
WHY WAGES DON'T FALL
DURING A RECESSION

TRUMAN F. BEWLEY

HARVARD UNIVERSITY PRESS
Cambridge, Massachusetts, and London, England 1999

1

INTRODUCTION

Why have money wages and salaries seldom declined during post-World War II recessions in the United States and abroad, despite high unemployment and intense competition for jobs? Instead, market pay rates continue to rise during downturns, albeit at a slower rate than during economic booms. Why don't labor markets behave like competitive commodity markets, where prices fall or even plunge when supply exceeds demand? Why do few firms avoid layoffs by cutting pay and lowering product prices so as to increase sales? How can the frequency of layoffs be reconciled with the movement within the business community to treat workers humanely? The failure of pay rates to fall is termed wage stickiness, downward wage rigidity, or simply wage rigidity, and has puzzled economists for years.

I have sought explanations by interviewing more than 300 businesspeople, labor leaders, counselors of the unemployed, and business consultants in the Northeast of the United States during the recession of the early 1990s. It is unusual for economists to do surveys of any sort, and I undertook this one because I could think of no other way to answer my questions. From the interviews, I conclude that wage rigidity stems from a desire to encourage loyalty, a motive that superficially seems incompatible with layoffs. My findings support none of the existing economic theories of wage rigidity, except those emphasizing the impact of pay cuts on morale. Other theories fail in part because they are based on the unrealistic psychological assumptions that people's abilities do not depend on their state of mind and that they are rational in the simplistic sense that they maximize a utility that depends only on their own consumption and working conditions, not on the welfare of others. Wage rigidity is the product of more complicated employee behavior, in the face of which manager reluctance to cut pay is rational. Worker behavior, however

is not always completely rational, though reasonable and understandable. A model that captures the essence of wage rigidity must take into account the capacity of employees to identify with their firm and to internalize its objectives. This internalization and workers' mood have a strong impact on job performance and call for material, moral, and symbolic reciprocation from company leadership.

1.1 Clarification of the Question

The issue of wage rigidity should not be confused with the questions of why unemployment exists at all and why it increases during recessions. All students of the labor market accept as normal a minimal frictional level of unemployment. What is hard to understand is why increases in unemployment have so little impact on labor compensation.

Wage rigidity does not necessarily help explain why unemployment rises during recessions, for wage flexibility might not prevent it from doing so. Though a single firm might employ more workers at lower wages, it does not follow that the same is true for all firms together. This conclusion would involve the fallacy of composition, for, as John Maynard Keynes (1936, chap. 19) pointed out, wage cuts in response to high unemployment could have indirect effects that would decrease real aggregate demand and increase unemployment. For instance, deflation would increase the burden of debt, and sustained deflation could raise real interest rates enough to discourage investment.

1.2 The Topic Is Controversial

It is hard to take a detached view of wage rigidity, because it requires facing unpleasant truths; the world is imperfect, people often suffer for no good reason, and it is not clear what to do about it. Furthermore, the subject is central to intense controversy between Keynesian and neoclassical macroeconomists over whether government economic policy should be used to stabilize aggregate income and employment. The debate has important political implications, creating a highly charged atmosphere that colors thinking about wage formation and unemployment. Economists on the Keynesian side of the dispute claim that wage rigidity is confirmed by statistical evidence that pay rates almost never fall. They say that labor markets do not automatically adjust to eliminate excess unemployment and that joblessness is a grave misfortune forced on people, most of whom want to work, even at wages lower than those earned previously. Keynesians think recessions are caused by declines in aggregate demand and should be cured by expansionary economic policy.¹

Those on the neoclassical side of the controversy assert that wage rigidity

is an illusion, that wages and salaries are flexible, and that labor markets always clear.² A central tenet of these beliefs is that the existing rate of unemployment is the optimal outcome of market forces and should not or cannot be affected by government policy. The view is that during recessions pay rates fall below reservation levels, or the minimum at which people are willing to work, and this decline causes workers to leave their jobs and become unemployed. Neoclassicists also assert that anyone can find some job quickly and that people remain unemployed because they want higher pay than is available to them.³ There are two main explanations of this behavior. According to the intertemporal substitution theory of Lucas and Rapping (1969), workers stop working because they want to enjoy leisure while it is unusually cheap in terms of forgone wage income. The market misperceptions theory of the same authors and of Lucas (1972) asserts that people leave their jobs to have more time to look for better-paying ones at other firms, believing incorrectly that pay is low only where they work. Opinions vary among neoclassicists as to the causes of recessions. One idea is that restrictive monetary policy precipitates decline. Another idea, associated with the names of David Lilien (1982) and Stephen Davis, John Haltiwanger, and Scott Schuh (1996), is that recessions are caused by waves of shocks, including productivity declines, to particular firms or industrial sectors. A similar idea, the real business cycle theory, is that economywide declines in productivity produce recessions. Some say that predictable macroeconomic policy cannot influence employment, but that only surprise changes in aggregate demand do so, and then just temporarily by producing unexpected fluctuations in nominal pay rates that confuse workers about the relation of their pay to what is normal. According to the neoclassical view, wage rates appear to be rigid downward only because workers prevent wages from declining noticeably by withdrawing labor as soon as they fall below reservation levels.⁴ A current opinion is that unemployment always nearly equals a "natural rate" that fluctuates and cannot be influenced by macroeconomic policy. According to this view, chronically high unemployment, such as the rates of 12 percent or more seen in the mid-1990s in some European countries, is due solely to microeconomic problems, including generous unemployment insurance benefits, labor market rigidities, or the low skills of the unemployed.

Neoclassical economists describe unemployment as voluntary, whereas Keynesians claim it is involuntary. Although this difference is the core of the debate, I try to avoid applying the word "voluntary" to joblessness, for doing so leads to a logical ambiguity. If someone points a gun at you and says, "Give me your money," you surrender it voluntarily, but only because the alternative is dreadful. In the same sense, people are voluntarily unemployed if they would rather look for a job paying what they think they deserve than take a lower-

paying one. Keynesians have tried to avoid this difficulty by labeling as involuntarily unemployed people who cannot find work at the prevailing wage at jobs they are willing and qualified to take. This definition involves a distracting difficulty; usually there is no such thing as a prevailing wage, because pay rates for the same kind of work vary widely among employers. This fiction is irrelevant, however, for in recessions most jobless people have difficulty locating any opening at the level of their qualifications, even one that pays little for their kind of work, and many gladly accept any position they find. They are likely to be rejected for positions at a level somewhat lower than that for which they are qualified, for reasons I will explain. They might have less difficulty finding and being accepted for very low level positions, but they would probably be ill-advised to take such jobs, unless faced with penury. Many do offer to take positions at a much lower level than they held previously, and those who refuse to do so could be said to be voluntarily unemployed because greater flexibility might enable them to find work more quickly. Nevertheless, they are involuntarily unemployed according to the everyday meaning of these words and according to the spirit of Keynes's usage. Judging from historical experience and what I have learned of the labor market, it seems probable that if aggregate demand for labor increased during a recession, fewer people would be jobless. This is the heart of what Keynes meant by "involuntary unemployment." I do not discuss the probable impact of changes in aggregate demand, however, for this question has to do with how the whole economic system functions and I gathered information only on microeconomic details. Among these were the flexibility of the unemployed and the difficulty they experienced finding work.

1.3 The Unsatisfactory State of Knowledge

Neoclassicists espouse a set of closely related logical explanations of why unemployment increases during recessions and why wages seem rigid downward, while actually flexible. Keynesians, in contrast, embrace no single theory of wage rigidity. Moreover, the lack of an accepted explanation is one of the neoclassicists' most persuasive arguments; labor markets must clear, because there is no compelling reason for them not to do so. They say that a convincing theory ought at least to explain why workers and employers do not exploit the gains from trade that would exist were there excess unemployment, for if workers were willing to work at wages at which employers could profitably employ them, then the two sides would agree on terms for exchanging labor for money.

Despite intense interest in wage rigidity, current understanding of the subject is highly unsatisfactory. Many explanations have been proposed, any or all of which could be valid, since they are not mutually contradictory. But little is

known about which theories are correct or under what conditions. Such an abundance of competing, unrefuted theories indicates ignorance, not knowledge; greater abundance makes each theory less likely.⁵

1.4 Examples of Explanations

Some examples should give an idea of the variety of theories. Those of Lucas and Rapping (1969) and Lucas (1972) described above are accepted by many economists. Their models are inconsistent with the fact that during recessions layoffs increase and quits decline, and McLaughlin (1990, 1991) has suggested a way of resolving the discrepancy regarding layoffs. He assumes that firms offer workers a choice between layoff and continued work at lower pay, layoff occurring only when the lower pay is refused. Layoffs are compatible with the Lucas-Rapping model, because they are regarded as quits following refusal of pay cuts.

Search models start from the realistic observation that buyers and sellers of labor have imperfect knowledge of trading opportunities, so that both sides must spend time and money to find or fill jobs. In these models, unemployment is a natural consequence of delays in the matching of people to open positions. In most of the models, wage rates are the outcome either of bargaining between employers and individual employees or are set unilaterally by employers and adjusted so as to ensure an adequate supply of labor. There are two types of job search models, those which attribute downward wage rigidity to job hunters' misperceptions of market wages, and those of the transactions type in which there are no misperceptions, wages are completely flexible, and unemployment results from adjustment to change.⁶ The Lucas-Rapping misperceptions model is of the first type.⁷

Another theory is that of Keynes (1936, chap. 2), who suggested that workers are so concerned about the relation of their wages to those of workers at other firms that no company dares cut pay. Resistance to wage cuts can be avoided only if all firms in an economy cut wages simultaneously so as to preserve traditional wage differentials. Since such reductions would be difficult to coordinate, nominal wages are rigid downward. Real wages can fall without meeting the same opposition, however, for inflation affects all workers, preserving wage differentials.

According to the implicit contract theory of Baily (1974), Gordon (1974), and Azariadis (1975), firms obtain labor most cheaply by guaranteeing that real wages will never decline, that is, by providing insurance against wage decline. If real wages do fall, employers have to pay more in the future than they would have otherwise in order to obtain or retain labor.

In the shirking theory of Eaton and White (1983) and Shapiro and Stiglitz

(1984), firms monitor workers randomly, firing those whose performance does not meet some standard. The higher are wages and salaries, the greater is the cost to workers of job loss, and hence the greater is their incentive to meet the standard. Increased pay, therefore, makes discipline more effective, raises productivity, and reduces labor monitoring costs. The theory implies that there must be some unemployment when the economy is in equilibrium, for if there were none any worker could find another job immediately after being dismissed, so that firms could control their workers only by paying them more than they would receive elsewhere. There can be no equilibrium under such conditions; very low unemployment would lead to rapid wage inflation as each firm tried to pay more than its labor market competitors. This theory does not necessarily imply either real or nominal downward wage rigidity, since higher unemployment increases the cost to workers of dismissal and so makes it possible to reduce pay. Wage flexibility may be slight, however. The theory is interesting, mainly because it typifies economists' thinking about how managers control workers. It is often assumed in economic models that managers order subordinates around and control them through threats, financial incentives, and monitoring.

According to the insider-outsider theory of Lindbeck and Snower (1988b), the resistance of an inner group of workers within each firm is responsible for wage rigidity. This group, the insiders, consists of employees who are well protected from layoff because of high seniority or special skills. They have no interest in giving up pay in order to save the jobs of recently hired employees, who are outsiders, or in order to encourage the hiring of unemployed workers. Insiders are able to prevent the firm from replacing employees with cheaper unemployed workers, who are outsiders.

A completely different explanation of wage rigidity is provided by the morale model of Solow (1979), espoused and elaborated by Akerlof (1982) and Akerlof and Yellen (1988, 1990). These authors assume that pay rates have a positive effect on productivity through their impact on morale. When setting pay, companies weigh labor costs against the impact of pay on productivity, and this trade-off determines a profit-maximizing wage. The models of Solow and of Akerlof and Yellen allow morale to depend on the wage level or on changes in it. If morale depends on the level, then the model can explain downward wage rigidity only if the profit maximizing wage exceeds what the firm must pay in order to recruit and retain workers. If morale depends on wage changes, then the theory can explain downward rigidity at any pay level.

1.5 The Dearth of Information

Each of these theories of wage rigidity depends critically on specific assumptions about the motives and behavior of workers or employers. Some of the suppo-

sitions may seem to conflict with common sense or to stem from the natural human tendency to blame the less fortunate for their suffering, yet to reject the theories on these grounds alone would be to base conclusions on opinion rather than on careful observation. Unfortunately, there are almost no published sources that give information detailed enough for verification of the assumptions in question. For instance, I found no reference showing whether McLaughlin is correct in asserting that firms offer employees the alternative of a pay cut before laying them off, though it is perhaps revealing that accounts of layoff procedures and surveys of unemployed people contain no reference to such proposals. The economics literature contains a great many tests of models of wage rigidity and unemployment, but most of these rely on indirect evidence, for economists usually obtain information only from introspection and from surveys made by public institutions that provide data on easily quantifiable variables. Surveys relevant to labor economics typically report wages, salaries, employment status, and perhaps the size and profitability of firms and the age, sex, and race of workers: insufficient information for discriminating among theories of wage rigidity. The feature that stands out in these data is that recessions increase unemployment and retard but do not stop the rate of growth of nominal wage averages. This is the pattern expressed by the Phillips curve, which shows the historical negative relation between unemployment and wage inflation (Phillips, 1958). This relation is roughly consistent with all well-known theories, since they were designed to explain it; often it is the main evidence cited to support them. The usual surveys contain other information, but every facet of the data seems consistent with several theories.

Nor is it possible to choose among the theories by appealing to introspection, which is used to justify both particular forms of irrational behavior and the assumption that economic behavior is rational. Many theories of wage rigidity assume that behavior is rational, so that they cannot be distinguished on these grounds. In any case, the rationality assumption is too general to be much help. It is a useful unifying principle in economics, but the implications of rationality depend on the conditions constraining decision makers, and often knowledge of these is precisely what is lacking. Furthermore, it may be naive to presume that all important forms of behavior are rational. Introspection does not necessarily give insight into the behavior of workers and employers, who react to complex and stressful circumstances at work that may be unfamiliar to academic economists, as they were to me. Indeed, some theories of wage determination reflect a mistaken extrapolation of the conditions of university life to business. Since standard sources of information do not provide a sound basis for choosing among explanations of wage rigidity, these will remain speculative until more enlightening sources are found.

1.6 The Utility of More Information

It is important to understand the mechanisms generating wage rigidity, both for scientific understanding and for the analysis of economic policy. In considering unemployment, it is vital to know to what extent it is a social evil. Joblessness is not a serious matter, if the models of Lucas and Rapping and of McLaughlin are accurate. They assume that layoff and unemployment are chosen by workers when wages fall or fall below expectation, so that the unemployed are just as well off as they would have been had they stayed with their jobs at their old pay or at the somewhat lower pay the employer offered. Layoff and unemployment are potentially much more harmful if they are forced on workers, as they are in Keynesian models. Another reason for understanding the microeconomic mechanisms creating unemployment is that they are critical for discovering how to reduce it. If the models of Lucas, Rapping, and McLaughlin are correct, then unemployment falls only when wages rise above expected levels, and since the public makes intelligent predictions about policy, it is hard for the government systematically to create the surprise increases in wages needed to reduce joblessness (Lucas, 1978). But suppose that during recessions wages do not decline and unemployment increases, because firms reduce employment as product demand declines. Then, it is possible that unemployment could be manipulated and even maintained at a low level by appropriate control of aggregate demand, and inflation would occur only if unemployment fell too low.

It is easy to think of policy implications of other microeconomic assumptions about unemployment. Keynes's theory of wage rigidity implies that workers would accept simultaneous wage cuts by most firms, if only the government could arrange them. The shirking model implies that laws increasing the disciplinary powers of firms would reduce wages and lead to a lower equilibrium level of unemployment. Whereas Keynes's model implies nominal wage rigidity, the implicit contract model implies that real wages are downwardly rigid, which in turn implies that employment might not respond to increases in aggregate demand.

1.7 Should Assumptions Be Verified Directly?

In "The Methodology of Positive Economics," Milton Friedman (1953) argues that it is pointless to verify a theory's assumptions directly. He asserts that the purpose of theories is to summarize principles that are useful for prediction and that the adequacy of predictions should be judged in terms of the uses made of them. A hypothesis should be accepted if its predictions are accurate enough to be useful. The realism of assumptions and the truth of a hypothesis are irrelevant to its acceptability. Friedman argues that assumptions help spec-

ify the circumstances in which a theory will predict successfully and help express principles chosen from the distracting details of reality. Such principles are needed, because reality is so confusing that it is impossible to perceive it without the aid of theories and their assumptions. Friedman stresses the distinction between the phenomena to be explained and other empirical information and he asserts that the sole usefulness of a theory is its ability to predict the phenomena to be explained. Although good tests of a hypothesis use data having to do with these phenomena, these data should have been unknown when it was formulated. Data should be used to verify the implications rather than the assumptions of a theory, where the words "prediction" and "implication" are used almost interchangeably. He argues that a theory should be designed so that its assumptions have to do with data other than the phenomena to be explained. It is impossible for assumptions to be accurate, since they must approximate and simplify reality in order to express simple rules useful for prediction. Carrying the argument a little further than he does, one might say that just as false premises can imply true conclusions, so false assumptions can lead to accurate predictions and hence the truth or accuracy of assumptions is irrelevant to the goal of a theory, which is prediction.

Friedman anticipates certain objections to his reasoning made later by Koopmans (1957). He recognizes that a theory's assumptions and implications cannot be distinguished logically, since the assumptions are themselves implications, one implication of a statement being the statement itself. Furthermore, hypotheses may be recast so that assumptions in one formulation are implications in another, deduced from still other assumptions. He even grants that among the tests of a theory is its ability to predict information bearing on assumptions.

Friedman handles these objections with his argument that a theory is judged only by its success in predicting data having to do with the phenomena to be explained. If a theory is formulated properly, its assumptions do not have to do with these phenomena, and weight is attached to direct empirical tests of assumptions, only if the data used relate to the phenomena to be explained.

To clarify his meaning, Friedman gives three examples. One is the physical law that the acceleration of a body is constant if it falls freely in a vacuum. He maintains that the applicability of this theory depends on whether its predictions are of practical use, not on how much air friction there is. Another example is the assumption that leaves are distributed on a tree so as to maximize sunlight exposure. He points out that the pertinence of the assumption depends on its predictions about leaf distribution, not on whether the leaves deliberately locate themselves so as to maximize light-gathering. His third example is the assumption that expert billiard players shoot as if they calculated precisely the ball's trajectory. Here, whether the theory applies depends on how well it pre-

dicts the path of the ball, not on whether billiard players actually make the calculations.

Friedman's argument is coherent if theories are understood merely to be instruments for making pragmatic predictions. Boland (1979) and Caldwell (1982) label this view an instrumentalist one. In the philosophy of science, this view contrasts with that of realists, who believe that there is a true theory, and conventionalists, who suspend judgment about whether there is a true theory, but agree to speak as if there were one. Realists believe that all aspects of theories, including assumptions, should be tested for consistency with empirical reality (Popper, 1957). For them, to ignore data about assumptions is to throw away useful information.

I find it difficult to choose between the three points of view, but believe that Friedman favors a too restricted use of instrumentalism. In my opinion, the weak point of his argument is his narrow view of the phenomena to be explained. From an instrumentalist perspective, we should want as complete as possible an explanation of the causes of these phenomena, so that we can learn what will happen when the circumstances creating the phenomena change, because of policy intervention or for other reasons. If we cannot make such changes experimentally, we cannot check the predictions we make against data and so cannot include them among the known phenomena to be explained. A false or unrealistic set of assumptions might by accident perfectly predict the known phenomena, but prove treacherous when conditions change. For this reason, it seems wisest to check as many aspects as possible of the mechanism generating the phenomena to be explained, including the realism of assumptions. Assumptions should be realistic in the sense that they correspond approximately to conditions under which the theory is applied, and how good the approximation should be depends on the mechanisms posited by the theory.⁸

There is a near circularity in Friedman's methodology, for the only evidence supporting explanations of the phenomena to be explained are these phenomena themselves. The circularity may be seen by considering the following nonsensical example. Someone tells you, "The stock market is rising, because *Rapacitas*, the demon of greed, is loose in the population." You answer, "How do I know this demon exists? What happens to him when the market falls?" The reply is, "Isn't the market rising and doesn't it always rise when *Rapacitas* is loose? When the market falls, his brother *Timiditas*, the demon of fear, takes his place."⁹

The circularity of the argument in this example is not a defect according to a literal interpretation of Friedman's reasoning, if stock market fluctuations are the only phenomenon to be explained. The flaws in the example's argument are not just that it names as a cause something that is ill defined and unobservable or that the theory cannot predict turning points in stock market av-

erages. No theory does the latter, so that this deficiency is not much of a reproach. Another flaw is that the phenomenon selected to be explained is too narrow. We would like to have a theory that explains the effects of interest rates, inflation, economic forecasts, and other influences. There is empirical data on the impact of these factors, and we would make it harder to dream up a nonsense explanation by including their impact in the set of phenomena to be explained. A stronger check of a theory is consistency with data that are distinct from and unrelated to the phenomena to be explained, data Friedman would exclude and that are, as Friedman says, the kind of data used to test assumptions. It becomes especially important to use such data when you want to use the theory to predict phenomena previously unobserved, such as the impact on the stock market of a form of economic policy never tried before.

Friedman introduces aesthetic criteria, such as simplicity and fruitfulness, for choosing among theories that predict equally well the phenomena to be explained. These criteria are so subjective, however, that they may not guard effectively against believing what you want to believe, a key function of methodology. A pet theory might be declared to be fruitful because its principles explain many things, though little evidence supports these explanations and they may even be nonsensical. A theory should be called fruitful only if the other explanations it inspires are not contradicted by existing evidence. We see that fruitfulness does not provide an independent criterion for eliminating extraneous explanations; we cannot define fruitfulness without deciding what it means for theories to be consistent with observation.

The shift from the Ptolemaic to the Copernican view of the solar system is often cited as showing the power of aesthetic criteria. The Copernican model is certainly simpler and more fruitful; it may have helped stimulate Newton's theory of gravity. This example does not, however, bear on the choice of assumptions, for the Copernican and the Ptolemaic models have the same assumptions that are assertions about reality, namely, that the sun and the planets are bodies in relative motion. Both models are mathematical representations of the same phenomena, the movements of the planets. The models differ only in form; one uses a sun-centered coordinate system and the other, in effect, uses an earth-centered one. The choice of coordinate system is a mathematical device, not an assumption.

The task of this book is to explain wage rigidity, which has to do with unemployment and increases in it during recessions. For the purposes I have in mind, there is little interest in being able to predict these phenomena in the sense of mimicking time series of wages and unemployment. It seems reasonable to hope that a successful explanation of wage rigidity would contribute to understanding the extent of the welfare loss associated with unemployment and

what can be done to reduce it. These issues are ultimate goals beyond those of this book. Many theories of wage rigidity and unemployment include partial answers to these questions as part of their assumptions, so that the phenomena of real interest, from my point of view, are described in the theories' assumptions. For instance, Lucas concludes that increased unemployment during recessions implies little welfare loss and that recessions can be minimized by avoiding unpredictable changes in monetary policy, and these judgments are immediate consequences of his assumptions that unemployed people choose to be unemployed because of unanticipated declines in wage rates (Lucas 1972, 1978). Theories of wage rigidity and unemployment typically are tested, however, as if their goal were to predict data on wages and unemployment.¹⁰ If these predictions were the focus of inquiry, then it would hardly matter which of the many theories was used, since they all fit these data almost equally well. For this reason, Lucas's policy conclusions are not strongly supported by the fact that his model matches the historical relations among wages, employment, and other variables. Good support can come only from information that distinguishes his microeconomic assumptions from others yielding different policy recommendations.

A fanciful example may illustrate the danger of taking too narrow a view of instrumentalism. You are an explorer seeking contact with the Dafs, an isolated tribe about which almost nothing is known. You observe one of their villages through binoculars from far away, because these people might attack foreigners. You observe that every morning on sunny days, men wearing bright yellow hats stand in backyards and make sweeping gestures toward the sky. You conclude that these men are priests of a sun-worshipping religion, and therefore must be leaders. This theory predicts well the behavior of the men observed and even explains the color of their hats. When you finally arrange a meeting with some Dafs, you meet a few men with yellow hats and a few other plainer people. Believing the first to be leaders, you offer them presents, at which point all the Dafs are outraged and assault you. What you have not observed is that yellow hats mark slaves, who throw grain to the household chickens in the yard on sunny days and inside on rainy ones. The leaders are the plainer people. It would have been worth your while not to settle for treating arm waving and wearing yellow hats as the phenomena to be explained, but to test your assumptions about behavior by taking risks to sneak up closer and see precisely what the men were up to.

Spying on the Dafs would be not only a way of testing a theory, but would gather information that might make possible the formulation of a new one. My hope in studying the labor market has been to seek inspiration for new theories as well as to test existing ones. I have sought not to oppose economic theory, but to improve it.

1.8 Problems with Surveys

Despite negative attitudes in economics toward direct microeconomic investigations, a few economists have sought information about the causes of wage rigidity and unemployment by interviewing employers, and I have done the same, going further than previous investigators. Surveys involve serious difficulties, which perhaps make it understandable that few economists undertake them.¹¹ It can be hard to get businesspeople to cooperate, and this obstacle makes random sampling of companies nearly infeasible. The 15 to 35 percent of a randomly selected group who agree to participate is probably a very biased sample, for I learned from experience that those who agree most readily are often the least interesting, striving to make a good impression and revealing little. Other major difficulties are that respondents may hide or falsify information, may not understand their own intentions, and may have little incentive to give accurate answers, points that Fritz Machlup (1946) stresses. Consequently, the data collected are often thought of as "touchy-feely" rather than "hard."

I tried to avoid bias through my methods of sampling and interviewing, which I describe in Chapter 2. I believe this study could be repeated, and is therefore scientific in the sense of being approximately replicable. I have checked results by comparing my observations with other information, including U.S. statistics on wages and employment, econometric studies, surveys of various sorts, and experimental work. Some of the more important work has been done by researchers in fields other than economics, especially in psychology. Other surveys of businesspeople and of the unemployed are consistent with my own, as are experimental results. Some variation, of course, is to be expected. My conclusions are no doubt influenced by the time and place of the interviews, and by current fashions in management and in counseling. Nevertheless, there is reason to believe the basic results will apply for some time, for wage rigidity has been an enduring phenomenon. It was even mentioned by Malthus in 1798.¹² What is universal in the findings can be detected only by more field work in other regions.

There remains the question of whether the statements of businesspeople should be taken seriously, even if all say the same things to all investigators. Though probably few people lie systematically, it would be reassuring to check what they say. One check is to compare their statements with what they do, but I could not rely on this method, since I had few ways to verify actions. A control I did use was the relation between circumstances and claimed motivation. For instance, suppose that employers say that a condition, such as high labor turnover, causes a certain business policy. If this policy occurs in companies with

high turnover and not in companies with low turnover, then the conclusion that turnover explains the policy is more convincing.

1.9 The Value of Surveys

Even if no controls were available, it would be presumptuous to ignore the testimony of people who make economic decisions and observe and participate in economic life. To do so would be to make economics a religion rather than a responsible analysis of experience. Good instincts about a subject can be developed only by contact with the phenomena studied.

Unfortunately, attitudes in the economics profession discourage the use of information other than well-known data. Among these attitudes are those articulated by Friedman and discussed above. An important one is the belief that nothing is seen if one looks too closely, that the forest is missed for the trees. It is true that remote views are useful, that aerial photographs disclose forest diseases. But it is also true that to understand forests, you should know something about trees.¹³ A related idea is that knowledge can be obtained only through the systematic testing of hypotheses and that unstructured observation and checking of facts are pointless or even anti-intellectual, because reality is only imagined and nothing can be seen without a theory in mind. This view leaves unanswered the question of how to discover good hypotheses, one of the goals of this project.

Still another attitude discouraging surveys is the belief that people do not know their own motives. In order to defend this view, economists imitate Friedman and cite metaphors, such as that of baseball players, who catch a ball without mathematically calculating its trajectory, though they behave as if they do so. This argument may justify assuming that people act unconsciously in conformity with some complicated model, but it is also true that if you want to learn how to play baseball, you would do better consulting a Ted Williams or a Casey Stengel than a mathematician.¹⁴ The businesspeople I talked to were articulate, had obviously thought a great deal about management problems, were able to analyze them clearly, and said they learned a great deal of what they knew from other managers. An outsider should also be able to learn from them how to run a business, and it is precisely knowledge of the personnel problems businesspeople face that is lacking in studies of wage rigidity. Although insight might be gained by interpreting such knowledge in terms of an optimization model, and I make suggestions in this direction, the main goal of this project is to uncover empirically the circumstances that give rise to wage rigidity.

Related to the idea that motives are unconscious is the notion that people may not be aware of principles governing their own behavior and that these

are the proper objects of study. An example is implicit contracting; employers may not realize that they are insuring workers' income in exchange for lower pay. Such principles are, of course, precisely what any researcher should look for, and it is possible that people are not conscious of them. However, in order to discover or verify them, it is necessary to study behavior, and it is hard to know how to study business behavior without talking to people in business. I also doubt that people would be unaware of something as obvious as implicit insurance contracts, if they were real.

Skeptical attitudes lead some economists to treat economic life as almost unknowable, like distant galaxies, though it goes on all around us. In economics, variables not appearing in well-known data sets are often labeled as unobservable, even when they could be observed in principle. For instance, Heckman and MaCurdy (1988) argue that the assumption of labor market clearance cannot be contradicted because standard data sets do not contain key variables applying to unemployed job seekers, such as nonmarket opportunities for the use of time and job offers received and refused.¹⁵ Although Heckman and MaCurdy admit the possibility of observing these variables, they stress that the unobservables make labor market equilibrium an irrefutable tautology. They do not point out that observation, difficult as it may be, is the obvious way to dispel the ambiguities. Someone just has to do the work. Why should economics differ from other sciences, where researchers spend much of their time collecting data? It is not healthy for a science to isolate itself from its subject of study, especially for a field that is highly contentious and where reality is constantly changing. In economics, it is all too easy to believe what one wants to believe, since theories become intertwined with political values and truly pertinent information is hard to obtain.

The utility of surveys is illustrated by my own experience. I have long wondered about wage rigidity. I was puzzled that firms and workers do not agree to cut wages and salaries during recessions when sales decline, for such cuts should benefit both sides by reducing costs and layoffs. I wasted years inventing theories describing impediments to pay cuts. Although none of my theories was convincing, I liked best one in which agreements to cut pay are difficult to negotiate because management and labor do not have the same information or opinions about the prospects of the company and the effects of pay reduction. I discounted surveys reporting that employers avoid cutting pay because of bad effects on morale, for this explanation seemed too easy and left too many questions unanswered. Why do not the obvious benefits of saving jobs overcome the bad effects on morale? Do not layoffs also hurt morale? Cannot the impact of pay cuts on morale be diminished by having them occur automatically when company sales or general economic conditions decline? Why do employers care about morale, since they can gain cooperation by threatening to fire workers

who do not perform well? Why not hire new workers during recessions at a reduced pay scale? Frustrated, I turned to interviews, and the first of these shook my prejudices. I learned from an officer of a medium-sized, nonunion manufacturing firm that cutting pay would have almost no impact on company employment, that hiring new workers at reduced pay would antagonize them, that reducing the pay of existing employees was nearly unthinkable because of the impact on worker attitudes, and that the advantage of layoffs over pay reduction was that they "get the misery out the door." Furthermore, attitudes were important for performance, and by attending to them the company was able to maintain a loyal and productive work force despite paying considerably less than its main competitors in the labor market. Only gradually, after hearing similar comments many times, did I concede that what was said should be taken seriously. The origin of my puzzlement and mistaken theorizing had been incorrect intuition. For instance, I believed that an individual firm could save a significant number of jobs by reducing pay. This is seldom true, and the firms for which it is true are precisely the ones most likely to cut pay.

1.10 The Scope of the Study

This inquiry is intended to be exploratory, touching on many issues in order to test existing theories, to seek new hypotheses, and to see the overall shape of phenomena associated with wage rigidity and unemployment. I used an inductive approach, which had the advantage of revealing unsuspected phenomena and relationships. However, the resulting breadth of the study increased the difficulty of analyzing the data and of obtaining interviews, and consequently many of my conclusions are tentative, though I am confident of the main ones. The data were difficult to organize and interpret, because interviews covered varying topics and did not contain answers to standardized questions asked of all respondents. Partly for this reason, I avoid statistical methods. Some businesspeople refused to be interviewed because they felt that in loose conversation they might inadvertently divulge confidential information, a fear that was justified; some respondents did disclose matters that I thought it best not to record. The difficulty of obtaining interviews reduced the value of random sampling and led me to arrange many interviews through personal contacts. There is, I believe, a trade-off in fieldwork between the randomness of a sample and the quality of the interviews. Respondents were most informative when they talked freely and the discussion wandered, but businesspeople were reluctant to grant such interviews. I found that many of my best interviews were those arranged through intermediaries, who put respondents at ease. A sample obtained through friends and acquaintances, however, is hardly random.

1.11 Implications for Future Research

Surveys need not be as unstructured as this one. Narrowly focused inquiries can use a fixed list of specific questions, so that statistical methods can be applied to the responses. Businesspeople probably would be more willing to cooperate with such studies than they were with this one, making possible random sampling.¹⁶ My hope is that economists will eventually make many narrow studies that will yield clear and firm conclusions.

The survey method could be used to shed light on urgent issues tangentially related to this inquiry, such as why large firms pay more for labor than small ones, why interindustry pay differentials are so large, why the distribution of earned income is becoming more unequal in many wealthy countries, whether higher unemployment in Europe than America is due to institutional differences or to more expansionary macroeconomic policies in the United States, whether wages are more flexible in Japan than in the United States, and if so why the difference exists and does Japan's greater flexibility explain its lower unemployment rate.

1.12 Findings

The main result of this inquiry is a commonsense explanation of downward wage rigidity in the private sector. The theory, summarized in the first section of Chapter 21, is an elaboration of that of Solow (1979) and Akerlof (1982). The investigation also yielded knowledge of the wage-setting process that may prove useful in thinking about macroeconomic policy. Unfortunately, it is not possible to go far in analyzing policy, because its effects depend on the interaction of wages and prices and little is known about price determination. Finally, I came upon hypotheses and phenomena that merit further exploration.

Though my findings contradict most theories of wage rigidity, the contradiction is not absolute, for each probably applies at any moment to certain individuals. Fieldwork makes obvious the enormous variety in the economic world. Many important decisions have no single correct resolution, because they depend on imponderables and on personal judgment. I interpret theories to be valid only if they apply often enough to be useful.

The results indicate that labor is in excess supply during recessions, so that the Keynesian side of the macroeconomic debate is the more accurate view, a view that contradicts the principle widely used in macroeconomics that economic equilibria maximize a social welfare function and so solve a dynamic optimization problem. This principle cannot apply even approximately if large amounts of labor are wasted. My observations are consistent with the Keynesian view that recessions are caused by declines in aggregate demand, though I

did not focus on this question. Firms had layoffs because of financial setbacks, technical improvements, and declines in product demand, never because of declines in productivity. Wage demands were not a cause of unemployment. On the contrary, many unemployed workers became excessively flexible, in the eyes both of people who counseled them and of employers. As are all explanations of wage rigidity, mine is consistent with the Phillips curve.

1.13 Two Novel Findings

Two phenomena I observed, though commonplace in the business world, have not previously been well documented. These are the shunning of overqualified job applicants and behavioral differences between the primary and secondary sectors of the labor market. Job applicants are said to be overqualified if they are suitable for substantially better jobs than the one applied for, where "better" means better paid, more interesting, or with more responsibility. Unemployed overqualified applicants were common in the Northeast during the recession. The label applies to semiskilled and skilled manual workers as well as to technicians, professionals, and managers. Many employers were reluctant to hire overqualified applicants, because of concern they would quit to take better positions or might be unhappy or a threat to their supervisors (sections 15.2 and 17.6).

I find it helpful to distinguish between primary- and secondary-sector jobs. The latter are short-term positions that are often part-time, whereas primary-sector jobs are long-term and full-time.¹⁷ Examples of secondary-sector workers are waiters and waitresses, floor crews in fast-food restaurants, sales clerks in most stores, taxi drivers, security guards, janitors, consultants, many telemarketers, and temporary, interim, or contract workers. Primary-sector personnel include most factory, clerical, and secretarial workers, technical, professional, and managerial employees with permanent positions, and salespeople in stores and restaurants with regular customers whom the staff should know on a first-name basis. Secondary-sector positions have high turnover because hiring and training costs are too low to make it worthwhile for firms to pay high enough wages to reduce quitting. Although pay is usually lower in the secondary than in the primary sector, this is not always so. Consultants are sometimes well paid, for instance, and some clerks and factory workers are poorly paid.

In contrast to the primary sector, in the secondary sector I found that the pay of new hires was more flexible downward and employers were more willing to hire overqualified workers, an attitude that made the sector a refuge for unemployed workers ready to take stopgap jobs (sections 17.2, 17.3, and 17.6). Low hiring and training costs here reduced concern about turnover among the overqualified. The greater flexibility of hiring pay derived from the lesser im-

portance of internal pay equity. In both sectors, the pay of existing employees was rigid downward, but in the secondary sector the pay of new hires was less tied to that of existing employees. In contract labor, there are no comparable existing employees, for temporary workers usually do not relate their pay to that of permanent employees. In other secondary-sector jobs, confusion caused by part-time schedules and high turnover makes it difficult for workers to get to know one another and to learn one another's pay, and there is less resentment of pay inequities because jobs are seldom taken seriously as careers.

1.14 Goals

In this book I seek to give an accurate impression of what I observed, including the frequency of particular views, their variety, the intensity and manner with which they were expressed, and their logical structure. As I describe the findings, I relate them to published evidence, usually in chapter appendices. Only at the end do I integrate the material and relate it to economic theory. Naturally, the data can be interpreted in more than one way, and I believe I give enough information to enable readers to draw their own conclusions. Although I would like to influence their views, my main hope is to demonstrate that despite the considerable resources required, it is well worth the effort to get out of the office and face economic reality rather than invent it.

WHERE TO FROM HERE?

While I was engaged in this project, ideas for more narrowly focused data collection occurred to me, a few of which I sketch here. If studies are sufficiently narrow, it should be possible to sample randomly and to use statistical techniques (sections 1.10 and 1.11).

22.1 Statistical Background

Little is known about the incidence and magnitude of pay cuts in the general population of companies in good and bad times. Even less is known about downward flexibility of the pay of new hires. Given sufficient resources, it would be simple to collect data on these matters from a fixed panel of companies over a period of years. Such data would make possible a careful comparison of hiring pay flexibility in the primary and secondary sectors. In order to pinpoint the causes of pay cuts, it would be essential to know the condition of companies at the time they cut pay, that is, whether they were suffering financially, had recently had layoffs, or had recently been taken over by new owners. In measuring hiring pay, it would be important to measure the pay of new hires for specific jobs and with specific qualifications and experience. Otherwise, the data might be contaminated by changes in the quality of the new hires relative to the types of positions filled. The quality of new hires for given positions probably increases during recessions.

As has been noted, household income surveys show a large number of year-to-year pay reductions (Appendix 11A, section 11.1). It is hard to evaluate these findings because the data may be inaccurate and because they do not reveal the reason for pay decreases; they may result from new job assignments, demotions, loss of overtime, or cuts in base pay. More reliable conclusions

might be reached by following a sample of, say, 500 households for several years, and asking precise questions about income, its sources, and changes in it. These responses should be checked with employers for accuracy.

There is almost no information available on the frequency with which firms use various types of punishment. Firms would probably be reluctant to share such data, unless they were collected in a way making it impossible to trace the source, for firms would worry about their possible use in employee suits. The data would help in assessing the shirking model.

Although some authors assume that the unemployed seldom take low-level stopgap jobs (for example, McCormick, 1990), I believe many do so, and other surveys support this conclusion. Evidence could be gathered by following a fixed panel of unemployed workers. The issue is important as it bears on the flexibility of the unemployed.

22.2 Tests of Hypotheses

According to the standard scientific method, the best strategy for testing a theory is to attack it from its weakest side, using the data most likely to contradict its implications. In economics, it is often difficult to perform powerful tests of theories, for typically little or no data are available bearing on key assumptions and implications. Investigators could make better tests if they collected their own data, and some ideas for such tests follow.

THE SHIRKING MODEL

One test of the shirking theory would be to conduct a questionnaire survey of workers in a few firms, asking them if they believed they would be fired for poor performance and if fear of firing motivated them. It would be important to learn the views of supervisors and managers as well on these matters. The theory would be contradicted if the fear of firing were not an important motivator. It would be interesting to see if the theory applies to unskilled labor filling short-term jobs; I believe it may.

Another test would be to find a few prosperous nonunion businesses, located in an area that had recently changed from a situation of labor shortage to one of extreme labor surplus and where laid-off employees had little hope of finding new work at wages close to those earned previously. Such situations arise when large plants using specific skills shut down in isolated areas. If the shirking model were valid, wage rates within the prosperous firms would decline.

KEYNES'S RELATIVE WAGE THEORY

Keynes's relative wage theory could be tested by finding a group of neighboring firms where workers knew a great deal about pay levels at all the firms. This

knowledge could be tested by means of questionnaires filled out by employees and from information provided by the firms about actual pay levels. Keynes's theory predicts that pay should remain downwardly rigid even during a severe downturn. I conducted a limited test of this sort, and found that pay was somewhat flexible in businesses where mutual knowledge of pay was great, contrary to the theory (section 7.5). Contrasting data could be provided by an isolated group of neighboring firms where workers knew almost nothing about pay outside their own workplace. The theory predicts that pay would fall at these firms during a recession.

THE LUCAS-RAPPING MODELS

The Lucas-Rapping theories could be tested by following a panel of unemployed workers during a recession, establishing how and why they lost their jobs, whether they preferred home production or nonmarket opportunities to jobs paying current wages, and whether they rejected or failed to investigate job opportunities because they paid too little. The study might be similar to that of Daniel (1990), while focusing more on the questions just listed. Daniel found that the jobless had few nonmarket opportunities and a great deal of difficulty finding work.

As was mentioned earlier, the Lucas-Rapping models predict that the number of quits into unemployment increase during recessions (section 20.1). No U.S. government data exist on such quits, as far as I know. The Bureau of Labor Statistics estimates the stock of currently unemployed who had quit their jobs, not the flow of quitters into unemployment. The cyclical behavior of this flow could be measured by following a representative panel of workers. Appropriate data have been collected since 1968 by the Michigan Panel Study of Income Dynamics, but it seems that no one has totaled up the number of quits into unemployment for every year of the survey.¹

THE IMPLICIT CONTRACT MODEL

An obvious test of the implicit contract model is to compare the pay of firms that have cut pay in the past with that of similar firms that never cut pay. According to the theory, the pay cutters would be obliged to pay more, and enough more to discourage pay cutting. Another test would be to see if companies give raises to offset increases in wage taxes. The theory predicts that firms do so, if they are risk neutral. My study took place shortly after the introduction of a new state wage income tax in Connecticut, and I found that firms were unwilling to make up for the tax increase (section 10.3).

THE ELASTICITY OF DEMAND FOR LABOR, AND THE SAVINGS GENERATED BY LAYOFFS

Reasons employers gave for preferring layoffs to pay cuts were that pay cuts save few jobs and layoffs save a great deal of money. A questionnaire survey

could determine whether these beliefs are widely held. It might be possible to estimate firms' actual elasticity of demand for labor by comparing labor use before and after pay cuts. More insight could be gained by investigating in detail the actual cost structure of a few firms, in order to calculate the impact of a hypothetical pay cut on marginal variable costs. The effect of a change in marginal costs on sales and employment would depend on the firm's pricing policy and on the elasticity of demand for its products, topics that are probably harder to study than wage rigidity. Knowledge of a firm's cost structure would also make possible accurate estimation of savings from layoffs.

22.3 Search for New Hypotheses

Explorations in search of hypotheses would most likely entail extensive interaction with economic decision makers and with observers of economic life.

WHY DO PAY CUTS HURT MORALE?

My conclusions about why pay cuts damage morale require elaboration, for my information comes from managers, who may not fully understand subordinates' feelings. It might be revealing to question workers about their reactions to hypothetical and actual pay cuts. I am aware of only two field studies of actual pay cuts, those of Greenberg (1989, 1990).

WHAT HAPPENS WHEN MORALE CRUMBLES?

Although many managers told vivid stories of having dealt with collapsing morale, such incidents have never been studied, as far as I know. It might be difficult to gain access to firms suffering from poor morale, for companies do not want to make their problems public. If cooperation could be obtained, both managers and workers should be questioned about their feelings and their understanding of the situation. Information should also be gathered on changes in productivity, turnover, and the ease of recruitment, since managers said these were strongly affected by morale. Such a study would be useful for management science as well as for understanding wage rigidity.

WHY ARE FIRMS RISK AVERSE?

Although theoretical speculations abound, little is known about why firms are reluctant to borrow money to cover losses and why they so dislike having temporary declines in profits. More information could be obtained by interviewing financial officers of firms as well as people who provide capital, such as lending officers of banks and insurance companies, and rich investors. Stock and bond brokers should be able to provide useful insights about the willingness of investors to advance money to companies in difficulty. The crucial issue is

why companies reduce investment expenditures in response to financial pressure.

ORGANIZATIONAL SLACK

It is not clear how organizational slack can be recognized and measured and why it exists when it does. These questions could be clarified by careful study of the use of labor in one or a few firms over the course of a business cycle. Key objectives would be the identification of waste and establishing the reasons for it.

WHY DO STRIKES OCCUR?

Despite many studies of labor negotiations and strikes, the causes of strikes remain unclear. A great deal might be learned by following actual negotiations and interviewing participants on both sides. Previous studies have not concentrated on the main issues of interest to economists, namely, the role of asymmetries of information and why the two sides do not avoid strikes.

22.4 Conclusion

The subject of economics has an enormous impact on everyone's life, and yet the discipline lacks the status of a real science, follows rather than leads ideological trends, and sometimes indulges in fanciful theoretical representations of reality. Many branches of economics are not anchored in empirical knowledge, probably because the subject originated as part of moral philosophy and is still regarded as having to do more with thinking than with observation. This attitude is compatible with the field's dependence on easily accessible statistical data, which, though essential, are also inadequate. Often it is not clear what they measure, and without this knowledge they can be used to support almost any contention. How can the unemployment rate be interpreted without knowledge of what it means to be unemployed? What sense can be made of wage data without knowing the impact on workers of pay raises and cuts? Empirical knowledge means systematic experience with the object of study, and this can be had only by taking responsibility for data collection. Unfortunately, the gathering of economic data is considered to be hardly within the scope of academic inquiry. This attitude is in conflict with the view expressed in texts on economic history that great progress in science and technology occurred only when thinking people came in contact with the practical realities of production and when pragmatic attitudes took precedence over religious and ideological ones. I do not claim to have made great discoveries or that my observations are unimpeachable, but I believe I have shown that much can be learned by confronting economic reality.