Volume Title: Business Cycles: Theory, History, Indicators, and Forecasting

Volume Author/Editor: Victor Zarnowitz

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-97890-7

Volume URL: http://www.nber.org/books/zarn92-1

Conference Date: n/a

Publication Date: January 1992

Chapter Title: Cyclical Aspects of Cost and Price Movements

Chapter Author: Victor Zarnowitz

Chapter URL: http://www.nber.org/chapters/c10375

Chapter pages in book: (p. 125 - 163)

4 Cyclical Aspects of Cost and Price Movements

4.1 Background and Concepts

4.1.1 Introduction

As argued in chapter 3, there is sufficient evidence to establish that business contractions were on the whole shorter and milder after 1948 than before World Wars I and II. A comparison of standard historical estimates and recent data also suggests that the general price changes were more volatile and more cyclical in the earlier periods.[1] Inflation has prevailed since 1950, and generally procyclical movements in the rate of inflation have replaced the earlier procyclical fluctuations in the price level. Thus, the age of *reduced* cyclical instability was also an age of *increased* downward rigidity of prices (and, it will be shown, money wages). This presents a problem because the classical and still prevailing view in the profession is that stickier wages and prices make the real economy *more* unstable.

In large part, the voluminous and rapidly expanding literature on "microfoundations" of macroeconomics consists of attempts to rationalize why wages do not adjust promptly and fully to changes in the demand for labor, and prices to changes in the demand for output. But great differences in the degree of price and wage flexibility (or stickiness) are found in historical studies and cross-sectional and cross-country comparisons (see, e.g., Sachs 1979,

1. This is strongly supported by the available long wholesale price series (Zarnowitz and Moore 1986, sec. 9.10; chapter 3, fig. 3.2 and table 3.1, in this volume). Consumer prices have always been much less variable, and here the change in behavior is subject to doubt. The standard estimates for the GNP deflator may have greatly overstated the historical price volatility by using only wholesale and not consumer prices (Balke and Gordon 1989, pp. 71–75). However, even consumer prices responded with significant declines to some of the business contractions of the past (see, e.g., Zarnowitz and Moore 1986, fig. 1), whereas after 1949 neither producer nor consumer price levels did.

1980; Branson and Rotemberg 1980; Gordon 1982a, 1983; Bruno and Sachs 1985; P. Andersen 1989). Apparently, business cycles coexist with a broad range of behavior in labor and product markets, though they may influence, and be influenced by, changes in that behavior. It is these interactions that should ultimately be of primary interest to students of macroeconomic performance and policy, not a postulated dichotomy of flexprice and fixprice models.

For good reasons, however, little is known as yet about what causes what here. Macroeconomics deals with a limited number of aggregate variables and potential sources of overall growth and instability; it finds it exceedingly difficult to deal with an evolving economy and heterogeneity of market firms and participants. Yet it is useful to ask how the main contemporary hypotheses bearing on business cycles and broadly defined price flexibility are related to each other. I will attempt to answer this question by selectively reviewing work in this area and assessing its major implications and problems in the light of statistical and historical evidence.

4.1.2 Hypotheses of Wage and Price Inertia

For a long time Keynesian macroeconomics focused on the inflexibility of nominal wages rather than prices as the main link between fluctuations in demand and output. But if it were typically the case that prices fall or rise less in business contractions while money wages continue to rise faster than prices, then real wages should tend to be countercyclical. This is simply not consistent with the evidence, as reviewed below.

Further, labor compensation is the largest component of both total income and cost of output, so high rigidity of wages would induce high rigidity of prices, which would at some point be inconsistent with the degree of efficiency in resource allocation expected of the market economy. The premise that nominal wages are unresponsive to cyclical demand fluctuations must be examined rather than maintained. It would seem more plausible to expect that money wage rates will on the average rise faster than the value of the marginal product of labor in periods of high employment and will rise more slowly or even decline in periods of low employment. The tendency for real wage changes to reflect productivity changes would then manifest itself mainly in the long run.

To be sure, a competitive firm will not knowingly pay real product wages in excess of the employee's marginal product, except transitionally and when it expects to recoup its short-term loss in the longer run. Similarly, the worker will not knowingly accept less than his or her marginal product, except under analogous circumstances. So any discrepancies between wages and productivity must be due to any or all of the following: (1) lack of knowledge by the firm, the worker, or both of what the relevant marginal products currently are; (2) uncertainty about the economic conditions facing the firm and the workers;

and (3) departures from competition, mainly in the direction of bargaining between large oligopolistic companies and large labor unions.

Conditions (1) and (2) are probably of long standing. However, productivity is particularly difficult to assess in various service industries, such as finance, insurance, trade, health, and education, and also in professions and sales, clerical, technical, and managerial occupations. Many of these areas have gained greatly in relative importance in recent times. Some large shifts of type (3) certainly occurred as well, but so did important countermovements in the direction of more global competition.

Research confirms the common observation that few labor markets are of the competitive auction type, where nominal wage adjustments clear the market continuously. Explicit contracts negotiated by unions and corporations determine wages of organized labor over staggered periods of 1–3 years. Implicit contracts whose durations run over years fix wages and salaries of career job holders. Here current compensation is an installment payment on a long-term employment relationship. It is often in the best interest of both employer and employee to maintain such an association on multivariate terms that are not very responsive to changes in general economic conditions considered transitory.[2]

Contractual arrangements between buyer and seller also exist in product markets, at least for limited time periods, and they tend to be associated with low frequency and small average size of price changes. But some prices remain unchanged over long time periods even where no such contracts exist (e.g., newstand prices of magazines; see Cecchetti 1986). There is evidence that product prices in many industries, particularly those that are more concentrated, change only infrequently.[3]

All this explains why the emphasis on wage rigidity gave way to more attention to price rigidity in recent years. The shift suggests the general notion that the main reason why workers lose jobs in recessions is not that real wages rise but that prices do not fall even though sales do. Also, it recognizes that, up to a point, prices can decline despite temporarily sticky wages when allowed by some combination of gains in productivity and reductions in nonlabor costs and profit margins.

4.1.3 Real and Nominal Rigidities

In the business cycle context, it is particularly important to ask how changes in the aggregate demand for real output (Q) influence the levels of

2. For a critical review of the writings on implicit contracts, see Rosen 1985. On the "stylized facts" and theories of the labor market, see Parsons 1986; Stiglitz 1986; and Kniesner and Goldsmith 1987.

3. See Carlton 1986. Other comprehensive studies of price adjustment and surveys of the large and growing literature include R. J. Gordon 1981, 1983; Blanchard 1987a and 1987b; Rotemberg 1987; Ball, Mankiw, and Romer 1988.

nominal prices and wages (P, W). If P does not react promptly or fully to a change in Q, at any given level of W, there is some real price rigidity. Analogously, if W does not react promptly or fully to a change in Q, at any P, there is some real wage rigidity. These partial adjustments involve changes in price-over-wage markups or real wages. They are complemented by interactions between changes in P and W, given Q. If P does not react promptly or fully to a change in W, there is some nominal price rigidity; in the opposite case, there is some nominal wage rigidity.

One way to formalize this double dichotomy and isolate the parameters that embody each of these rigidities, as implemented in Blanchard 1987a, is to estimate and analyze equations of the type

$$p = \mu w + \alpha q, \tag{1}$$

and

$$w = \eta p + \beta q. \tag{2}$$

Here q might be represented by output in equation (1) and by employment in equation (2). Each variable (p, w, and q) involves a lag polynomial applied to the logarithms of levels, or alternatively to log differences (e.g., $p = a(L)P$ or $p = a'(L)DP$). The smaller the values of μ and η, the greater the nominal price and wage rigidities, and the smaller the values of α and β, the greater the real price and wage rigidities. Unless $\mu = 1$, p does not adjust fully to changes in w, and the same applies, mutatis mutandis, to η, α, and β. Depending on the time frame and type of analysis, the adjustments can differ with respect to both how complete (or partial) and how fast (or slow) they are.[4]

To close this simple model, it is necessary to make q a function of some exogenous factors. Keynes attributed business cycles to shifts in *real* aggregate demand, due mainly to instability of investment. In contrast, contemporary models that bear the broad Keynesian label relate the cycles to shifts in *nominal* aggregate demand, due mainly to instability of money supply. Given this orientation, q is often taken to be proportional to real balances ($m - p$), where m is the log of the nominal money stock. Realistically, however, q should be thought of as depending on both monetary factors and nonmonetary ones (e.g., shifts in government spending and taxes, exports).

There are many good reasons for some degree of inertia in the response of relative wages and prices to changes in real demand. Some but not all of these explanations require major deviations from the basic competitive model. On the other hand, the problem of nominal rigidities simply does not arise under perfect competition. Suppose a sudden tightening of monetary policy causes

4. Blanchard (1987a) constrains the estimates so that p and w react to each other fully in the long run, but this restriction may not be appropriate (see comment by Sims on his paper).

a fall in aggregate demand, while no change occurs in technology or preferences. Under the classical competitive paradigm, since the equilibrium structure of prices remains unaltered, all nominal prices must decline proportionately. It is only in this way that the previous equilibrium levels of real money balances and real economic activity can be restored (or, rather, maintained: the process is executed promptly and smoothly by continuous price changes directed from open auction markets at atomistic agents).

4.2 Why Are Wages Not More Responsive to Declines in Demand?

4.2.1 Relative Wages and Efficiency Wages

For Keynes (1936, chs. 2 and 19), the idea that marks a key departure from "classical economics" is that labor is an aggregate consisting of groups that, whether or not formally organized, have as their main objective the protection of their *relative* real wages. Each group watches the others it compares itself with and bargains to obtain a possible advantage but at least not to fall behind. Thus, in a depression any group will resist an attempt to reduce its money wage. Although a uniform across-the-board wage cut may then be acceptable to labor, there is no practical way to engineer it in a diversified economy in which there is some dispersion of market power. The stickiness of relative earnings thus translates into a stickiness of nominal wage rates in general. Moreover, such partial reductions in these rates as come about are apt to do more harm than good, because they contribute to the spread of deflationary expectations and inhibit spending. In contrast, reflationary policies, by raising the price level and so lowering real wages overall, can provide the needed adjustments without disturbing the wage structure in ways to which labor objects.

If relative wages are what matters to groups of workers, then the net productivity of labor should be a positive function of these wages. In a sufficiently strong form, this relationship implies that a decline of real wages below a certain norm is associated with higher, not lower, unit labor costs. Here the old emphasis on the dependence of wages on productivity is replaced by a new emphasis on the reverse causation, the dependence of productivity on wages. Stiglitz (1986, pp. 182–92; 1987) provides the model for this analysis; for criticisms and references, see Katz 1986 and Murphy and Topel 1989.

Suppose that as the real wage of the ith labor group, ω_i, increases up to ω_i^*, the productivity of the group, λ_i, rises more than proportionately, while a further increase in ω_i is associated with less than proportionate gains in λ_i. Hence the cost per effective unit of labor service is U-shaped with a minimum at ω_i^*. The firm that employs ith labor will pay the "efficiency wage" ω_i^* with the intention of minimizing unit labor costs; it will not extend employment to

people who offer their services at a lower wage, because it believes that this would reduce productivity and raise costs. The unemployed cannot underbid the "insiders" if they are viewed as sufficiently less productive or likely to raise the effective costs of labor to the firm.

Efficiency wages could be paid for any or all of several reasons, but problems arise for each of these variants

1. Let the perceived relative wage determine each employee's effort and indeed willingness to work. This may be viewed as an individualistic version of Keynes's notion that any labor group is averse to a reduction in its nominal wage, which is inevitably a relative wage cut. A self-validating convention of this kind could in principle generate a pervasive downward stickiness of wages in a decentralized market economy. But historically money wages fell in depressions, and wage concessions are not uncommon even during shorter and less severe contractions (as in the early 1980s). Considerations of fairness may well be very important, but economic self-interest should be stronger still. High unemployment will increase the value of holding a job and the incentive to work hard. Should prices fall more than wages, as they did in the early 1930s, these motivations will be further enhanced, and the role of relative wages could well be greatly diminished.

2. In the modern versions of the theory, it is not labor's resistance to lower wages that is stressed but rather management's strategy to give workers incentives in the form of relatively high earnings and benefits. The idea is that this is required to attract and keep high-quality employees and have them work at high levels of effort. Well-paid workers are less prone to search, quit, or join a union and exercise more bargaining power collectively. As a result, the firm will benefit from lower labor turnover costs. But the high-wage option is not the only one that is available here, and it need not be optimal; for example, offers of greater training and promotion opportunities or greater job security may be more attractive to some firms and workers.

3. The efficiency wages ω_i^* are set by firms with some (perhaps only informationally based and short-run) monopsony or oligopsony powers in the labor markets. It is not clear how important empirically such conditions are. In the resulting interior solution, small deviations of wages from long-term equilibrium cause only second-order losses to the firm, so that nonadjustment of wages to minor shocks is "near-rational" (Akerlof and Yellen 1985). Here the effect of efficiency wages is to decrease the fluctuations in wages and increase those in employment.

4. Shirking and negligence can cause large damages to firms, particularly where many workers perform sensitive tasks using complex machinery and equipment (so can fraud and stealing, which may be more diffuse). The detection of such misconduct is taken to be costly in the so-called primary (high-wage) sector of the labor market, but not in the secondary (low-wage) sector. Hence, the primary jobs pay a premium above the market-clearing wage for secondary jobs, so as to deter shirking, even if the workers are all alike in

ability (Shapiro and Stiglitz 1984; Bulow and Summers 1986). Critics point out that superior market-clearing solutions for the shirking problem exist in the form of explicit or implicit contracts that involve performance bonds posted by employees.[5]

4.2.2 Implications for Unemployment and Some Evidence

To rationalize involuntary unemployment in the sense of excess supply of labor, a theory must rule out the existence of a sufficiently attractive and expandable flexible-wage sector. In the model with dual labor markets, workers in the secondary market have menial jobs with low prospects and high turnover. This acts as a reputational barrier against the employment of these workers in the primary market. The unemployment in this scenario is mainly frictional and long term, due to the high incidence of quits and separations in the secondary sector and to the job rationing and queues in the primary sector. Static models of this type provide no explanation of cyclical unemployment. They generally produce an overall tendency for real wages to be sticky or even constant (Solow 1979).

Efficiency wages, like the market-clearing wages they exceed, are unobservable; hence there is, understandably, no direct knowledge of their empirical importance. It is also difficult to define the primary and secondary sectors, since industries and occupations generally overlap both. The level of wages paid perhaps constitutes the main measurable distinction between the two sectors (Doeringer and Piore 1971; Lang and Leonard 1987).

Consequently, the evidence bearing on these theories is indirect; it is also limited and mixed. The observed relatively large and stable interindustry wage differentials may provide support for the efficient wage models (as argued in Krueger and Summers 1987, 1988) or they may be due to unobserved quality differences across workers (Murphy and Topel 1987, 1989).[6] Some labor market statistics appear to contradict the hypothesis that hiring from a queue of long-waiting jobless searchers is particularly important in the primary as distinct from the secondary labor market.[7] Also, it is commonly observed that people look for better jobs while employed and do not deem it

5. Direct and complete bonding is apparently rare, perhaps because of limited ability or willingness to borrow, mistrust between the parties, etc., but there are practical alternatives observed in long-term employment relationships. Workers may in effect buy responsible high-paying jobs gradually, by initially accepting low wages for simple, easily monitored tasks and then proving themselves and advancing along the job ladders within the firm. Participation in costs of specific training programs and nonvested pension plans can also be interpreted as forms of bonding. On the role of seniority wages to solve the incentive problem, see Lazear 1981.

6. It is agreed that heterogeneity of labor, differentials compensating for nonwage aspects of the industry, and institutional factors such as unionization account for a substantial part of the wage structure, but the unexplained remainder appears to be large.

7 According to data for men who changed employers and industry or occupation between any successive years in 1977–83, 23.7% of new hires in high-wage industries and 29.5% in low-wage industries experienced unemployment spells (averaging 3.0 and 3.6 weeks, respectively) between jobs (see Murphy and Topel 1988, table 2 and text).

necessary to quit in order to search more efficiently. It is unsatisfactory for a theory to rely on assuming the opposite.

Shifts in the aggregate demand for output and in the derived demand for labor presumably have less effect on the efficiency wage than on the completely flexible market-clearing wage. This suggests that employment and unemployment should be more demand sensitive, and probably more cyclical, in the primary than in the secondary sector. Other considerations point in the same direction. Primary workers have large amounts of human capital, much of it firm specific; their spells of joblessness are likely to be infrequent, concentrated in bad times, but also long because they are motivated and financially able to search extensively. Secondary workers have less human capital and are less attached to their jobs, which have relatively low pay and promise; hence they are unemployed much more often in both good and bad times, but generally over shorter periods because their incentives and opportunities to search are limited.

It is difficult to test for such dichotomies because well-matched and sufficiently detailed time-series data on unemployment and wages are lacking. What is readily established is that both unemployment and average hourly earnings tend to be relatively high in some major industries (durable manufactures, mining, construction, transportation, and public utilities) and low in others (finance, insurance, and real estate; trade; and services). In industries where unemployment is on average higher, it also tends to fluctuate more between business cycle turns and be more volatile. Where earnings are on average higher, they also tend to be more variable over time. Measures of cyclicality and volatility of unemployment numbers and rates are positively correlated with both the average levels and standard deviations of earnings by industry.[8]

These findings are broadly consistent with the notion that the high-wage sectors should display high cyclical sensitivity, but they also point to several problems and ambiguities. Should the positive relationship between the level and variability of average hourly earnings be confirmed by a detailed analysis, this would be difficult to reconcile with the hypothesis that higher wages are the sticky efficiency wages. The high-wage industries may be associated with more cyclical unemployment for reasons other than those advanced by the sticky-wage theory, including the reverse-causation argument that higher wages are paid in part because of the greater instability of employment. Moreover, there are some well-known factors that help explain the interindustry

8. For data on unemployment and average hourly earnings in private nonagricultural industries, 1964–82 (from U.S. Bureau of Labor Statistics, various bulletins), the correlations between the mean levels and standard deviations are .94 for earnings and .82 for the unemployment rates. The correlations across these measures between earnings and unemployment rates and also between cyclical amplitudes of unemployment and the average levels or variability of earnings have a range of .62–.75. For the underlying detailed measures, see Zarnowitz 1989a, table 2.

differences in sensitivity to business cycles, notably the durability and storability of products.

4.2.3 The Role of Long-Term Labor Contracts

Explicit employment arrangements or contracts cover a large proportion of workers who stay on the job, particularly those organized in labor unions. Presumably much more widespread, however, are implicit contracts that benefit the firm and its work force and are therefore, as long as they are so perceived, complied with by both, despite the absence of express, legally enforceable agreements. For example, firms routinely extend employment and grant raises from year to year, especially to employees on salaries. There is evidence that long-tenure career jobs are indeed very important in the U.S. labor markets today; for example, as shown in Hall 1982, about 40% of workers aged 30 and above are likely to hold jobs that will last twenty years or more. Current compensation for work of this type may be viewed as an installment payment on a long-term employer-employee association, and as such its influence on current employment is apt to be small.

If workers are generally more risk averse than businessmen and have less access to capital markets, they may in effect wish to buy from firms insurance against income decreases in bad times in return for smaller increases in good times (Baily 1974; Azariadis 1975). In the same implicit bargain, the firm is conceded much latitude in varying the volume of employment over time (not surprisingly, given the normal functions and requirements of business and the realization that fluctuations in the demand for the firm's output must be accepted as exogenous). Both the firm and labor gain if hours and effort rise (decline) when there is more (less) work to be done. Convincing arguments attribute considerable efficiency at the microlevel to contracts that include a variety of risk-sharing arrangements and such devices as temporary layoffs, overtime pay, and cost-of-living adjustments (see in particular Hall 1980b).

But even optimal contracts are constrained by inevitable gaps and inaccuracies of information and cannot allow for all contingencies that are hidden from foresight in the world of uncertainty. Thus, the early implicit contract models noted above can account for sticky real wages but not cyclical unemployment. It is not clear why workers should be so strongly concerned about wage fluctuations but not about the risk of job losses during a recession. The contract theories face various problems related to the observability and verifiability of the events covered, the complexity and enforcement of the contracts, moral hazard, and adverse selection. In particular, the more complicated models with asymmetric information depend on special assumptions on how firms and workers differ with respect to what they know and what their attitudes to risk are (Grossman and Hart 1983). Some of the more questionable premises are that *generally* firms are risk averse, workers are uninformed about the state of the economy, and contracts are effectively explicit but of short duration.

Coordination or externality problems are inherent in systems with imperfect

information and incomplete markets. The reason is simple: any change determined by decentralized decisions of parties to a contract or transaction concerns a relative or real price, since it is made given other prices. Thus nominal rigidities can arise from real rigidities, no matter what degree of microefficiency in real terms the implicit contracting may achieve.

It might seem that institutional means or contractual contingency provisions are readily available to reduce the nominal stickiness, at least in wages, by indexation. But informational lags and errors do not allow a complete wage indexation in practice. Moreover, to the extent that indexation does produce an approximately fixed real wage, it aggravates the effects of real shocks even as it reduces those of nominal shocks, so that, for example, the economy may suffer less from demand disturbances but more from supply disturbances (Gray 1976; Fischer 1977c). Also, if indexation is believed to reduce government's interest in pursuing consistent counterinflationary policies, the case for it is at least politically weakened (Fischer and Summers 1989).[9]

I conclude that the theory of optimal contracting cannot explain cyclical fluctuations in employment and unemployment. Indeed logically, the more efficient the contracts are in foreseeing or providing for contingencies, the less room they should leave for persistent departures from full employment. But inefficient labor contracts that incorporate informational asymmetries imply underemployment or work sharing rather than more layoffs than hires during business contractions. Simpler models with relative wage contracting (Taylor 1983; Summers 1988) may prove more helpful on nominal rigidities and unemployment but their consistency with rational expectations and optimizing behavior continues to be questioned.

4.2.4 A Summary of Wage Hypotheses

Table 4.1 provides for the reader's convenience a synopsis of the attempts to explain why wages are or appear to be sticky. After the preceding discussion, only two additional points remain to be made.

The first hypothesis is that real wage movements clear the labor market but are small because the supply of labor is highly elastic. This is consistent with the simple competitive model and assumed as a part of the recent theoretical framework of "real business cycles." It is contradicted by most tests of intertemporal substitution of leisure see (chapter 2, sec. 2.4.3) and generally by microestimates of labor supply elasticities (Pencavel 1986). Some reported results based on aggregate data show large elasticities (Kennan 1988) but they depend critically on prior assumptions and are very imprecise (Taylor 1988).[10]

9. General indexation extending to prices as well as wages presents still other difficult problems such as the choice of a commodity standard or numeraire in proposed monetary reforms (Blanchard 1987b).

10. Long time series reflecting the decrease in hours also indicate an inelastic or slightly negatively sloped supply, but postwar time series show substantially higher elasticities, especially for countries other than the United States (see Greenwald and Stiglitz 1988, pp. 223–29). In part,

Table 4.1 **An Outline of Hypotheses about the Apparent Stickiness of Wages**

Hypothesis	Critique
1. Supply of labor is highly elastic to transitory changes in the current and expected real return on work effort.	Tests do not support the idea that labor supply responds sensitively to current and expected changes in real wages and real interest rates. Evidence from panel data suggests that total labor supply elasticities are small.
2. Long-term implicit labor contracts ensure flexibility of employment for firms and stable real wages for workers.	Whether motivated by workers' risk aversion or informational asymmetries, the contract theory contributes to an explanation of sticky wages but not unemployment. Efficiency of the contracts is questioned.
3. Unions gain for their members wages above the competitive equilibrium, which results in unemployment among nonmembers.	The ability of unions to obtain and keep high wages depends on the degree of control they have over the entry into the work force. The hypothesis attracts more interest in Europe than in the United States because of the greater power of European unions, or "insiders."
4. Each labor group (unionized or not) protects, and in particular resists reductions in, its relative wage. This induces stickiness in wages generally.	A wage cut for any individual group may be resented as unfair and depress worker morale. Coordinated wage cuts across a decentralied economy are not practicable. A possible reason for nominal as well as real wage inertia. Criticized mainly for lack of roots in optimizing behavior.
5. Efficiency wages above the competitive equilibrium are paid by firms to attract and keep high-quality, high-effort workers and so minimize unit labor costs.	There are good reasons for productivity to rise with real wages, but the validity of the assumed form of this relation is uncertain. The links to relative wages and menu cost of wage change have some implications for cyclical unemployment, but most versions and elements of this hypothesis have none.

Note: See text for more detail and references.

Bargaining between a labor union with a fixed number of members and a sectoral monopolist can produce a sticky negotiated real wage in the presence of fluctuations in demand and employment, assuming that the unemployed members have an alternative source of income and that there is little variation over the cycle in the wage elasticity of labor demand. Let this situation prevail in the primary sector, while the competitive secondary sector offers much poorer jobs at a much lower flexible wage. Then unemployment will consist of those on temporary layoffs from the primary sector and those who wait for jobs in that sector. A fall in demand during a recession will widen the intersectoral wage differential. The models that give these results are worked out in McDonald and Solow 1981, 1985. Hypothesis 3 in table 4.1 states the relationship between unions and relative wage rigidity in general terms.

---

these differences are reconciled by the finding that the supply of labor for secondary workers is very elastic.

4.3 Why Prices Are Sticky; or, How Markets Clear

4.3.1 Competition and Costs

A combination of highly elastic supply and fluctuating demand produces a condition of relative price stability in equilibrium, which could be mistaken for a disequilibrium phenomenon of real price rigidity. This apparently simple and attractive theoretical possibility deserves to be considered.

The competitive firm's short-run supply function is horizontal at the minimum of its average variable cost (the threshold of profitable production) but elsewhere equals the firm's marginal cost (MC) curve, which slopes upward because of diminishing returns. The firm has strong incentives to operate as close to the capacity output (where MC rises steeply) as the price will allow. The supply curve for a competitive industry as a whole will tend to be still steeper because an expansion of industry output is apt to raise input prices and so shift upward the MC curves of all firms concerned. This argues against the hypothesis suggested.

However, it is not only the short-run supply responses that matter here. Since business cycles consist of movements measured in quarters and years, during which all factors of production eventually become variable, the analysis should include longer-run responses as well, where the supply elasticities will tend to be greater. Moreover, even over short periods of time firms will see some price changes as more "permanent" than others, and may differ in these assessments, so that what is being observed represents a changing mix of short-run and longer adjustments. On simple assumptions, notably of open access to given technology, the intermediate-run industry supply curves would probably be relatively flat in the middle output region and increasingly upward sloping at higher rates of production, which drive up MC.

It follows that the simple microeconomics of a competitive industry is not necessarily inconsistent with the story that high supply elasticities prevail for medium time horizons relevant for the analysis of business cycles. What is inconsistent with this hypothesis is the notion that sufficiently large demand-induced fluctuations in output can coexist with stable prices. When high levels of production are reached in a business expansion, overtime hours of work at extra pay rates become common and prices of many inputs rise sharply; at low output rates in a business contraction, the average workweek is cut and some input prices fall. These prices largely originate in competitive markets. They include the spot market prices of raw industrial materials known as "sensitive" and also wages of some categories of unskilled and unorganized workers.

The microeconomic argument in this section abstracts from changes in money and the general price level. It cannot explain why comprehensive price indexes seem to have responded less to decreases than to increases in demand during the recent predominantly inflationary period, in contrast to a more symmetric cyclical behavior in earlier times. The issue of the source of any

upward or downward price inertia is still open (Cagan 1979; Okun 1981). Indeed, studies of micro data sets on transaction prices find no support so far for the existence of either bias (Stigler and Kindahl 1970; Carlton 1986).[11] What they do find, however, is that many industrial product prices remain unchanged over long time intervals that must overlap business cycle phases. Thus the evidence in Carlton 1986, 1987 shows average durations of price rigidity exceeding one year in most cases and two or even three years in some.

4.3.2 Nonprice Market Clearing

Evidence of *long* periods of rigidity in prices of numerous diverse products cannot be reconciled with a competitive model in which price always adjusts promptly to equate supply to demand. It is simply improbable that both demand and supply schedules do not move or that their shifts persistently offset each other so as to leave the price unchanged. Some common observations are also inconsistent with the model of market clearing through price change only: at times there is unintended inventory accumulation of goods held for sale, and at times there are queues for goods that are in short supply (so either sellers or buyers are unsatisfied). These considerations lead naturally to the question of what mechanisms may help clear markets in addition to price, and how they are related to the competitive model.

I suggested a long time ago that the endogenously determined delivery period variable (k) is an important equilibrating device (Zarnowitz 1962). Prompter delivery enhances the value of the product and so raises the quantity demanded (ordered) per unit of time; it also raises the average cost of the quantity supplied. This implies, as a condition for profit maximization by the firm, that

$$dP - dc = -[(D_k/D_p) - C_k]dk = 0, \tag{3}$$

where $D_p = \partial D/\partial P < 0$, $D_k = \partial D/\partial k < 0$, and $C_k = \partial C/\partial k < 0$ based on the demand function $q^d = D(P, k)$ and the cost function $c = C(q^s, k)$. For any quantity demanded, say q_1^d, there is an indifference curve in the price/delivery period (P, k) space; and for the same quantity supplied, q_1^s, there is an indifference curve in the cost/delivery period (c, k) space. Joint optimization of P and k requires that the slopes of these loci be equal, so that

$$dP/dk = dc/dk \quad \text{and} \quad C_k = -(D_k/D_p) \tag{4}$$

in conformity to equation (3).

In this simple model, as long as there is some substitutability between P

11. Surprisingly little price-theoretic work seems to have been done in this area. It includes some studies which show that the asymmetry can result from the pricing behavior of profit-maximizing firms with some market power, nondecreasing MC functions, and nonincreasing demand elasticity (Kuran 1983, 1986) or from differential effects of inventory and backlog adjustments (see secs. 4.3.2 and 4.3.3 below).

and k, the fluctuations of demand will in general be met partly by changes in price and partly by changes in delivery period, inducing a positive correlation between the two variables. Any responsiveness of k should make P more stable over time, but a variety of outcomes is possible, reflecting the characteristics of the product and the state of the market.

Even a physically identical commodity differs economically to the user depending on the date of its availability. An impatient buyer will pay a little more for an earlier purchase but a sufficiently large differential over the expected or stated future price will cause postponement or placement of an order for future delivery. If cost or demand changes today affect both present and future prices, intertemporal substitution in the use of producer and consumer goods may increase the equilibrating role of delivery lags and reduce that of current prices. Thus when demand is elastic with respect to P but inelastic with respect to k, more of the market clearing will be accomplished by variations over time in k and less by variations in P.

Some industries take or solicit orders in advance of production and hold large backlogs of unfilled orders and small or no inventories of finished goods. Other industries make goods and hold them in finished-goods inventories for sale. Products that are customized and particularly costly or risky to store tend to be made to order. There is much evidence that production to order is very important, especially in manufacturing, and that it is associated with large fluctuations in backlog-shipment ratios (a measure of k) and relatively small variability of prices (Zarnowitz 1973). Large procyclical movements are also observed in the index of "vendor performance" based on data on delivery lags reported by purchasing managers.

Market-determined changes in delivery periods and unfilled orders need not be indicative of noncompetitive behavior causing excess rigidities in P. Instead, their principal explanation, consistent with a high degree of competition, may be simply that when buyers choose or consent to wait, smaller or slower price fluctuations are needed to clear the markets (Zarnowitz 1973; Carlton 1979, 1983, 1987). Demand is often more likely to be sensitive to changes in P than to changes in k.[12]

Where inventories of finished goods are held, they absorb the impact of temporary shifts in demand at least partly, thereby permitting avoidance, reduction, or deferral of price changes. However, the production-smoothing (buffer) role of inventories seems to be in the main limited to short unexpected demand shocks. For inventory investment to act as a strong buffer against longer fluctuations, it would have to be countercyclical, but instead it is on the whole procyclical and destabilizing as demonstrated, for instance, by the fact that GNP varies more over the cycle than final sales.

12. However, there is some evidence that unfilled orders tend to be larger relative to shipments in highly concentrated industries than in less concentrated industries (Zarnowitz 1962). This suggests to Scherer (1980, pp. 195–96) that variations in order backlogs are particularly important in oligopolistic industries.

4.3.3 Some Implications of Market Power

The conventional analysis of imperfect or monopolistic competition is limited to price-quantity solutions for given cost functions: the firm faces a downward-sloping demand curve and sets P above MC according to the optimal markup $P/MC = \eta/(1 + \eta)$, where η = elasticity of demand. This is a static theory that assumes full knowledge of demand and cost functions that firms in dynamic environments do not possess. Reliably estimating the current and future values of η is difficult and may not be feasible when demand is subject to unpredictable fluctuations or when uncertainty obscures the effects of the firm's own price adjustments on the behavior of other market participants. If the firm knows that P>MC, it will always wish to sell more at the current price but it may be reluctant to reduce price lest this should fail to stimulate sales sufficiently and so result in lower profits.[13]

Suppose the firm sees the uncertainty associated with price variations as greater than the uncertainty associated with output variations, which are to some degree reduced by changes in delivery lags and order backlogs or by changes in inventories. Then its preferred strategy would be to use price adjustments sparingly and conservatively and to rely more on quantity adjustments in response to fluctuations in demand that may prove transitory.[14]

Several recent studies (Hall 1986, 1988; Shapiro 1987; Domowitz, Hubbard, and Petersen 1988) present estimates of markup ratios P/MC that, for many diverse industries, substantially exceed the competitive benchmark of unity. Hall argues that spatial separation and product differentiation result in a wide dispersion of market power but not of extra profit opportunities. This is so because high setup, advertising, and fixed costs create much persistent extra capacity, absorb excess profits, and discourage new entry. In this view, firms operate most of the time above their (large) minimum efficient scales of production but below the maximum efficient scales, that is, where their MC curves are perhaps just slightly rising or nearly horizontal. So the observed stickiness of prices is seen as reflecting the hypothesized stability of costs under a regime of monopolistic competition. If both MC and η were (approximately) constant, firms would have little to gain from active pricing policies.

13. For a firm with some market power, linear demand and cost functions, and a homogeneous product, there will be more downward than upward price rigidity if the cost of shortages and stockouts exceeds the cost of inventory holding, and vice versa (see Amihud and Mendelson 1983; also Reagan 1982). This is an interesting notion but it remains to be seen how it would fare under less restrictive assumptions and with recognition of variable delivery periods and the distinction between made-to-stock and made-to-order goods. Thus, since backlogs always exist in production to order, they are not just an excess-demand phenomenon, and so their treatment as "negative inventories" is much too limited and potentially misleading.

14. For a development of this argument, see Zarnowitz 1962, p. 390; and 1973, pp. 301–5. For a recent formulation of the hypothesis that managers tend to be more uncertain about the consequences of their pricing and wage decisions than about those of their output and employment decisions, see Greenwald and Stiglitz 1989 (where inventory changes are considered but variable delivery lags are not). Even when quantity adjustments are otherwise costlier than price adjustments, the uncertainty differential may cause firms to prefer the former to the latter.

Data for manufacturing indicate that plant capacities are indeed designed to meet peak demands and are often much underutilized in cyclical slowdowns and recessions, even though the minimum efficient scales of operation are typically not so large. Early postwar and interwar estimates generally showed short-run MC to be declining or constant and long-run average cost functions to be falling and hence suggestive of economies of scale. But these regressions, being based to a large extent on observations of low levels of output, do not rule out rising costs in higher operational ranges, except probably for public utilities and railroads—industries long regulated as "natural monopolies." [15] Such microstudies can be informative mainly about long-run tendencies in the selected industries. About general dynamics of the behavior of costs and related variables, there is more to be learned from work on cyclical movements of monthly and quarterly aggregate time series.

4.3.4 Cyclical Changes in Cost, Productivity, and Profits

Several stylized facts established by studies in business cycles can be interpreted to favor the idea that MC is rising. In addition to the already noted cyclical sensitivity of the average workweek in manufacturing, there are the short but persistent lags of employment and the longer lags of inventories and unfilled orders, which suggest the presence of important costs of adjustment in labor input and production, respectively.

Indexes compiled and analyzed by Hultgren (1960, 1965) show that labor cost per unit of output and total unit cost in manufacturing, 1947–61, often declined in the late stages of contractions and regularly declined during recoveries and then rose as expansions matured and at times also rose well after a downturn. [16] This observed pattern is in a general sense consistent with the theory of U-shaped curves, since it implies that unit costs fall in the lower range of output and rise in the higher range, at least near the capacity levels. But it suggests additional dynamic elements in that the costs appear to depend on the rates of change as well as levels of production.

A decline in unit labor cost takes place when output per hour of work rises faster than the average hourly compensation; a rise occurs when the opposite happens. The former condition tends to prevail toward the end of a recession and notably during a recovery and high expansion; the latter condition tends to prevail in slowdowns and downswings. Labor productivity is basically procyclical and leading; its largest gains come at cyclical upturns and in recoveries when real growth rates are usually very high. On the other hand, wages

15. The verdict of a comprehensive survey paper by Walters (1963) was that this literature fails to convincingly refute the classical hypothesis of U-shaped cost curves; for an opposite view, see Simon 1979. Critics also charged that the techniques used had a bias toward linearity, which was disputed by the authors (Johnston 1960). Outside the regulated industries, the results are in fact rather meager and mixed. However, one should also note the findings from estimated production functions, which tend to provide broader support for constant returns to scale (Jorgenson 1974).

16. For railroads, a much longer record studied by Hultgren showed a stronger tendency for cost to be inversely related to output (in units of passenger and freight traffic).

often increase most in late stages of expansion and near cyclical downturns when labor markets become tight and output growth slackens. It is then that unit labor costs (ULC) begin to increase faster than the level of prices of industrial products. Total unit costs (UC) usually rise even more because of the developing shortages of raw and intermediate materials—inputs whose prices are known as particularly sensitive. As a result, markups on ULC and UC (P/ULC and P/UC), tend to turn down even while sales and output continue to grow. Profit margins, as inferred from aggregate accounting data, turn down with a similar early timing.

Figure 4.1 shows the average movements of ULC, the P/ULC ratio, output per hour, and the corporate profit margin before and during the recessions and recoveries of 1948–80.[17] Rates of change rather than levels are used for labor productivity and unit costs so as to eliminate the large but uneven upward trends in these variables. The resulting patterns demonstrate the strong inverse relationship between the cyclical change in average productivity of labor and the concurrent change in average costs of labor (see the left-hand half of the chart). The markup and profit margin variables are approximately stationary and here level data are used. The two series show nearly parallel cyclical movements most of the time, but the relative amplitudes are greater for the corporate profit margin than for the P/ULC ratio. Profitability responds directly to changes in productivity and inversely to changes in cost.

For the pre-1948 period, comparable data are not available but there is a strong presumption that the underlying relationships have a long history. In 1913 Mitchell predicted the sequence of changes in these variables without the benefit of data (see chapter 2, sec. 2.3.6 above). Hultgren (1950, 1960) documented with monthly series for 1932–58 that labor productivity led and ULC lagged in their procyclical movements. Consistent evidence is presented by Bernanke and Powell (1986) for 1923–29 and 1954–82 and by Klein and Moore (1985) for several major industrial countries.[18] An old and widely accepted explanation of why employment varies less than output is that firms hoard labor in sluggish times to preserve their human capital investments and because capacity and input adjustments take time and speeding them up is costly, particularly under uncertainty about demand. This hypothesis can also account for the tendency of output per hour to lead at business cycle turns.

The support that Mitchell's theory gets from the data, including the

17. The data are quarterly, mostly for the nonfarm business sector, from the U.S. Bureau of Labor Statistics. The BLS series, published since 1972, begin in 1948; they cover costs and profits per unit of output and also comparable prices received.

The series used in figure 4.1 contain 132 observations each. The mean duration of the six recessions in 1948–80 was 11 months, the distance between the vertical peak and trough lines in the graphs. For each recession-recovery sequence, each of the series was indexed to its level in the quarter of the business cycle peak and expressed in terms of two-quarter smoothed changes (following Moore and Cullity 1983, pp. 255–61 and fig. 16–5). The graphs in figure 4.1 represent averages of the resulting patterns. Measures of intercycle dispersion and inspection of the individual patterns indicate that these averages are fairly representative.

18. See also Fair 1969 and previous references in this section and chapter 2, sec. 2.3.6.

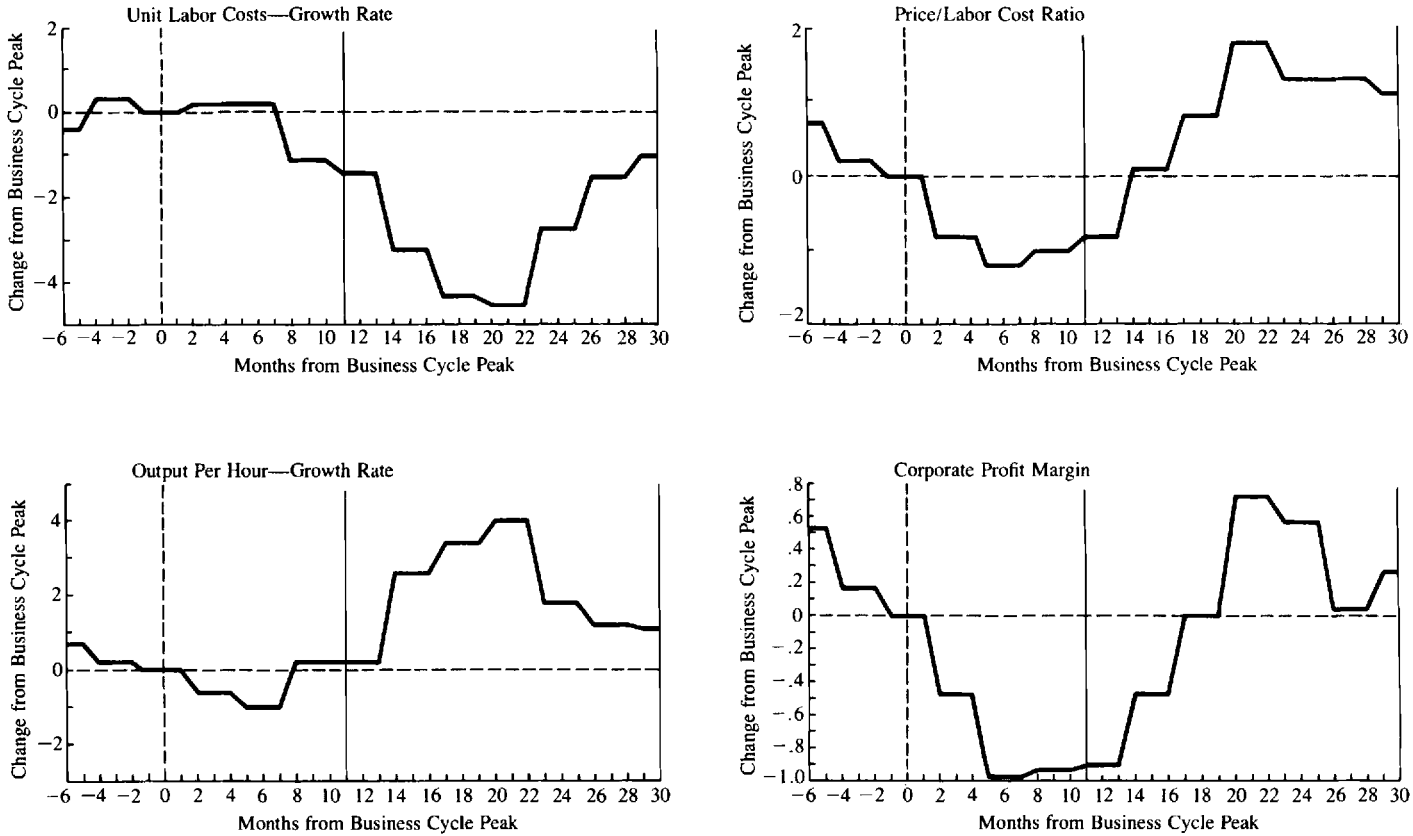

Fig. 4.1 Patterns of behavior of unit labor cost, markup, labor productivity, and profit margin during recession and recovery, 1948–80

Source: BLS data, which cover the nonfarm business sector, except that the corporate profit margin series refers to all nonfinancial corporations.

evidence for the prewar periods, when prices were relatively flexible, suggests that cyclical variations in costs, margins, and profits deserve much more attention than they have received in recent literature and that the stickiness of prices perhaps deserves much less. It should also be promising to reexamine Mitchell's approach in light of current ideas (and vice versa).[19]

4.3.5 A Summary of Price Hypotheses

The attempts to explain price inertia are summarized in table 4.2. Two of the five listed hypotheses require some additional discussion. I then close this section with a summary of a recent appraisal of the theories of price stickiness by business executives.

In oligopoly (hypothesis point 3 in the table), the few sellers know that their pricing decisions are interdependent and act accordingly. Suppose each firm expects its rivals to match any price cut it would make but not a price hike, which means that each firm faces a demand curve that is kinked at the current price. This describes a situation in which the price would not change in response to moderate shifts in cost; it does not explain what the price is or how it was determined. Much of the time oligopolists may simply play it safe and settle for the implicit collusive outcome of the game. For example, it pays them to raise their prices simultaneously as soon as they recognize shared increases in cost or demand. Undercutting the collusive price can be individually rewarding, but the risks of a price war are large (see Green and Porter 1984; Porter 1985). If each firm is reluctant to move first on the price front, long periods of price rigidity are likely; if price leadership emerges, intermittent industry-wide price changes would be the expected result. The argument, then, reduces to the already familiar theme that commitments to maintain relative prices produce nominal rigidities as firms refrain from initiating price changes. Empirically, the macroeconomic importance of the oligopoly case is at least uncertain (for an attempt at a generalization, see Woglom 1982).

A hypothesis that gained more attention recently relies on monopolistic competition where firms produce differentiated products and have some market power in price setting. Firms have profit functions that are flat at the top,

19. Mitchell viewed business cycles as sustained by self-generated stresses and imbalances (as did other prominent scholars of his era), whereas most contemporary writers assume that macroeconomic fluctuations are caused by exogenous shocks. Thus, Mitchell would agree with the modern "real business cycle" theorists that changes in productivity are procyclical and very important, but not that they are exogenous, generalized shocks to technology that on their own "drive" business cycles. He would clearly disagree with the treatment of the average price-cost margin as close to constant (which is often found in recent models of wage and price rigidities). Recent regression estimates of industry P/MC markups find mostly procyclical behavior but yield mixed results (Domowitz, Hubbard, and Petersen 1986, 1988; Morrison 1988). This suggests that different market structures can produce different cyclical behavior, which may well be true (e.g., if η fluctuated procyclically, P/MC should be countercyclical—but there is no evidence that either this condition or its opposite prevails). However, these studies use short annual series, which cannot reveal leads and lags of less than a year's duration and may not be adequate to determine the cyclicality of the markups.

Table 4.2 **An Outline of Hypotheses about the Apparent Stickiness of Prices**

Hypothesis	Critique
1. Marginal cost (MC) tends to be flat and stable over time; hence so is the competitive price P = MC.	Competitive auction markets cleared only by price are costly to organize and maintain. Most markets deviate from this model. MC should be rising for the firm in competitive equilibrium.
2. $P >$ MC by a stable margin because of monopolistic competition with flat and stable MC and constant elasticity of demand.	MC should be rising at least near capacity (where overtime payments are common) and is unlikely to be constant over the large range of output covered by business cycles (BC). Average cost and profit margins vary during BCs, in part in offsetting patterns.
3. Sticky oligopoly prices reflect kinked demand curves and/or desire to avoid costly price warfare.	Oligopoly prices may be sticky in response to *moderate* cost and demand changes, given the uncertainty about rivals' reactions, but large changes can induce the few sellers to alter price simultaneously.
4. Nonprice market-clearing mechanisms reduce the frequency of price changes under various types of market structure.	Procyclical movements in delivery lags demonstrably help absorb changes in demand, but the price-smoothing role of changes in inventory is not clear. Factors that may favor nonprice instruments: intertemporal substitution, product heterogeneity, advertising, search costs, and steady customer relations.
5. Each price change has a fixed menu cost, which is not worth incurring because the loss of profit from nonadjustment of price is second-order small.	The direct menu costs are probably very small and unlikely to prevent price adjustments in response to *large* changes in demand. The less elastic MC and the more elastic demand, the larger are the costs required to be effective. Results are uncertain if prices are initially not in equilibrium. Indirect costs, such as customer goodwill, may matter more.

Note: See text for more detail and references.

with zero derivatives with respect to price at the optimum; hence small deviations from the equilibrium price involve only small (second-order) reductions in realized profits. A firm that believes its price is set about right in relative terms has little incentive to change it, even when the "menu cost" of doing so is small. Yet the aggregate result of the individually rational decisions to keep prices unchanged is price-level stickiness, which can have large (first-order) macroeconomic costs. Suppose reduced monetary growth lowers aggregate demand: if all firms reduced their prices even slightly, they would collectively benefit from an increase in real money balances and a decrease in the level of interest rates.[20]

20. On the other hand, a small increase in monetary growth that would leave prices unchanged (because of menu costs) would have a positive welfare effect in this model by raising output. The reason for this is that equilibrium output in this imperfectly competitive economy is suboptimal compared with the competitive model because of an aggregate demand externality (Blanchard and Kiyotaki 1987).

The last model in table 4.2 sums up the menu cost hypothesis in its simplest form. In short, this model raises the possibility that weak price rigidities (and wage rigidities; see Akerlof and Yellen 1985) may be associated with strong fluctuations in output and employment (Mankiw 1985; Parkin 1986). But the applicability of the idea is still to be established. The costs of producing new menus or price tags seem trivial, but it is worth noting that price changes, unlike quantity (output, inventory, backlog) changes, must be publicized. The costs of calculational inconvenience to customers and possible goodwill losses may be much more important (Okun 1981; McCallum 1986). If the total menu costs are small, they may not prevent price adjustments in response to large shifts in aggregate demand. The less elastic the labor supply and MC, and the more elastic the demand curves, the larger are the required menu costs. Finally, the dynamics of the process, about which little is known, depends on whether prices are initially in equilibrium, on the timing and staggering of pricing decisions, and on potentially important nonprice adjustments. Multiple equilibria are possible, which reflects how individual prices interact and react to changes in the price level (Rotemberg 1987; Blanchard 1987b).

In a very recent study (Blinder 1991), business executives were carefully interviewed about (1) the extent of the observed price stickiness and (2) the perceived validity of different theories of infrequent price changes. According to the preliminary results of this well-designed research project, 55% of the 72 companies interviewed so far "repriced no more often than once a year" (p. 93). Thus, many prices are indeed "sticky" in the sense of showing relatively infrequent changes. They respond to shifts in demand and cost conditions with lags averaging 3–4 months.

When ranked by the firms, the hypothesis of variable delivery lags (nonprice market clearing) scored best. "Coordination failure"—the idea that firms wait for others to alter price first—ranked second, cost-based pricing, implicit contracts, and explicit nominal contracts ranked third, fourth, and fifth, respectively. The menu costs hypothesis (cost of the price adjustment) scored still lower, followed by the notion of procyclical demand elasticity. Five other hypotheses received more rejections than support. They include the theory that inventories rather than prices are adjusted in response to demand changes and the theory that both marginal costs and markups are constant over the business cycle.

4.4 Historical Evidence and Current Theories

The presently popular "rigidity" theories pay little if any attention to historical changes in the behavior of wages and prices. This section takes a look at these changes and asks what lessons history holds for a comparative appraisal of the extant hypotheses. Some international differences are also considered.

4.4.1 Nominal and Real Wages in Business Cycles

Measures of aspects of cyclical movements in several historical time series of wage earnings and rates in the prewar, interwar, and postwar periods are presented in table 4.3. The average annual money earnings from wages declined in about half of the business contractions of 1860–1914 and in all of those of 1920–38, according to the data compiled in Phelps Brown 1968 (lines 1–2). In contrast, they kept rising through the period 1945–60, which witnessed four moderate or mild recessions (line 3). Data for 1889–1914 from Rees 1961 show that peaks and troughs in annual earnings matched nearly two thirds of the like business cycle turns of the period, but those in hourly earnings fewer than half (lines 7 and 8). However, hourly earnings score high on cyclical conformity in the interwar period, according to both the annual series from Rees 1960 and the monthly series from Creamer 1956 (lines 9 and 11). The same applies to an index of money wage rates presented in Creamer 1950 (line 12).

The conclusion is that most of the major business downturns and some of the minor ones have historically been associated with declines in nominal wage earnings.[21] The comprehensive study by Phelps Brown, which goes back to 1860 and also covers France, Germany, Sweden, and the United Kingdom, supplies much additional evidence that money wages basically followed a procyclical pattern of movement in the eight decades preceding World War II, though skipping many smaller fluctuations.

However, it is clear from all available data that no cyclical declines at all occurred in any of the comprehensive money wage indicators for the United States after 1945. Such long and sustained expansions in wage rates and incomes, hitherto unprecedented, prevailed concurrently in other major industrialized and market-oriented countries as well, reflecting both the generally rising prosperity and inflation of the postwar era.

The real wage series used in Table 4.3 are estimated by dividing the money wage series by cost of living or consumer price indexes. Before 1914, more business cycle turns were matched by like turns in deflated annual and hourly earnings than in the corresponding nominal data, but the real series had more irregular timing and smaller amplitudes (cf. lines 1 and 4, and lines 8 and 10). In 1920–38, deflated annual earnings moved in sympathy with all major business contractions and recoveries but skipped the smaller cycles, which the money earnings did not (cf. lines 2 and 5).

In 1945–60 the real series declined only between 1945 and 1947, lagging behind the business contraction of 1944–46, which marked the reconversion

21. For example, the annual Rees series on hourly earnings shows downward movements in 1893–95, 1896–98, 1903–4, 1907–8, 1913–14, 1920–22, 1924–25, 1929–33, and 1937–39 (see Rees 1961, pp. 33–34; 1960, pp. 2–3). Earlier data in Douglas 1930 are broadly consistent with these results. The monthly series declined strongly but with long lags in response to the depressions of 1920–21 and 1929–33 and very mildly and sluggishly during and/or after the other business contractions of the interwar period (Creamer 1950, 1956, ch. 5).

from the war to the peacetime economy (cf. lines 3 and 6). The currently available quarterly Bureau of Labor Statistics (BLS) data on the real hourly compensation of production (nonsupervisory) workers on private nonagricultural payrolls provide strong, consistent evidence that these real wages did in fact keep rising throughout the 1947–60 period with its four recessions (line 13). However, the same series shows much less growth and more cyclicality in recent times; it flattened in 1969–70 and declined mildly but with considerable persistence in 1973–74, 1978–81, and 1983–84, matching five of the eight business cycle turns of the period since 1961 (line 14).

No detrending or differencing operations are needed to answer the question whether wages had declines corresponding to the contractions as dated by the NBER. Questions about the degree of responsiveness of real wages to changes in employment require more subtle analytic techniques. Studies of this problem, which are generally restricted by the data to very recent and short segments of time, produced mixed results. One reason is that the composition of the work force varies, in part systematically, over the business cycle. Expansions generate more overtime income, and those who change jobs and new entrants have more procyclical and on the average lower wages than those who stay on the same job. Hence the net effect of aggregation is a countercyclical bias, as shown in Bils 1985.

Note also that real wages as discussed so far are money wages deflated by the consumer price index (CPI), not by the producer price index (PPI) of the current output of labor. The wage in terms of consumables is of prime interest to workers, whereas the wage in terms of the products of work (product wage rate) is of prime interest to firms. Producer prices tend to vary more than consumer prices; if money wages are less flexible than PPI but more flexible than CPI, then the PPI-deflated real wages could turn out to be countercyclical or acyclical and the CPI-deflated real wages procyclical.[22]

Bernanke and Powell (1986) find that real wages (CPI-deflated) were weakly procyclical, lagged output significantly in the prewar (1923–39) period, and more nearly coincided with output or even led it in some industries in the postwar (1954–82) period. Product wages (PPI-deflated) led more often, particularly in the recent era, and had larger and more erratic variations and some countercyclical tendencies.[23] Both real and product wages were more serially persistent and less cyclically variable in the postwar than in the prewar period.

Money wages show long and pronounced upward trends, so an analysis of

22. See also Tsiang 1947 for an early analysis of the U.S. interwar data that suggests a similar distinction between the behavior of real wages and that of product wages.

23. The results for product wages presented in Sachs 1980 are generally similar. Geary and Kennan (1982) report that product wages are not significantly associated with employment in the United States and other OECD countries, but this may be due to the noisiness of these relationships in the time domain (as noted by Bernanke and Powell, who find stronger indications of cyclicality from the frequency domain).

Table 4.3 **Money and Real Wage Earnings and Rates: Measures of Cyclical Conformity, Timing, and Amplitudes, 1860–1987**

		No. of B.C. Turns[a]		No. of Timing Observations[b]			Average Lead (−) or Lag (+)[c]		Average Percentage Amplitude[d]	
	Period (1)	Covered (2)	Matched (3)	Leads (4)	Coincidences (5)	Lags (6)	Peaks (7)	Troughs (8)	Expansions (9)	Contractions (10)
					Average Annual Money Wage Earnings (Phelps Brown)[e]					
1	1860–1914	28	14	0	11	3	0.7(1.2)	0.1(0.4)	26(12)	−9(6)
2	1920–1938	10	10	0	8	2	0(0)	0.4(0.6)	18(20)	−13(13)
3	1945–1960	8	0	0	0	0	n.o.	n.o.	n.o.	n.o.
					Average Annual Real Wage Earnings (Phelps Brown)[f]					
4	1860–1914	28	15	5	7	3	0.5(1.9)	−0.4(0.8)	22(17)	−9(12)
5	1920–1938	10	6	0	6	0	0(0)	0(0)	33(17)	−6(5)
6	1945–1960	8	2	0	0	2	+1	+1	44	−8
					Average Annual Money Earnings, Manufacturing (Rees)[g]					
7	1889–1914	16	10	0	10	0	0(0)	0(0)	16(8)	−5(6)
					Average Hourly Money Earnings, Manufacturing (Rees)[h]					
8	1889–1914	16	6	0	6	0	0(0)	0(0)	15(9)	−4(3)
9	1920–1939	10	8	0	3	5	0.2(0.5)	+1(0)	19(18)	−9(10)
					Average Hourly Real Earnings, Manufacturing (Rees)[i]					
10	1890–1914	16	10	1	6	3	0.2(1.0)	0.5(1.0)	10(5)	−2(1)
					Average Hourly Money Earnings, Manufacturing (Creamer)[j]					
11	1921–1938	9	9	0	0	9	0.9(0.4)	0.6(0.5)	13(18)	−5(12)

					Index of Money Wage Rates, Manufacturing (Creamer)[k]					
12	1920–1935	8	8	0	0	8	0.6(0.2)	0.9(0.6)	4(6)	−7(8)
					Index of Real Average Hourly Compensation, Nonfarm Business Sector (BLS)[l]					
13	1947–1960	8	0	0	0	0	n.o.	n.o.	n.o.	n.o.
14	1947–1987	16	5	3	0	2	−0.8(1.1)	0(0.4)	4(2)	−3(3)

Note: n.o. = no observations.

[a]Business cycle turns (peaks and troughs) as dated by the NBER.

[b]Lines 1–10: Based on the annual business cycle (reference) chronology of NBER because each of these series is annual. Lines 11–12: Based on monthly series and the monthly NBER reference chronology. Line 13: Based on a quarterly series and the quarterly NBER reference chronology.

[c]Expressed in (fractions of) years. Entries not in parentheses are means; entries in parentheses are standard deviations.

[d]Based on specific cycle movements (expansions and contractions) in the series covered. All such movements are included, regardless of whether or not they correspond to the business cycle phases in the NBER chronology. The averages refer to the total expansion and contraction amplitudes in percentages. Entries not in parentheses are means; entries in parentheses are standard deviations.

[e]Based on the current-dollar data in Phelps Brown 1968, app. 3, tables on the last five pages (corresponding to pp. 448–52), col. 1 in each case. Coverage: 1860–1914, manufacturing; 1920–38, bituminous coal mines, class I steam railroads; 1945–60, manufacturing, mining, communications, contract construction, railroads and bus lines, gas and electric utilities. Computed from hourly and weekly earnings by taking into account changes in the average numbers of working hours and weeks per year. Sources include Long 1960 and Rees 1961 (see Phelps Brown 1968, app. 2).

[f]Based on the "index of wage-earnings in composite units of consumables" in the tables referred to in n. *e*. Money wages deflated with the cost-of-living or consumer price indexes. The entries in line 6, cols. 7–10, are single observations.

[g]Based on data in current dollars, per full-time equivalent worker, in Rees 1961, p. 33 (table 10, col. 1).

[h]Line 8: Based on data in current dollars in Rees 1961, p. 33 (table 10, col. 5). Line 9: Based on data in current dollars in Rees 1960, p. 3 (table 1, col. 1).

[i]Based on data in 1914 dollars in Rees 1961, p. 120 (table 44, col. 1).

[j]Based on data in Creamer 1956, p. 40 (table 8, col. 2) and p. 48 (table 11, col. 2) for the entries in columns 3–8 and 9–10, respectively.

[k]Based on data in Creamer 1950, p. 7 (table 1, cols. 3 and 7) and pp. 43–45 (table A) for the entries in columns 3–8 and 9–10, respectively. Creamer used monthly data on changes in wage rates compiled by the U.S. Bureau of Labor Statistics (BLS) from reports in its establishment and payroll sample for 1919–35.

[l]Based on monthly BLS data taken from *Business Conditions Digest*, October 1987, p. 104, and other issues.

the cyclical behavior of their first differences or deviations from trend must complement the level analysis. Looking at differences between rates of change (last year of expansion minus contraction) in the average hourly compensation in manufacturing, Sachs (1980) finds that they were negative in 13 of 17 business cycles between 1890 and 1975.[24] Thus a deceleration of growth or disinflation in wages marked most of the business cycle downturns. Absolute declines in wages, however, occurred only on five of these occasions, all of them before World War II. According to these annual data, wage inflation actually accelerated in the recessions of 1969–70 and 1973–75.

In sum, the study of historical data makes it clear that the inflexibility of money wages is not a universal law (as some of the recent literature would suggest) but is essentially a phenomenon of the post–World War II period. (Much the same statement can be made about the stickiness of the price level in the face of aggregate demand changes, as demonstrated below.) Real wages do not conform closely to business cycles but do fluctuate in a predominantly procyclical fashion. Before 1914 and between the world wars, these movements tended to be somewhat less frequent and smaller than those of nominal wages, but the situation has been reversed in the recent era of persistent price and wage inflation.

4.4.2 Trends and Cycles in Producer and Consumer Price Indexes

The main fact about the recent evolution of both wages and prices is that they ceased declining during the postwar business recessions; that is, disinflation replaced deflation. The evidence is substantial and uncontradicted (Cagan 1975, 1979; Zarnowitz and Moore 1986; chapter 3, this volume). The following is a brief qualitative summary of the findings that are relevant here (for numerical detail, see Zarnowitz 1989b, table 2).

The first 150 years of U.S. history can be divided into three periods marked by inflation (associated to a large extent with wars) alternating with three periods of deflation. The cyclical conformity of prices, as indicated by the percentage of business cycle peaks and troughs matched by like turns in the wholesale price index (WPI), was considerable throughout and, interestingly, on average higher in the deflationary than in the inflationary periods. When the trend in WPI was up, the index rose strongly in expansions and fell weakly in contractions; when the trend was down, the movements in the index were likewise procyclical, only tilted in the opposite direction. Also, in the former periods the year-to-year increases in the index were more frequent than the decreases, whereas in the latter periods the opposite was the case. For the era 1789–1932 as a whole, the decreases were nearly as frequent as the increases.

24. The Great Depression and World War II years are excluded from these measures. Prices and wages fell in 1929–33 but rose strongly through the rest of the 1930s despite high unemployment. This is widely viewed as anomalous and attributable to special factors, namely the New Deal legislation, the support and growth of unions, and later the wartime administrative controls.

The record for the CPI is somewhat shorter but it too leaves no doubt about the basically procyclical behavior of prices between 1820 and 1932, even on an annual basis. Again, the distinction between the inflationary and deflationary periods shows up clearly in the data. The frequency of decreases in somewhat less here than for WPI.

In 1932–52, a very turbulent age of depression and wars, prices embarked on a long upward trend but they still declined in two of the three business contractions, the exception being the recession of 1945, which marked the economy's reconversion to peacetime. However, in the three following decades the annual price indexes ceased altogether to decline in response to recessions. Indeed, in 1952–82 for the first time both WPI and CPI rose on average more in the years of contraction than in those of expansion.

Monthly data contain more noise than the annual data but also more information, particularly on persistence and timing. In 1953–64, 35% of the monthly changes in WPI were effectively zero; in 1964–76, 19%; and in 1976–88, only 5%.[25] Of course, as inflation intensified, there was a steady rise in the share of index increases (from 39% of the time in 1953–64 to 74% in 1976–88), but the share of decreases exceeded 20% in both 1953–64 and 1976–88 (it fell to 12% in 1964–76). However, these decreases were sporadic and did not amount to cyclical movements matching the business contractions after the mid-1950s. Indeed, in the 1970s the largest rises in the WPI occurred during the recessions associated with the huge oil price shocks. The recessions did cause some reductions in the rate of inflation but only with considerable lags. The monthly CPI series shows an even greater preponderance of increases and generally higher inflation rates.

4.4.3 Trends in Unionization and Competition

Wage and price rigidities are by no means universal but they are treated as such in most contemporary models that try to explain them. We should favor any theory that can account not simply for the existence of the rigidities but also for their apparent increase in recent times.

Several of the relevant hypotheses assume deviations from the competitive model in the form of negotiated wages and price setting by firms. Did the degree of monopoly—of market power of labor and business—rise historically and produce reductions in wage and price flexibility? There is no clear and established answer to this broad question, but the issue cannot be avoided. My reading of the literature and evidence suggests that (*a*) unionization probably did contribute to the changed behavior of wages after the Great Contraction of 1929–33 and in the first two or three post–World II decades and (*b*) it is rather doubtful that a general increase in the degree of monopoly occurred,

25. It should be noted that the early annual index series carry no decimals. The monthly series for both WPI and CPI carry one decimal each.

although this probably does not rule out the possibility that noncompetitive practices or conduct became more common.

Unionization

The sensitivity of wages to cyclical changes in unemployment and related labor market conditions is smaller in the union than in the nonunion sector so that recessions generally increase the pay advantage of union workers (Rees 1973, chs. 10 and 16; D. J. B. Mitchell 1980, chs. 4 and 6). The rapid growth in unionization, from 6% of the labor force in 1933 to 27% in 1953, was naturally associated with large increases in the role, scale, and costs of collective bargaining. Impasses and strikes can be particularly costly to both parties in a dispute; therefore, long, often 3-year, union contracts became common in the postwar period.

However, union membership in the United States has fallen as a share of the total labor force since the mid-1950s and in absolute numbers since the mid-1970s, despite a rise in the public sector. Employer resistance to unions stiffened as growth rates of output and labor productivity declined and competition intensified internationally (Freeman 1988; Reder 1988). Wages and benefits negotiated by major unions grew in 1968–79 at annual rates averaging nearly 10% for the first year and 7.5% for the second and third contract years, according to BLS data. During the recession year of 1982, these rates fell rapidly to about 3% over the life of the contracts, and they remained there (occasionally even lower) in the ensuing expansion. Correspondingly, real average hourly earnings and compensation of production workers on private payrolls grew only weakly and discontinuously in the 1980s.

This remarkable development surely reflected the loss of economic and political power of the unions but also other, partly related and temporary factors: disinflation, import penetration, and declines or slow growth of some of the older domestic industries. At the same time, the civilian unemployment rate in the United States fell from 10.8% in December 1982 to 5.0% in March 1989. In contrast, unemployment rose and remained high in Europe, where unions are generally much more powerful, despite the concurrent business expansion. Theories that blame insiders for high real wage rigidity and persistent unemployment gained considerable popularity there, and they are indeed consistent with many, though by no means all, of the relevant facts.[26]

26. Lindbeck and Snower (1986) argue that current employees ("insiders") have the power to make it unprofitable for the firm to hire potential employees ("outsiders") at lower wages. Thus the burden of unemployment falls on the outsiders, while the insiders, who may be protected by union membership, enjoy both the stability of indexed incomes and a high degree of job security. A more general theory of hysteresis or high persistence of unemployment is developed in Blanchard and Summers 1986, 1987b; for a different view, see Gordon 1987. As noted in Andersen 1989, the union wage theory does not explain why some of the Nordic countries, where the share of unionized labor is very high, had relatively low unemployment.

Concentration and Market Power

Formidable conceptual, measurement, and aggregation problems beset all attempts to estimate the extent (and, even more, the trend) of competition versus monopoly in the economy. Most of the evidence comes from concentration ratios, particularly for manufacturing, but the relationship between these measures and the degree of monopoly depends on the size of the domestic market, the importance of foreign competition, the availability of substitutes, and the extent of collusion. Views on the issue have long ranged from the popular but undocumented belief that competition is steadily and perhaps strongly declining to the more cautious and supportable suggestion that there is much stability over time in the aggregate concentration and no clear unidirectional trend.[27] The share of the 100 largest corporations in total manufacturing assets moved up from 34% to 42% in 1927–33, down to 38% in 1941, up from 37% to 46% in 1947–57, and up again from near 45% to 48% in 1966–71 (Scherer 1980, fig. 3.1, p. 48). There is much agreement that close approaches to both pure competition and pure monopoly are rare, but some authors stress the prevalence of broadly defined or "workable" competition, and others stress that of monopolistic competition and oligopoly (cf. Nutter 1951, p. 44, and Scherer 1980, p. 67).

If it were well established that market power increased and competitive price taking gave way to price setting under conditions of imperfect competition, models that rely on monopolistic competition and oligopoly should have a better chance to explain the apparent rise in price inertia. (As shown in Carlton 1986, the average length of spells of price rigidity is an increasing function of industry concentration.) But there is no convincing evidence that greater monopoly power is what actually distinguishes the last 40 years from the earlier era. Large corporations setting prices have been around for a long time. The increasing globalization of markets is an important postwar trend that presumably has the net effect of raising the levels of competition and the importance of changes in relative prices and wages.

4.4.4 How to Explain the Observed Changes

The expected consequence of labor's experience in the depressed 1930s would be an increase in workers' aversion to the risk of job losses. This should have given rise to greater demand for employment insurance than for wage stabilization. But labor contracts commonly specify basic income to be paid

27. For an early study that encouraged the former thesis, see Berle and Means 1932; for the latter position, see Stigler 1949 and Nutter 1951. Stigler's judgment was that "competition declined moderately from the Civil War to the end of the nineteenth century, and thereafter increased moderately" (p. 54). Nutter's estimate of the share of manufacturing output accounted for by monopolistic industries (those with four-firm concentration ratios of 50 and higher) was 33% for 1895–1904, and Scherer (1980, p. 68) obtains approximately the same figure for 1963 (the corresponding proportions were 24% in 1947, 30% in 1954 and 1958, and 29% in 1972).

for work and fringe benefits; they do not ensure tenure. In the postwar era, the cyclical variability of money wages has practically vanished, but real wages have stagnated recently for many workers. Moreover, unemployment continues to have large cyclical swings and it even has a definite upward trend. It is difficult to see how the implicit contract theory can explain these phenomena.

The idea that above-equilibrium real wages are required to attract productive and loyal workers should apply better to prosperous than depressed times, to large established companies with internal labor markets than small firms with uncertain prospects, and to high-paying career jobs than low-paying menial jobs. It is possible that the applicability of the efficiency wage theory increased during the postwar ear, which was one of long phases of expansion, with strong growth of career employment in the corporate sector and elsewhere (government, professions). But this is presently only a vague speculation; in fact, there is as yet little tested knowledge to bear on the validity of models of this type.

Much the same applies to the menu cost theory of pricing. The overall costs of changing prices not having been measured, we simply do not know what they are, how they vary across time and space, and how they compare with the costs of changing quantities (of output, inventories, backlogs). We do know that stable prices reduce communication costs to sellers and shopping costs to buyers, which is conducive to long-term associations between the transacting parties. Perhaps the importance of such associations in the economy has increased in recent times but, again, the facts of the matter are yet to be established.

Long and staggered union contracts probably played an important role in reducing the cyclical variability and increasing the persistence of wages during the post-Depression era when the unions were on the rise. Such contracts can generate complex distributed lags in wages and prices (see chapter 2, section 2.4.7). But in the second half of the postwar period the share of organized labor in the United States, never very high, was falling. And outside North America, including countries where unions are very strong, wage negotiations are mostly annual and/or are centralized or simultaneous (Andersen 1989, pp. 43–44).

A comprehensive theory that can explain the reasons for both the persistence and the evolution of cycles remains a major unaccomplished task. But when the observed changes that are likely to have contributed to the relative stabilization of the economy in recent times are taken as given, it can be seen that they are probably also responsible in part for the reduction in the cyclicality of wages and prices. Prices as well as quantities tend to be cyclically less sensitive for services than for goods (Moore 1983, ch. 12). Hence the large rise in the relative importance of services will have had the effect of making the overall cost and price indexes more sticky. As the needs and preferences of consumers and producers grew more diversified and technologies to satisfy them were being developed, the share of output made to order may well have

increased, too. If so, the result would be a greater role for delivery lag adjustments and a smaller one for price adjustments (perhaps also some reductions in the overall levels of manufacturers' finished inventories and short-term demand uncertainty).

The persistence of inflation in the last four decades should have rationally promoted indexation of incomes and increased the frequency of individual prices rises; that is, it probably worked to reduce nominal wage and price rigidities on the upward side but may have raised them on the downward side. But all this was apparently consistent with much inertia in the micro price data.

Concerning relative price adjustments, the effects of government policies were partly stabilizing and partly interfering and destabilizing. To the extent that the experience of greater macroeconomic stability generates expectations of more of the same, the probability that business recessions will remain relatively short and mild increases. Because buyers see less reason to reduce spending, sellers see less reason to reduce prices of inputs and outputs. But the optimism can be reversed by shocks or side developments such as the accelerated inflation in the 1970s, and both positive and negative expectations can be temporarily self-fulfilling.

In sum, it is not coincidental that the cyclicality of wages and prices declined during the last half-century even while business cycles moderated and inflation prevailed. But it is still possible that more flexibility would have helped. Between 1966 and 1980 ULC in manufacturing, which used to fall during business contractions, experienced only retardations as money wages rose strongly despite the adverse supply shocks and a decline in the growth rates of gross productivity of labor (whose procyclical and leading behavior pattern, however, remained unchanged). So new concern was expressed about the stickiness of wages, which was linked by some authors to the stagflation of the 1970s (Moore and Cullity 1983; Haberler 1988).

Yet the classical view that more flexibility is always desirable is not universally accepted. In what follows an opposite hypothesis will be considered.

4.5 Price Expectations and Interest Rates

4.5.1 Expectational and Distributional Effects of Deflations and Inflations

A macroeconomic model represents flexible prices by a steep, and sticky prices by a flat, aggregate supply (AS) curve. If the short-term fluctuations in aggregate demand (AD) are treated as given, the flatter the AS curve is, the more real activity will fluctuate. Hence, by this simple argument, price flexibility is necessarily stabilizing.

It is possible, however, that the variation of AD itself is not independent of the flexibility of wages and prices in general but rather is an increasing function of that flexibility. For this to be so, the expected rate of price change (p^e)

must have a strong positive effect on AD, and the actual price level (P) must have a weak negative effect. If P declines in a recession, this raises the real values of the money stock and net wealth of the public and hence acts to reverse the decline. But if at the same time p^e turns negative (i.e., a deflation is widely expected), then this would tend to depress AD by causing postponement of purchases.[28]

A similar distinction between level and expected-change effects applies to nominal wages. A reduction in the current wage level relative to the prospective future wage level would mean lower production costs now and hence encourage employment and investment, but the anticipation that wage rates generally will continue to decline would depress expected incomes, demand, and profitability and result in deferment of consumption and investment.[29]

A point-counterpoint discussion of these arguments led Keynes (1936, ch. 19) to the conclusion that on balance, there is little to recommend gradual reductions in money wages as a cure against depression. If prices fell in step with wages, no advantage of lower real costs would accrue to the employers; if prices were destabilized and fell more, conditions would deteriorate further. Decentralized wage bargaining can alter the relative money wages, not the overall level of real wages.

Moreover, it has long been recognized that when debt contracts are set in nominal terms, deflation worsens the financial position of debtors by raising the real value of their liabilities (Fisher 1933; Keynes 1936; Hart 1938; Minsky 1975, 1977). Such contracts certainly exist on a large scale in the United States and many other countries. True, the assets of creditors increase in real value at the same time, but the two effects need not and probably do not offset each other completely (Tobin 1975, 1980). Debtors spend more than creditors relative to their respective incomes and wealth (this is why the former borrow and the latter lend in the first place). Debtors who face increasing threats of insolvency are apt to curtail their spending sharply, and those who actually go bankrupt directly cause losses to creditors. In the 1930s, the results probably included a sharp rise in the real costs of financial intermediation, impeding the functioning of credit markets and depressing macroeconomic activity (Bernanke 1981, 1983). Simulations in Caskey and Fazzari 1987 suggest that a large real-debt effect can make greater wage and price flexibility seriously destabilizing.[30]

28. In the early statement, Patinkin (1948) contrasted the roles of P and p^e under the headings of static and dynamic analysis, respectively, and noted that the expectational effect, though presumably temporary, could well prove the stronger in the short run (see also Tobin 1975).

29. Other possible consequences of wage cuts include, on the positive side, lower interest rates (through reduced money demand), higher net exports (through reduced costs and greater competitiveness in an open economy), and possibly more business optimism. On the negative side, redistributional effects detrimental to workers and debtors would be likely to diminish aggregate demand.

30. Note that the argument about the depressant effect of a rise in the real burden of nominal debt and the attendant redistribution of wealth relies on the current movement of P and not p^e. A

In the absence of a general deflation, relative price movements combined with nominal debt contracting can seriously hurt some industries, occupations, or sectors, but these are partial difficulties that normally remain contained. The case in point is the experience of debt-burdened farmers after the 1981–82 recession when prices of their products and land stopped rising or fell.[31]

Jut as unanticipated deflation shifts income from debtors to creditors, unanticipated inflation does the opposite. The associated wealth redistribution effects are probably substantial, but individual losses and gains largely offset each other so that the costs to society vary and are difficult to determine (Fischer and Modigliani 1978). If money wages are sticky, unexpected price-level rises would reduce real wages and increase output. If there are no nominal rigidities but people misperceive the absolute for relative price rises, then unanticipated inflation would also increase output by raising labor supply.[32]

Large inflations or deflations of monetary origin are likely to obstruct, rather than promote, the flexibility of relative prices that is instrumental to the proper functioning of the market economy. Actual wage and price flexibility being always limited, deflations often aggravated the severe depressions of the past, and inflations (such as the long and uneven one of the 1970s) often contributed stresses and imbalances to expansions.

History suggests that what matters is not so much the direction but rather the size and variability of the general price movement. Large and volatile changes in the price level are difficult to predict, generate uncertainty, and adversely affect real economic activity. Evidence for the 1970s, when both inflation and unemployment rates were rising and inflation was increasingly variable, is consistent with this hypothesis (see, e.g., M. Friedman 1977; Makin 1982; also chapter 17). So are, a fortiori, the data for the short but severe depression of 1920–21, which was preceded by a large inflation and accompanied for a year by the most conspicuously rapid deflation on record for the United States (Friedman and Schwartz 1963b, ch. 5). On the other hand, one can find examples of moderate deflation or inflation coexisting with either high or low real growth rates.[33]

---

failure of foresight is still required, but only one that occurred earlier, at the time when the debts had been contracted for. The debtors must have failed to anticipate the downturn and decline in *P.* which is easy to explain if these were rare events.

31. The industry and regional troubles associated with the oil price decline of 1986 constitute another example.

32. Analogously, surprise deflation has negative effects on real aggregate economic activity in both those Keynesian theories that have nominal contracts with limited or delayed wage adjustments and those neoclassical theories that assume flexible wages and prices but also incomplete information.

33. M. Friedman and Schwartz (1963a, pp. 41–42) describe the post–Civil War period 1865–73 as one of deflation in product prices but not money wages, a mild rise in the stock of money, and a high average rate of increases in real income. Zarnowitz and Moore (1986, p. 553) compare some periods of low or negative inflation and relatively high real growth (e.g., 1923–29) with

4.5.2 Models with Potentially Destabilizing Price and Interest Expectations

Taylor (1986a) holds that the postwar improvement in economic stability occurred *despite* the greater rigidity of wages and prices, mainly due to policy-induced reductions in nominal GNP shocks. Monetary authorities react to rising inflation with tighter policies; this depresses output growth, which in time reduces inflation. The result has been smaller variances of the growth rates of output and prices but also greater persistence (serial correlation) of the fluctuations in both variables.

In contrast, De Long and Summers (1986a, 1986b) argue that at least some of the postwar moderation of the cycle is due precisely to the fact that money wages and prices have become more sticky and persistent. They assume that the short-run nominal interest rate (i) enters the LM equation, and the real rate ($r^e = i - p^e$) enters the IS equation. IS and LM together determine aggregate demand, while Taylor's model of overlapping wage contracts is used on the supply side. There is a serially correlated demand shift term but no wage or price shock. The model implies that increased price flexibility is destabilizing at the margin over a considerable range of assumed parameter values. Again, the reason is that output reacts more elastically to p^e than to P, but the channel is r^e.[34]

However, as shown by Driskill and Sheffrin (1986), increasing price flexibility is stabilizing at the margin in systems where fluctuations are driven by supply shocks (as in contract models with shocks to nominal wages; see Taylor 1979, 1980a). Hence greater price flexibility may simultaneously increase the output variance generated by demand shocks and reduce the output variance generated by supply shocks. The former effect will prevail over the latter in a model in which demand shocks are serially correlated and supply shocks are not. But the postwar era witnessed some large supply (or price adjustment) shocks of considerable persistence. It is clear that generally the results will depend on the mix and correlations of the demand and supply shifts.

Further, the assumption of a policy rule that links money supply to the interest rate leaves no room for any stabilizing role of monetary policy. Yet, as shown in chapter 3, monetary growth was much less volatile in the post–World War II era than in the earlier periods and was, moreover, free of the previously experienced phases of negative values. Policy may deserve some credit for this development, even if it is to a large extent endogenous (cf. H.

---

others of high inflation and lower growth (e.g., 1969–81). When inflationary trends prevail, business expansions tend to be longer but not necessarily stronger (Zarnowitz and Moore 1986, pp. 525–31).

34. If people expect disinflation or deflation (i.e., p^e falls or becomes negative), then so long as i does not adjust fully, r^e will increase, lowering real investment and output. Correspondingly, when p^e rises, r^e will fall, which raises economic activity. Hence, more price variability causes more output instability.

Grossman 1986). The stabilizing potential of flexible prices would be much enhanced under conditions of stable growth of nominal demand in the will-o'-the-wisp model with optimal monetary policy (S. King 1988).

In basic analytical terms, the short-run relationship between i and p (two endogenous, simultaneously determined variables) is unstable because it depends on the nature of the shifts in, or disturbances to, the system as a whole. This argument is sufficient to show that there is no simple and definite way to relate changes in r^e to economic fluctuations. Still, even a highly stylized attempt at some quantifiable linkage should be worthwhile. In times when persistent inflation is recognized and expected to continue undiminished, the simple price adjustment equation

$$p = p_{-1} + \frac{b(Q - Q^*)}{Q^*}, \tag{5}$$

may not be a bad assumption.[35] (Here Q is actual and Q^* is potential output; subscripts t are omitted.) A dynamic version of the demand function for money is approximately

$$m - p = kq - \frac{h\Delta i}{i}, \tag{6}$$

where $\Delta i = i - i_{-1}$. Following the derivation in Baily 1978 (pp. 42–45), consider an AD shock that raises Q above Q^* by a fraction q as money and prices increase at the equilibrium rate $\bar{m} = \bar{p}$. Then, from equation (5),

$$p = \bar{p} + bq, \tag{7}$$

and equations (6) and (7) imply that

$$\frac{\Delta i}{i} = q\left(\frac{k + b}{h}\right). \tag{8}$$

Suppose that the deviation of p from $\bar{p}$ in equation (7) is anticipated. Then, differencing the definitional equation $r^e = i - p^e$ and using equation (8) results in

$$\Delta r^e = \Delta i - bq. \tag{9}$$

And combining equations (8) and (9) gives

$$\Delta r^e = q\left(\frac{ik + ib - bh}{h}\right). \tag{10}$$

When equation (10) is solved for alternative values of the parameters (table 4.4), several relationships emerge. First, when inflation accelerates strongly

35. Survey data for 1959–76 indicate that forecasts of inflation have been on the average closely related to the most recent observed values of inflation (Zarnowitz 1979).

Table 4.4 **Hypothetical Responses of the Real Interest Rate to a Demand Disturbance**

A. $i=.06$, $k=0.7$, $h=0.2$		B. $i=.06$, $b=0.2$			C. $k=0.7$, $b=0.2$, $h=0.2$	
b	Δr^e	k	h	Δr^e	i	Δr^e
0	.21	0.7	0.2	.07	.02	−.11
0.2	.07	1.0	0.2	.16	.04	−.02
0.4	−.07	1.5	0.2	.31	.06	.07
0.6	−.21	0.7	0.1	.34	.08	.16
0.8	−.35	0.7	0.2	.07	.10	.25
1.0	−.49	0.7	0.5	−.09	.12	.54

Note: Based on the formula $\Delta r^e/q = i\,(k/h + b/h - b)$; see Baily 1978 and text for the derivation. A disturbance to AD is assumed to raise output from equilibrium by 1% ($q = 1$). Δr^e = change in real interest; k = income elasticity of the demand for money; h = interest elasticity of the demand for money; b = coefficient of the GNP gap in the price adjustment equation; i = nominal interest rate.

in response to excess demand as measured by the GNP gap (i.e., b in eq. [5] is large), Δr^e falls, given the values of i, k, and h. So in this sense greater price flexibility can be destabilizing in the short run (see pt. A of the table). Second, given i and b, Δr^e increases for higher values of the income elasticity of the demand for money (k) and decreases for higher values of the income elasticity of the demand for money (h). When h is as large as 0.5, changes in r^e are destabilizing (table 4.4, pt. B). Third, at values of k, b, and h that seem reasonable (0.7, 0.2, and 0.2, respectively), Δr^e is negative at relatively low values of the nominal interest rate (i) (pt. C). The meaning of this is not clear. The presence of i in equation (10) makes Δr^e dependent on money growth even in the long run, but the proper focus of this analysis is certainly on moderate short-run movements.

The examples suggest that r^e will change only by small amounts in response to a rise or fall of 1% in total output: for plausible parameter readings, perhaps at most by 30 or 40 basis points, most likely less.[36] This seems to be a sensible result that is consistent with much that is known about the behavior of interest rates and price expectations.

On the other hand, the De Long–Summers hypothesis implies that changes in r^e strongly influence aggregate economic activity, presumably through their effects on real private investment broadly defined (I). Since these changes are generally small, this requires a high elasticity of I with respect to r^e. Although such a major role for r^e is consistent with theory, it has not found much systematic support in the data. Most tests show relatively weak effects of cost of

36. Baily's own preferred figures correspond to line 2 in part A of table 4.4.

capital or the real interest rate on business and household real capital outlays, and some show no significant net effects at all.[37]

4.5.3 Interest and Inflation in History

What the *expected* inflation rates and real interest rates are at any time is always very uncertain, as indicated by the diversity of estimates of p^e obtained from different sources or by different methods (surveys of consumers and forecasters; inferences from financial and commodity market data; statistical implementation of the rational expectations model). Historically, the response of nominal interest rates to movements in the rate of price change has been quite varied and mostly weak.

Before World War II, inflation and deflation alternated in peacetime expansions and contractions, allowing for longer trends; short-term interest rates moved procyclically, like wholesale prices but with longer lags, and long-term rates had much smaller fluctuations (Zarnowitz and Moore 1986, pp. 553–65). Short rates moved in broad sympathy to the very large waves of inflation and deflation during World War I and thereafter through the slump of 1929–33, but even then their adjustment was lagged and incomplete. The ex post, observable real rate, $r = i - p$, was dominated much of the time by short erratic variations; it shows little association with output and other cyclical variables but a strong negative correlation with inflation (Mishkin 1981b).

The evidence is consistent with the hypothesis that most of the time people viewed the price movement at transitory and limited, so that their expectations smoothed out much of it. Hence, the changes in p^e tended to be small and not persistent so that the r^e and i values were on the whole close.[38]

A markedly tighter relation between i and p arises only in the second half of the 1960s and through the 1970s, a period during which the persistence of inflation must have become public knowledge. This is shown clearly by the graphs and correlations in M. Friedman and Schwartz 1982 (pp. 527–46; see also Summers 1983, pp. 216–25). The real rate rose from negative to high positive values during the great inflation of this period, which must have largely reflected the expectational adjustments of i to p. It should be noted, however, that the relationship weakened again during the following years of disinflation: r remained high in 1981–85 as i declined much less than p did.

The De Long–Summers hypothesis relies on the "nonadjustment" of nominal interest rates; that is, it requires that i respond only sluggishly and incom-

37. The literature is voluminous. For surveys and references see Jorgenson 1971; P. K. Clark 1979; Chirinko 1988. See also the estimates and critique of investment equations in Gordon and Veitch 1986.

38. That this was so looks particularly plausible for the pre–World War I era of the gold standard, when the long-term rate of price change was near zero (which, incidentally, helps explain the shape and relatively good fit of the original Phillips curve for the United Kingdom; see Phillips 1958 and Barro 1987, ch. 16).

pletely to changes in p^e. If the response is sufficiently strong, the change in r^e will be too small to have much of a destabilizing effect. Thus, the increase in the promptness and size of interest adjustments during the period of accelerated inflation in the late 1960s and the 1970s should have reduced the applicability of the hypothesis. And, according to the theoretical argument already noted, the same applies to the effects of the supply shocks, which also occurred in the 1970s.

To conclude, the validity of the hypothesis that stickier wages and prices had a moderating influence on postwar business cycles remains open. It was neither shown to draw much support from the data nor invalidated. The theoretical possibility of price flexibility being destabilizing clearly exists, but the historical importance of such a condition is uncertain.

A more limited hypothesis, which I think is favored by general considerations and evidence, is that only large changes in p^e and r^e associated with major deflations, inflations, and disinflations have much destabilizing potential; small and moderate changes, which are far more common, matter little.

4.6 Concluding Observations

The sensitivity of wage and price levels to cyclical declines was generally greater before than after World War II. The recent era was also one of more moderate fluctuations in real economic activity and more persistent inflation. Important structural, institutional, and policy changes contributed to all three of these concurrent developments. These linkages elude the current analytical models, which either disregard the evidence on the stickiness of industrial input and output prices or treat this stickiness as invariable, ubiquitous, and necessarily symmetrical.

Considerations of lesser risk and cost to long-term transaction partners may favor nonprice market-clearing devices even under conditions of substantial competition, as illustrated by the importance of changes in delivery periods and quality and availability of products. Relative wage cuts are disliked by nonunionized as well as unionized workers. Uncertainties about the consequences of price changes can deter them even if the direct costs involved are small.

Imperfections of competition, information, and markets make some wage and price rigidities inevitable. The departures from flexibility need not always be destabilizing; indeed, protracted and anticipated wage declines can aggravate demand contractions. But large wage or other cost increases outrunning productivity gains can also worsen the economy's condition. The strong hypothesis that the increased rigidities actually reduced the instability throughout the postwar era does not score convincingly against alternatives. The counterarguments draw support from the increased responsiveness of interest rates to changes in expected inflation and the importance of supply shocks in the 1970s.

The destabilizing potential of general price movements is probably nonlinear: great for large and rapid changes and negligible for small and slow changes. Major deflations of the past had demonstrably strong and adverse expectational and distributional effects. As the recent inflation accelerated, it grew increasingly volatile, generated much uncertainty and popular discontent, and led to policy interventions that had disturbing consequences of their own. Moderate fluctuations in the price level or the rate of inflation have not been shown to be necessarily detrimental to reasonably steady growth in real activity.