocatively call alchemy. The real-time experiment can be regarded as an exercise in financial alchemy. Chapter 16 examines the limitations of social science.

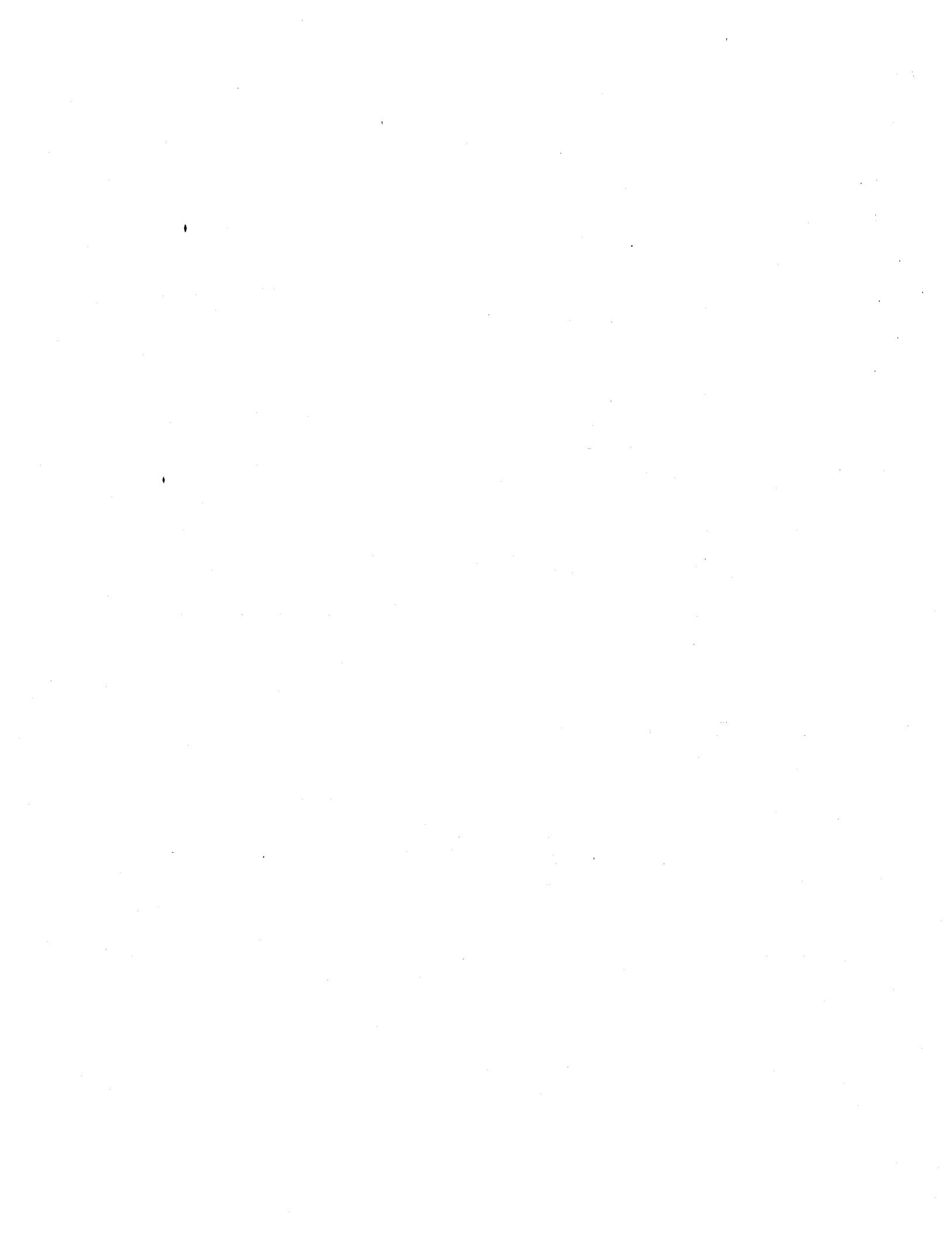
Part V seeks to provide some prescriptions for economic policy. Chapter 17 examines the relative merits of free markets and regulation, and Chapter 18 argues in favor of an international central bank. Since both the market mechanism and the attempts to regulate it are inherently flawed, it may be argued that all attempts at systemic reform are doomed to failure. I reject that argument in Chapter 19.

In the Epilogue I explore the implications of the concept of reflexivity outside the sphere of finance and in a final flight of fancy I attempt to provide my own answers to some age-old metaphysical questions.

Since my thinking has evolved in the course of writing, it may be helpful for the reader to know when the various chapters were written. The first part of the book was completed before I embarked on the real-time experiment in August 1985. Moreover, Chapters 5–9, which deal with recent history, preceded in time Chapter 4, which outlines the concept of the credit and regulatory cycle. Chapter 4 incorporates discoveries I made in the course of writing; that is why it is so tentative in character.

I should like to emphasize that this book is not meant to provide a practical guide to getting rich in the stock market. Most of what I know is in the book, at least in theoretical form. I have not kept anything deliberately hidden. But the chain of reasoning operates in the opposite direction: I am not trying to explain how to use my approach to make money; rather, I am using my experiences in the financial markets to develop an approach to the

study of historical processes in general and the present historical moment in particular. If I did not believe that my investment activities can serve that purpose, I would not want to write about them. As long as I am actively engaged in business, I would be better off to keep them a trade secret. But I would value it much more highly than any business success if I could contribute to an understanding of the world in which we live or, better yet, if I could help to preserve the economic and political system that has allowed me to flourish as a participant.



PART ONE

THEORY

1

THE THEORY OF REFLEXIVITY

Anti-equilibrium

Economic theory is devoted to the study of equilibrium positions. The concept of an equilibrium is very useful. It allows us to focus on the final outcome rather than on the process that leads up to it. But the concept is also very deceptive. It has the aura of something empirical: since the adjustment process is supposed to lead to an equilibrium, an equilibrium position seems somehow implicit in our observations. That is not true. Equilibrium itself has rarely been observed in real life—market prices have a notorious habit of fluctuating. The process that can be observed is supposed to move toward an equilibrium. Why is it that the equilibrium is never reached? It is true that market participants adjust to market prices but they may be adjusting to a constantly moving target. In that case, calling the participants' behavior an adjustment process may be a misnomer and equilibrium theory becomes irrelevant to the real world.

Equilibrium is the product of an axiomatic system. Economic theory is constructed like logic or mathematics: it is based on certain postulates and all of its conclusions are derived from them by logical manipulation. The possibility that equilibrium is never reached need not invalidate the logical construction, but when a hypothetical equilibrium is presented as a model of reality a significant distortion is introduced. If we lived in a world in which the angles of a triangle did not add up to 180 degrees, Euclidean geometry would constitute such a misleading model.

The crowning achievement of the axiomatic approach is the theory of perfect competition. Although it was first propounded nearly two hundred years ago, it has never been superseded; only the method of analysis has been refined. The theory holds that under certain specified circumstances the unrestrained pursuit of self-interest leads to the optimum allocation of resources. The equilibrium point is reached when each firm produces at a level where its marginal cost equals the market price and each consumer buys an amount whose marginal "utility" equals the market price. Analysis shows that the equilibrium position maximizes the benefit of all participants, provided no individual buyer or seller can influence market prices. It is this line of argument that has served as the theoretical underpinning for the laissez-faire policies of the nineteenth century, and it is also the basis of the current belief in the "magic of the marketplace."

Let us examine the main assumptions of the theory of perfect competition. Those that are spelled out include perfect knowledge; homogeneous and divisible products; and a large enough number of participants so that no single participant can influence the market price.

The assumption of perfect knowledge is suspect because understanding a situation in which one participates cannot qualify as knowledge. That was the assumption that I found so unacceptable as a student. I have no doubt that classical economists used the assumption in exactly that sense in which I found it objectionable because nineteenth-century thinkers were less aware of the limitations of knowledge than we are today. As the epistemological problems began to surface, exponents of the theory found that they could get by using a more modest word: information. In its modern formulation the theory merely postulates perfect information.¹

Unfortunately, this assumption is not quite sufficient to support the construction of the theory. To make up for the deficiency, modern economists resorted to an ingenious device: they insisted that the demand and supply curves should be taken as given. They did not present this as a postulate; rather, they based their claim on methodological grounds. They argued that the task of economics is to study the relationship between supply and demand and not either by itself. Demand may be a suitable subject for psychologists, supply may be the province of engineers or management scientists; both are beyond the scope of economics.² Therefore, both must be taken as given.

Yet, if we stop to ask what it means that the conditions of supply and demand are independently given, it is clear that an additional assumption has been introduced. Otherwise, where would those curves come from? We are dealing with an assumption disguised as a methodological device. Participants are supposed to choose between alternatives in accordance with their scale of preferences. The unspoken assumption is that the participants know what those preferences and alternatives are.

As I shall try to show, this assumption is untenable. The shape of the supply and demand curves cannot be taken as independently given, because both of them incorporate the participants' expectations about events that are shaped by their own expectations.

Nowhere is the role of expectations more clearly visible than in financial markets. Buy and sell decisions are based on expectations about future prices, and future prices, in turn, are contingent on present buy and sell decisions. To speak of supply and demand as if they were determined by forces that are independent of the market participants' expectations is quite misleading. The situation is not quite so clear-cut in the case of commodities, where supply is largely dependent on production and demand on consumption. But the price that determines the amounts produced and consumed is not necessarily the present price. On the contrary, market participants are more likely to be guided by future prices, either as expressed in futures markets or as anticipated by themselves. In either case, it is inappropriate to speak of independently given supply and demand curves because both curves incorporate the participants' expectations about future prices.

The very idea that events in the marketplace may affect the shape of the demand and supply curves seems incongruous to those who have been reared on classical economics. The demand and supply curves are supposed to determine the market price. If they were themselves subject to market influences, prices would cease to be uniquely determined. Instead of equilibrium, we would be left with fluctuating prices. This would be a devastating state of affairs. All the conclusions of economic theory would lose their relevance to the real world.

It is to prevent this outcome that the methodological device that treats the supply and demand curves as independently given was introduced. Yet there is something insidious about using a meth-

odological device to obscure an assumption that would be untenable if it were spelled out. To preserve the integrity of economic theory as an axiomatic system, its assumptions ought to be explicitly stated. We may then conclude that economic theory is no more relevant to the real world than non-Euclidean geometry, but at least we would know where we stand. Instead, we have been deceived by a methodological subterfuge. The demand and supply curves are presented in textbooks as though they were grounded in empirical evidence. But there is scant evidence for independently given demand and supply curves. Anyone who trades in markets where prices are continuously changing knows that participants are very much influenced by market developments. Rising prices often attract buyers and vice versa. How could self-reinforcing trends persist if supply and demand curves were independent of market prices? Yet, even a cursory look at commodity, stock, and currency markets confirms that trends are the rule rather than the exception.

The theory of perfect competition could be defended by arguing that the trends we can observe in commodity and financial markets are merely temporary aberrations which will be eliminated in the long run by the "fundamental" forces of supply and demand. It should be remembered that the theory of perfect competition does not claim to define the path of adjustment; it merely analyzes the situation after all the adjustments have taken place. The trouble with the argument is that there can be no assurance that "fundamental" forces will correct "speculative" excesses. It is just as possible that speculation will alter the supposedly fundamental conditions of supply and demand.

In the normal course of events, a speculative price rise provokes countervailing forces: supply is increased and demand reduced, and the temporary excess is corrected with the passage of time. But there are exceptions. In foreign exchange, for example, a sustained price movement can be self-validating, because of its impact on domestic price levels. The same is true in the stock market where the performance of a stock may affect the performance of the company in question in a number of ways. And in examining the recent history of international lending we shall find that excessive lending first increased the borrowing capacity of debtor countries, as measured by their debt ratios, and then, when the banks wanted to be repaid, the debtor countries' ability to do so evaporated. Generally speaking, we shall find that the expansion

and contraction of credit can affect the debtors' ability and willingness to pay.

Are these exceptions that confirm the rule, or do they necessitate a revision of accepted theory? The answer depends on the frequency and severity of their occurrence. If we are dealing with an isolated instance, we can treat it as a paradox; but if one incident follows another, we must question the theory.

I contend that such paradoxical behavior is typical of all financial markets that serve as a discounting mechanism for future developments, notably stock markets, foreign exchange markets, banking, and all forms of credit. Microeconomic theory may continue to ignore it, because there are large areas of economic activity where it occurs only occasionally or not at all; but we cannot expect to understand macroeconomic developments without taking the phenomenon into account. A world of fluctuating exchange rates and large-scale capital movements is characterized by vicious and benign circles in which the "normal" pattern of causation, as defined by classical economics, seems to be reversed: market developments dictate the evolution of the conditions of supply and demand, not the other way around.

If the process of adjustment does not lead to an equilibrium, what happens to the conclusions of economic theory? The answer is that they remain valid as deductions but they lose their relevance to the real world. If we want to understand the real world, we must divert our gaze from a hypothetical final outcome and concentrate our attention on the process of change that we can observe all around us.

This will require a radical shift in our thinking. A process of change is much more difficult to understand than a static equilibrium. We shall have to revise many of our preconceived ideas about the kind of understanding that is attainable and satisfy ourselves with conclusions that are far less definite than those that economic theory sought to provide.

The Problem of Imperfect Understanding

The understanding of the actual course of events, as distinct from a hypothetical equilibrium, poses problems that have not been properly appreciated. The problems arise because participants

base their decisions on an inherently imperfect understanding of the situation in which they participate. There are two related sets of problems to be considered: the imperfect understanding of the participant and the imperfect understanding of the social scientist. We must be careful not to confuse the two.

In this section I shall try to explain why the participants' understanding is inherently imperfect. In the next section I shall examine why the imperfect understanding of the participants poses difficulties for the social sciences.

The imperfect understanding of the participant is a difficult concept to define and an even more difficult one to work with. I shall try to approach it by comparing the position of the participant with that of a natural scientist. (I have to specify a natural scientist, because social scientists are confronted with special problems arising out of the imperfect understanding of participants. These will be dealt with in the next section.) The purpose of the comparison is to establish a standard in terms of which the understanding of the participant can be called imperfect. What makes the comparison tricky is that the natural scientist is not capable of perfect understanding either. Far from it. As Karl Popper has shown,³ it is a cardinal principle of scientific method that perfect knowledge is not attainable. Scientists work by constantly testing plausible hypotheses and propounding new ones. If they did not treat all conclusions as provisional and subject to improvement, natural science could not have reached its present state of development and it could not progress any further. Although it is far from perfect, the knowledge attained by natural scientists provides a standard in terms of which the participants' understanding can be called imperfect.

Natural scientists have one great advantage over participants: they deal with phenomena that occur independently of what anybody says or thinks about them. The phenomena belong to one universe, the scientists' statements to another. The phenomena then serve as an independent, objective criterion by which the truth or validity of scientific statements can be judged. Statements that correspond to the facts are true; those that do not are false. To the extent that the correspondence can be established, the scientist's understanding qualifies as knowledge. We do not need to go into the various difficulties that stand in the way of establishing the correspondence. The important point is that scientists have an objective criterion at their disposal.

By contrast, the situation to which the participants' thinking relates is not independently given: it is contingent on their own decisions. As an objective criterion for establishing the truth or validity of the participants' views, it is deficient. It does provide a criterion of sorts: some expectations are validated by subsequent events, others are not. But the process of validation leaves something to be desired: one can never be sure whether it is the expectation that corresponds to the subsequent event or the subsequent event that conforms to the expectation. The segregation between thoughts and events that prevails in natural science is simply missing.

Thinking plays a dual role. On the one hand, participants seek to understand the situation in which they participate; on the other, their understanding serves as the basis of decisions which influence the course of events. The two roles interfere with each other. Neither role is performed as well as it could be if it were performed separately. If the course of events were independent of the participants' decisions, the participants' understanding could equal that of a natural scientist; and if participants could base their decisions on knowledge, however provisional, the results of their actions would have a better chance of corresponding to their intentions. As it is, participants act on the basis of imperfect understanding and the course of events bears the imprint of that imperfection.

In a milder form, the lack of separation between the subject matter and the act of thinking may also arise in natural science. The most celebrated example is in quantum physics where the act of observation interferes with the observed phenomenon. It has given rise to Heisenberg's uncertainty principle which, in effect, establishes a limit to the scientist's ability to attain knowledge. But in natural science the problem occurs only at the limit, whereas for the participant it is at the very center of his thinking. For one thing, the scientist makes a deliberate attempt not to interfere with his subject matter, whereas the participant's primary purpose is to mold the situation in which he participates to his own satisfaction. More important, in quantum physics it is only the act of observation which interferes with the subject matter, not the theory of uncertainty, whereas in the case of thinking participants their own thoughts form part of the subject matter to which they relate. The positive accomplishments of natural science are confined to the area where thinking and events are effec-

tively segregated. When events have thinking participants that area shrinks to the vanishing point.

The Problem of the Social Sciences

We are now in a position to examine the problems of the social sciences. Again, there are two distinct issues to be considered. One relates to the subject matter, the other to the observer.

Scientific method is designed to deal with facts; but, as we have seen, events which have thinking participants do not consist of facts alone. The participant's thinking plays a causal role, yet it does not correspond to the facts for the simple reason that it does not relate to facts. Participants have to deal with a situation that is contingent on their own decisions; their thinking constitutes an indispensable ingredient in that situation. Whether we treat it as a fact of a special kind or something other than a fact, the participants' thinking introduces an element of uncertainty into the subject matter. This element is absent in the natural sciences. As we have seen, there is some similarity between the uncertainty introduced by the participants' thinking and Heisenberg's uncertainty principle in quantum physics but, as we shall soon see, the parallel is misleading.

Now for the role of the scientific observer: it is much easier to maintain the required segregation between facts and statements when the subject matter itself does not contain any statements, observations, or thoughts of any kind.

Most discussions about the shortcomings of the social sciences have focused on the second issue. Expressions like "self-fulfilling prophecies" or "self-defeating experiments" are widely used but usually they relate to the would-be scientist. Yet it is the self-influencing character of the participants' thinking that is responsible for the element of uncertainty (or indeterminacy) I mentioned before. The difficulties of scientific observation pale into insignificance when compared with the indeterminacy of the subject matter. The indeterminacy would remain even if all the problems relating to the observer were resolved, whereas the problems of the observer can be directly attributed to the indeterminacy of the subject matter. Thus the problem of the social sciences is not merely methodological but inherent in the subject matter.

The undue emphasis on the role of the scientific observer can

be attributed to the false analogy with Heisenberg's uncertainty principle. I am no expert in quantum physics but, as I understand it, the principle holds that the mass and velocity of quantum particles cannot be measured at the same time because the act of measurement interferes with the object that is being measured. In this case, the element of uncertainty is introduced by the outside observer. (Whether the behavior of quantum particles is inherently random is a separate question.) The parallel with the social sciences is misleading because in the latter case the indeterminacy (uncertainty) is caused by the participants. Only if quantum particles behaved as thinking participants would the analogy hold.

I shall try to reverse the discussion to its proper order: the uncertainty in the subject matter first, the role of the scientist second.

The idea that there is a fundamental difference in the subject matter of the natural and social sciences has not been generally recognized. On the contrary, Karl Popper, whom I otherwise admire, enunciated what he called the "doctrine of the unity of science";⁴ that is, the same methods and criteria apply in both the natural and social sciences. Although it has not been universally accepted, it has not been conclusively refuted either. I shall try to do so here.

In order to appreciate the problem posed by thinking participants, let us take a closer look at the way scientific method operates. For this purpose I am invoking Karl Popper's scheme of scientific method, described in technical terms as the "deductive-nomological," or "D-N," model. Like every model, it presents a simplified and idealized version of a more complex reality, but exactly because it is simple and elegant it suits my purpose very well.

The model is built on three kinds of statements: specific initial conditions, specific final conditions, and generalizations of universal validity. Combining a set of generalizations with known initial conditions yields predictions; combining them with known final conditions provides explanations; and matching known initial with known final conditions serves as testing for the generalizations involved. It can be seen that there is a symmetry between predictions and explanations; they are logically reversible. Testing is different, because no amount of testing can prove that a generalization is universally valid. Scientific theories

can only be falsified, never verified. The asymmetry between verification and falsification and the symmetry between prediction and explanation are the two crucial features of Popper's scheme.

The model works only if certain conditions are fulfilled. An essential condition is that the content of the statements should exist in total isolation from the statements that are made about them; only then does the content provide an independent criterion for judging the truth or validity of the statements that relate to it.* Other essential requirements are that the initial and final conditions should consist of facts which are amenable to scientific observation and the generalizations should have universal validity. That is, if a given set of conditions recurred, it would have to be followed or preceded by the same set of conditions as before. It can be seen that the requirement of universal validity defines not only the nature of scientific laws but also the character of initial and final conditions: they must consist of observable facts governed by universal laws. It is this requirement that is so difficult to meet when a situation has thinking participants.

What constitutes scientific observation is a matter of debate that we need not enter into here. Clearly, a single observation by a single scientist is not admissible. Exactly because the correspondence between facts and statements is so difficult to establish, science is a collective enterprise where the work of each scientist has to be open to control and criticism by others.

The interaction between scientists is governed by certain conventions. These conventions are neither clearly defined nor permanently fixed. They derive their authority from the fact that they produce desired results. Individual scientists often find the conventions quite onerous and try various shortcuts in order to attain a desired result. Only because the shortcuts do not work do the conventions of scientific method continue to prevail.

Perhaps the most outstanding example of the observer trying to impose his will on his subject matter is the attempt to convert base metal into gold. Alchemists struggled long and hard until they were finally persuaded to abandon their enterprise by their

* Interestingly, this condition holds not only for the facts that constitute the initial and final conditions but also for the generalizations that connect them. The laws of nature must prevail whether they are recognized or not. This is the basis of Popper's abortive idea of a "third world" of objective thoughts (in *Objective Knowledge*, New York: Oxford University Press, 1972).

lack of success. The failure was inevitable because the behavior of base metals is governed by laws of universal validity which cannot be modified by any statements, incantations, or rituals.

Let us now consider the behavior of human beings. Do they obey universally valid laws that can be formulated in accordance with the D-N model? Undoubtedly, there are many aspects of human behavior, from birth to death and in between, which are amenable to the same treatment as other natural phenomena. But there is one aspect of human behavior which seems to exhibit characteristics different from those of the phenomena which form the subject matter of natural science: the decision-making process. Decisions are based on an imperfect understanding of the situation. How does such a situation provide the initial and final conditions which are supposed to be connected according to laws of universal validity? Do those conditions include or exclude the participants' thinking? If thinking is included, the conditions are not amenable to scientific observation, because only the effects of the participants' thinking can be observed, not the process itself. If the thinking process is excluded and only its effects are admitted as evidence, the universal validity of scientific generalizations is destroyed because a given set of conditions is not necessarily preceded or succeeded by the same set every time: the sequence of events is influenced by the participants' thinking, yet there is a lack of correspondence between the participants' thinking and the course of events. In either case, the D-N model breaks down.

This may not be the end of the world, but it is a serious blow to scientific method. The method has been so successful that we find it hard to believe that there should be a large and vital area beyond its scope. Natural science has also encountered limitations in the form of Heisenberg's uncertainty principle but the limits were reached only after a number of impressive accomplishments—the uncertainty principle itself counts as one of the great discoveries of natural science. In social science we encounter difficulties even before we get started: the imperfect understanding of the participants is incompatible with the D-N model.

This conclusion is so devastating that every effort has been made to escape it. To review the various attempts would take a whole book—and an interesting one at that. I shall confine my attention to economic theory, which constitutes one of the most ingenious attempts, and in some ways the most effective one, to escape the problems connected with imperfect understanding. It

simply assumes away the problems by erecting a hypothetical system in which the participants' decisions are fully determined by the available information. This approach yields conclusions which meet some of the requirements of the D-N model. For instance, the theory of perfect competition qualifies as universally valid and—at least in principle—it can be used to explain and predict conditions with equal force. The theory fails, however, when it comes to testing, leaving in doubt the relevance of the hypothetical to actual conditions.

Social scientists have gone to great lengths trying to maintain the unity of method but with remarkably little success. Their endeavors have yielded little more than a parody of natural science. In a sense, the attempt to impose the methods of natural science on social phenomena is comparable to the efforts of alchemists who sought to apply the methods of magic to the field of natural science. But while the failure of the alchemists was well-nigh total, social scientists have managed to make a considerable impact on their subject matter. Situations which have thinking participants may be impervious to the methods of natural science, but they are susceptible to the methods of alchemy. The thinking of participants, exactly because it is not governed by reality, is easily influenced by theories. In the field of natural phenomena, scientific method is effective only when its theories are valid; but in social, political, and economic matters, theories can be effective without being valid. Whereas alchemy has failed as natural science, social science can succeed as alchemy.

This brings us to an examination of the relationship between the scientist and his subject matter. As we have seen, the D-N model requires that scientists keep their statements and observations rigorously segregated from the subject matter to which they relate; only then can the subject matter fulfill its function and serve as an objective criterion for judging the truth or validity of scientific statements. The conventions of scientific method are designed to maintain the required segregation.

In natural science the conventions are effective because the scientist's thinking is, in fact, distinct from its subject matter. The scientist can influence the subject matter only by actions, not by thoughts, and the scientist's actions are guided by the same laws as all other natural phenomena. Specifically, nothing the scientist can do will turn base metals into gold. The scientist may gain some personal advantage by flouting the conventions of science,

but the advantage is achieved only by a deception which is liable to be revealed by those who abide by the conventions.

Social phenomena are different. The imperfect understanding of the participant interferes with the proper functioning of the D-N model. This has far-reaching implications for the conventions of scientific method. It limits the results that can be produced by observing the conventions and, what is worse, it opens the way to attaining worthwhile results by transgressing them. There is much to be gained by pretending to abide by the conventions of scientific method without actually doing so. Natural science is held in great esteem: a theory that claims to be scientific can influence the gullible public much better than one which frankly admits its political or ideological bias. I only need to mention Marxism and psychoanalysis as typical examples; but laissez-faire capitalism, with its reliance on the theory of perfect competition, is also a case in point. It is noteworthy that both Marx and Freud were vocal in protesting their scientific status and based many of their conclusions on the authority they derived from being "scientific." Once this point sinks in, the very expression "social science" becomes suspect. It is a magic word employed by social alchemists in their effort to impose their will on their subject matter by incantation.

How can the "true" practitioners of scientific method protect themselves against such malpractice? It seems to me that there is only one way out: to deprive social science of the status it enjoys on account of natural science. Social science ought to be recognized as a false metaphor.

That does not mean that we must give up the pursuit of truth in exploring social phenomena. It means only that the pursuit of truth requires us to recognize that the D-N model does not apply to situations with thinking participants. We must abandon the doctrine of the unity of method and cease the slavish imitation of natural science.

The D-N model has been so successful in natural science that it has come to be equated with scientific method. Ironically, the model has been largely superseded in modern natural science; the social sciences, however, are still trying to compete with the accomplishments of nineteenth-century natural science. The attempt is in vain because, as we have seen, the subject matter does not fit the D-N mold. But, as quantum physics has shown, scientific method is not necessarily confined to the D-N model: statis-

tical, probabilistic generalizations may prove more fruitful. Nor should we ignore the possibility of developing novel approaches which have no counterpart in natural science. Given the differences in subject matter, there ought to be differences in the method of study.

I shall explore a novel approach in this book, but before I start, I want to make sure that it will not be judged by the standards of the D-N model. A world of imperfect understanding does not lend itself to generalizations which can be used to explain and to predict specific events. The symmetry between explanation and prediction prevails only in the absence of thinking participants. Otherwise, predictions must always be conditioned on the participants' perceptions; thus they cannot have the finality which they enjoy in the D-N model. On the other hand, past events are just as final as in the D-N model; thus, explanation turns out to be an easier task than prediction. Once we abandon the constraint that predictions and explanations are logically reversible, we can build a theoretical framework which is appropriate to the subject matter. Unfortunately, the theory cannot be tested in the same way as those which fit into Popper's logical structure. That is not to say that testing must be abandoned; on the contrary, as long as we remain concerned with understanding things as they are, we ought to insist on testing our views. We need to develop novel ways of testing. I shall do so in the real-time experiment that frankly admits, and exploits, the alchemical potential in the study of social phenomena (Part III).

The Participants' Bias

My approach is to tackle the problem of imperfect understanding head on. What makes the participants' understanding imperfect is that their thinking affects the situation to which it relates. The causal role played by the participants' thinking has no counterpart in the phenomena studied by natural scientists. It is obviously not the only force shaping the course of events, but it is a force which is unique to events that have thinking participants. Hence it deserves to take center stage.

As we have seen, imperfect understanding is a very difficult concept to work with. We have established that there is a lack of correspondence between the participant's thinking and the situa-

tion to which it relates; but the lack of correspondence is difficult to define, let alone measure. The participant's thinking is part of the situation to which it relates and the very idea of correspondence is inappropriate to describing a relationship between a part and a whole. The idea was imported from natural science, where facts and statements belong to separate universes and from philosophy, where correspondence serves as the criterion of truth. The analogy does not apply to the participant who is, by definition, part of the situation that he is trying to understand. We can speak of a lack of correspondence, but we cannot define that to which the participant's understanding fails to correspond because it simply does not exist. To simplify matters, I shall speak of an inherent bias in the participants' thinking. Since the bias is inherent, the unbiased is unattainable.

There is, however, a norm in the outside world against which the participants' bias can be measured. Although there is no reality independent of the participants' perception, there is a reality that is dependent on it. In other words, there is a sequence of events that actually occurs and that sequence reflects the participants' behavior. The actual course of events is likely to differ from the participants' expectations and the divergence can be taken as an indication of the participants' bias. Unfortunately, it can be taken only as an indication—not as the full measure of the bias—because the actual course of events already incorporates the effects of the participants' thinking. Thus the participants' bias finds expression both in the divergence between outcome and expectations and in the actual course of events. A phenomenon that is partially observable and partially submerged in the course of events does not lend itself readily to scientific investigation. We can now appreciate why economists were so anxious to eliminate it from their theories. We shall make it the focal point of our investigation.

The Concept of Reflexivity

The connection between the participants' thinking and the situation in which they participate can be broken up into two functional relationships. I call the participants' efforts to understand the situation the cognitive or passive function and the impact of their thinking on the real world the participating or active func-

Theory

tion. In the cognitive function, the participants' perceptions depend on the situation; in the participating function, the situation is influenced by the participants' perceptions. It can be seen that the two functions work in opposite directions: in the cognitive function the independent variable is the situation; in the participating function it is the participants' thinking.

There are many cases where one or the other function can be observed in isolation but there are also instances where they are both operating at the same time. An obvious example of the cognitive function is when someone learns from experience. Examples of the participating function are to be found in textbooks of economics where the participants apply a given set of preferences to a given set of opportunities and in the process determine prices.

When both functions operate at the same time, they interfere with each other. Functions need an independent variable in order to produce a determinate result, but in this case the independent variable of one function is the dependent variable of the other. Instead of a determinate result, we have an interplay in which both the situation and the participants' views are dependent variables so that an initial change precipitates further changes both in the situation and in the participants' views. I call this interaction "reflexivity," using the word as the French do when they describe a verb whose subject and object are the same. Using simple mathematics, reflexivity can be depicted as a pair of recursive functions:

$$\begin{aligned} y &= f(x) \quad \text{cognitive function} \\ x &= \phi(y) \quad \text{participating function} \end{aligned}$$

Therefore,

$$\begin{aligned} y &= f[\phi(y)] \\ x &= \phi[f(x)] \end{aligned}$$

This is the theoretical foundation of my approach. The two recursive functions do not produce an equilibrium but a never-ending process of change. The process is fundamentally different from the processes that are studied by natural science. There, one set of facts follows another without any interference from thoughts or perceptions (although in quantum physics, observa-

tion introduces uncertainty). When a situation has thinking participants, the sequence of events does not lead directly from one set of facts to the next; rather, it connects facts to perceptions and perceptions to facts in a shoelace pattern. Thus, the concept of reflexivity yields a "shoelace" theory of history.

It must be recognized that the shoelace theory is a kind of dialectic. It can be interpreted as a synthesis of Hegel's dialectic of ideas and Marx's dialectical materialism. Instead of either thoughts or material conditions evolving in a dialectic fashion on their own, it is the interplay between the two that produces a dialectic process. The only reason I do not use the word more prominently is that I do not want to be burdened by the excess luggage that comes with it. I find Hegel obscure, and Marx propounded a deterministic theory of history that is diametrically opposed to my own view.

The historical process, as I see it, is open ended. Its main driving force is the participants' bias. To be sure, it is not the only force at work, but it is a force that is unique to the historical process and sets it apart from the processes studied by natural science. Biological evolution is attributed to genetic mutation; I contend that historical processes are shaped by the misconceptions of the participants. I would even go as far as to say that the ideas that make history consist of fertile fallacies. A fertile fallacy is originally conceived as an insight; only when it is translated into reality do its shortcomings become apparent; it then begets another fertile fallacy that is antithetical to it, and so it goes. Each fallacy provides a new experience and, to the extent that people learn from experience, the process can be described as progress. Fallacy is, of course, too strong a word, but it is helpful in directing attention in the right direction: to the participants' bias.

I shall not pursue the subject further here, but it is obvious that the concept of reflexivity, as described here, has implications far beyond the range of topics tackled in this book.

Reflexivity Versus Equilibrium

Returning to economic theory, it can be argued that it is the participants' bias that renders the equilibrium position unattainable. The target toward which the adjustment process leads incorporates a bias, and the bias may shift in the process. When that

happens, the process aims not at an equilibrium but at a moving target.

To put matters into perspective, we may classify events into two categories: humdrum, everyday events that are correctly anticipated by the participants and do not provoke a change in their perceptions, and unique, historical events that affect the participants' bias and lead to further changes. The first kind of event is susceptible to equilibrium analysis, the second is not: it can be understood only as part of a historical process.

In everyday events, only the participating function is operative; the cognitive function is given. In the case of unique, historic developments, both functions operate simultaneously so that neither the participants' views nor the situation to which they relate remain the same as they were before. That is what justifies describing such developments as historic.

It should be emphasized that my definition of historical change involves a tautology. First, I classify events according to their effect on the participants' bias: those that alter the participant's bias are historic and those that do not are humdrum. I then claim that it is changes in the participants' bias that qualify a course of events as historical.

Tautologies can be useful, provided they are recognized as such. In this case it helps to put equilibrium analysis into the proper perspective. I have defined historical change as an interplay between the cognitive and participating functions. What makes the change historical is that it affects both the course of events and the participants' perceptions so that the next event cannot be a mere repetition of the one that preceded it.

Equilibrium analysis eliminates historical change by assuming away the cognitive function. The supply and demand curves utilized by economic theory are expressions of the participating function only. The cognitive function is replaced by the assumption of perfect knowledge. If the cognitive function were operating, events in the marketplace could alter the shape of the demand and supply curves, and the equilibrium studied by economists need never be reached.

How significant is the omission of the cognitive function? In other words, how significant is the distortion introduced by neglecting the participants' bias?

In microeconomic analysis, the distortion is negligible and the participants' bias can be accounted for easily. As a first step, the

participants' bias can be taken as given: that provides a static equilibrium. To make the analysis more dynamic, changes in the participants' bias can be added piecemeal, expressed as changes in consumer habits or production methods. All that is obscured by this piecemeal approach is the possible connection between the various changes in the conditions of supply and demand, but that omission does not invalidate the conclusions microeconomic analysis seeks to establish.

When it comes to financial markets, the distortion is more serious. The participants' bias is an element in determining prices and no important market development leaves the participants' bias unaffected. The search for an equilibrium price turns out to be a wild goose chase and theories about the equilibrium price can themselves become a fertile source of bias. To paraphrase J. P. Morgan, financial markets will continue to fluctuate.

In trying to deal with macroeconomic developments, equilibrium analysis is totally inappropriate. Nothing could be further removed from reality than the assumption that participants base their decisions on perfect knowledge. People are groping to anticipate the future with the help of whatever guideposts they can establish. The outcome tends to diverge from expectations, leading to constantly changing expectations and constantly changing outcomes. The process is reflexive.

In his *General Theory of Employment, Interest and Money*, Keynes managed to show that full employment is a special case; if we could develop a general theory of reflexivity, equilibrium would appear as a special case. While it will hardly qualify as a general theory, I shall try to explore the role of reflexivity in financial markets; that means that I shall try to interpret their functioning as a historical process.

2

REFLEXIVITY IN THE STOCK MARKET

In trying to develop a theory of reflexivity, I shall start with the stock market. For one thing, it is the market I am most familiar with: I have been a professional investor for more than twenty-five years. For another, the stock market provides an excellent laboratory for testing theories: changes are expressed in quantitative terms and the data are easily accessible. Even the participants' views are usually available in the form of brokers' reports. Most important, I have actually tested my theory in the stock market and I have some interesting case studies to present.

As I mentioned in the introduction, I did not develop my ideas on reflexivity in connection with my activities in the stock market. The theory of reflexivity started out as abstract philosophical speculation and only gradually did I discover its relevance to the behavior of stock prices. I was singularly unsuccessful in formulating my theory at the level of abstraction at which I conceived it: my failure as a philosopher stands in stark contrast with my career as an investment professional. I hope that by presenting my ideas in the reverse order from the one in which I arrived at them I may be able to avoid getting lost in arcane abstractions.

There is yet another reason why the stock market may provide the best entry point for the study of reflexive phenomena. The stock market comes as close to meeting the criteria of perfect competition as any market: a central marketplace, homogeneous products, low transaction and transportation costs, instant communications, a large enough crowd of participants to ensure that no individual can influence market prices in the ordinary course of events, and special rules for insider transactions as well as

special safeguards to provide all participants with access to relevant information. What more can one ask for? If there is any place where the theory of perfect competition ought to be translated into practice, it is in the stock market.

Yet there is little empirical evidence of an equilibrium or even a tendency for prices to move toward an equilibrium. The concept of an equilibrium seems irrelevant at best and misleading at worst. The evidence shows persistent fluctuations, whatever length of time is chosen as the period of observation. Admittedly, the underlying conditions that are supposed to be reflected in stock prices are also constantly changing, but it is difficult to establish any firm relationship between changes in stock prices and changes in underlying conditions. Whatever relationship can be established has to be imputed rather than observed. I intend to use the theory of reflexivity to criticize the preoccupation of economic theory with the equilibrium position. What better example could I find than the stock market?

Existing theories about the behavior of stock prices are remarkably inadequate. They are of so little value to the practitioner that I am not even fully familiar with them. The fact that I could get by without them speaks for itself.

Generally, theories fall into two categories: fundamentalist and technical. More recently, the random walk theory has come into vogue; this theory holds that the market fully discounts all future developments so that the individual participant's chances of over- or underperforming the market as a whole are even. This line of argument has served as theoretical justification for the increasing number of institutions that invest their money in index funds. The theory is manifestly false—I have disproved it by consistently outperforming the averages over a period of twelve years. Institutions may be well advised to invest in index funds rather than making specific investment decisions, but the reason is to be found in their substandard performance, not in the impossibility of outperforming the averages.

Technical analysis studies market patterns and the demand and supply of stocks. It has undoubtedly merit in predicting probabilities but not the actual course of events. For the purposes of this discussion it is of no particular interest, because it has little theoretical foundation other than the assertions that stock prices are determined by their supply and demand and that past experience is relevant in predicting the future.

Fundamental analysis is more interesting because it is an out-

growth of equilibrium theory. Stocks are supposed to have a true or fundamental value as distinct from their current market price. The fundamental value of a stock may be defined either in relation to the earning power of the underlying assets or in relation to the fundamental value of other stocks. In either case, the market price of a stock is supposed to tend toward its fundamental value over a period of time so that the analysis of fundamental values provides a useful guide to investment decisions.

The important point about this approach is that the connection between stock prices and the companies whose stocks are traded is assumed to be in one direction. The fortunes of the companies determine—however belatedly—the relative values of the various stocks traded in the stock market. The possibility that stock market developments may affect the fortunes of the companies is left out of account. There is a clear parallel with the theory of price where the indifference curve determines the relative amounts consumed, and the possibility that the market may influence the indifference curve is disregarded. The parallel is not accidental: the fundamentalist approach is based on the theory of price. But the omission is more glaring in the stock market than in other markets. Stock market valuations have a direct way of influencing underlying values: through the issue and repurchase of shares and options and through corporate transactions of all kinds—mergers, acquisitions, going public, going private, and so on. There are also more subtle ways in which stock prices may influence the standing of a company: credit rating, consumer acceptance, management credibility, etc. The influence of these factors on stock prices is, of course, fully recognized; it is the influence of stock prices on these factors that is so strangely ignored by the fundamentalist approach.

If there are any glaring discrepancies between prevailing stock prices and fundamental values, they are attributed to future developments in the companies concerned that are not yet known but are correctly anticipated by the stock market. Movements in stock prices are believed to precede the developments that subsequently justify them. How future developments ought to be discounted is the subject of an ongoing debate, but it is presumed that the market is doing the job correctly even if the correct method cannot be theoretically established. This point of view follows naturally from the theory of perfect competition. It is summed up in the assertion that "the market is always right." The

assertion is generally accepted, even by people who do not put much faith in fundamental analysis.

I take a totally opposite point of view. I do not accept the proposition that stock prices are a passive reflection of underlying values, nor do I accept the proposition that the reflection tends to correspond to the underlying value. I contend that market valuations are always distorted; moreover—and this is the crucial departure from equilibrium theory—the distortions can affect the underlying values. Stock prices are not merely passive reflections; they are active ingredients in a process in which both stock prices and the fortunes of the companies whose stocks are traded are determined. In other words, I regard changes in stock prices as part of a historical process and I focus on the discrepancy between the participants' expectations and the actual course of events as a causal factor in that process.

To explain the process, I take the discrepancy as my starting point. I do not rule out the possibility that events may actually correspond to people's expectations, but I treat it as a limiting case. Translating this assertion into market terms, I claim that market participants are always biased in one way or another. I do not deny that markets have a predictive or anticipating power that seems uncanny at times, but I argue that it can be explained by the influence that the participants' bias has on the course of events. For instance, the stock market is generally believed to anticipate recessions; it would be more correct to say that it can help to precipitate them. Thus I replace the assertion that markets are always right with two others:

1. Markets are always biased in one direction or another.
2. Markets can influence the events that they anticipate.

The combination of these two assertions explains why markets may so often appear to anticipate events correctly.

Using the participants' bias as our starting point, we can try to build a model of the interaction between the participants' views and the situation in which they participate. What makes the analysis so difficult is that the participants' views are part of the situation to which they relate. To make any sense of such a complex situation, we need to simplify it. I introduced a simplifying concept when I spoke of the participants' bias. Now I want to take

the argument a step further and introduce the concept of a prevailing bias.

Markets have many participants, whose views are bound to differ. I shall assume that many of the individual biases cancel each other out, leaving what I call the "prevailing bias." This assumption is not appropriate to all historical processes but it does apply to the stock market and to other markets as well. What makes the procedure of aggregating individual perceptions legitimate is that they can be related to a common denominator, namely, stock prices. In other historical processes, the participants' views are too diffuse to be aggregated and the concept of a prevailing bias becomes little more than a metaphor. In these cases a different model may be needed, but in the stock market the participants' bias finds expression in purchases and sales. Other things being equal, a positive bias leads to rising stock prices and a negative one to falling prices. Thus the prevailing bias is an observable phenomenon.

Other things are, of course, never equal. We need to know a little more about those "other things" in order to build our model. At this point I shall introduce a second simplifying concept. I shall postulate an "underlying trend" that influences the movement of stock prices whether it is recognized by investors or not. The influence on stock prices will, of course, vary, depending on the market participants' views. The trend in stock prices can then be envisioned as a composite of the "underlying trend" and the "prevailing bias."

How do these two factors interact? It will be recalled that there are two connections at play: the participating and the cognitive functions. The underlying trend influences the participants' perceptions through the cognitive function; the resulting change in perceptions affects the situation through the participating function. In the case of the stock market, the primary impact is on stock prices. The change in stock prices may, in turn, affect both the participants' bias and the underlying trend.

We have here a reflexive relationship in which stock prices are determined by two factors—underlying trend and prevailing bias—both of which are, in turn, influenced by stock prices. The interplay between stock prices and the other two factors has no constant: what is supposed to be the independent variable in one function is the dependent variable in the other. Without a constant, there is no tendency toward equilibrium. The sequence of

events is best interpreted as a process of historical change in which none of the variables—stock prices, underlying trend, and prevailing bias—remains as it was before. In the typical sequence the three variables reinforce each other first in one direction and then in the other in a pattern that is known, in its simplest form, as boom and bust.

First, we must start with some definitions. When stock prices reinforce the underlying trend, we shall call the trend self-reinforcing; when they work in the opposite direction, self-correcting. The same terminology holds for the prevailing bias: it can be self-reinforcing or self-correcting. It is important to realize what these terms mean. When a trend is reinforced, it accelerates. When the bias is reinforced, the divergence between expectations and the actual course of future stock prices gets wider and, conversely, when it is self-correcting, the divergence gets narrower. As far as stock prices are concerned, we shall describe them simply as rising and falling. When the prevailing bias helps to raise prices we shall call it positive; when it works in the opposite direction, negative. Thus rising prices are reinforced by a positive bias and falling prices by a negative one. In a boom/bust sequence we would expect to find at least one stretch where rising prices are reinforced by a positive bias and another where falling prices are reinforced by a negative bias. There must also be a point where the underlying trend and the prevailing bias combine to reverse the trend in stock prices.

Let us now try to build a rudimentary model of boom and bust. We start with an underlying trend that is not yet recognized—although a prevailing bias that is not yet reflected in stock prices is also conceivable. Thus, the prevailing bias is negative to start with. When the market participants recognize the trend, this change in perceptions will affect stock prices. The change in stock prices may or may not affect the underlying trend. In the latter case, there is little more to discuss. In the former case we have the beginning of a self-reinforcing process.

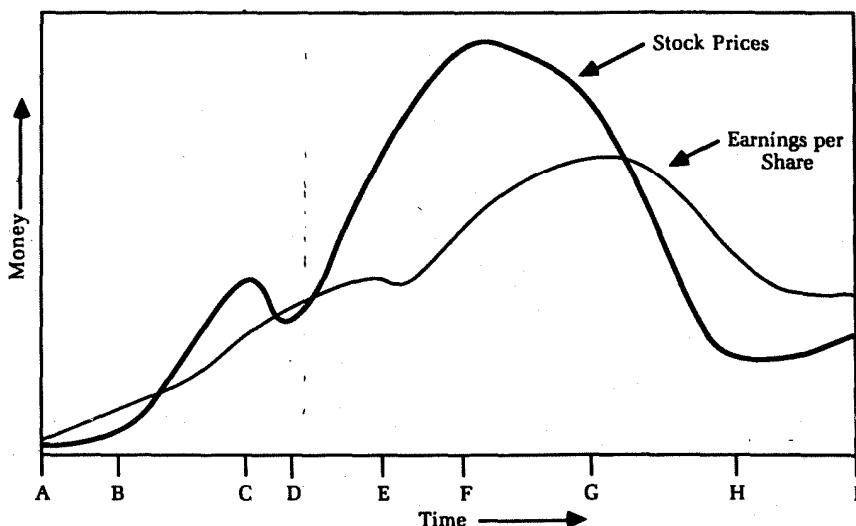
The enhanced trend will affect the prevailing bias in one of two ways: it will lead to the expectation of further acceleration or to the expectation of a correction. In the latter case, the underlying trend may or may not survive the correction in stock prices. In the former case, a positive bias develops causing a further rise in stock prices and a further acceleration in the underlying trend. As long as the bias is self-reinforcing, expectations rise even faster

than stock prices. The underlying trend becomes increasingly influenced by stock prices and the rise in stock prices becomes increasingly dependent on the prevailing bias, so that both the underlying trend and the prevailing bias become increasingly vulnerable. Eventually, the trend in prices cannot sustain prevailing expectations and a correction sets in. Disappointed expectations have a negative effect on stock prices, and faltering stock prices weaken the underlying trend. If the underlying trend has become overly dependent on stock prices, the correction may turn into a total reversal. In that case, stock prices fall, the underlying trend is reversed, and expectations fall even further. In this way, a self-reinforcing process gets started in the opposite direction. Eventually, the downturn also reaches a climax and reverses itself.

Typically, a self-reinforcing process undergoes orderly corrections in the early stages, and, if it survives them, the bias tends to be reinforced, and is less easily shaken. When the process is advanced, corrections become scarcer and the danger of a climactic reversal greater.

I have sketched out a typical boom/bust sequence. It can be illustrated by two curves that follow more or less the same direction. One represents stock prices, and the other, earnings per share. It would be natural to envision the earnings curve as a measure of the underlying trend, and the divergence between the two curves as an indication of the underlying bias. The relationship is much more complex. The earnings curve incorporates not only the underlying trend but also the influence of stock prices on that trend; the prevailing bias is expressed only partially by the divergence between the two curves—partially it is already reflected in those curves. Concepts that are only partially observable are extremely difficult to work with; that is why we have chosen variables that can be observed and quantified—although, as we shall see later, the quantification of earnings per share can be very misleading. For present purposes, we shall assume that the “fundamentals” in which investors are interested are properly measured by earnings per share.

A typical path for the two curves may be as follows. (See the figure below.) At first, recognition of an underlying trend is lagging but the trend is strong enough to manifest itself in earnings per share (AB). When the underlying trend is finally recognized, it is reinforced by rising expectations (BC). Doubts arise, but the trend survives. Alternatively, the trend waivers but reasserts it-



self. Such testing may be repeated several times, but here I show it only once (CD). Eventually, conviction develops and it is no longer shaken by a setback in the earning trend (DE). Expectations become excessive, and fail to be sustained by reality (EF). The bias is recognized as such and expectations are lowered (FG). Stock prices lose their last prop and plunge (G). The underlying trend is reversed, reinforcing the decline (GH). Eventually, the pessimism becomes overdone and the market stabilizes (HI).

It should be emphasized that this is only one possible path that results from the interplay of a single underlying trend and a prevailing bias. There could be more than one trend at work and the prevailing bias could have many nuances, so that the sequence of events might require a totally different representation.

A few words about the theoretical construction of the model may be in order. We are interested in the interplay between the participants' bias and the actual course of events. Yet the participants' bias is not directly represented in our model; both curves denote the actual course of events. The prevailing bias is partially incorporated in the two curves and partially denoted by the divergence between them.

The great merit of this construction is that it uses variables that can be quantified. Stock prices serve as a convenient proxy for the situation to which the participants' bias relates. In other historical processes the situation that is interconnected with the parti-

pants' perception by the cognitive and participating functions is more difficult to identify and impossible to quantify. It is the availability of a convenient proxy that renders the stock market such a useful laboratory for studying reflexivity.

Unfortunately, the model offers only a partial explanation of how stock prices are determined. The concept of an underlying trend has been introduced as a placeholder term, to denote changes in the "fundamentals." What the fundamentals are has not been defined. Even the question of how the fundamentals are to be measured has not been answered. Earnings, dividends, asset value, free cash flow: all these yardsticks are relevant, as well as many others, but the relative weight given to each depends on the investors' judgments and is therefore subject to their bias. We may use earnings per share for purposes of illustration, but that merely begs the question. It is a question security analysts have been struggling with for a long time. We do not need to answer it here in order to develop a theory of reflexivity.

Without knowing anything about the fundamentals we can make some worthwhile generalizations. The first generalization is that stock prices must have some effect on the fundamentals (whatever they are), in order to create a boom/bust pattern. Sometimes the connection is direct, as in the examples I shall use in this chapter, but generally it is indirect. Often it makes its effect felt through a political process, such as changes in taxation, or regulation, or attitudes to saving and investment.

It is possible to have a reflexive connection between stock prices and the prevailing bias even if the fundamentals remain unaffected, but the connection becomes interesting only if the fundamentals are also involved. Without a change in fundamentals, the prevailing bias is likely to be corrected in short order, as we can observe in the daily fluctuations of stock prices. It would be quite in order to ignore the bias as mere noise. That is what the theory of perfect competition and the fundamentalist approach to security analysis have done. By contrast, when the fundamentals are affected, the bias cannot be left out of account without serious distortion, because the bias gives rise to a self-reinforcing/self-defeating process that leaves neither stock prices, nor the fundamentals, nor the participants' views the same as they were before.

The second generalization is that there is bound to be a flaw in the participants' perception of the fundamentals. The flaw may

not be apparent in the early stages but it is likely to manifest itself later on. When it does, it sets the stage for a reversal in the prevailing bias. If the change in bias reverses the underlying trend a self-reinforcing process is set in motion in the opposite direction. What the flaw is and how and when it is likely to manifest itself are the keys to understanding boom/bust sequences.

The model I presented above is built on these two generalizations. It hardly needs emphasizing how crude the model is. Nevertheless, it is valuable in identifying the crucial features of a typical boom/bust sequence. These are the unrecognized trend; the beginning of a self-reinforcing process; the successful test; the growing conviction, resulting in a widening divergence between reality and expectations; the flaw in perceptions; the climax; a self-reinforcing process in the opposite direction. Just by identifying these features we can gain some insight into the behavior of stock prices. We cannot expect much more from our rudimentary model.

In any case, a reflexive model cannot take the place of fundamental analysis: all it can do is to provide an ingredient that is missing from it. In principle, the two approaches could be reconciled. Fundamental analysis seeks to establish how underlying values are reflected in stock prices, whereas the theory of reflexivity shows how stock prices can influence underlying values. One provides a static picture, the other a dynamic one.

A theory that offers a partial explanation of stock price movements can be very useful to the investor especially if it illuminates a relationship that other investors fail to grasp. Investors operate with limited funds and limited intelligence: they do not need to know everything. As long as they understand something better than others, they have an edge. The trouble with any kind of specialized knowledge is that one's area of expertise may not be especially interesting, but the theory of reflexivity serves to identify historically significant price movements, so it goes right to the heart of the matter.

The rudimentary model I have outlined above has proved extremely rewarding in my career as an investor. That may seem surprising because the model is so simple and it fits a well-trodden stock market pattern so well that one would expect every investor to be familiar with it. Yet, that is not the case. Why? Part of the answer must be that market participants have been mis-

guided by a different theoretical construction, one derived from classical economics and, even more important, from the natural sciences. The ingrained attitude is that stock prices are the passive reflection of some underlying reality and not an active ingredient in the historical process. This view is simply false. It is remarkable that the error has still not been fully recognized. Nevertheless, investors do recognize the sequence I have described and do respond to it, only they respond later than someone who is working with the appropriate model and is on the lookout for the crucial features that define the shape of the price curve. That is what has given me my edge.

The first time I used the model systematically was in the conglomerate boom of the late 1960s. It enabled me to make money both on the way up and on the way down.

The key to the conglomerate boom was a prevailing misconception among investors. Investors had come to value growth in per-share earnings and failed to discriminate about the way the earnings growth was accomplished. A few companies learned to produce earnings growth through acquisitions. Once the market started to reward them for their performance, their task became easier because they could offer their own highly priced stock in acquiring other companies.

In theory, the process works as follows. Let us assume that all of the companies involved have the same intrinsic growth in earnings but the stock of the acquiring company sells at twice the earnings multiple of the acquired ones; if the acquiring company manages to double its size, its earnings per share jump by 50%, and its growth rate increases accordingly.

In practice, the early conglomerates started out with high intrinsic growth rates that were rewarded by high multiples. Several of the pathbreakers were high-technology companies with a strong defense component whose managements recognized that their historic growth rate could not be sustained indefinitely: Textron, Teledyne, Ling-Temco-Vought (later LTV), to mention a few. They started to acquire more mundane companies, but, as their per-share earnings growth accelerated, the multiples expanded instead of contracting. Their success attracted imitators and later on even the most humdrum companies could attain a high multiple by going on an acquisition spree. For instance, the bulk of Ogden's earnings was derived from trading scrap metal; nevertheless, the stock sold at more than twenty times earnings at its peak.

Eventually, a company could achieve a high multiple just by promising to put it to good use by making acquisitions.

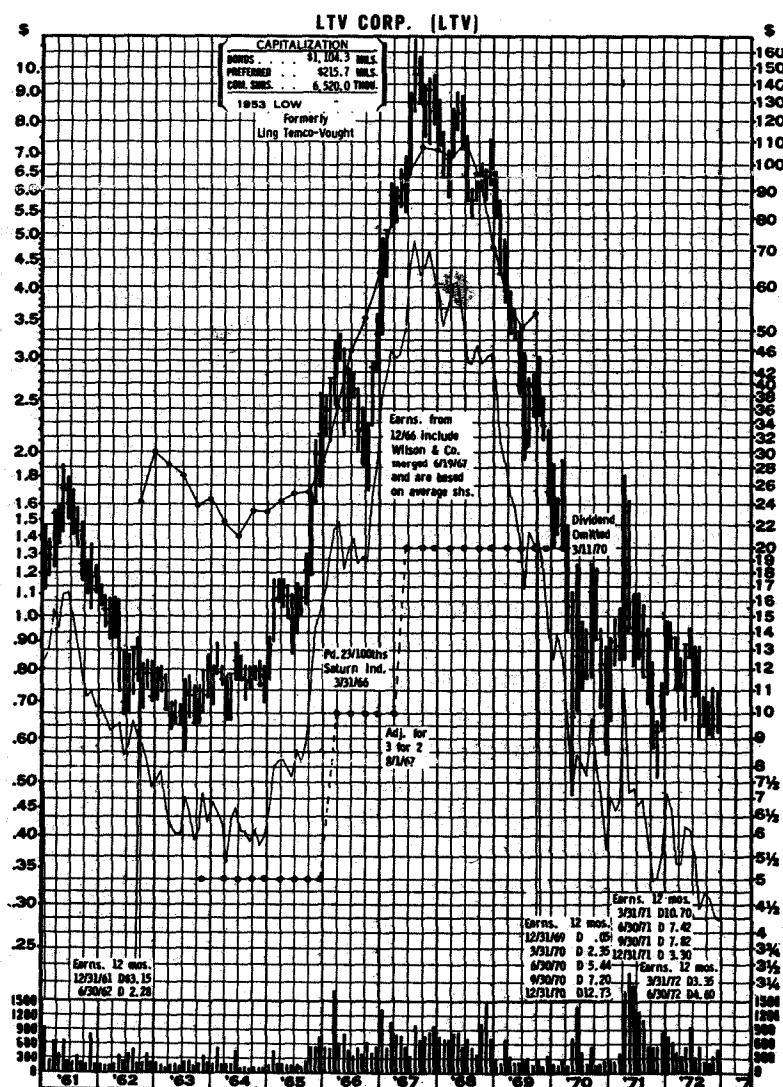
Managements developed special accounting techniques that enhanced the impact of acquisitions. They also introduced changes in the acquired companies, streamlining operations, disposing of assets, and generally focusing on the bottom line, but these changes were less significant than the impact on per-share earnings of the acquisitions themselves.

Investors responded like Indians to firewater. At first, the record of each company was judged on its own merit, but gradually conglomerates became recognized as a group. A new breed of investors emerged, the so-called go-go fund managers, or gun-slingers, who developed a special affinity with the managements of conglomerates. Direct lines of communication were opened between them and conglomerates would place so-called "letter stock" directly with investors. Eventually, conglomerates learned to manage their stock prices as well as their earnings.

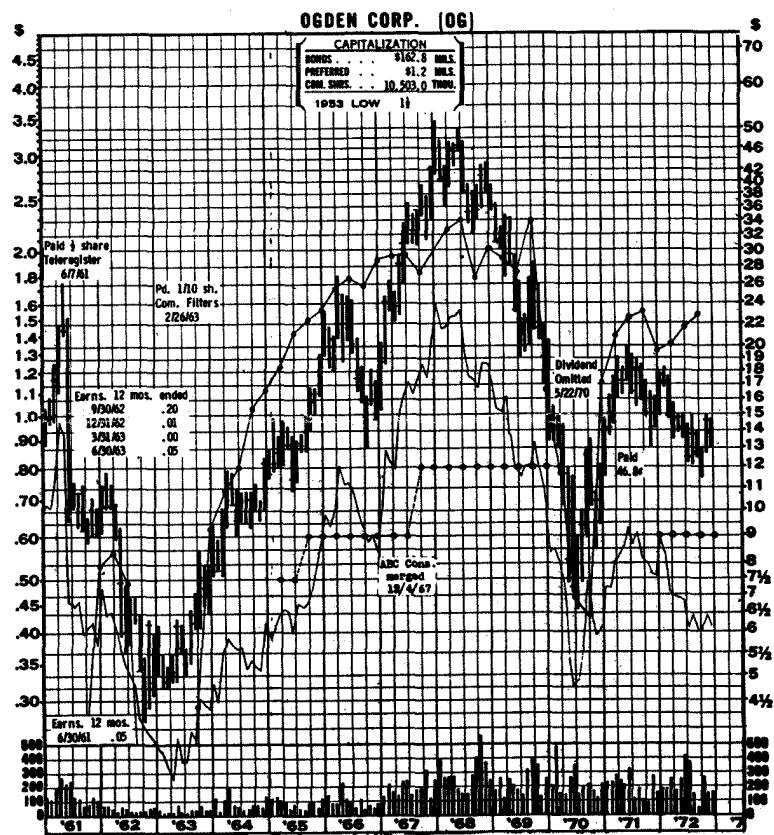
Events followed the sequence described in my model. Multiples expanded and eventually reality could not sustain expectations. More and more people realized the misconception on which the boom rested even as they continued to play the game. Acquisitions had to get larger and larger in order to maintain the momentum, and in the end they ran into the limits of size. The climactic event was the attempt by Saul Steinberg to acquire Chemical Bank: it was fought and defeated by the establishment.

When stock prices started to fall, the decline fed on itself. The favorable impact of acquisitions on per-share earnings diminished and eventually it became impractical to make new acquisitions. The internal problems that had been swept under the carpet during the period of rapid external growth began to surface. Earnings reports revealed unpleasant surprises. Investors became disillusioned and managements went through their own crises: after the heady days of success, few were willing to buckle down to the burdens of day-to-day management. The situation was aggravated by recession, and many of the high-flying conglomerates literally disintegrated. Investors were prepared to believe the worst and in some cases the worst actually occurred. In others, reality turned out to be better than expectations and eventually the situation stabilized with the surviving companies, often under new management, slowly working themselves out from under the debris.

The conglomerate boom is particularly well suited to serve as an illustration of my rudimentary model because the "fundamentals" are readily quantified. Investors based their valuations on reported per-share earnings. However meaningless the figures, they provide charts that closely conform to my theoretical prototype. Here they are:

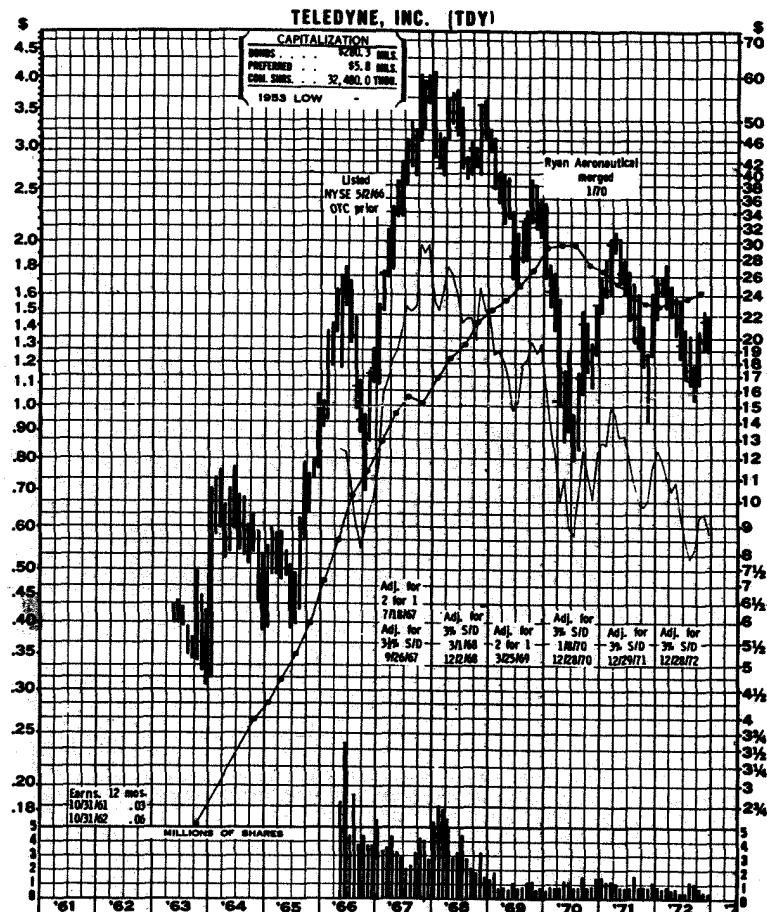


Courtesy of Securities Research Company, a Division of Babson-United Investment Advisors, Inc.,
208 Newbury St., Boston, MA 02116.



Courtesy of Securities Research Company, a Division of Babson-United Investment Advisors, Inc.,
208 Newbury St., Boston, MA 02116.

Theory



Courtesy of Securities Research Company, a Division of Babson-United Investment Advisors, Inc.,
208 Newbury St., Boston, MA 02116.

My best documented encounter with a boom/bust sequence is that of Real Estate Investment Trusts. REITs, as they are called, are a special corporate form brought into existence by legislation. Their key feature is that they can distribute their income free of corporate taxation, provided they distribute all the income they receive. The opportunity created by this legislation remained largely unexploited until 1969 when a large number of REITs were founded. I was present at the creation and, fresh from my experience with conglomerates, recognized their boom/bust potential. I published a research report whose key part reads as follows:

THE CASE FOR MORTGAGE TRUSTS
(February 1970)

THE CONCEPT

Superficially, mortgage trusts seem to resemble mutual funds designed to provide high current yields. But the analogy is misleading. The true attraction of mortgage trusts lies in their ability to generate capital gains for their shareholders by selling additional shares at a premium over book value. If a trust with a book value of \$10 and a 12% return on equity doubles its equity by selling additional shares at \$20, the book value jumps to \$13.33 and per-share earnings go from \$1.20 to \$1.60.

Investors are willing to pay a premium because of the high yield and the expectation of per-share earnings growth. The higher the premium, the easier it is for the trust to fulfill this expectation. The process is a self-reinforcing one. Once it gets under way, the trust can show a steady growth in per-share earnings despite the fact that it distributes practically all its earnings as dividends. Investors who participate in the process early enough can enjoy the compound benefits of a high return on equity, a rising book value, and a rising premium over book value.

ANALYTICAL APPROACH

The conventional method of security analysis is to try and predict the future course of earnings and then to estimate the price that investors may be willing to pay for those earnings. This method is inappropriate to the analysis of mortgage trusts because the price that investors are willing to pay for the shares is an important factor in determining the future course of earnings.

Instead of predicting future earnings and valuations separately, we shall try to predict the future course of the entire self-reinforcing process. We shall identify three major factors which reinforce each other and we shall sketch out a scenario of the probable course of development. The three factors are:

1. The effective rate of return on the mortgage trust's capital
2. The rate of growth of the mortgage trust's size
3. Investor recognition, i.e., the multiple investors are willing to pay for a given rate of growth in per-share earnings

THE SCENARIO

Act One: At present, the effective yield on construction loans is at an optimum. Not only are interest rates high but losses are at a relatively low level. There is a pent-up demand for housing and new houses readily find buyers. There is a shortage of funds so that the projects which do get off the ground are economically well justified. Builders who are still in business are more substantial and more reliable than at the tail end of a boom. Moreover, they do their utmost to complete the construction phase as fast as possible because money is so expensive. Shortages of labor and material do cause defaults and delays but rising costs permit mortgage trusts to liquidate their commitments without loss.

Money is tight and alternative sources of interim financing are in short supply. Investor recognition of the mortgage trust concept has progressed far enough to permit the formation of new trusts and the rapid expansion of existing ones. The self-reinforcing process gets under way.

Act Two: If and when inflation abates, the effective yield on construction loans will decline. On the other hand, there will be a housing boom and bank credit will be available at advantageous rates. With higher leverage, the rate of return on equity can be maintained despite a lower effective yield. With a growing market and growing investor recognition, the premium over book value may continue to increase. Mortgage trusts are likely to take full advantage of the premium and show a rapid rise in both size and per-share earnings. Since entry into the field is unrestricted, the number of mortgage trusts will also increase.

Act Three: The self-reinforcing process will continue until mortgage trusts have captured a significant part of the construction loan market. Increasing competition will then force them to take greater risks. Construction activity itself will have become more speculative and bad loans will increase. Eventually, the housing boom will slacken off and housing surpluses will appear in various parts of the country, accompanied by a slack real estate market and temporary declines in real estate prices. At this point, some of the mortgage trusts will be bound to have a large number of delinquent loans in their portfolios and the banks will panic and demand that their lines of credit be paid off.

Act Four: Investor disappointment will affect the valuation of the group, and a lower premium coupled with slower growth will in

turn reduce the per-share earnings progression. The multiple will decline and the group will go through a shakeout period. After the shakeout, the industry will have attained maturity: there will be few new entries, regulations may be introduced, and existing trusts will settle down to a more moderate growth.

EVALUATION

The shakeout is a long time away. Before it occurs, mortgage trusts will have grown manifold in size and mortgage trust shares will have shown tremendous gains. It is not a danger that should deter investors at the present time.

The only real danger at present is that the self-reinforcing process may not get under way at all. In a really serious stock market decline investors may be unwilling to pay any premium even for a 12% return on equity. We doubt that such conditions would arise; we are more inclined to expect an environment in which a 12% return is more exceptional than it has been recently and in which the self-reinforcing processes of the last few years, notably conglomerates and computer leasing companies, are going through their shakeout period. In such an environment there should be enough money available for a self-reinforcing process which is just starting, especially if it is the only game in town.

If the process fails to get under way, investors would find downside protection in the book value. The new trusts are coming to the market at book value plus underwriting commission (usually 10%). Most recently formed trusts are selling at a premium which is still modest. It will be recalled that when their assets are fully employed in interim loans, mortgage trusts can earn 11% on their book without leverage and 12% with a 1:1 leverage. A modest premium over book value would seem justified even in the absence of growth.

If the self-reinforcing process does get under way, shareholders in well-managed mortgage trusts should enjoy the compound benefits of a high return on equity, a rising book value, and a rising premium over book value for the next few years. The capital gains potential is of the same order of magnitude as at the beginning of other self-reinforcing processes in recent stock market history.

My report had an interesting history. It came at a time when go-go fund managers had suffered severe losses in the collapse of the conglomerates. Since they were entitled to share in the profits but did not have to share in the losses of the funds they managed,

they were inclined to grasp at anything that held out the prospect of a quick profit. They instinctively understood how a self-reinforcing process works since they had just participated in one and they were anxious to play. The report found a tremendous response whose extent I realized only when I received a telephone call from a bank in Cleveland asking for a new copy because theirs had gone through so many Xerox incarnations that it was no longer legible. There were only a few mortgage trusts in existence at the time but the shares were so eagerly sought after that they nearly doubled in price in the space of a month or so. Demand generated supply and there was a flood of new issues coming to market. When it became clear that the stream of new mortgage trusts was inexhaustible, prices fell almost as rapidly as they had gone up. Obviously the readers of the report failed to take into account the ease of entry and their mistake was corrected in short order. Nevertheless their enthusiastic reception helped to get the self-reinforcing process described in the report under way. Subsequent events took the course outlined in the report. Mortgage trust shares enjoyed a boom that was not as violent as the one that followed the publication of the report but turned out to be more enduring.

I had invested heavily in mortgage trusts and took some profits when the reception of my study exceeded my expectations. But I was sufficiently carried away by my own success to be caught in the downdraft with a significant inventory. I hung on and even increased my positions. I followed the industry closely for a year or so and sold my holdings with good profits. Then I lost touch with the group until a few years later when problems began to surface. I was tempted to establish a short position but was handicapped in that I was no longer familiar with the terrain. Nevertheless, when I reread the report I had written several years earlier, I was persuaded by my own prediction. I decided to sell the group short more or less indiscriminately. Moreover, as the shares fell I maintained the same level of exposure by selling additional shares short. My original prediction was fulfilled and most REITs went broke. The result was that I reaped more than 100% profit on my short positions—a seeming impossibility since the maximum profit on a short position is 100%. (The explanation is that I kept on selling additional shares.)

Self-reinforcing/self-defeating cycles like the conglomerate boom and the REITs do not occur every day. There are long fallow

periods when the specialist in such cycles remains unemployed. He need not starve, however. The divergence between underlying trends and investor recognition persists at all times and the astute investor can take advantage of it. New industries arise, or old ones come back into favor. Typically, they are inadequately followed at first. For instance, when defense spending started to rise in the early 1970s after a long decline, there were only two or three analysts left who followed the industry, although it still represented a substantial portion of the economy. Those who were left were too demoralized to recognize the beginning of a major new trend. That was a wonderful time to invest in defense stocks. There were high-tech defense stocks that had never been visited by an analyst, like E-Systems, Inc., and well-established companies that had fallen on evil days trying to diversify out of defense, like Sanders Associates, or getting caught up in scandals trying to sell airplanes through bribery, like Northrop and Lockheed.

In the case of defense stocks, there was no self-reinforcing process involved but investor recognition certainly helped the price of the stocks. Actually, it is a rare case where the investors' bias leaves the fundamentals totally unaffected. Even with defense stocks the prevailing bias played a role, but it was a negative one. Lockheed had to be bailed out by the government and companies like Sanders Associates had to restructure their debt by offering convertible bonds at prices that turned out to be very low in retrospect. Only when the negative bias was corrected was there very little further feedback: companies had little need for additional capital, and managements, having been burned once before, were leery of diversifying out of defense. There were exceptions, like United Aircraft, but investors' bias never turned positive enough to allow a self-reinforcing process to get under way: many of United Aircraft's acquisitions were for cash and those that involved stock did not enhance earnings significantly. The result was a larger, diversified company, but no boom and bust in the stock.

Perhaps the most interesting case of negative bias occurred in technology stocks. After the stock market debacle of 1974, investors were leery of any company that needed to raise equity capital from outside sources. Distributed data processing was in the early stages of its development. New companies like Datapoint and Four-Phase were in the vanguard with IBM lagging badly. The market was practically exploding but these little companies were

hamstrung by their inability to raise capital. The stocks were selling at very low multiples of anticipated earnings and the main argument against them was that they would not be able to grow fast enough to meet the demand and eventually IBM would move into the market. The argument turned out to be valid, but not before these companies became large and prosperous and investors became eager to throw money at them at high multiples. Those who had been willing to fight the negative bias were amply rewarded.

As the various niches occupied by these small companies converged to form a large market, most of them were absorbed by larger companies and those that stayed independent fell on evil days. Datapoint is currently looking for a home at a much reduced multiple. Four-Phase was recently acquired by Motorola, which proceeded to lose its shirt on it. If the initial market reaction to distributed data-processing companies had been more positive, it is possible that some of the early starters might have grown fast enough to survive, just as the earlier wave of microcomputer manufacturers did spawn a few enduring companies like Digital Equipment and Data General.

The negative bias of the 1975-1976 period gave way to the opposite extreme. It found expression in a venture capital boom that culminated in the second quarter of 1983. The sequence of events is not as clear cut as in the case of REITs, but that is only because high technology is not as homogeneous an industry. The same reflexive interaction between stock prices, prevailing bias, and fundamentals can be observed but much more specific knowledge is required to trace the course of events.

The availability of venture capital on attractive terms led to a proliferation of new ventures. Every new company needed equipment, as well as inventory, so that electronic equipment manufacturers enjoyed a boom, and so did the manufacturers of products and components. The electronics industry is a large customer of its own products so that the boom was self-reinforcing. But the proliferation of companies intensified competition. Industry leaders lost their market position as a new generation of products was introduced because the individuals who were responsible for developing them left their companies and set up new ones. Instead of companies growing in step with their industry, the industry grew by the multiplication of companies. Investors failed to recognize this trend; as a result, technology stocks in general and new issues in particular became substantially overvalued.

The new issue boom culminated in the second quarter of 1983. When prices started to decline, fewer issues could be sold and eventually venture capitalists became less venturesome. As fewer companies were formed and the existing ones depleted their cash, the market for technology products softened. Competition intensified and profit margins deteriorated. The process started to feed on itself and the low point has probably still not been reached.*

The venture capital boom was not the only cause of the subsequent shakeout—the strong dollar and the rise of Japanese competition were at least as important—but it is clear that stock prices had an impact on the “fundamentals” in both directions.

What distinguishes the conglomerate and REIT sequences from the venture capital boom is that in the first two cases the underlying trend itself was based on the exploitation of the investors’ bias while in the third it was not. In the case of conglomerates the idea was to acquire other companies with inflated paper; in the case of REITs the idea was equity leveraging. The idea behind the latest generation of technology products had nothing to do with the stock market.

To understand the ups and downs of technology stocks we must know something about the underlying trends in technology; in the case of conglomerates and REITs we need to know little else than the theory of reflexivity.

It is important to realize, however, that knowing everything about underlying trends in technology is not sufficient to explain the ups and downs of technology stocks: the reflexive interaction between underlying trends, prevailing biases, and stock prices also needs to be understood. Combining the two kinds of understanding is extremely difficult. Those who want to be familiar with technology must follow the industry closely and continuously; those who want to exploit the divergence between perception and reality must move from group to group. Most technology experts are ignorant of reflexivity and tend to remain fully invested at all times. Their popularity and influence wax and wane in a reflexive fashion. After the recent decline in technology stocks, a new breed of analysts seems to be emerging who are overly sensitive to the importance of investors’ perceptions. After a decent interval it may be once again profitable to go against the prevailing bias and invest in technology stocks on the basis of fundamental trends.

* P.S., February 1987: No longer true after the current explosive rally.

I have always had a lot of difficulty investing in technology stocks because of the specialized knowledge required. Finally, I managed to gain a good insight into the computer industry during the 1975–1976 period and profited from the prevailing negative bias. I held on to my positions for a few years but then I sold them and lost my grip on the industry. In 1981 I made the mistake of not participating in a venture capital fund operated by one of the most successful venture capitalists of the period in the belief that the boom could not last long enough to allow investors to exit in time. In this, I was undoubtedly influenced by misgivings about the larger picture. In any event, his investors realized a good profit in 1983. By that time I was totally out of touch with technology stocks and the boom passed me by.

Even the conglomerate and REIT sequences were not totally self-contained. Extraneous developments, such as the level of economic activity, regulation, or specific events (e.g., the attempted takeover of Chemical Bank), played a crucial role in the conglomerate boom. In less "pure" sequences the importance of outside influences is even greater.

We are currently in the midst of another self-reinforcing/self-defeating cycle that will go down in history as the mergermania of the 1980s. Instead of inflated paper, it is cash that serves as the currency. The scale of transactions already dwarfs the conglomerate boom. Mergermania is but an element in a much larger ongoing historical drama whose ramifications reach far beyond the stock market and involve politics, exchange rates, monetary and fiscal policies, quirks of taxation, international capital movements, and many other developments.

I shall make an attempt at unraveling the ongoing historical drama, but that is not as simple as analyzing a more or less self-contained boom/bust sequence. The larger picture is full of reflexive interactions as well as nonreflexive fundamental trends. We need a more complex model that allows for the transition from one boom/bust sequence to another and for the coexistence of several reflexive processes at the same time.

Before I embark on such an ambitious project, I want to examine another market that is characterized by vicious and benign circles: the currency market.

3

REFLEXIVITY IN THE CURRENCY MARKET*

While reflexive interactions are intermittent in the stock market, they are continuous in the market for currencies. I shall try to show that freely floating exchange rates are inherently unstable; moreover, the instability is cumulative so that the eventual breakdown of a freely floating exchange rate system is virtually assured.

The traditional view of the currency market is that it tends toward equilibrium. An overvalued exchange rate encourages imports and discourages exports until equilibrium is reestablished. Similarly, an improvement in competitive position is reflected in an appreciating exchange rate that reduces the trade surplus so that equilibrium is again reestablished. Speculation cannot disrupt the trend toward equilibrium—if speculators anticipate the future correctly, they accelerate the trend; if they misjudge it, they will be penalized by the underlying trend that may be delayed but will inexorably assert itself.

Experience since floating exchange rates were introduced in 1973 has disproved this view. Instead of fundamentals determining exchange rates, exchange rates have found a way of influencing the fundamentals. For instance, a strong exchange rate discourages inflation: wages remain stable and the price of imports falls. When exports have a large import component, a country can remain competitive almost indefinitely in spite of a steady

* This chapter was written in April/May 1985 and revised in December 1986.

appreciation of its currency, as Germany demonstrated in the 1970s.

The fact is that the relationship between the domestic inflation rate and the international exchange rate is not unidirectional but circular. Changes in one may precede changes in the other, but it does not make sense to describe one as the cause and the other as the effect because they mutually reinforce each other. It is more appropriate to speak of a vicious circle in which the currency depreciates and inflation accelerates or of a benign circle where the opposite happens.

Vicious and benign circles are a far cry from equilibrium. Nevertheless, they could produce a state of affairs akin to equilibrium if the reflexive, mutually self-reinforcing relationship could be sustained indefinitely. But that is not the case. The self-reinforcing process tends to become more vulnerable the longer it lasts and eventually it is bound to reverse itself, setting in motion a self-reinforcing process in the opposite direction. A complete cycle is characterized by wide fluctuations not only in the exchange rate but also in interest rates, inflation, and/or the level of economic activity.

The participants' bias introduces an element of instability into the system. If the system had an innate tendency toward equilibrium the participants' bias could not disrupt it; at worst, it could introduce some random, short-term fluctuations. But when the causal connections are reflexive, the participants' bias may engender, sustain, or destroy a vicious or benign circle. Moreover, the prevailing bias takes on a life of its own as one of the constituent parts in a circular relationship. It finds expression in speculative capital movements that may serve as a counterweight to an imbalance in trade, allowing a trade surplus or deficit to exceed, both in size and in duration, the level that could have been sustained in its absence. When that happens speculation becomes a destabilizing influence.

International capital movements tend to follow a self-reinforcing/self-defeating pattern similar to the one we identified in the stock market. But the model we used for stock price movements cannot be applied to currency markets without substantial modifications. In the stock market we focused on the reflexive relationship between two variables: stock prices and a single underlying trend. We were trying to build the simplest possible model and we were willing to simplify a much more complex reality to serve

our purpose. In the currency market we cannot get by with two variables; even the simplest model will need seven or eight. We have selected four rates and four quantities, namely:

- e nominal exchange rate (number of foreign currency units for one domestic currency unit; $\uparrow e$ = strengthening)
- i nominal interest rate
- p domestic versus foreign price level ($\uparrow p$ = increase in domestic prices faster than in foreign prices and vice versa)
- v level of economic activity
- N nonspeculative capital flow } \uparrow = increased outflow
- S speculative capital flow } \downarrow = increased inflow
- T trade balance } \uparrow = surplus
- B government budget } \downarrow = deficit

Our task is to establish how these variables relate to each other. We shall not attempt to explore all the relationships but only those that are necessary to build simple models. In other words, we are not aiming at a general theory, only at a partial explanation of currency movements. Our focus is the exchange rate and we bring in the other variables only when we need them. We shall not quantify any of the variables but only indicate direction (\uparrow, \downarrow) or order of magnitude ($>, <$).

Before we start, two general observations can be made. One is that relationships tend to be circular; that is, variables can serve as both cause and effect in relation to other variables. We shall denote the causal connection by a horizontal arrow (\rightarrow). The other point is that the relationship of the variables need not be internally consistent; it is the inconsistencies that make that entire situation move in a certain direction, creating vicious or benign circles. Equilibrium would require internal consistency; historical change does not. Describing historical change in terms of vicious and benign circles is, of course, merely a figure of speech. A circular movement between component parts when the entire system is in motion could also be described as a spiral. Moreover, what is benign and what is vicious are in the eye of the beholder.

Exchange rates are determined by the demand and supply of currencies. For present purposes, we can group the various factors that constitute demand and supply under three headings: trade, nonspeculative capital transactions, and speculative capital

transactions. This gives us the simplest model of a freely floating exchange rate system:

$$(\downarrow T + \uparrow N + \uparrow S) \rightarrow \downarrow e$$

In other words, the sum of the currency transactions under the three headings determines the direction of the exchange rate.

Our primary interest is in investigating the role that the participants' bias plays in exchange rate movements. To facilitate the investigation, we shall assume that the bias finds expression only in speculative capital transactions (S), while trade (T) and non-speculative capital flows (N) are independent of expectations: they constitute the "fundamentals." In reality, the "fundamentals" are also influenced by the participants' expectations about the future course of exchange rates. Trade figures are notoriously distorted by leads and lags in payment, not to mention the effect of expectations on the inventory policy of exporters and importers. As far as capital movements are concerned, perhaps the only transaction that is totally independent of expectations is the payment of interest on accumulated debt; the reinvestment of interest receipts already qualifies as a speculative transaction. The repatriation of bank debt from less developed countries is probably best described as nonspeculative, although speculative considerations come into play if and when the assets are redeployed. What about direct investment? If managements were interested only in the total rate of return, it ought to be classified as speculative, but often there are overriding industrial considerations. It can be seen that there are many gradations between speculative and nonspeculative transactions; but we do not do any great violence to reality by putting them into these two broad categories.

We shall focus on speculative capital transactions because that is where the participants' bias finds expression. Speculative capital moves in search of the highest total return. Total return has three elements: the interest rate differential, the exchange rate differential, and the capital appreciation in local currency. Since the third element varies from case to case we can propose the following general rule: speculative capital is attracted by rising exchange rates and rising interest rates.

$$\uparrow(e + i) \rightarrow \downarrow S$$

Of the two, exchange rates are by far the more important. It does not take much of a decline in the currency to render the total return negative. By the same token, when an appreciating currency also offers an interest rate advantage, the total return exceeds anything that a holder of financial assets could expect in the normal course of events.

That is not to say that interest rate differentials are unimportant; but much of their importance lies in their effect on exchange rates and that depends on the participants' perceptions. There are times when relative interest rates seem to be a major influence; at other times they are totally disregarded. For instance, from 1982 to 1986 capital was attracted to the currency with the highest interest rate, namely, the dollar, but in the late 1970s Switzerland could not arrest the influx of capital even by imposing negative interest rates. Moreover, perceptions about the importance of interest rates are often wrong. For instance, until November 1984 the strength of the dollar was widely attributed to high interest rates in the United States. When interest rates declined without the dollar weakening this view was discredited and the dollar went through the roof.

Expectations about exchange rates play the same role in currency markets as expectations about stock prices do in the stock market: they constitute the paramount consideration for those who are motivated by the total rate of return. In the stock market that covers practically all investors, in currency markets all speculative transactions.

In the stock market we used a model that focused on stock prices and disregarded dividend income. No great distortion was involved because in the kind of boom/bust sequences we were considering stock price movements far outweigh dividend income. Similar conditions prevail in currency markets: expectations about future exchange rates constitute the main motivation in speculative capital transactions.

The major difference between the stock market and the currency market seems to be the role played by the fundamentals. We have seen that the "fundamentals" were rather nebulous even in the case of stocks but at least we had no reason to doubt that stock prices were somehow connected to the fundamentals. In the case of currencies the trade balance is clearly the most important fundamental factor, yet the dollar strengthened between 1982 and 1985 while the trade balance of the United States was deteriorat-

ing. It would seem that the fundamentals are even less relevant in determining price trends than in the stock market. We do not need to look far afield for an explanation: it is to be found in the relative importance of speculative capital movements.

As we have seen, speculative capital is motivated primarily by expectations about future exchange rates. To the extent that exchange rates are dominated by speculative capital transfers, they are purely reflexive: expectations relate to expectations and the prevailing bias can validate itself almost indefinitely. The situation is highly unstable: if the opposite bias prevailed, it could also validate itself. The greater the relative importance of speculation, the more unstable the system becomes: the total rate of return can flip-flop with every change in the prevailing bias.

In our discussion of the stock market we identified certain sequences such as the conglomerate boom where the prevailing bias formed an important part of the underlying trend, but we concluded that such pure examples of reflexivity are exceptional. By contrast, in a system of freely fluctuating exchange rates reflexivity constitutes the rule. Of course, there is no such thing as a purely reflexive situation. Speculation is only one of the factors that determine exchange rates and the other factors must also be taken into account in formulating one's expectations. Thus, expectations cannot be totally capricious: they must be rooted in something other than themselves. How a prevailing bias becomes established and, even more important, how it is reversed are the most important questions confronting us.

There are no universally valid answers. Reflexive processes tend to follow a certain pattern. In the early stages, the trend has to be self-reinforcing, otherwise the process aborts. As the trend extends, it becomes increasingly vulnerable because the fundamentals such as trade and interest payments move against the trend, in accordance with the precepts of classical analysis, and the trend becomes increasingly dependent on the prevailing bias. Eventually a turning point is reached and, in a full-fledged sequence, a self-reinforcing process starts operating in the opposite direction.

Within this general pattern each sequence is unique. It is the characteristic feature of a reflexive process that neither the participants' perceptions nor the situation to which they relate remain unaffected by it. It follows that no sequence can repeat itself. Not even the variables that interact in a circular fashion need be the

same; certainly they will not carry the same weight on different occasions.

We have had two major reflexive moves in the dollar since the breakdown of the Bretton Woods system and at least that many in sterling. It is instructive to compare the two big moves in the dollar because the interaction between the trade balance and capital movements was radically different in the two instances.

In the late 1970s the dollar got progressively weaker, especially against the continental currencies, while in the 1980s it got progressively stronger. We shall call the first move Carter's vicious circle and the second Reagan's benign circle. We can build simple models to show how different the two trends were.

In the case of Germany in the late 1970s, the German mark was strong ($\uparrow e$). Speculative purchases played a major role in making it stronger ($\downarrow S$) and sustaining the benign circle. Germany started with a trade surplus and the strength of the currency helped to keep the price level down. Since exports had a large import content the real exchange rate, as opposed to the nominal, remained more or less stable ($\downarrow(ep)$) and the effect on the trade balance was negligible ($\downarrow T$). With the speculative inflow predominating ($\downarrow S > \downarrow T$), the benign circle was self-reinforcing:

$$\uparrow e \rightarrow \downarrow p \rightarrow \downarrow(ep) \rightarrow (\downarrow T < \downarrow S) \rightarrow \uparrow e$$

The fact that the rate of currency appreciation exceeded the interest rate differential made it very profitable to hold German marks, so that the speculative inflow was both self-reinforcing and self-validating.

What was a benign circle for Germany was a vicious circle for the United States. As the exchange rate depreciated, inflation accelerated. Despite rising nominal interest rates, real rates remained very low, if not negative. Various measures were tried to compensate for the outflow of capital, of which the issue of so-called Carter bonds denominated in German marks and Swiss francs was the most dramatic, but nothing seemed to work until the Federal Reserve embraced a strict monetarist policy. Then came the election of Ronald Reagan to the presidency and the dollar embarked on a sustained rise.

During Reagan's benign circle the strong dollar caused a sharp deterioration in the trade balance of the United States. In contrast to Germany in the late 1970s, the United States did not have a

trade surplus to start with. Moreover, the appreciation in the currency was not matched by inflation rate differentials. The inflation rate declined in the United States but it remained low in other countries as well. As a consequence, the United States developed an unprecedented trade deficit as well as an unprecedented interest rate differential in favor of the dollar. It was extremely attractive to hold dollars as long as the dollar remained firm, and the dollar remained firm as long as the deficit on current account was fully matched by a surplus on capital account. In our notation:

$$(\uparrow e + \uparrow i) \rightarrow (\downarrow S > \downarrow T) \rightarrow \uparrow e \rightarrow (\downarrow S > \downarrow T)$$

The models are obviously oversimplified. We shall explore Reagan's benign circle in greater depth later. The point we are trying to make here is that different sequences have totally different structures. In the case of Germany in the late 1970s the appreciation of the currency was sustained by the inflation rate differential and the balance of trade was largely unaffected. Reagan's benign circle was sustained by a differential in interest rates rather than inflation rates and there was an ever-growing trade deficit which was matched by an ever-growing inflow of capital. While in the first case it was possible to claim some kind of equilibrium, in the second case the disequilibrium was palpable. The inflow of capital depended on a strong dollar and a strong dollar depended on an ever-rising inflow of capital which carried with it ever-rising interest and repayment obligations ($\uparrow N$). It was obvious that the benign circle could not be sustained indefinitely. Yet, while it lasted, any currency speculator who dared to fight the trend had to pay dearly for it. Speculation did not serve to reestablish equilibrium. On the contrary, it reinforced the trend and thereby increased the disequilibrium, which would eventually have to be corrected.

Although each self-reinforcing circle is unique, we can make some universally valid generalizations about freely fluctuating exchange rates. First, the relative importance of speculative transactions tends to increase during the lifetime of a self-reinforcing trend. Second, the prevailing bias is a trend-following one and the longer the trend persists, the stronger the bias becomes. The third is simply that once a trend is established it tends to persist and to run its full course; when the turn finally comes, it tends to

set into motion a self-reinforcing process in the opposite direction. In other words, currencies tend to move in large waves, with each move lasting several years.

These three tendencies are mutually self-validating. It is the growth in speculative capital flows moving in a trend-following fashion that makes the trend so persistent; it is the persistence of the trend that makes a trend-following bias so rewarding; and it is the rewards reaped by speculation that attract increasing amounts of capital.

The longer a benign circle lasts, the more attractive it is to hold financial assets in the appreciating currency and the more important the exchange rate becomes in calculating total return. Those who are inclined to fight the trend are progressively eliminated and in the end only trend followers survive as active participants. As speculation gains in importance, other factors lose their influence. There is nothing to guide speculators but the market itself, and the market is dominated by trend followers. These considerations explain how the dollar could continue to appreciate in the face of an ever-rising trade deficit. Eventually, a crossover point would have been reached, even without the intervention of the authorities, when the inflow of speculative funds could not keep pace with the trade deficit and with rising interest obligations, and the trend would have been reversed. Since the predominating bias is trend following, speculative capital would then have started moving in the opposite direction. If and when that happened, the reversal could easily have accelerated into a free fall. For one thing, speculation and "fundamental" flows would then have worked in the same direction. Even more important, when a change in trend is recognized, the volume of speculative transactions is likely to undergo a dramatic, not to say catastrophic, increase. While a trend persists, speculative flows are incremental; but a reversal involves not only the current flow but also the accumulated stock of speculative capital. The longer the trend has persisted, the larger the accumulation. There are, of course, mitigating circumstances. One is that market participants are likely to recognize a change in trend only gradually. The other is that the authorities are bound to be aware of the danger and do something to prevent a crash. How the drama actually unfolded will be the subject of a later chapter. Here we are trying to establish a general proposition.

Taking the three generalizations together, it can be asserted that

speculation is progressively destabilizing. The destabilizing effect arises not because the speculative capital flows must be eventually reversed but exactly because they need not be reversed until much later. If they had to be reversed in short order, capital transactions would provide a welcome cushion for making the adjustment process less painful. If they need not be reversed, the participants get to depend on them so that eventually when the turn comes the adjustment becomes that much more painful.

It is quite likely that the generalization about the progressive accumulation of hot money holds true not only within a cycle but also from one cycle to another, although the history of fluctuating exchange rates is too short to provide reliable evidence. It has certainly been true so far—the size of speculative capital movements was far greater in Reagan's benign circle than it was during Carter's vicious circle. Empirical studies of the 1930s also showed a cumulative growth in "hot money" movements,¹ although circumstances were somewhat different because currencies were not freely floating.

We can see why hot money should continue to accumulate as long as real interest rates are high and the return on physical investments low: keeping capital in liquid form in an appreciating currency is more rewarding than investing it in physical assets. What is needed to give the generalization universal validity is an argument that would show that fluctuating exchange rates are associated with high returns on financial assets and low returns on physical investments. Let me try. We have seen that hot money can earn exceptional returns if it gets the trend right; since it sets the trend, that is likely to be the case. Physical assets represent the opposite side of the coin: they cannot move to take advantage of the trend. The tradable goods sector is bound to suffer when a currency appreciates. Of course, a depreciating currency brings windfall profits to exporters, but having been hurt before, exporters are loath to invest on the basis of a temporary advantage: they prefer to hold their profits in financial assets, contributing to the growth of hot money. The process can be most clearly observed in the United Kingdom, where exporters refused to expand when sterling fell below \$1.10 in 1985, despite record profits. How right they were! Sterling rose above \$1.50 by April 1986. Thus, both an appreciating currency and a depreciating currency discourage physical investment and foster the accumulation of "hot money."

We can attempt yet another tentative generalization. When a long-term trend loses its momentum, short-term volatility tends to rise. It is easy to see why that should be so: the trend-following crowd is disoriented. The generalization is tentative because it is based on inadequate evidence. It certainly was true when the dollar reversed its trend in 1985.²

If these generalizations are indeed valid, the eventual demise of a system of freely fluctuating exchange rates is inevitable. Fluctuations become so wild that either the system has to be modified by some kind of government intervention or it is bound to break down. Currency markets thus provide the best support for my contention that financial markets are inherently unstable. There is no built-in tendency toward equilibrium: to the extent that we need stability we must introduce it by deliberate policy measures.

These conclusions may not strike the reader as particularly revolutionary at the present time, but they certainly contradicted the prevailing wisdom at the time they were written in April/May 1985. There was widespread malaise about the instability of exchange rates, but belief in the magic of the market was still running strong, and the famous Plaza agreement in September 1985 came as something of a shock to market participants. Even today, there is no theoretical underpinning for the contention that a freely floating exchange rate system is cumulatively destabilizing. That is what I hope to have provided here.

I have been speculating in currencies ever since they started floating, but I have failed to make money on a consistent basis. On balance, I traded profitably through 1980 and then chalked up losses between 1981 and 1985. My approach has been tentative, based more on intuition than on conviction. By temperament, I have always been more interested in picking the turning point than in following a trend. I managed to catch both the rise and fall of European currencies against the dollar until 1981, but I traded myself out of my positions too soon. Having lost the trend, I found it too demeaning to start following the trend-followers; I tried to pick the reversal point instead—needless to say, without success. I had some temporary profits in the early part of 1984, but I gave them all back. I was again engaged in a speculation against the dollar at the time I wrote this chapter (April/May 1985). Writing it has undoubtedly helped to clarify my thoughts.

The real-time experiment recorded in Part III may be regarded

as a practical test of the theory propounded here. Admittedly the theory is far too abstract to be of much use in making concrete predictions. Specifically, the turning point cannot be determined until it has actually occurred. But, as we shall see, the theory can be very useful in interpreting events as they unfold.

4

THE CREDIT AND REGULATORY CYCLE *

There seems to be a special affinity between reflexivity and credit. That is hardly surprising: credit depends on expectations; expectations involve bias; hence credit is one of the main avenues that permit bias to play a causal role in the course of events. But there is more to it. Credit seems to be associated with a particular kind of reflexive pattern that is known as boom and bust. The pattern is asymmetrical: the boom is drawn out and accelerates gradually; the bust is sudden and often catastrophic. By contrast, when credit is not an essential ingredient in a reflexive process, the pattern tends to be more symmetrical. For instance, in the currency market it does not seem to make much difference whether the dollar is rising or falling: the exchange rate seems to follow a wavelike pattern.

I believe the asymmetry arises out of a reflexive connection between loan and collateral. In this context I give collateral a very broad definition: it will denote whatever determines the creditworthiness of a debtor, whether it is actually pledged or not. It may mean a piece of property or an expected future stream of income; in either case, it is something on which the lender is willing to place a value. Valuation is supposed to be a passive relationship in which the value reflects the underlying asset; but in this case it involves a positive act: a loan is made. The act of lending may affect the collateral value: that is the connection that gives rise to a reflexive process.

* Written in August 1985.

It will be recalled that we have analyzed reflexivity as two connections working in opposite directions: the "normal" connection where a value is placed on future events, as in the stock market or banking—we have called it the cognitive function; and a "perverse" connection in which expectations affect that which is expected—we have called it the participating function. The participating function is perverse because its effect is not always felt, and when it does operate its influence is so difficult to disentangle that it tends to go unrecognized. The prevailing view of how financial markets operate tends to leave the participating function out of account. For instance, in the international lending boom, bankers did not recognize that the debt ratios of borrowing countries were favorably influenced by their own lending activity. In the conglomerate boom, investors did not realize that per-share earnings growth can be affected by the valuation they place on it. At present, most people do not realize that the erosion of collateral values can depress the economy.

The act of lending usually stimulates economic activity. It enables the borrower to consume more than he would otherwise, or to invest in productive assets. There are exceptions, to be sure: if the assets in question are not physical but financial ones, the effect is not necessarily stimulative. By the same token, debt service has a depressing impact. Resources that would otherwise be devoted to consumption or the creation of a future stream of income are withdrawn. As the total amount of debt outstanding accumulates, the portion that has to be utilized for debt service increases. It is only net new lending that stimulates, and total new lending has to keep rising in order to keep net new lending stable.

The connection between lending and economic activity is far from straightforward (that is, in fact, the best justification for the monetarists' preoccupation with money supply, to the neglect of credit). The major difficulty is that credit need not be involved in the physical production or consumption of goods and services; it may be used for purely financial purposes. In this case, its influence on economic activity becomes problematic. For purposes of this discussion it may be helpful to distinguish between a "real" economy and a "financial" economy. Economic activity takes place in the "real" economy, while the extension and repayment of credit occur in the "financial" economy. The reflexive interaction between the act of lending and the value of the collateral may

then connect the "real" and the "financial" economy or it may be confined to the "financial" economy. Here we shall focus on the first case.

A strong economy tends to enhance the asset values and income streams that serve to determine creditworthiness. In the early stages of a reflexive process of credit expansion the amount of credit involved is relatively small so that its impact on collateral values is negligible. That is why the expansionary phase is slow to start with and credit remains soundly based at first. But as the amount of debt accumulates, total lending increases in importance and begins to have an appreciable effect on collateral values. The process continues until a point is reached where total credit cannot increase fast enough to continue stimulating the economy. By that time, collateral values have become greatly dependent on the stimulative effect of new lending and, as new lending fails to accelerate, collateral values begin to decline. The erosion of collateral values has a depressing effect on economic activity, which in turn reinforces the erosion of collateral values. Since the collateral has been pretty fully utilized at that point, a decline may precipitate the liquidation of loans, which in turn may make the decline more precipitous. That is the anatomy of a typical boom and bust.

Booms and busts are not symmetrical because, at the inception of a boom, both the volume of credit and the value of the collateral are at a minimum; at the time of the bust, both are at a maximum. But there is another factor at play. The liquidation of loans takes time; the faster it has to be accomplished, the greater the effect on the value of the collateral. In a bust, the reflexive interaction between loans and collateral becomes compressed within a very short time frame and the consequences can be catastrophic. It is the sudden liquidation of accumulated positions that gives a bust such a different shape from the preceding boom.

It can be seen that the boom/bust sequence is a particular variant of reflexivity. Booms can arise whenever there is a two-way connection between values and the act of valuation. The act of valuation takes many forms. In the stock market, it is equity that is valued; in banking, it is collateral. It is possible, although unlikely, that a boom could be generated without any credit expansion. The two examples we studied in the stock market, the REIT and conglomerate booms, could, in theory, have unfolded without the stocks being used as collateral, although in practice there

was a lot of credit involved. In the absence of credit the reversal would be a more gradual process. The contraction would not be a mirror image of the expansion for the reason mentioned earlier—the reflexive element in valuations is greater at the time of reversal than at the inception of a trend—but the compression that is characteristic of busts would be absent.

Both the boom/bust pattern and its explanation are almost too obvious to be interesting. The amazing thing is that the reflexive connection between lending and collateral has not been generally recognized. There is an enormous literature on the trade cycle, but I have not seen much awareness of the reflexive relationship described here. Moreover, the trade cycles that are generally discussed in textbooks differ in duration from the credit cycle I am discussing here: they are short-term fluctuations within a larger pattern. There is an awareness of a larger cycle, usually referred to as the Kondratieff wave, but it has never been "scientifically" explained. At present, there is much concern that we may be approaching another recession but the general assumption is that we are dealing with a recession just like any other; the fact that we are in the declining phase of the larger cycle is usually left out of account. I contend that all previous recessions since the end of World War II occurred while credit was expanding, while the one we may or may not be facing now would occur when borrowing capacity in the real economy is contracting. This creates a situation that has no precedent in recent history.

Exactly where we are in the larger cycle is difficult to determine. I must confess I have been confused on the issue since 1982. The reason for my confusion is that while the boom has clearly run out of steam, the bust has not taken place.

Busts can be very disruptive, especially if the liquidation of collateral causes a sudden compression of credit. The consequences are so unpleasant that strenuous efforts are made to avoid them. The institution of central banking has evolved in a continuing attempt to prevent sudden, catastrophic contractions in credit. Since a panic is hard to arrest once it has started, prevention is best practiced in the expansionary phase. That is why the role of central banks has gradually expanded to include the regulation of the money supply. That is also why organized financial markets regulate the ratio of collateral to credit.

Until now, the authorities have been able to prevent a bust. We find ourselves in a twilight zone where the "normal" process of

credit expansion culminated long ago but the "normal" process of credit contraction has been prevented by the authorities. We are in uncharted territory because the actions of the authorities have no precedent.

The fact that banks and organized financial markets are regulated complicates the course of events tremendously. Financial history is best interpreted as a reflexive process in which there are two sets of participants instead of one: competitors and regulators.

Such a system is much more complex than the case we studied in the stock market. There, the regulatory environment was more or less fixed: it was the backdrop against which the drama was acted out. Here, the regulatory environment is an integral part of the process.

It is important to realize that the regulators are also participants. There is a natural tendency to regard them as superhuman beings who somehow stand outside and above the economic process and intervene only when the participants have made a mess of it. That is not the case. They also are human, all too human. They operate with imperfect understanding and their activities have unintended consequences. Indeed, they seem to adjust to changing circumstances less well than those who are motivated by profit and loss, so that regulations are generally designed to prevent the last mishap, not the next one. The deficiencies of regulation tend to be more noticeable when conditions are rapidly changing, and conditions tend to change more rapidly when the economy is less regulated.

We begin to discern a reflexive relationship between the regulators and the economy they regulate. It gives rise to a process that takes place concurrently with the process of credit expansion and contraction and interacts with it. It is no wonder that the result is so complex and perplexing!

The regulatory cycle does not have the asymmetric character of the credit cycle. It seems to fit the wave pattern we developed for the currency market better than the boom and bust pattern of the credit cycle. Just as freely floating currencies tend to fluctuate between over- and undervaluation, so market economies tend to fluctuate between over- and underregulation. The length of the cycle seems to be correlated with the credit cycle and one can sense intuitively why that should be so. Credit expansion and contraction have much to do with changes in the economy,

which in turn have a bearing on the adequacy of regulations. Conversely, the regulatory environment influences not only how fast credit can expand but also how far. Clearly, there is a two-way connection between credit and regulation, but it is far from clear to me at this stage of the investigation what pattern, if any, the interaction follows. That is the main source of my confusion.

We have identified a credit cycle that follows a boom/bust pattern: a regulatory cycle that is more wavelike, and an interplay between the two whose pattern is unclear. There are, of course, many secular developments involved as well, some of which relate to credit, some to regulation, and some to both. We have mentioned that central banks tend to get stronger after each crisis; that is a secular development that renders each cycle unique. In the Great Depression both the banking system and the international trading system collapsed, making the contraction of credit and economic activity much more severe than it would have been otherwise. We can be certain that every effort will be made to avoid a similar collapse in this cycle. We did not dwell on the information revolution or on the improvements in transportation that have helped the development of an integrated world economy. The outcome of all these influences is a unique course of events that is easier to explain than to predict.

Looked at from this perspective, the entire postwar period is part of a large expansionary boom that is now well advanced and ripe for a bust. The bust has been avoided, however, by the intervention of the authorities at crucial moments. The interplay between government actions and the market mechanism has given rise to a unique constellation that I have called Reagan's Imperial Circle. We are now at the crucial moment where the Imperial Circle is beginning to unwind and the authorities must invent another solution to forestall a bust.

The same postwar period has also seen an almost complete swing from government regulation to unrestrained competition. We are now at a curious moment where the bias in favor of deregulation is still on the rise yet the need for government intervention in specific areas is beginning to reassert itself. The banking industry, for one, is already on the way to becoming more regulated.

One could try and write a history of the postwar era in these terms. The present credit cycle started after the end of World War II; the origins of the regulatory cycle go back even farther to the

New Deal, although the creation of the Bretton Woods system can be taken as the starting point as far as the world economy is concerned. The expansion that followed was intimately related to the removal of restrictions on international trade and investment. But international capital movements created problems for the Bretton Woods system that were never anticipated and remain unresolved.

I shall not attempt to present the complete story here. I shall start at the point where I became actively involved and I shall follow the path of my own involvement. This will give the investigation a more experimental character.

My experiences began after the breakdown of the system of fixed exchange rates in 1973. Relationships that had been fixed became subject to reflexive influences and my interest veered from specific companies and industries to macroeconomic processes. My study of "growth banks" in 1972 constituted a transition point, although I did not know it at the time.

With the passage of time I found the instability of macroeconomic trends increasingly disturbing in both a subjective and an objective sense and I decided to distance myself from active investing in 1981. After the crisis of 1982 I made a theoretical study of the international debt problem. I was under the mistaken impression that the crisis of 1982 constituted the climax in the process of credit expansion. I thought that the authorities were not doing enough to prevent a bust; I failed to realize that they were doing too much. They actually kept credit expanding, albeit on a more unsound basis than ever. The United States replaced less developed countries as the "borrower of last resort" and commercial banks tried to grow their way out of the loans made to less developed countries by aggressively expanding in other directions. This led to another series of crises in 1984 that constituted the real turning point for the banking and thrift industries. We are now suffering from the aftereffects of that climax. The United States government continues to borrow on an ever-increasing scale, but here, too, a turning point is at hand. The dollar has begun to decline and foreigners are going to be repaid in a depreciating currency. Perhaps the last great engine of credit creation that is still going full blast is in the stock market where mergermania is at its peak; but it does not have a stimulating effect on the real economy.

The theoretical framework of an interrelated credit and regula-

tory cycle became somewhat clearer to me in the course of writing this book, although I cannot claim that the process of clarification is complete. Nevertheless, I felt it was appropriate to summarize it at this point. I shall now use the theoretical framework outlined here to explain the course of events since 1972. The reader should be advised that the writing of the explanation predates the formulation of this admittedly tentative theoretical framework.

P. S., DECEMBER 1986

Having completed the explanation, I engaged in a real-time experiment from August 1985 to the end of 1986, trying to predict the evolution of the credit and regulatory cycle. I came to a strange conclusion. It appears that the cycle got stuck in 1982. If it had not been for the successful intervention of the monetary authorities, the international debt crisis would have led to a collapse of the banking system. As it is, the collapse has been avoided, but the genuine trend reversal which it would have ushered in also failed to take place. We now live in a system where we continually go to the brink and then recoil when we see the abyss opening up at our feet. The cohesion we manage to muster in the sight of a disaster tends to disintegrate as soon as the danger recedes, and the process then repeats itself in different forms. We can observe it in international lending; in the U.S. budget deficit; in the international monetary system; in OPEC; in the banking system; in financial markets; and 1987 will undoubtedly be the year when protectionism drives the international trade system to the brink of collapse but probably not beyond it.