What Is New-Keynesian Economics?

By ROBERT J. GORDON
Northwestern University and National Bureau of Economic Research

This research was supported by the National Science Foundation. I am grateful to George Williams for help with the data and to Steven Allen, Laurence Ball, Olivier J. Blanchard, Timothy Bresnahan, Charles Calomiris, Dennis Carlton, Robert Chirinko, Russell Cooper, Stanley Fischer, Herschel Grossman, R. Glenn Hubbard, David Laidler, John Leahy, Assar Lindbeck, N. Gregory Mankiw, David Romer, Julio Rotemberg, Dennis Snower, John Taylor, Andrew Weiss, and two anonymous referees for comments on one or more earlier drafts. I am also indebted to Michael Parkin and Edmund S. Phelps for helping to establish the etymology of the phrase “new-Keynesian.”

I. Introduction

A. Background

In the late 1970s it appeared that the U.S. macroeconomic landscape was being swept by a new-classical tide, and that Keynesian economics had become an isolated backwater. In fact there is still a widespread impression that the best and brightest young macroeconomists almost uniformly marched under the new-classical banner as the decade of the 1980s began. Yet it is now apparent that the rumors of the death of Keynesian economics were greatly exaggerated. Building on foundations laid in the late 1970s by Stanley Fischer (1977a) and Edmund Phelps and John Taylor (1977), a large number of authors, young and middle-aged alike, in the past decade have produced an outpouring of research within the Keynesian tradition that attempts to build the microeconomic foundations of wage and price stickiness. The adjective new-Keynesian nicely juxtaposes this body of research with its arch-opposite, the new-classical approach.

The strongest written statement of the dominance of new-classical macroeconomics among the younger generation is by Alan Blinder: “By about 1980, it was hard to find an American academic macroeconomist under the age of 40 who professed to be a Keynesian. That was an astonishing intellectual turnabout in less than a decade—an intellectual revolution for sure . . . the young were recruited disproportionately into the new classical ranks. . . . By 1980 or so, the adage ‘there are no Keynesians under the age of 40’ was part of the folklore of the (American) economics profession” (1988, p. 278).

The label new-Keynesian should be attributed to Michael Parkin (1982), who has offered me the opinion that he originated the term new-Keynesian theory, not new-Keynesian macroeconomics. The term new-Keynesian theory was incorporated into a chapter subsection in Phelps (1985, p. 562) and “new-Keynesian model” in a chapter title in the fourth edition of my textbook (Gordon 1990), written in 1986. One of the first uses of the label new-Keynesian economics in a scholarly article is by Laurence Ball, N. Gregory Mankiw, and David Romer (1988). The word new rather than neo to describe the recent work in the classical tradition distinguishes it from what Paul Samuelson in the early postwar period called the neoclassical synthesis of old-Keynesian
This paper extracts the essential elements of new-Keynesian economics for an audience of professional economists who are not specialists in the microeconomic foundations of macroeconomics. There is no intention to survey comprehensively every notable paper in the field, but rather to sift the literature for the most important ideas and themes. One commentator has asserted that the new-Keynesian literature has provided too many explanations of wage and price stickiness, and so we apply tough standards to the major contributions, asking whether they make an essential contribution to an understanding of the adjustment of wages and prices. In short, our intent is to ask what is new and what is convincing in the large literature that collectively has become known as the new-Keynesian economics.

B. Main Themes

Like its precursor a decade ago (R. Gordon 1981), this paper differs from conventional surveys not just in its intent to sift and criticize rather than to provide a broad and evenhanded overview. It also contains a substantial empirical prologue before reaching the core material on new-Keynesian theory. The prologue (Parts II and III) argues that there are three different dimensions of price stickiness (which we will label the inertia, rate-of-change, and level effects). A brief survey of the emerging literature in the new empirical industrial organization, together with a new empirical time-series investigation of price adjustment across time and countries, reveals the essential fact that any satisfactory theory of price adjustment must explain the variability of price adjustment parameters across industries, across countries, and across historical intervals. We ultimately reach the verdict that much of new-Keynesian theory does not succeed in explaining these facts.

The prologue (Parts II and III) is followed by the core of the paper, the critical review of theoretical contributions in the new-Keynesian literature. The review is organized by recognizing two central distinctions, the first between price setting in product markets and wage setting in labor markets, and the second between nominal rigidity and real rigidity. The theoretical analysis in the paper is organized into a treatment of main themes and issues (Part IV), and discussions of nominal rigidity in the product market (Part V), real rigidity in the product market (Part VI), and models of labor market rigidity (Part VII), followed by a conclusion (Part VIII).

The task of new-Keynesian economics is to explain why changes in the aggregate price level are sticky, that is, why price changes do not mimic changes in nominal GNP. Sticky prices imply that real GNP is not an object of choice by individual workers and firms but rather is cast adrift as a residual. Thus new-Keynesian economics is about the choices of monopolistically competitive firms that set their individual prices and accept the level of real sales as a constraint, in contrast to new-classical economics in which competitive price-taking firms make choices about output.

Why do changes in the aggregate price level fail to mimic changes in nominal GNP? Two main themes emerge from the theoretical review, (1) the reasons for the absence of nominal GNP indexation of individual prices, and (2) the reasons why, in the absence of such indexation, individual prices fail fully to reflect changes in nominal GNP. Underlying the first theme is an essential element of any industrial economy—the role of

---

macroeconomics and classical microeconomics. In turn, the word new rather than neo is used for the recent work in the Keynesian tradition, so that it can be properly juxtaposed to the new-classical approach.
idiosyncratic elements of cost and demand. Firms care about the relation of their own price to their own marginal cost. But because idiosyncratic shocks cause their own costs and demand to evolve differently than nominal aggregate demand, firms have no reason to accept the risk involved in indexing their price to nominal aggregate demand. The absence of nominal GNP indexation opens the way for theories of real rigidity to explain the sources of nominal price stickiness.

The second theme is that, in the absence of nominal GNP indexation, changes in individual prices will respond to changes in individual marginal costs, not changes in nominal GNP. Thus the aggregate price level will be sticky unless firms expect changes in their own marginal costs to mimic changes in nominal GNP. Yet they have no such expectation. In the framework that I label the input-output approach, each of thousands of heterogeneous firms is enmeshed in a web of intricate supplier-demander relationships. The input-output element helps to explain why firms do not simply assume that marginal costs will move in parallel with aggregate nominal demand: Most firms do not know the identity of all of their suppliers, their suppliers’ suppliers, and so on. The input-output approach places equal emphasis on the purchase-material and labor-cost components of marginal cost and points to models of real rigidities in the labor market, including the efficiency wage and insider-outsider models, to help explain why prices are less flexible in some industries than in others.

An important empirical finding in Part III is that prices were sticky not just in the Great Depression and the postwar era, but long before World War I. This fact casts doubt on institutional sources of price and wage rigidity, for example, labor unions, and reinforces our emphasis on universal features of microeconomic structure. In our treatment, price and wage stickiness emerges from a core set of microeconomic elements that are timeless and placeless: a technology of transactions, heterogeneity of goods and factor inputs, imperfect competition, imperfect information, and imperfect capital markets. Because these core elements remove any incentive for individual agents to focus on nominal demand in making their own price-setting decisions, their presence supports the traditional view that Keynesian economics is fundamentally about the macroeconomic externalities of individual decisions and the coordination failure inherent in a free-market economy.

C. The Dichotomy Between Supply and Demand

With much ground to cover, there are many interesting topics in macroeconomics that cannot be treated here. The coverage is limited to the determinants of aggregate supply behavior, roughly, the division of a change in nominal GNP growth between changes in prices and output, and the role of wage stickiness (if any) in contributing to price stickiness. The entire demand side of the economy is omitted as beyond the scope of the paper. In particular, we pay no attention to the reasons why aggregate demand fluctuations exhibit positive serial correlation, nor to the respective role of monetary and nonmonetary demand disturbances in causing these fluctuations, nor to the significance of changes in the behavior of money demand and velocity that have occurred in the 1980s, nor to the merits of monetary rules, nor to the relative merits of monetary rules versus nominal GNP rules. These topics on the demand side can be omitted, simply because they are not at the heart of the conflict between new-Keynesian and new-classical macroeconomics.
Omission of the demand side from the scope of the paper leads us to skip over those contributions, sometimes classified as new-Keynesian, which emphasize credit rationing as a source of fluctuations in commodity demand and as a channel through which the influence of monetary policy is transmitted (see Olivier J. Blanchard and Fischer 1989, pp. 478–88). We also omit any treatment of feedbacks from changes in the parameters of aggregate price stickiness to the variance of aggregate nominal demand (see the debate between Taylor 1986; and Bradford DeLong and Lawrence Summers 1986). We take as a precedent for imposing a dichotomy between supply and demand, and for assuming nominal GNP to be exogenous, Robert Lucas' famous paper on the international output-inflation trade-off (1973), which assumed that nominal GNP was an exogenous random walk. In short, we are interested here in the price times output side of the quantity equation (MV = PQ), to the exclusion of the money times velocity side.

However, our focus here on nominal GNP rather than money helps to clarify one source of frequent misunderstanding in this area. New-Keynesian macroeconomics is not limited to the question "Why Does Money Affect Output?" If prices are sticky, then any change in nominal GNP will affect real output, no matter whether its source is a change in the nominal money supply or some autonomous movement of spending on consumption, investment, government purchases, or net exports. Further, nominal price stickiness opens the way for supply shocks, for example, a change in the relative price of oil, to create macroeconomic externalities that supplement the initial impact on output of the shock by induced demand feedbacks. The microeconomic theories surveyed in this paper apply equally to the broad question as to why demand disturbances in money and autonomous spending, as well as supply shocks, cause changes in real output.

II. The Three Dimensions of Wage and Price Stickiness

A. Price Stickiness in the Presence of Policy Feedback

A prerequisite for any theory purporting to explain wage and/or price stickiness is a demonstration that the phenomenon of stickiness exists in real-world data. In Part II we begin by defining three different dimensions of price stickiness.

---

3 I accept David Laidler's objection in correspondence that Keynesian economics is about more than wage and price stickiness and includes a treatment of "how the monetary system interferes with the coordination of inter-temporal choices." The new-Keynesian analysis of credit rationing and other failings of the monetary system is recognized as a legitimate research activity but falls outside the scope of this paper, which is delimited by the supply-demand dichotomy.

4 Nevertheless, we recognize the importance of feedback from price behavior to nominal GNP for the econometric estimation of price adjustment coefficients and devote considerable emphasis in Part II to the treatment of econometric bias that results from such feedback.

5 This is the title of the recent survey by Blanchard (1987a).
stickiness and distinguish between the essential role of price stickiness and the peripheral role of wage stickiness. The exposition is carried out insofar as possible with a set of identities, which clarify issues without imposing any theory at all.

By definition, the log of nominal GNP ($X$) must be divided between the log of the GNP deflator ($P$) and the log of real GNP ($Q$):

$$X = P + Q. \quad (1)$$

Reserving uppercase letters for logs of levels and lowercase letters for percentage changes per unit of time, we take the time derivative of (1) and obtain:

$$\dot{x} = \dot{p} + \dot{q}, \quad (2)$$

which states that any change in nominal GNP must be divided between a change in the aggregate price level and a change in real GNP. Next, we subtract from both sides of (2) the long-run equilibrium or natural growth rate of real GNP ($q^*$), and use a “hat” (â) to designate variables defined net of that trend growth rate of real output:

$$x - q^* = p + (q - q^*); \quad (3)$$

$$\dot{x} = \dot{p} + \dot{q}.$$

This states that an excess of nominal GNP growth over the long-run growth rate of real output ($\dot{x}$) must be accompanied by some combination of inflation ($\dot{p}$) and a deviation of real output from that same long-run growth rate ($\dot{q}$).

In many recessions and depressions over the course of the industrial era the economy has experienced a decline in output and employment that appears to have constrained employees to work fewer hours than they wished at the current real wage, and firms to produce less output than they wished at the current price. These episodes admit the possibility that actual output and long-run equilibrium output are two distinct concepts, implying in turn that the way is open to consider the meaning of price stickiness. For instance, if the rate of change of prices over the business cycle is always equal to some constant fraction ($\alpha$) of the excess nominal GNP movement, then business-cycle movements in real output ($\dot{q}$) must soak up the remaining fraction $(1 - \alpha)$:

$$p = \alpha \dot{x}, \quad (4)$$

$$\dot{q} = \dot{x} - p = (1 - \alpha) \dot{x}.$$

One concludes from (4) that an economy with relatively sticky prices (a small $\alpha$) must exhibit correspondingly large fluctuations in real output, as long as fluctuations in nominal demand ($\dot{x}$) are independent of the price stickiness parameter $\alpha$.

It is tempting to estimate a regression equation like either line of (4) to determine the degree of price stickiness ($\alpha$). But four crucial features of the economy—level effects, inertia effects, policy feedback, and supply shocks—are ignored in (4) and may invalidate any interpretation of an estimated value of $\alpha$ as representing a structural price-stickiness coefficient. The first problem is that (4) ignores level or Phillips-curve effects. It is possible for actual output to be growing at its long-run equilibrium growth rate (i.e., $\dot{x} = 0$) while being off its equilibrium growth path, that is, when there is a gap between the levels of actual and equilibrium output. The second problem is the possible presence of price inertia, as occurs when lagged variables (especially lagged inflation) enter into the determination of current inflation. We defer the introduction of level and inertia effects until the next section in order to concentrate on the other two basic problems with (4), which concern policy feedback and supply shocks.

The third problem is the possible presence of policy feedback from inflation to excess nominal GNP growth, as would
occur with a policy of monetary accommodation to price changes. Such feedback would be implied when the central bank attempts to peg or stabilize interest rates, or with a real bills doctrine in which bank loans automatically expand to meet the needs of trade. The fourth problem arises in the presence of autonomous supply shocks which shift the rate of price change up and down relative to that predicted by (4). We now consider a model in which the interaction between policy feedback and supply shocks becomes crucial in estimating the coefficient of price adjustment (\( \alpha \)) in an equation like (4). The subsequent results on coefficient bias apply to literally every empirical study that has attempted to relate price or output change to such endogenous variables as nominal or real GNP, the money supply, or unemployment.

Consider the two-equation model:

\[
p = \alpha \dot{x} + z
\]
\[
\dot{x} = \theta p + e,
\]

where \( z \) is the supply-shock term and \( e \) is the demand shock. The coefficient of policy feedback (\( \theta \)) would be positive if growth in the money supply responds positively to a contemporaneous change in the inflation rate.

It is easy to see that in a world with no supply shocks (\( z = 0 \)), policy accommodation makes no difference. Here we relegate the algebra to the source note in Table 1 and consider a numerical example with a 10 percent positive realization of \( e \), a price-adjustment parameter \( \alpha = 0.5 \), and a policy accommodation parameter \( \theta = 1.0 \). Then (5) is satisfied with the values \( p = 10, \dot{x} = 20, \) and \( \dot{q} = 10 \). A regression of \( p \) on \( \dot{x} \) for a sample period with no supply shocks will recover the correct value of \( \alpha \), \( 0.5 = 10/20 \). Despite policy feedback, we would correctly infer that the smaller the price adjustment coefficient, the larger the amplitude of output fluctuations in \( \dot{q} \). Intuitively, because in the absence of supply shocks price change depends only on nominal demand (\( \dot{x} \)), and any policy feedback simply “blows up” price and nominal demand change by the same proportion.

We cannot, however, recover the correct value of \( \alpha \) in the presence of supply shocks. With a supply shock \( z = 10 \) but no demand shock (\( e = 0 \)), and with the same values of \( \alpha \) and \( \theta \), (5) is satisfied for \( p = \dot{x} = 20 \) and \( \dot{q} = 0 \). If no “\( z \)” variable is included to capture the supply-shock effect, a simple regression of \( p \) on \( \dot{x} \) will recover an incorrect value of \( \alpha = 1 \). In general, as shown in the notes to Table 1, a regression of \( p \) on \( \dot{x} \) in a sample containing both demand and supply shocks will yield an upward biased estimate of the price-adjustment parameter \( \alpha \), the larger is the accommodation parameter (\( \theta \)) and the larger is the variance of supply shocks relative to demand shocks (\( \sigma^2_z/\sigma^2_e \)). The problem cannot be avoided by replacing nominal GNP change (\( \dot{x} \)) by real GNP change (\( \dot{q} \)) in the first equation in (5), because this would introduce a negative bias that works in reverse and is larger, the smaller the extent of policy accommodation.

Table 1 provides examples of the bias that will result in estimating the price-stickiness coefficient (\( \alpha \)), when excess nominal GNP growth (\( \dot{x} \)) or excess real GNP growth (\( \dot{q} \)) are used as the alternative explanatory demand growth variables. Columns 3 and 4 show that there is no bias in using \( \dot{x} \) with any degree of policy feedback or any importance of supply shocks, as long as both do not occur together. Using \( \dot{q} \) as the explanatory variable introduces a downward bias when there are supply shocks, even if there is no policy feedback. In intermediate situations, as on lines 4 and 5, estimates using alternatively \( \dot{x} \) and \( \dot{q} \) bracket the true
TABLE 1

RANGE OF ESTIMATED PRICE STICKINESS COEFFICIENTS
WHEN THE TRUE COEFFICIENT IS \( \alpha = 0.25 \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Policy Response Coefficient (( \theta ))</th>
<th>Relative Importance of Supply Shocks (( r ))</th>
<th>Estimated Value of ( \alpha ) When Regressor Is ( \hat{x} )</th>
<th>( \hat{q} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0.25</td>
<td>0.25</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>-0.50</td>
</tr>
<tr>
<td>4</td>
<td>0.5</td>
<td>0.25</td>
<td>0.35</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>1</td>
<td>0.60</td>
<td>-0.15</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.25</td>
<td>0.40</td>
<td>0.25</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0.63</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Sources by column: With reference to the model in equation (5) in the text, the estimated coefficient \( E(\alpha) \) in column (3) is the true coefficient \( \alpha \) plus the ratio of the covariance of \( \hat{x} \) with \( z \) divided by the variance of \( \hat{x} \):

\[
E(\alpha) = \alpha + [\theta(1 - \alpha)\theta]/[1 + \theta^2\theta],
\]

where \( \theta = \sigma^2z/\sigma^2e. \) When \( \hat{\theta} = \hat{x} - p \) is used as an alternative explanatory variable in the first equation of (5), instead of \( \hat{x} \), the equation estimated is \( p = \beta\hat{\theta}, \) and the estimated coefficient \( E(\beta) \) is

\[
E(\beta) = \beta - [(1 - \theta)(1 + \beta(1 - \theta))]/[1 + (1 - \theta)^2\theta].
\]

The coefficient shown in column (4) is the value of \( \alpha \) that would be calculated on the assumption that \( \beta \) is true, \( \alpha = \beta/(1 + \beta). \)

value. Using \( \hat{q} \) retrieves the correct coefficient only when policy feedback is complete, that is, when policy fully accommodates the supply shock, as in lines 6 and 7.

We reach five important conclusions from this analysis. First, to the extent that demand shocks have been substantially more important than supply shocks (at least prior to the oil-shock decade of the 1970s), the degree of price stickiness can be measured by the coefficient on excess nominal GNP change (\( \hat{x} \)) in a regression equation explaining price change, even if policy has partially or wholly accommodated price changes. Second, in view of the widespread view of Milton Friedman and Anna Schwartz (1963) and most commentators that monetary policy has been accommodative (i.e., procyclical rather than countercyclical), the price-adjustment coefficient is more likely to be upward than downward biased when the demand variable is \( \hat{x} \), thus overstating the extent of price flexibility and tilting the conclusions against the new-Keynesian view that prices are sticky and toward the new-classical view that prices are flexible. Third, in the presence of partial policy accommodation, equations with nominal and real GNP changes (\( \hat{x} \) and \( \hat{q} \)) as alternative demand variables will bracket the true coefficient of price stickiness. Fourth, any empirical study of price adjustment should attempt to find proxies for the supply shocks themselves, rather than allowing such shocks to remain hidden in the error term, in order to minimize these biases that occur in the presence of policy accommodation. Fifth, any study that does not control for supply shocks is likely to reach unreliable conclusions regarding the extent and/or sec-
ular change in price stickiness. For instance, a conclusion that prices had become more sticky since World War II could be subject to the criticism that prewar price adjustment coefficients are upward biased because of some combination of (a) greater prewar policy accommodation and (b) a higher prewar variance of unmeasured supply shocks.

B. Where the Phillips Curve and Price Inertia Fit In

One reason different authors disagree on historical changes in the extent of price stickiness is that authors have focused on different dimensions of stickiness. Thus far we have characterized price stickiness by a single parameter \( \alpha \), which denotes the marginal response of the rate of price change to a change in the excess growth rate of nominal GNP. Yet this relation between the change in prices and the change in demand stands in contrast to the relation between the change in prices and the level of demand, that is, the Phillips curve, that may come first to mind in connection with price adjustment. While the Phillips curve was originally developed (A. W. Phillips 1958) as an association between the change in nominal wage rates and the level of unemployment, it has become common to use the Phillips-curve terminology to label any relation between the rate of change of nominal prices or wages and the level of a utilization variable like the unemployment rate or detrended output. Here we focus on detrended output rather than unemployment and, because our interest is primarily in price rather than wage stickiness, we write a Phillips-curve relation for price change:

\[ p_t = \gamma \hat{Q}_t + z_t, \]  

where \( \hat{Q}_t \) is the log ratio of actual to natural output, and we indicate explicitly the time subscript that previously has been suppressed. The supply-shock term from (5) is included here in each subsequent price adjustment equation, in view of our previous conclusion that adjustment coefficients will be biased unless a careful attempt is made to control for supply shocks.

A third dimension of price stickiness is serial correlation, sometimes simply called inertia. A frequent specification of the postwar U.S. inflation process combines the Phillips curve and inertia:

\[ p_t = \lambda p_{t-1} + \gamma \hat{Q}_t + z_t. \]  

When the lagged inflation term is interpreted as a proxy for the expected rate of inflation \( (p_t^e) \), then (7) is called an expectational Phillips curve. Friedman's (1968) natural rate hypothesis (NRH) states that the coefficient on \( p_t^e \) in an expectational Phillips curve is unity,

\[ p_t = p_t^e + \gamma \hat{Q}_t + z_t. \]  

This expression is compatible with steady fully anticipated inflation when actual and natural output are equal \( (\hat{Q}_t = 0) \) and implies that inflation steadily accelerates whenever the log output ratio is positive.

But, as originally pointed out by Thomas Sargent (1971), the NRH does not imply that the coefficient \( \lambda \) in (7) must be unity. The coefficient on \( p_t^e \) in (8) could be unity, while at the same time rational agents could form their expectations of inflation by applying a coefficient \( \lambda \) below...
unity to lagged inflation, if this provided the best possible predictor. For instance, if inflation were a random walk the optimal predictor would be $\lambda = 1$, but if inflation were white noise, the optimal predictor would be $\lambda = 0$. By expressing the Phillips curve in form (7) rather than (8), we recognize that the coefficient $\lambda$ may vary in different times and places, depending on the nature of the inflation process. Further, (7) recognizes, as (8) does not, that there may be many reasons for serial dependence in the inflation rate, of which expectation formation is only one, and overlapping wage and price contracts may be among the others.

Blanchard (1987b) has stressed that there are two dimensions of price adjustment, corresponding to the two parameters $\lambda$ and $\gamma$ in (7). An equation like (7) implies that shocks to nominal aggregate demand cause the economy to travel through loops on a diagram plotting inflation ($p_t$) against the output ratio ($\dot{Q}_t$), and an economy with low values of $\lambda$ and $\gamma$ has “fat loops”; that is, it exhibits relatively large output fluctuations and only a slow incorporation of the change in nominal demand growth into the rate of inflation.

However, in addition to the two adjustment parameters in (7), we have already introduced a third parameter ($\alpha$) in (4) and (5), which measures the fraction of current excess nominal GNP change ($\dot{x}_t$) taking the form of price change. How are these parameters related? The connection when we add the explanatory variable contained in (4) to those already present in (7):  

$$ p_t = \lambda p_{t-1} + \alpha \dot{x}_t + \gamma \dot{Q}_t + z_t. \quad (9) $$

While the $\alpha \dot{x}_t$ term may appear to drop from the sky, in fact equation (9) can be interpreted simply as loosening the artificial restriction in (7) that allows only the current value of the log output ratio to enter. The more general form (9) allows both the current and one lagged value of the output ratio to enter as explanatory variables, as becomes transparent when we use the identity that $\dot{Q}_t = \dot{Q}_{t-1} + \dot{x}_t - p_t$ to rewrite (9) in either of two equivalent forms:

$$ p_t = \frac{1}{1-\alpha} [\lambda p_{t-1} + (\alpha + \gamma) \dot{Q}_t - \alpha \dot{Q}_{t-1} + z_t], \quad \text{or,} \quad (10a) $$

$$ p_t = \frac{1}{1-\alpha} [\lambda p_{t-1} + \alpha \dot{Q}_t + \gamma \dot{Q}_t + z_t]. \quad (10b) $$

Note that either (10a) or (10b) reduces to (7) when the $\alpha$ parameter is set equal to zero. If both the current and one lagged output term matter for the rate of price change, as in (10a), this implies in (10b) that the rate of change of prices is related to both the rate of change ($\dot{q}_t$) and the level ($\dot{Q}_t$) of output. The generalization of the Phillips-curve hypothesis contained in (9) and (10) illustrates that the same hypothesis of price adjustment can be expressed in several alternative forms, and that the extent of price change in response to a change in nominal demand depends not on a single parameter, but on the three parameters $\lambda$, $\alpha$, and $\gamma$.

9 Blanchard presents an equation like (7) in which the rate of wage change also appears, because he is interested in the speed of transmission of cost changes into price changes. But the same point applies to (7), where we are interested in the division of nominal demand changes between price changes and output changes.

10 The identity in the text, $\dot{Q}_t = \dot{Q}_{t-1} + \dot{x}_t - p_t$, is identical to the identity written as equation (3) above, in view of the fact that $\dot{q}_t$ (the rate of change of detrended output) is the same as $\dot{Q}_t - \dot{Q}_{t-1}$ (the change in the log ratio of actual to trend output).

11 Early precursors of (10a) and (10b), developed and originally published in 1972-73, are reprinted in David Laidler (1975, pp. 127, 140) and differ only in assuming that $\theta = 1$ and that $z_t = 0$.

12 The inclusion of both level and rate-of-change effects dates back to Richard Lipsey (1960), who aggregated a model with heterogeneous micro labor markets characterized by limited labor mobility between markets and showed that the rate of change of wages would depend on both the level and rate of change of the aggregate unemployment rate. In Lipsey’s model the economy exhibits counterclock-
C. Where Wages Fit In

Keynesian economics has traditionally been more concerned with wage rigidity than price rigidity. Yet our discussion to this point has made no mention of wages. This is fitting, because only price stickiness, not wage stickiness, is a necessary condition for business cycles in real output, given a particular path of nominal aggregate demand. There are no arithmetically necessary implications of nominal wage rigidity for the cyclical behavior of output or employment, because sufficient flexibility in profits could allow prices to be flexible (so that \( p \approx \alpha x \)), even if the nominal wage rate were absolutely fixed. Yet a world of highly flexible profits with completely rigid wages would have economic, if not arithmetical, implications. High profit volatility for any given firm would shift the firm’s securities out along the mean-variance schedule and raise the average cost of capital, thus creating pressure in two directions, toward an increase in the flexibility of wages and toward a decrease in the flexibility of prices, both of which would reduce the volatility of profits. In new-Keynesian economics there is no primacy to wage rigidity as contrasted with price rigidity, and thus no presumption that wages are less cyclically sensitive than prices. In fact, much of the research of the past half decade has been directed toward the microfoundations of price rigidity.

The nature of cyclical flexibility in real wages has always played a role in discussions of Keynesian economics, dating back to the debate involving John Dunlop (1938), Lorie Tarshis (1938), and John Maynard Keynes (1939). Even though these authors are known for the criticism of the countercyclical real wage assumption implicit in the General Theory, resulting from its assumption of price flexibility combined with nominal wage rigidity, it is less well known that Tarshis in 1939 soon recanted and provided evidence of a relatively strong negative correlation between average hourly earnings and total hours worked. Subsequently we shall examine new evidence on the cyclicality of real wages.

D. Rate of Change or Hysteresis Effects

Equations (9) and (10) imply that there may be three quite different types of price stickiness, indicated respectively by a relatively high value of the \( \lambda \) parameter, and by relatively low values of the \( \alpha \) and \( \gamma \) adjustment parameters. The role of the inertia parameter \( \lambda \) is straightforward, with a higher value of \( \lambda \) prolonging the duration of adjustment to changes in nominal demand, for any given values of the \( \alpha \) and \( \gamma \) parameters, and increasing the importance of overshooting and dynamic adjustment loops. The distinction between rate-of-change adjustment (\( \alpha \)) and level or Phillips-curve adjustment (\( \gamma \)) is clarified by examining extreme cases in which one or the other is absent. When there is no rate-of-change effect (\( \alpha = 0 \)) we are back in the simple Phillips-curve framework in which only the level of output matters. For any given values of the \( \lambda \) and \( \gamma \) parameters, the acceleration of inflation implied by an output ratio of +5 percent is the same, regardless of whether the output ratio is rapidly rising or rapidly falling.

The opposite extreme is of more interest, because it has been the focus of so much attention in the context of high Eu-

---

\(^{13}\) I am grateful to Robert Chirinko for providing me with a copy of the Tarshis (1939) note.
European unemployment in the 1980s. An economy lacking a level effect ($\gamma = 0$) is said to be characterized by hysteresis. Considerable theoretical work has recently emerged to explain hysteresis phenomena, particularly in the context of the insider-outsider model of employment reviewed below in Part VII.D. Whatever the theoretical explanation, the presence of hysteresis would have profound implications for both economic doctrine and policy.\(^{14}\) Friedman’s NRH posits a self-correction or level effect that automatically stabilizes output at its equilibrium value in the presence of steady nominal demand growth. With no level effect, the economy could settle down at any arbitrary distance from its equilibrium output path (with $q_t = 0$) and experience a constant rate of inflation, with no tendency for self-correction. And, if the NRH were abandoned, it would cast stabilization policy adrift from its previous mooring, the task of steering the economy toward a fixed natural rate ($Q_t = 0$), and open to the central implication of hysteresis that any level of detrended output or rate of unemployment, no matter how low or high, would be consistent with steady inflation (at a rate that depends on the history of both inflation and unemployment).

As we see below, the pattern of price adjustment described by hysteresis is not a novel phenomenon isolated to Europe in the 1980s, for the Phillips curve or “level effect” also vanished in the United States, the United Kingdom, and Germany during the interwar period.\(^{15}\) A key implication of (9) is that with hysteresis ($\gamma = 0$) and with $\lambda = 1 - \alpha$, the acceleration or deceleration of inflation, as well as the change in detrended output, depends only on the difference between $\hat{x}_t$ and $p_{t-1}$, that is, whether or not excess nominal GNP growth ratifies the inherited inflation rate:

\[
p_t - p_{t-1} = \alpha(\hat{x}_t - p_{t-1}) + z_t, \quad \text{and (11)}
\]

\[
Q_t - Q_{t-1} = (1 - \alpha)(\hat{x}_t - p_{t-1}) - z_t. \quad \text{(12)}
\]

In short, hysteresis implies that changes in both inflation and output are completely independent of the level of detrended output, and that an economy in the depths of a great depression can experience an acceleration of inflation, no matter how high the level of unemployment or low the level of detrended output, if excess nominal GNP growth exceeds last period’s inflation rate.

Empirical estimates of the general price-adjustment model in (9) and (10) can reveal the size of the three adjustment parameters ($\lambda$, $\alpha$, and $\gamma$) in different countries and historical eras. There remains the issue of which alternative specification in (9) and (10) is preferable for estimation. As argued earlier, in the presence of policy feedback and unmeasured supply shocks, the $\alpha$ adjustment parameter is likely to be overstated when nominal GNP change ($\hat{x}$) is included as in (9) and understated when real GNP change ($\hat{q}$) is used instead as in (10b). This suggests that estimates based alternatively on both forms are preferable, because they will “bracket” the true parameter.

We conclude from this discussion that three parameters are required to measure the degree of price stickiness: $\lambda$ measuring the extent of inertia, $\alpha$ measuring the rate-of-change or hysteresis effect, and $\gamma$ measuring the level or Phillips-curve effect. Any attempt to measure changes in the degree of stickiness over time, or differences among countries,
may be flawed if it omits any of these three parameters from empirical testing. Further, we have seen that policy accommodation of supply shocks can bias coefficients of price adjustment, and thus any adequate empirical investigation must make a careful attempt to control for supply shocks as well.

Part II began by stressing the most obvious implication of the identity linking changes in nominal demand, real output, and the price level. Changes in the price level must exactly mimic changes in nominal demand if business cycles in real output are to be avoided. Thus the requirements for perfect price flexibility are highly restrictive: In the context of equation (10) price changes can mimic changes in nominal demand only if $\alpha = 1$, $\lambda = 0$, and $\gamma = 0$. Thus any combination of a rate-of-change coefficient below unity, the presence of Phillips-curve level effects ($\gamma > 0$), or the presence of inertia effects is sufficient to generate business cycles. However, Lucas (1973) showed that Phillips-curve level effects could be derived in a business-cycle model in which markets clear; thus the absence of perfect price flexibility is not sufficient to distinguish between new-classical market-clearing models and new-Keynesian sticky-price models. Instead, the presence of price inertia ($\theta > 0$) is crucial for rejecting the new-classical interpretation and demonstrating the existence of price stickiness.\(^\text{16}\)

III. The Variety of Historical Experience

A. Diversity of Response Across Industries

Since well before the publication of Keynes' General Theory, for example, Mills (1927), industrial economists have been aware that the responsiveness of prices to changes in demand differs sharply across industries. The contrast between the flexibility of the prices of agricultural products, and the inflexibility of the prices of complex manufactured goods, was the point of departure of Gardiner Means' (1935) administered price hypothesis. In the Great Depression every farmer knew what Table 2 shows:

<table>
<thead>
<tr>
<th>Industry</th>
<th>Percentage Decline in Price</th>
<th>Percentage Decline in Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural implements</td>
<td>6</td>
<td>80</td>
</tr>
<tr>
<td>Motor vehicles</td>
<td>16</td>
<td>80</td>
</tr>
<tr>
<td>Textile products</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>Petroleum</td>
<td>56</td>
<td>20</td>
</tr>
<tr>
<td>Agricultural products</td>
<td>63</td>
<td>6</td>
</tr>
</tbody>
</table>

Source: Gardiner Means (1935, p. 8).

In an economic downturn the farmer was the victim of a highly unfavorable twist in relative prices, because the prices of agricultural products fell much more than those of many manufactured goods, especially the agricultural implements listed on the first line that represent one of the main purchased inputs in the farm sector. Within the spectrum of manufactured goods, crude products like textiles tended to exhibit more price flexibility than more finished products like tractors and automobiles.

Unfortunately, there are few empirical studies that document these differences systematically. George Stigler and James Kindahl (1970) collected prices from buyers for a large number of products, and these data were analyzed by Carlton (1986) to determine if there were any

\(^{16}\) R. Gordon (1982b) shows that the Lucas (1973) model can be nested in a general model of price adjustment like (9) and can be rejected in the presence of price inertia.
structural relations between seller and buyer characteristics and the degree of price rigidity. Carlton concludes that there is a significant degree of price rigidity: "It is not unusual in some industries for prices to individual buyers to remain unchanged for several years" (1986, p. 638). Unfortunately, however, neither Stigler and Kindahl nor Carlton show that, as Means suggested, the degree of "complexity" of a product is related to price rigidity. Although Carlton did try to measure complexity as well as other structural variables, he was able to find a significant positive correlation only between the concentration ratio for a product and the duration of its price rigidity (i.e., the number of months a price remains unchanged).^{17}

But it is important to stress another of Carlton's findings that may be of substantial importance in assessing the theories reviewed below. By no means are all prices rigid or do they remain unchanged for substantial periods of time: "The fixed costs of changing price at least to some buyers may be small. There are plenty of instances where small price changes occur" (Carlton 1986, p. 638). Specifically, "there are a significant number of price changes that one would consider small (i.e., less than 1 percent) for most commodities and transaction types." Industries where frequent price changes are common include plywood and nonferrous metals, and commodities with relatively long spells of rigid prices include steel, paper, chemicals, cement, and glass. Carlton's evidence that spells of price rigidity can be both short and long calls into question the generality of the oft-cited study by Stephen Cecchetti (1986) which provides evidence that newsstand prices of magazines can remain unchanged for years (see also Kashyap 1990). Carlton's finding that spells are sometimes short and price changes sometimes small would appear to call into question the theories of new-Keynesian economists based on "menu costs" of price changes, reviewed below in Part V.D. However, this apparent implication is subject to the caveat that if demand and/or supply shift permanently, then small price adjustments can produce large benefits and will be observed even if fixed costs are large (Carlton 1989a, p. 932).

There has been remarkably little interaction between new-Keynesian theory and the evidence provided in the emerging literature of the new empirical industrial organization (NEIO) recently surveyed by Timothy Bresnahan (1989). The overall conclusion is that there is "a great deal of market power, in the sense of price-cost margins, in some concentrated industries" (Bresnahan 1989, p. 1052). One could emphasize the words a great deal as supporting the emphasis by new-Keynesian theorists on models of monopolistic rather than perfect competition. Or one could emphasize the word "some" to point out that the world is made up of both monopolistic and competitive industries. But the matter is even more complex: One important theme of recent NEIO work is that pricing behavior can alternate between collusive monopolistic behavior and price wars in which a cartel temporarily collapses, implying that a given industry is characterized neither by exclusively monopolistic nor competitive behavior.

The theme of heterogeneity extends along other dimensions. Product differentiation is so pervasive that "there is almost no industry for which the position that there are more than 100 products is untenable: without putting more structure on the problem, the analyst could

---

^{17} I am grateful to Dennis Carlton for suggesting the wording of the last two sentences.
need to estimate literally thousands of elasticities” (Bresnahan 1989, p. 1045). Heterogeneity extends to pricing behavior across firms in a single industry. For instance, a study of airline competition found not only that concentration affects price in airline city-pair markets, but also that the identity of the competitors matters. Carlton stresses that a given seller can charge different prices to different customers and change them at different times, based on “a seller’s knowledge of his customers and on the optimality of non-price rationing” (Carlton 1989a). In a cross-section of industries, numerous dimensions of structure appear to vary together, including mass production, large-scale facilities, unionization, capital intensity, concentration, and cyclical price rigidity, all of which are more pronounced in the cyclically sensitive sectors of the economy, particularly durable goods.\(^\text{19}\)

As we shall see below, new-Keynesian theory has contributed relatively little to understanding these differences across industries, and as yet there has been virtually no research that attempts to test theories on a diversity of industrial data. We emphasize the numerous aspects of heterogeneity across and within industries to support several themes that emerge below, including the importance of idiosyncratic elements of product cost and demand that prevent firms from assuming, as in so many simple models, that their costs and product demand will mimic the behavior of nominal aggregate demand. Even so basic a distinction as Arthur Okun’s (1975, 1981) dichotomy between auction and customer markets rarely surfaces in new-Keynesian writing, much less in new-classical contributions. And the seminal work in understanding the coexistence of auction and customer markets has been contributed by microeconomists, especially Carlton (e.g., 1989b), who stresses that, because of the high costs of establishing auction markets, “there is no incentive for the efficient creation of markets” (p. 7).\(^\text{20}\)

B. Diversity of Response Across Time and Space

Just as challenging for theorists as the diversity of responses across industries at a particular time in a particular country is the diversity of responses across time and countries. Much of the empirical work in this area has been within the context of a debate over whether prices, wages, or both have become less flexible in the postwar U.S. as contrasted with various periods before the Great Depression (among these studies are Steven Allen 1989; Daniel Mitchell 1985; Jeffrey Sachs 1980; Schultze 1981, 1986; John Taylor 1986; R. Gordon 1980, 1982b). In related work Charles Schultze and others have examined differences in response coefficients over both time and space for the U.S. and several other major industrialized nations (see George Alogos-

---

\(^{18}\) Further evidence on the extent of product differentiation comes in detailed studies of international trade, showing the countries at the same stage of development both import and export goods within the same industrial categories. See Magnus Blomstrom, Robert Lipsey, and Lennart Ohlson (1989).

\(^{19}\) This point was suggested in a letter from Bresnahan, who describes these common features of cyclically sensitive industries as “some famous coincidences about industry structure.”

\(^{20}\) Particularly striking is Carlton’s example of the costs of running the futures markets in Chicago, consisting of large office buildings, expensive real estate, elaborate record keeping, and the large time cost of the many people involved. “A significant fraction of the economy of the city of Chicago is devoted to the making of markets. If a magic spell could be cast to make transactions costless, the Chicago economy would be devastated, at least in the short run. This emphasizes how far from costless the making of markets really is” (1989b, p. 6).
It is beyond the scope of this paper to track all the differences in data and specification that contribute to the variety of conclusions that these studies have reached; that would require a separate survey on this issue alone. Some of the disagreements, particularly about changes in cyclical behavior for the U.S., arise because authors often do not recognize that there are three dimensions to price and wage rigidity, as demonstrated in Part II. These are the degree of inertia or serial correlation (λ), the rate-of-change or hysteresis coefficient (α), and the level or Phillips-curve coefficient (γ). Here we provide a link between that classification scheme and historical data by presenting estimates of the three parameters based on price-adjustment equations (9) and (10) developed above. We address two issues, differences in the responsiveness of prices and wages over U.S. history, and differences in the responsiveness of prices over the period since 1870 for five major industrial nations (U.S., U.K., France, Germany, and Japan).

The empirical equations summarized in this paper are estimated only for nominal and real output data corresponding to the $\hat{x}_t$ and $\hat{Q}_t$ variables in the theoretical price-adjustment equation (9). There is no attempt to estimate alternative versions for other possible nominal and real demand variables, for example, the money supply or unemployment. Annual output data extend back much further than unemployment data—to 1855 for the U.K., 1870 for France, Germany, and the U.S., and 1885 for Japan. Wage-adjustment equations are illustrated only for the U.S., pending a careful study to determine whether wage data for other countries are consistent over time.\footnote{As we have been reminded by Steven Allen (1989), the standard prewar series on U.S. wages are for production workers in manufacturing and must be linked with a postwar series on manufacturing wages, not the wage index for the nonfarm private economy that is most often used in studies limited to the postwar period. Allen concludes after an exhaustive study that differences in measurement methods in either wage or output series do not change his conclusion that the cyclical sensitivity of wages was the same in the prewar and postwar periods.}

Numerous decisions must be made in the development of tests covering such a long span of history for these nations. These include the method of detrending and the development of proxy variables for the major supply shocks, a critical issue in view of the likely bias in coefficient estimates when supply shocks are left unmeasured. Another issue is the estimation of parameter shifts over subintervals of a long historical sample period. Details on the methodology and the regression estimates are provided in Appendix A, which shows that it is desirable to conduct the estimation with slightly transformed versions of (9) and (10b). This allows us to proceed directly to Tables 3 and 4, where the underlying parameters are unscrambled from the transformed equations and presented for different countries and historical eras.

The estimated parameters are provided for changes in prices, nominal wages, and real wages for the U.S. in Table 3. We are interested in the nature of changes in the three price- and wage-adjustment parameters over time, and also evidence on the hotly debated issue of the cyclical sensitivity of real wages. Following our analysis in Part II, two estimates of each parameter are provided. The left-hand element in each column is based on an adjustment equation in which excess nominal GNP growth is included, and this is likely to yield an up-
TABLE 3
ESTIMATED PRICE AND WAGE ADJUSTMENT PARAMETERS FOR THE U.S., 1873–1987

<table>
<thead>
<tr>
<th></th>
<th>Inertia Effect (λ)</th>
<th>Rate-of-Change Effect (α)</th>
<th>Level Effect (γ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td><strong>Price change</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1873–1914, 1923–29</td>
<td>0.29 (0.37)</td>
<td>0.17 (0.03)</td>
<td>0.28 (0.28)</td>
</tr>
<tr>
<td>1915–22</td>
<td>0.29 (0.43)</td>
<td>0.69 (0.43)</td>
<td>0.28 (0.17)</td>
</tr>
<tr>
<td>1930–53</td>
<td>0.25 (0.37)</td>
<td>0.29 (0.03)</td>
<td>0.09 (0.08)</td>
</tr>
<tr>
<td>1954–87</td>
<td>0.87 (1.01)</td>
<td>0.17 (0.03)</td>
<td>0.28 (0.28)</td>
</tr>
<tr>
<td><strong>Nominal wage change</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1873–1914, 1923–29</td>
<td>0.46 (0.51)</td>
<td>0.20 (0.07)</td>
<td>0.43 (0.40)</td>
</tr>
<tr>
<td>1915–22</td>
<td>0.46 (0.44)</td>
<td>0.73 (0.49)</td>
<td>0.43 (0.22)</td>
</tr>
<tr>
<td>1930–53</td>
<td>0.32 (0.37)</td>
<td>0.44 (0.33)</td>
<td>−0.01 (0.02)</td>
</tr>
<tr>
<td>1954–87</td>
<td>0.67 (0.86)</td>
<td>0.20 (0.07)</td>
<td>0.43 (0.40)</td>
</tr>
<tr>
<td><strong>Real wage change</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1873–1914, 1923–29</td>
<td>0.17 (0.14)</td>
<td>0.03 (0.04)</td>
<td>0.15 (0.12)</td>
</tr>
<tr>
<td>1915–22</td>
<td>0.17 (0.01)</td>
<td>0.04 (0.06)</td>
<td>0.15 (0.05)</td>
</tr>
<tr>
<td>1930–53</td>
<td>0.05 (0.00)</td>
<td>0.15 (0.30)</td>
<td>−0.10 (−0.06)</td>
</tr>
<tr>
<td>1954–87</td>
<td>−0.20 (−0.15)</td>
<td>0.03 (0.04)</td>
<td>0.15 (0.12)</td>
</tr>
</tbody>
</table>

Notes: Equation specifications and details are provided in Appendix A. The left parameter in each column comes from unscrambling the coefficients of equation (9') in Appendix A, the version containing excess nominal GNP growth (it) as the rate-of-change variable; the right parameter in parentheses () comes from unscrambling the coefficients of equation (9''), the version containing excess real GNP growth (4t) as the rate-of-change variable.

The single most striking finding in Table 3 is that neither prices nor wages were more sticky in 1954–87 than 1873–1914, as measured by the rate-of-change (α) and level (γ) coefficients. The sole change between pre–World War I and post–World War II was an increase in the inertia (λ) coefficient, and this increase was much greater for prices than wages. Between 1915 and 1953, however, there were substantial changes. The α parameter rose substantially during World War I, while the γ parameter virtually disappeared during 1930–53. When the estimated price-change parameters are subtracted from the wage-

22 This finding is consistent with that of Allen's careful (1989) study, which examines only wage behavior, not price behavior. Allen's specification is similar to mine and uses both unemployment and output gap data, but no nominal GNP data or supply shock proxies, and is thus subject to a bias in the unemployment or output coefficients toward zero. Allen's conclusion claims that his study finds similar behavior prewar and postwar, but his text reveals that he finds the same increase in the inertia effect (coefficients on lagged inflation) as is shown in Table 3.
TABLE 4
ESTIMATED PRICE ADJUSTMENT PARAMETERS FOR FIVE COUNTRIES, 1873-1986

<table>
<thead>
<tr>
<th></th>
<th>Inertia Effect (λ)</th>
<th>Rate-of-Change Effect (α)</th>
<th>Level Effect (γ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td><strong>U.S.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1873–1914, 1923–29</td>
<td>0.29 (0.37)</td>
<td>0.17 (0.03)</td>
<td>0.28 (0.28)</td>
</tr>
<tr>
<td>1915–22</td>
<td>0.29 (0.43)</td>
<td>0.69 (0.43)</td>
<td>0.28 (0.17)</td>
</tr>
<tr>
<td>1930–53</td>
<td>0.25 (0.37)</td>
<td>0.29 (0.03)</td>
<td>0.09 (0.08)</td>
</tr>
<tr>
<td>1954–87</td>
<td>0.87 (1.01)</td>
<td>0.17 (0.03)</td>
<td>0.28 (0.28)</td>
</tr>
<tr>
<td><strong>U.K.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1858–1914</td>
<td>0.24 (0.23)</td>
<td>0.43 (−0.06)</td>
<td>0.35 (0.38)</td>
</tr>
<tr>
<td>1915–22</td>
<td>0.24 (0.15)</td>
<td>0.60 (0.33)</td>
<td>0.35 (0.56)</td>
</tr>
<tr>
<td>1923–38</td>
<td>0.20 (0.18)</td>
<td>0.26 (0.17)</td>
<td>0.09 (0.07)</td>
</tr>
<tr>
<td>1960–86</td>
<td>0.57 (1.00)</td>
<td>0.43 (−0.06)</td>
<td>0.35 (0.38)</td>
</tr>
<tr>
<td><strong>France</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1873–1913</td>
<td>−0.20 (−0.38)</td>
<td>0.47 (0.10)</td>
<td>0.26 (0.03)</td>
</tr>
<tr>
<td>1925–38</td>
<td>0.15 (0.09)</td>
<td>0.47 (0.26)</td>
<td>0.26 (0.65)</td>
</tr>
<tr>
<td>1960–86</td>
<td>0.55 (0.40)</td>
<td>0.47 (0.10)</td>
<td>0.26 (0.03)</td>
</tr>
<tr>
<td><strong>Germany</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1873–1913</td>
<td>0.00 (0.30)</td>
<td>0.66 (−0.11)</td>
<td>0.21 (0.12)</td>
</tr>
<tr>
<td>1925–38</td>
<td>0.00 (0.00)</td>
<td>0.40 (0.38)</td>
<td>0.07 (0.07)</td>
</tr>
<tr>
<td>1960–86</td>
<td>0.73 (1.08)</td>
<td>0.33 (−0.11)</td>
<td>0.21 (0.12)</td>
</tr>
<tr>
<td><strong>Japan</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1888–1914, 1923–38</td>
<td>0.15 (0.70)</td>
<td>0.64 (0.02)</td>
<td>0.39 (0.34)</td>
</tr>
<tr>
<td>1915–22</td>
<td>0.15 (0.18)</td>
<td>0.87 (0.75)</td>
<td>0.39 (0.09)</td>
</tr>
<tr>
<td>1960–86</td>
<td>0.15 (0.50)</td>
<td>0.64 (0.29)</td>
<td>0.39 (0.24)</td>
</tr>
</tbody>
</table>

Notes: Sample period for the U.K. begins in 1958, for Japan begins in 1888, and for the U.S. ends in 1987. Equation specifications and details are provided in Appendix A. The left parameter in each column comes from unscrambling the coefficients of equation (9') in the appendix, the version containing excess nominal GNP growth (it) as the rate-of-change variable; the right parameter in parentheses () comes from unscrambling the coefficients of equation (9''), the version containing excess real GNP growth (it) as the rate-of-change variable.

change parameters, the results before 1930 and after 1953 suggest that real wages have a negligible rate-of-change effect but a substantial procyclical level effect. That is, a persistent economic boom causes steady upward pressure on the real wage, and a persistent recession does the reverse. However, this finding is subject to the qualification that the manufacturing wage data used here exaggerate the cyclical sensitivity of economy-wide rates. When the equations are reestimated for the postwar 1954–87 period alone with the fixed-weight nonfarm wage index replacing the manufacturing wage index, the cyclical sensitivity of real wages drops to zero.23

Table 4 compares the results for U.S. prices with similar price equations for the other countries. Again the most striking finding is that the α and γ parameters were the same before World War I and after World War II in the U.K., France, and Japan, with a decline in the α coefficient only in Germany (and one may question the linking of German data over 23The respective parameter estimates for 1954–87 in an equation for the change in the real wage are, with manufacturing wage data, \( \lambda = -0.33, \alpha = 0.12, \gamma = -0.12 \). With the fixed-weight nonfarm wage index (spliced to the employment cost index in 1975), the parameters are \( \lambda = 0.01, \alpha = -0.07, \gamma = -0.02 \).
periods when its borders were so different). In every country but Japan the inertia effect was much higher after World War II than before World War I. The U.K. and Japan duplicate the jump in the $\alpha$ coefficient already observed for the U.S. during the 1915–22 interval, and both the U.K. and Germany exhibit a substantial decline in the $\gamma$ coefficient during the interwar period.

**C. Implications**

Is the aggregate price level highly flexible, mimicking changes in excess nominal GNP growth? Or does the aggregate price level live a life of its own, bearing little relation to excess nominal GNP growth and thus allowing those nominal changes to create business cycles in real output? The conclusion from Tables 3 and 4 is that both these statements are true. And many in-between responses have been observed as well.

At one extreme is the very high rate-of-change coefficient for Japan throughout, and for the U.S. and the U.K. during World War I and its aftermath. Figure 1 plots the 1886–1914 data for Japan and shows how closely price changes track excess nominal GNP changes. The figure also exhibits cycles in the log output ratio that are small relative to the large amplitude of nominal GNP changes. We have argued above that the best estimates of the adjustment parameters are given by the average of the two estimates shown in each column in Tables 3 and 4. On the basis of these averages, it is quite apparent in Table 4 that the U.S. has the smallest rate-of-change parameter ($\alpha$) and Japan the largest, both before World War I and after World War II, with the other countries arrayed in between. The postwar U.S. also contrasts starkly with Japan in its strong inertia effect. The top frame of Figure 2 shows how loose is the relation between inflation and excess nominal GNP growth in the postwar U.S., and how large is the amplitude of output cycles relative to nominal GNP growth cycles. Again basing conclusions on the average of the two figures in each column, inertia effects in all countries were negligible before World War II.

These results demonstrate the strong diversity of aggregate price-adjustment behavior that has occurred across time and across countries. The variety of historical responses of price changes to nominal demand changes raises questions that new-Keynesian theorists have barely begun to address. Perhaps the most widely noted empirical test thus far devised by new-Keynesian economists (Laurence Ball, N. Gregory Mankiw and David Romer 1988) takes as its point of departure Lucas' (1973) demonstration that the Phillips curve becomes steeper with a higher variance of the growth rate.

Excess Nominal GNP

GNP Deflator


Figure 2. Inflation, Adjusted Nominal GNP Growth, and the Output Ratio, United States, 1950-89

of nominal demand. Ball, Mankiw, and Romer show that menu-cost theory supports the Lucas correlation but also makes the additional prediction that increases in the mean growth rate of nominal demand should steepen the Phillips curve, because with staggered price setting an increase in the mean inflation rate increases the frequency of price changes. Thus far, their empirical work in support of this theoretical prediction has been subject to substantial criticism. In relation to our empirical results of Tables 3 and 4, either the Lucas or the Ball, Mankiw, and Romer approach can help to explain why prices became more flexible during World War I but contribute little or nothing to an understanding of the other main findings: the similar level and rate-of-change effects before World War I and after World War II, despite the higher variance of nominal demand in the earlier period; the disappearance of the level effect in the Depression years; the emergence of inertia after World War II; and the differences in price flexibility among the five countries.

D. Empirical Research and the Revival of Keynesian Economics

Theories are often judged on their ability to explain time-series data on aggregate variables. This is clearly evident in the interaction of events and ideas in the past two decades. Theories have risen and fallen in acceptance in accord with the correspondence of their predictions with the evolution of actual events in the macroeconomy. To gain perspective on the development of new-Keynesian economics, we need to understand what went wrong with the old-Keynesian economics. Our emphasis here is the empirical failure of the Keynesian paradigm of the 1960s, and the elements that contributed to the empirical revival of the Keynesian approach in the 1980s. We concentrate on empirical aspects of the contest between new-classical and new-Keynesian economics, and we limit the scope of the paper by omitting any theoretical critique of either the Lucas imperfect-information (Mark I) approach or the real-business-cycle (Mark II) variant of new-classical macroeconomics.

See the numerous criticisms of the paper by Ball, Mankiw, and Romer (1988) contained in the discussant comments by George Akerlof, Andrew Rose, and Janet Yellen, as well as by Christopher Sims.

The Lucas (1972, 1973) imperfect information approach (Mark I) is now widely viewed as unconvincing, because it is undermined by the availability of information on the aggregate price level and money supply over a much shorter time period than the duration of the average business cycle. Major contributions to real-business-cycle theory (Mark II) include Finn Kydland and Edward Prescott (1982) and Prescott (1986). A generally supportive survey is provided by Charles Plosser (1989), and critical surveys include Gregory Mankiw (1989) and Bennett McCalum (1989).
In the 1960s Keynesian economists incorporated into their theoretical and econometric models an exploitable negative long-run Phillips-curve trade-off between inflation and unemployment. The acceleration of inflation after 1965, together with the positive correlation between inflation and unemployment observed during much of the 1970s, caused the mid-1960s Keynesian orthodoxy to unravel. In flowery language that amounted to a simultaneous declaration of war and announcement of victory, Lucas and Thomas Sargent (1978, pp. 49–50) described “the task which faces contemporary students of the business cycle [as] that of sorting through the wreckage . . . of that remarkable intellectual event called the Keynesian Revolution. . . .” It is not widely recognized that the empirical reconstruction of Keynesian economics occurred prior to the wave of theoretical work that is now most commonly associated with the term new-Keynesian economics. Lucas and Sargent were only partly right. Yes, the predictions of the late 1960s were incorrect, but incorrect forecasts do not provide de facto proof that a doctrine’s theoretical underpinnings are fundamentally flawed. The essential element of Keynesian doctrine is non-market-clearing, which in turn requires the gradual adjustment of prices. The 1960s version of the Phillips relation combined three elements, (1) gradual price adjustment, (2) a long-run trade-off, and (3) a closed-economy, demand-only approach with no role for import prices or supply shocks. Yet only (1) is necessary to maintain the essence of the Keynesian paradigm, non-market-clearing. The other two elements, (2) and (3), were ephemeral empirical results, based mainly on the 15 or 20 years of U.S. postwar data, that revealed more of the short time horizon and closed-economy mentality of the first generation of econometric model builders than any fundamental weakness of the non-market-clearing approach.

The long-run trade-off result was abandoned within five years of Friedman’s presidential address.26 This allowed the gradual-adjustment property of the 1960s-style wage and price equations to be combined with the long-run neutrality property advocated by Friedman. The effects of supply shocks, including the relative prices of oil and imports, were absorbed into the U.S. Phillips-curve framework in my work of the mid-1970s, which was developed alongside the work by David Laidler, Michael Parkin, and their collaborators in the open-economy setting of the U.K.27 The result was an econometric analogy to the dynamic aggregate demand and supply model that was introduced with the 1978–79 publication of a new generation of economic principles and intermediate macro textbooks.28 Now a single reduced-form econometric equation for price change, like those summarized in Tables 3 and 4 above, could incorporate the effects of gradual adjustment, of demand shocks that created a temporary positive correlation between inflation and output, and of supply shocks that created a temporary negative correlation. By the end of the 1970s the supply side of the economy had been opened up to outside influences, and the list of relevant supply shocks for the U.S. had grown to include not only price controls and oil shocks, 26 Simultaneous work by me (R. Gordon 1972) and by Otto Eckstein and Roger Brinner (1972) showed how postwar wage and price data could be made consistent with long-run neutrality.


28 As author Alan Blinder described the aggregate demand and supply model as developed in his own textbook, “now the Marshallian scissors come in a giant economy size.”
but also changes in non-oil import prices, exchange rates, tax rates, and the minimum wage. 29

This so-called gradual-adjustment price-change equation is completely non-structural and as such is in principle highly vulnerable to the Lucas (1976) critique. We have seen in Tables 3 and 4 that coefficients of price adjustment are subject to substantial change when there are major changes in the economic environment, as in World War I or the Great Depression. The sharp U.S. disinflation of the 1980s posed a formidable challenge which the empirical price-adjustment equation could have failed but did not. A central implication of the resuscitated 1980-vintage empirical Phillips curves, the value of the sacrifice ratio of lost output required to achieve a permanent deceleration of inflation, turned out to be surprisingly close to predictions made in advance. This suggests that, at least in the U.S., the substantial changes in price-adjustment parameters observed in Tables 3 and 4 to have occurred in previous historical eras have been largely absent in the postwar U.S. setting. 30

The empirical stability and predictive success of the resuscitated U.S. Phillips curve is highly ironic in view of the inflammatory language used by Lucas and Sargent. If anything lay smoldering in “wreckage” in the mid-1980s, it was the few abortive attempts to estimate price equations within the framework of Mark I new-classical macroeconomics, particularly those by Robert Barro (1977a, 1978; Barro and Mark Rush 1980). So strongly was price inertia embedded in the U.S. data that Barro could explain price movements only by entering a distributed lag of between four and six years of monetary surprises that themselves only lasted a single quarter. Why agents should be reacting with a four-year lag to a one-quarter monetary surprise was never explained. The attraction for the economics profession of the empirical versions of Mark I new-classical macro, like the theoretical versions, was undermined by the discrepancy between the time lags involved in data dissemination, measured in days or weeks, as contrasted to the lags of price changes in response to nominal demand shocks, measured in years or half-decades.

IV. New-Keynesian Theory: Common Features

A. Essential Features of Keynesian Economics

The essential feature of Keynesian macroeconomics is the absence of continuous market clearing. Thus a Keynesian model is by definition a non-market-clearing model, one in which prices fail to adjust rapidly enough to clear markets within some relatively short period of time. Common to almost all Keynesian models is the prediction that in response to a decline in nominal demand, the aggregate price level will decline less than proportionately over a substantial time period, during which the actual price level is above the equilibrium level of real output. The fact that the price level is too high means that the subequilibrium level of
output actually produced is not chosen voluntarily by firms and workers, but rather is imposed on them as a constraint. It is the decline in nominal demand together with the absence of full price adjustment that causes the economic system itself to impose the constraint on each agent; nominal demand is insufficient to generate adequate real sales at the actual price level. Each agent faces a constraint that is indirectly a result of its own failure to reduce sufficiently its price and this points to a coordination failure as a central ingredient in the description of Keynesian price stickiness.

So many people now refer to new-classical models as equilibrium business-cycle models that the word equilibrium has been co-opted as meaning the opposite of the term Keynesian. This leads some commentators to label an approach that is the opposite of equilibrium economics as disequilibrium economics. In one sense this is mere semantics; it does not matter whether we describe the U.S. in 1932 or Europe in the mid-1980s as being in a state of disequilibrium or low-employment equilibrium. However, the adjective disequilibrium is not helpful, as it conveys “a failure of agents to realize perceived gains from trade” (to use Robert Barro’s provocative 1979 phrase). Rather, it is best to regard the core feature of Keynesian economics as the gradual adjustment of prices and its corollary, that output and employment are not choice variables.

In contrast to new-classical equilibrium models, with their price-taking firms (“yeoman barbers”) making voluntary choices of the output level, Keynesian non-market-clearing models turn the role of prices and output on their head, with demand-taking firms making voluntary choices of the price level. Thus price-setting behavior is the essence of Keynesian economics. Any attempt to imbed it in microeconomic foundations must begin from monopolistic or imperfect competition, not perfect competition, because Keynesian agents are inherently price setters, not price takers.

A central theme of both new-classical and new-Keynesian macroeconomics is that accurate empirical predictions are necessary but not sufficient conditions of an acceptable theory. In addition, a theory must have microeconomic foundations in the behavior of utility-maximizing and profit-maximizing individual agents. The search for tractable analytic models to form the micro foundations often leads analysts astray, causing them to lose sight of the forest as they construct their single exquisitely proportioned tree. Almost all new-classical theory is conducted in the analytically convenient setting of “representative agent models,” where one can move back and forth between the individual agent and the aggregate economy simply by adding or removing i subscripts, without having to consider such analytically inconvenient issues as coordination failures or the speed of price adjustment. Professional microeconomists, as distinguished from macroeconomists who dabble in microeconomic modeling, find the failure to confront aggregation seriously to be the most critical flaw of representative agent modeling. A surprising number of new-Keynesian models share in common the neglect of aggregation; the aggregate economy is simply the representative agent multiplied by n. Accordingly, we shall find unsatisfactory those new-Keynesian models that neglect aggregation issues, and we shall emphasize the central role of interactions among agents, including coordination failures, macroeconomic externalities, and producer-supplier relations.

31 John Pencavel suggests to me that this critical view by microeconomists is widespread.
B. Micro Agents, Macro Spillovers, and Coordination Failures

The development of new-Keynesian economics in the past decade has primarily involved the search for rigorous and convincing models of wage and/or price stickiness based on maximizing behavior and rational expectations. The ground rules of this search are commonly accepted. The key ingredient in the now-abandoned Mark I new-classical approach was not rational expectations, but rather the assumption of continuous market clearing, as is evident in the labels new-classical macroeconomics or equilibrium macroeconomics. Most new-Keynesian models combine rational expectations with maximizing behavior at the level of the individual agent. Any attempt to build a model based on irrational behavior or submaximizing behavior is viewed as cheating. No new-Keynesian wants to build a model with agents that Barro (1979) could criticize as failing "to realize perceived gains from trade." So the game is to tease a failure of macro markets to clear from a starting point of rational expectations and the maximization of profits and individual welfare at the micro level. In short, effects of changes in nominal aggregate demand on real output and employment are derived in models characterized by equilibria in which each individual agent takes only those actions that make him better off and in which no agent foregoes an opportunity to take advantage of a "gain from trade."

The recent development of microfoundations for wage and price stickiness does not, of course, represent the first attempt to develop micro underpinnings for Keynesian economics. The work of Friedman and Franco Modigliani on consumption, Dale Jorgenson on investment, and William Baumol and James Tobin on the demand for money were all based on profit-maximizing behavior at the micro level. But all this work was carried out within a partial equilibrium framework, assuming in particular that both real income and the price level were given. A useful distinction can be made between micro theorizing at the level of individual demand and supply functions, and micro analysis of the market mechanisms (especially the price system) whereby the actions of maximizing agents are coordinated. Even before the advent of new-classical economics, the work of Robert Clower (1965) and Axel Leijonhufvud (1968) stressed interactions and spillovers among markets and argued that the nexus of research should shift from a partial to a general equilibrium setting.

An interesting aspect of recent U.S. new-Keynesian research is the near-total lack of interest in the general equilibrium properties of non-market-clearing models. That effort is viewed as having reached a quick dead end after the insights yielded in the pioneering work of Barro and Herschel Grossman (1971, 1976), building on the earlier contributions of Don Patinkin (1965), Clower, and Leijonhufvud. Explaining sticky wages and/or prices is viewed as a tough task, and no one is prepared to anticipate its achievement by examining broader theoretical implications. The disdain shown by new-Keynesian theorists for the work of Barro and Grossman, and the latter evolution of that line of research in the hands of Malinvaud, Muellbauer and Portes, Benassy, Grandmont, and others—notably all Europeans—is understandable in light of the primacy of micro foundations models as the prerequisite

---

32 I am grateful to David Laidler for suggesting this distinction.
33 In defense of the new-Keynesian approach, Andrew Weiss has suggested to me that "we have to solve the partial equilibrium problems first; these also are the most interesting."
for macro discourse. Nevertheless I find that even the most perceptive new-Keynesian commentators tend to forget the central message of these models.

This message is that spillovers between markets imply that the failure of one market to clear imposes constraints on agents in other markets. Most notably, when firms in a recession experience a decline in sales at the going price, this excess supply of commodities "spills over" into a decline in labor demand at the going real wage. In this light, I am sometimes surprised to read otherwise sensible commentators refer to the inconsistency of one or another new-Keynesian explanation with microeconomic evidence on the elasticity of labor supply. Such evidence is simply irrelevant for Keynesian macroeconomics. In a genuinely Keynesian model, agents are not in a position to choose the amount they work or produce as output varies over the cycle, and so the constrained amount that they do work or produce cannot be interpreted as tracing movements along a choice-theoretic labor supply curve or production function.

Much existing new-Keynesian theorizing is riddled with inconsistencies as a result of its neglect of constraints and spillovers, and its focus on single markets, one at a time, in a partial equilibrium framework. For instance, several of the most prominent models of price determination in the presence of adjustment costs limit the source of price stickiness to the product market; they often assume a perfectly competitive labor market in which workers slide up and down their labor supply curves, indifferent between economic states that offer relatively large and small amounts of leisure. Such models stand Keynesian economics on its head, because any satisfactory explanation of business cycles that warrants the label "Keynesian" must incorporate not just price stickiness, but in addition some element that explains the evident unhappiness of the employed in recessions and depressions. Further, such models fail to explain why the adjustment costs that lead to price stickiness do not in parallel imply wage stickiness.

One important exception to this neglect of macroeconomic constraints and spillovers is the seminal work of Russell Cooper and Andrew John (1988) on macroeconomic coordination failures. In several new-classical models in which agents set output, they show that spillovers and strategic complementarities can arise at the levels of preferences and technology or in the organization of transactions. They reach the same conclusion as Barro and Grossman (without making the connection) that macroeconomic quantities belong in microeconomic choice functions. Almost alone among recent American authors in the new-Keynesian tradition, Cooper and John cite Jean-Pascal Benassy's fixed-price models (1975, 1982) and conclude for such models that "strategic complementarity is a distinguishing element of models with Keynesian features" (1988, p. 461).

The contribution of Cooper and John reaffirms the traditional view (see particularly Leijonhufvud 1981) that coordination failures represent the core problem in macroeconomics. In response to a nominal demand change, no single private agent has an incentive to move its price exactly in proportion unless it believes that all other agents will do likewise, and will do so without delay. In Tobin's example,

---

34 Research on general disequilibrium or fixed-price models appears to have become a specialized European activity in macroeconomics, with near-total invisibility in a recent survey I conducted of first-year graduate macro reading lists at the top ten American economics departments.
No one can see the spectacle in the theater or stadium if everyone stands, but who has the incentive to obey a general admonition to sit down. When the teacher tells her grade school class there will be no picnic unless all gum-chewing ceases, would any rational child who shares the general liking of gum stop? Threats against everybody in general addressed to nobody in particular rarely work. (1989, p. 15)

The same point can be put differently: Rational microeconomic agents care about the relation of their own price to their own costs, not to aggregate nominal demand. Unless a single agent believes that the actions of all other agents will make its marginal costs mimic the behavior of nominal demand with minimal lags, the aggregate price level cannot mimic nominal demand, and Keynesian output fluctuations result.

A notable limitation of most recent formal models related to coordination failures, including Cooper and John and Steven Durlauf (1989, esp. p. 110), is a classical setting of competitive output setters, rather than a Keynesian world of monopolistic price setters. In Durlauf’s words,

> the hallmark of this class of theories is the compatibility of different levels of real activity with the same microeconomic specification of individual firms and consumers. The key source of the multiplicity of long-run equilibriums is the positive effect that high production by some set of agents has on the decision of others to produce.

This approach, based in part on seminal research by Peter Diamond (1982, 1984), essentially concerns the cyclical behavior of productivity, the positive response of which is claimed to reflect “thick markets” as a result of “positive complementarities.” However, this has little to do with the essential Keynesian coordination failure, the absence of incentives for price-setting agents to move their individual prices in tandem with aggregate nominal demand rather than individual marginal cost.35

C. Real Rigidities, Nominal Rigidities, and the Indexation Puzzle

Two central distinctions are required as a preliminary to any summary of recent new-Keynesian work. The first is between price setting in product markets and wage setting in labor markets. The second distinction is between nominal rigidity and real rigidity.

The necessary condition for non-market-clearing is a barrier to the full adjustment of nominal prices, that is, something that prevents movements in nominal prices that are equiproportionate to movements in nominal demand. However, some of the new-Keynesian theories explain real rigidities as the stickiness of a wage relative to another wage, of a wage relative to a price, or of a price relative to another price. Explanations of real rigidities in product markets include customer markets, inventory models, and theories of markups under imperfect competition, while those of labor markets include implicit contracts, efficiency wages, and insider-outside models. But theories of real rigidities are subject to the criticism that they do not explain nominal rigidity, because nothing prevents each individual agent from indexing its nominal price to nominal aggregate demand.

There is surprisingly little discussion in recent new-Keynesian papers of optimal indexation nor of the relation between the absence of full indexation and the sources of nominal rigidities. Jo Anna Gray (1976), Fischer (1977b), and others showed in the mid-1970s that full CPI indexation is not optimal in the presence of supply shocks. Intuitively, no agent

---

35 In related work Peter Howitt (1986) calls this effect a “thin market externality.”
can afford an indexed contract that rigidifies real wages and relative prices if supply disturbances continuously shift the optimal relative price. However, Gray’s argument only supports indexation to a mix of price indexes preferred by firms and workers, not zero indexation. Failing to index is tantamount to linking the prices and wages of individual agents to a price index whose value is constant, and this becomes increasingly irrational as the inflation rate increases.  

Further, full indexation of the wage rate to nominal GNP escapes most of the theoretical objections to CPI indexation, because nominal GNP indexation leaves the price level free to move to equate the real wage to the marginal product of labor. Adopting our previous notation with lowercase letters representing growth rates, the condition necessary for labor’s share in national income to remain constant is that the growth rate of the real wage \((w - p)\) equals the growth rate of labor’s average product \((q - n)\):

\[
\begin{align*}
  w - p &= q - n, \\
  w &= p + q - n = x - n.
\end{align*}
\]

Thus the growth rate of the nominal wage rate should be indexed to the growth rate of nominal GNP per unit of labor input \((x - n)\). If an adverse supply shock reduces labor’s average product, then such indexation allows the needed reduction in the real wage, whether nominal GNP growth remains constant and the rate of inflation increases, or whether the inflation rate remains constant and the growth rate of nominal GNP decelerates.

Fischer (1986, pp. 152–53; 263–69) has pondered why indexation to the price level is so often incomplete, and why we so rarely observe indexed contracts contingent on other variables (whether economy-wide like nominal GNP or idiosyncratic variables like firm sales or profits). The primary barrier to indexation may be the costs of making contracts more complicated, particularly when it is recognized that there are conflicts along at least two dimensions. First is the Gray-Fischer point that the presence of aggregate supply shocks makes incomplete indexation optimal, and second is the presence of firm-specific shocks that create a conflict between the general market basket to which workers would prefer to index, and the firm-specific variables to which firms would prefer to index. Parties to a contract may differ not only in their objective functions, but also in their perceptions of the relative importance of aggregate demand shocks, aggregate supply shocks, and firm-specific shocks, and these perceptions may change continuously, requiring that the form of indexation in each new contract be negotiated from scratch.

As we review sources of rigidities in Keynesian models, we shall return to the issue of nominal GNP indexation. Are the nominal rigidities adequate to explain the real-world absence of such indexation? How are the two distinctions, product versus labor market and real versus nominal rigidities, related to each other? We begin our inquiry by reviewing models of nominal price rigidity in product markets, beginning with the elementary example of a textbook monopolist. This example implies that the response of price to a shift in demand is conditional not just on the elasticity of demand and the shape of the marginal cost curve, but crucially on the shift (if any) of marginal cost in response to demand. Thus product and labor market rigidities are complementary and may be of equal importance.

\footnote{McCallum (1986, p. 409) argues that linking to a constant price index instead of to the CPI would be chosen only by those agents whose most preferred index is negatively correlated to the CPI.}
V. The Search for Structure: Nominal Price Rigidity in the Product Market

A. The Textbook Monopolist and the Behavior of Marginal Cost

The behavior of a textbook monopolist is part of relative price theory, and therefore would appear to belong in our subsequent discussion of real rigidities. However, the monopolist model has been used to derive the conditions under which costs of adjustment create a barrier to changes in nominal prices. This explains the connection between theories of nominal price stickiness and the traditional partial equilibrium analysis of a price-setting monopolist illustrated in Figure 3. Note that two special assumptions are made in drawing Figure 3, that the demand curves are linear and that the marginal cost curve is horizontal. Implications of dropping both of these assumptions are discussed shortly.

In response to a shift in the demand curve from $D_0$ to $D_1$, quantity will change unless there is an equiproportionate shift in nominal marginal revenue and nominal marginal cost at the original level of real output. The implied marginal cost curve that maintains a constant level of output ($Q_0$) is labeled “Required” $MC_1$. If, following the decline in demand, marginal cost drops instantly to the “Required” $MC_1$ line, then the intersection of $MR$ and $MC$ will drop from $E$ to $G$, and the price will fall by exactly the vertical displacement of the demand curve, from $P_0$ to $P_2$. Any source of incomplete adjustment in marginal cost can then explain an incomplete adjustment of price. For instance, if the marginal cost schedule remains fixed at $MC_0$, the intersection of $MR$ and $MC$ occurs at point $F$, the new price is $P_1$, and the new quantity is $Q_1$.

When we loosen the two special assumptions incorporated into Figure 3, we alter the path of the price level at output levels away from the initial level $Q_0$ but not the basic conclusion about the required drop in marginal cost for the economy to remain at $Q_0$. For instance, replacing the special assumption of a linear demand curve with a demand curve of constant elasticity, while retaining the assumption of a horizontal $MC$ schedule, point $C$ would lie directly to the left of point $A$, and the price level would remain fixed in the presence of any leftward movement of the demand curve. Second, replacing the horizontal $MC$ schedule with a positively sloped $MC$ schedule going through point $D$ would move points $C$ and $F$ down and to the right, thus increasing the response of the price level to the decline in demand and correspondingly reducing the output response.

This analysis suggests that the primary reason for sticky price adjustment is the sticky adjustment of marginal cost. This would appear to place the analysis of cost stickiness at the top of the new-Keynesian research agenda. From the standpoint of the aggregate economy, the most important cost component is labor cost,
suggesting the familiar idea from the old Keynesian economics that wage inflexibility is the key element in price stickiness. However, from the standpoint of the individual firm, labor cost may be less important than purchased materials as a component of cost, and this recognition elevates to the top of the research agenda, along with wage determination, the formation of expectations by individual producers about the prices of purchased materials.

While Figure 3 identifies the rigidity of marginal cost as the key ingredient in price stickiness, it also leaves open a role for direct barriers to the adjustment of price to the profit-maximizing level for the monopolist, that is, to point C in the case of a fixed MC schedule or to point D in the case of a fully flexible MC schedule. Point B represents a price above the profit-maximizing levels C or D, and could be explained by costs of adjustment of the price level emphasized by new-Keynesian theorists under the general heading of menu costs or by old-Keynesian economists who emphasized rules of thumb like fixed markups of price over long-run average cost. If the price level is predetermined at point B, while marginal costs are predetermined along the schedule MC₀, output and employment may vary up and down in response to variation in product demand without a change in the real product wage.

This analysis of Figure 3 helps to organize our treatment of recent new-Keynesian research on the sources of price stickiness. First we examine the studies of direct barriers to price adjustment, independent of the behavior of marginal cost, which cause the price to deviate from the price that would be set by a profit-maximizing monopolist who has no costs of adjustment to consider. These direct barriers may be subdivided into two categories, (1) state-dependent rules, which call for price changes if the optimal price strays outside of boundaries determined by menu costs of price adjustment, and (2) time-dependent rules, which call for price changes at fixed and predetermined intervals written into contracts and are in turn presumably based on the costs of negotiating new contracts at more frequent intervals. This branch of new-Keynesian economics reinterprets these rules as profit-maximizing when menu-type or negotiation-type costs of adjustment are taken into account. Then we turn to sources of stickiness in marginal cost, both in prices of purchased materials and in wages.

It should be clear from this analysis that the labor market and product market may be equally important in contributing sources of price rigidity. There has been some tendency to stress product markets relatively more in recent research and to search for some source of nominal rigidity for prices in the form of state-dependent or time-dependent rules. Yet it is clear from the monopolist example that any source of nominal rigidity will do: A menu cost for wage adjustment will make marginal cost sticky and indirectly create a source of nominal price stickiness, even if costs of adjusting prices are completely absent.

B. The Representative-Agent Model of Monopolistic Competition

It is clear from Figure 3 that the mere presence of monopolistic competition does not create a presumption of price stickiness. Some ingredient must be introduced either as a direct barrier to instantaneous price adjustment or as a source of sticky marginal cost. In the recent new-Keynesian literature this point is most often made in the context of a simple model of a representative-agent monopolist developed by Blanchard and Nobuhiro Kiyotaki (1987) and described
as the canonical model of monopolistic competition by Fischer (1988).\textsuperscript{37} There are \( n \) identical producer-consumers producing goods that are imperfect substitutes, and there are no purchased materials. Nominal aggregate demand depends only on the nominal money supply. Marginal cost consists of the marginal disutility of production for each producer-consumer. The canonical model describes the determinants of output and of the desired relative price \((P_i/P)\). With constant returns in production and a constant marginal disutility of work, the model is equivalent to Figure 3 above with a flat \( MC \) schedule and a constant-elasticity demand curve. The producer reacts to changes in demand by changing output while leaving the relative price constant.

Only with an upward sloping \( MC \) schedule (due either to decreasing returns to labor in production or to an increasing marginal disutility of work) does the producer desire to change the relative price in response to a shift in demand. However, because there is complete symmetry across producers, relative prices must all be equal to unity. An attempt to decrease relative price in response to a decline in demand leads to a decrease in all nominal prices and in the aggregate price level, and this adjustment of the aggregate price level continues until all relative prices are back to unity. Money is completely neutral. The only element introduced by monopolistic competition is a declining marginal revenue schedule, which means that in equilibrium (with \( P_i/P = 1 \)) price is above marginal cost rather than equal to marginal cost, and output is lower than in competitive equilibrium.


There is no role for sticky marginal cost in the Blanchard and Kiyotaki "pure" model of monopolistic competition, because the imposition of symmetry across identical representative-agent producers has the effect of implicitly indexing both the relative price \((P_i)\) and marginal cost to the aggregate price level, which in turn depends only on the nominal money supply. Thus the new-Keynesian theorists recognize that they must go beyond the mere introduction of monopolistic competition in order to locate the sources of price stickiness. One route is to study direct barriers to nominal price adjustment in the form of state-dependent or time-dependent rules. The other direction is to study the sources of sticky marginal cost.

C. \( S,s \) State-dependent Pricing Rules

The new menu-cost literature owes its origins to a paper by Barro (1972) on the \( S,s \) approach to price adjustment by a profit-maximizing monopolist who faces a lump-sum cost of adjusting prices. The common theme linking the older \( S,s \) literature and the newer menu-cost literature is that price setters do not change price every time the desired price level changes, but only when the desired level deviates by more than a particular percentage from the current price. In the \( S,s \) literature the width of the percent band is arbitrarily given, while in the menu-cost literature the width, while also given, is presumed to be "small" and ultimately capable of being explained by particular adjustment costs. For expository purposes these contributions may be discussed together, because they both concern barriers to the adjustment of nominal prices and share the common theme of a percentage band within which the price remains fixed.

The basic \( S,s \) result is derived for a monopolist facing a stochastic additive shift in its demand curve taking the form
of a random walk without drift. The optimal strategy for the monopolist is shown to be the selection of "floor" and "ceiling" bands, with the price remaining constant when the shift is within the bands and changing fully to the new desired level when the shift is outside the bands. The width of the band, expressed as a percentage of the current price, depends positively on the cost of a price change and inversely on the opportunity cost of not changing, which in turn depends on the slopes of the demand and cost functions.\(^{38}\)

This result is specific to a demand disturbance that is modeled as a random walk, so that changes in the disturbance are serially independent, and as yet optimal rules have not been derived for more general processes in which the changes in the disturbance are serially correlated (as surely they must be in view of serial correlation in changes in nominal demand evident in Figures 2 and 3). Instead, most of the extensions of the S,s approach concern inflation which is at a sufficiently rapid rate that the price level cannot decrease, so the choice problem is simplified to choosing the timing of price increases. Eytan Sheshinski and Yoram Weiss (1977, 1983) show that the S,s approach carries over to inflation; now the price is increased at any point when it sinks below the optimal price by an amount exceeding a lower s band.

Andrew Caplin and Daniel Spulber (1987) have investigated the implications of aggregating S,s behavior from the level of the individual to the aggregate economy. Their striking result is that one-sided S,s rules (as are appropriate in an economy with an inflationary bias) do not lead to price stickiness or the non-neutrality of money. If firms face both local and aggregate shocks, their price changes will be independent and staggered across time. But when they do increase their individual price, they will raise it sufficiently to boost the aggregate price level by the full amount of the aggregate shock. For example, if demand increases in a series of one-unit steps, and adjustment costs limit individual firms to a price increase every fourth step, then that individual price increase will be four units and will increase the aggregate price level by one unit.

The Caplin-Spulber result is contingent on an unrealistic assumption, that the desired price follows a continuous and monotone path. A more general model, which reverses their main conclusion, has been developed by Caplin and John Leahy (1989). Their main point is that when the monetary shocks are two-sided, that is, when money can go both up and down, without any monotonic tendency in a single direction, there can be long periods in which the aggregate price level does not change in response to monetary disturbances. Intuitively, money is neutral only when the economy continuously hits an upper S or lower s band, but a more general stochastic process for money may leave it inside both bands for substantial periods during which there is no incentive for any agent to change its nominal price.

A difficulty in the S,s literature is that for analytical tractability all firms are identical, and thus have price increases of equal size that differ only in timing. This is belied by virtually all evidence on cross-section pricing behavior, including the differing cyclical responsiveness of prices across industries in the Great Depression (shown in Table 2). This evidence suggests that elements beyond simple state-dependent pricing rules must lie behind observed price behavior at the micro level.

---

\(^{38}\) The original result was derived by Barro (1972) and is restated by Blanchard and Fischer (1989, pp. 402-05).
D. The Menu-Cost Insight and Its Limitations

Taking the $S,s$ literature as a point of departure, what new insights have been contributed by the menu-cost literature developed in the mid-1980s? The menu-cost approach began defensively in response to those critics who argued that costs of changing nominal prices are much too small to justify output fluctuations of the size observed in the U.S. Its key insight is that second-order adjustment costs may have first-order social consequences, simply because profit functions are flat on top. Following a change in demand there may be little difference in the firm's profit if it does or does not adjust its price, and thus even small menu costs may potentially dissuade the firm from price adjustment. Yet the social consequence of such a failure to adjust price may be large swings in output.

The proponents of the menu-cost approach are quick to admit that this widely used label is misleading. Included among the nominal costs of price adjustment are not just the literal application of the label to changing prices on menus, lists, catalogs, and other printed material, but more generally the entire range of costs that managers must incur whenever nominal prices are changed. Meetings, phone calls, and trips to renegotiate with suppliers all fall under the rubric of menu costs. Included in this more general definition of menu costs is Okun's (1981) analysis of the product market. Okun explains the reluctance of firms to shift from FIFO (first in first out) pricing policies to the more economically rational behavior of replacement cost pricing as a consequence of the perceived costs of delegating pricing authority to lower levels of management, in contrast to general FIFO-type rules of thumb that save these costs of delegation even if they lead to pricing decisions that may be otherwise suboptimal. All these physical costs of printing, negotiating, and delegating are doubtless present in the real world of business, although one can quibble with their importance. Costs of negotiating are also a key ingredient that motivates staggered contracting, a time-dependent rule considered in the next section.

Whatever the nature of the menu costs, the analysis may be presented in terms of Figure 3 above, which already provides the ingredients necessary to illustrate the point initially made by George Akerlof and Janet Yellen (1985) and Mankiw (1985). Following Mankiw, we examine the situation in which demand has declined in Figure 3 from $D_0$ to $D_1$ and marginal cost has declined from $MC_0$ to "Required $MC_1$." The optimal price-output point is at $D$, and we ask what difference is made if the firm leaves its price unchanged at $P_0$. Figure 4 copies the new demand curve and shows the same points $B$ and $D$ as in Figure 3. The difference between profits at points $D$ and $B$ is shown by the rectangles $T - R$. However, at point $B$ total surplus is smaller by the area $S + T$ than at point $D$. But the firm will only reduce price if the extra profit $T - R$ exceeds the menu cost. Mankiw shows that as the price elasticity of demand varies from ten to two, the ratio of the social cost to the profit increment varies from 23 to 200. His results, as do Figures 3 and 4, assume that the marginal cost schedule is flat. In general, the flatter is the marginal cost schedule, the smaller are the menu costs needed to make the firm's fixed-

---

39 Laurence Ball disputes this interpretation and claims that "the central point [of recent work] is that nominal rigidity has negative externalities because it exacerbates fluctuations in real aggregate demand." But this is clear as a matter of definition (see equation (4) in Section II.A), is common to any theory of price stickiness, and has nothing to do with the particular contributions of recent work.
price decision optimal and hence to create an output response from a change in nominal aggregate demand.

At least four important criticisms of the menu-cost approach may be offered. Taken together, they make a strong case against this core contribution of the new-Keynesian macroeconomics.

1. A consideration of symmetry brings the basic conclusion into question. If the failure to reduce price in response to a demand reduction makes output too low, then the failure to raise price in response to a demand increase makes output too high. Yet, starting from an initial profit-maximizing equilibrium level of output like $Q_0$ under monopolistic competition, society gains from additional output because price is above marginal cost. Hence the menu-cost model fails to prove its point: Social costs in recessions are balanced by social gains in booms. Any cost from price rigidity must involve increasing the variance of output, not changing its mean, and hence are likely to be second-order, just as the costs of changing price are second-order. One cannot conclude that sticky prices necessarily reduce welfare, for the comparison of two second-order effects turns on model-dependent comparisons of parameter values.\(^{40}\) This argument of Laurence Ball and David Romer (1989a) greatly weakens the appeal of the menu-cost approach, although their own model implies that the second-order social costs can be much larger than the costs of changing price.

2. But independent of the Ball-Romer symmetry argument, the menu-cost approach seems flawed from the start, because it considers only costs of price adjustment and totally ignores costs of output adjustment. This places its assumptions in diametric opposition to other important branches of macroeconomics, such as Tobin’s $q$ theory of investment based on time-dependent physical costs of changing the capital stock, or the production-smoothing theory of inventory behavior based on the assumption that a smooth rather than variable production minimizes cost. Costs of output adjustment raise the cost of not changing price and tilt the firm’s decision toward price flexibility; whether costs of output adjustment raise the social costs of aggregate output fluctuations depends on the relative size of the private and social costs.

3. The two-period comparison of Figure 4 neglects the calculus of costs and benefits in future periods. The proper setting is dynamic, as with the analogous question of the “sacrifice ratio” in the form of a temporary aggregate output loss required to achieve a permanent reduction of the aggregate inflation rate. The proper comparison is between the one-shot menu cost and the present value of the infinite stream of losses by maintaining the price (and output) levels at different values than the social optimum. The Ball-Romer symmetry argument vitiates the force of this criticism, because the

\(^{40}\) This important point credited to Ball and Romer, whose paper was written in 1986, is summarized and endorsed by Rotemberg (1987, pp. 83–85).
infinite stream of losses when the price is set too high is balanced by a similar stream of gains when the price is set too low.

4. Like the S,s approach, the menu-cost approach fails to explain why prices of some products are more flexible than others over the business cycle. The failing of the Mankiw version illustrated in Figure 4 is in this regard similar to that of the canonical Blanchard-Kiyotaki model described earlier; marginal cost is simply assumed to move in proportion with demand. Once we consider the many layers of heterogeneity of products and industries studied in the industrial organization literature of Part III.A, we recognize that no individual firm can assume its marginal cost will be perfectly correlated with aggregate demand. Subsequently this will lead us to the input-output table as an essential component in the description of price stickiness and will reinforce our previous point that the failure to consider heterogeneity and aggregation issues is a central flaw in representative agent modeling.

Thus we return to the original objection to models of nominal rigidity based on adjustment costs. Any satisfactory model of price rigidity must be able to cope with the Great Depression, yet the magnitude of demand shifts between 1929 and 1933 would seem to swamp any reasonable guess as to the magnitude of S,s bands or menu costs. And one does not have to dip into history to doubt the relevance of such adjustment costs. Everyone has witnessed the fast-changing price tags in the produce section of the neighborhood supermarket: There seems to be nothing to prevent the price of a pint of strawberries from moving from $1.89 to $0.59 to $1.89 in successive weeks. Carlton’s evidence shows not just that prices can jump by large amounts in successive weeks, but by small amounts in successive months. John Roberts, David Stockton, and Charles Struckmeyer (1989) have found in time-series data for 20 two-digit U.S. industries over 1958–83 that the adjustment of nominal prices to nominal labor and materials costs takes place extremely rapidly. For four industries, 90 percent of price adjustment occurs within the month, and for no industry is the first-month adjustment lower than 45 percent. This provides strong evidence that the menu-cost approach is on the wrong track, and that the key issues concern the stickiness of both wages and materials costs, not final goods prices. The mere fact of imperfect competition and price tags appears to be quite compatible with nominal flexibility.

E. Time-dependent Rules and Staggered Contracts

The last element to be considered in new-Keynesian explanations of nominal price rigidity is the staggered contract model. As noted above, new-Keynesian economics can be said to begin with the Fischer (1977a) and Phelps-Taylor (1977) models of staggered contracts, which emphasized wage contracts. More recently, models of staggered price contracts have been developed by Blanchard (1983, 1986) and Ball and Romer (1989b), among others. These models investigate the implications of staggered overlapping price-setting intervals of constant length, and in the case of Ball-Romer investigate the conditions necessary for firms to engage in staggering.

The staggering and overlapping of intervals in which individual prices or wages are fixed introduces a critical element of realism into new-Keynesian economics. As shown by Blanchard (1983), a change in nominal demand can affect output for a period that exceeds the length of the interval during which prices are predetermined, which we will call the contract interval even though there is no necessity that explicit or implicit contracts be involved.
Consider contracts of $n$ months, with a fraction $1/n$ of agents resetting contracts each month. The firms that reset their price in the first month following a demand shift move their price not to the optimum given the level of nominal demand, but to the optimum contingent on the fact that the other outstanding contracts cause a fraction $(n - 1)/n$ of the aggregate price level to be preset. Any firm adjusting all the way would cause a suboptimal divergence of its relative price from the optimum level, given the stickiness of other prices.

The study of staggered prices takes as its point of departure that the length of predetermination of prices reflects a balancing of the costs of adjusting prices and the opportunity costs of nonadjustment, just as in the $S_s$ model. Because this decision need be made only by the profit-maximizing price-setting monopolist, there is no need for an actual contract with another agent. There has been little attention to the nature of the adjustment costs or in particular their variation across industries or over time, which is unfortunate as this might provide the element that is missing in so many new-Keynesian models, the ability to explain cross-time and cross-industry differences in price behavior. In particular, there is no element in the theory that would explain why the rate-of-change ($\alpha$) coefficient of price adjustment increased in most countries during World War I, or why the inertia ($\lambda$) effect increased in most countries except Japan after World War II. As a separate criticism, there has been no attempt to introduce explicit indexation into these staggered contract models, which contain no element to explain why firms do not predetermine real prices for a time interval (presumably to save on management decision costs) and then index the nominal price to nominal GNP.

Instead, attention has been concentrated on the question of why there is staggering rather than complete synchronization. Ball and Romer (1989b) show that staggering is a stable equilibrium if there are firm-specific shocks that arrive at different times for different firms. However, they show that synchronization can also be an equilibrium: "Multiple equilibria are possible because there is an incentive for synchronized price setters to remain bunched, but not for staggered price setters to move toward synchronization" (1989b, p. 193). There seems to be a debate as to whether firm-specific shocks are sufficient to guarantee a staggering equilibrium, but only in the context of simple models in which firms can choose only to change price at odd or even dates. A more general setting in which some prices are changed weekly and others yearly, and in which the yearly price changers have 365 possible dates on which to change, destroys the argument that each individual price setter still has an incentive to "bunch" price changes at the same time, because there is no such thing as the "same time." 41

VI. Sources of Real Rigidity in the Product Market

A. Customer Markets

The analysis of nominal price rigidity in Part V treats only the first quadrant of our two-by-two matrix defined along the

41 For this reason I find unconvincing the skepticism of Blanchard and Fischer (1989, p. 401) that it is possible to derive stable staggered contracting, as when they write "the introduction of stochastic idiosyncratic shocks does not make staggering more likely." Their argument is carried out entirely within an either-or choice between even and odd dates of price changing, and they show that a 50-50 equilibrium with the same number of firms choosing each date is unstable, because the slightest tilt in either direction gives all the other firms an incentive to shift. But with uneven frequency of shocks and a large number of possible dates of changing, the incentive to shift disappears. If my optimal frequency of price change is weekly, the fact that there is bunching with more price changes on January 1 than any of the other 364 days of the year does not lead me to limit my price changes to once a year on January 1.
dimensions of product versus labor market and nominal versus real rigidity. We turn now to models of real rigidity, that is, models that explain why real wages or prices are unresponsive to changes in economic activity. In the product market, a model of real rigidity explains why a firm would choose to hold its relative price or price-cost margin constant. In the context of the canonical Blanchard-Kiyotaki monopoly model, this occurs with a constant marginal disutility of work and constant returns to labor. In the textbook monopoly model of Figure 3, the price-cost margin is fixed if the marginal cost schedule is horizontal and the elasticity of demand is constant.

Recall that our discussion of indexation in Part IV.C introduced a basic objection to all models of real rigidity. No matter how rigid is the real wage or real price, what prevents the nominal price and wage from being indexed to nominal GNP? At that point we asked whether nominal rigidities were sufficiently important to be able to explain the absence of nominal GNP indexation. Thus far we have concluded that nominal rigidities based on $S,s$ or menu-cost models are not convincing, while time-dependent rules in the form of staggered price-setting intervals are completely compatible with any form of indexation. Thus a critical test for the theories of real price and wage rigidity is whether they stand up to the indexation criticism.

Perhaps the earliest prominent model of real rigidity in product markets explains why customers do not respond instantaneously to changes in real prices, that is, in the price charged by one firm relative to others. Arthur Okun (1975), building on the work of Armen Alchian (1969) as well as Phelps and Sidney Winter (1970), popularized the distinction between auction and customer markets. The former are perfectly competitive. But, in the latter, costly search makes customers willing to pay a premium to do business with customary suppliers, and intertemporal comparison shopping discourages firms from changing prices in response to short-run changes in demand in order to avoid giving customers an incentive to begin exploring. Okun argues that his customer-search model explains markup pricing practices based on full costs. Customers appear willing to accept as fair an increase in price based on a permanent increase in cost, whereas transitory events, whether an increase in demand or a reduction in productivity, are not generally expected to last long enough to justify price increases.

Okun’s approach has several critical defects. He argues that price increases based on cost are perceived as fair, while cost increases based on demand are not so perceived. The case for customer dissatisfaction is difficult to argue, because any loss of goodwill created by a price increase in a boom would be balanced by a gain of goodwill created by a price decrease in a recession. Okun seems to be thinking of an inflationary world in which price changes are one-sided, as in the trend-inflation $S,s$ literature. Further, the fairness explanation leaves open the determination of fair behavior, and in fact what is perceived to be fair may just reflect whatever behavior may be normal, for whatever reason. Thus Okun’s approach has an element of circularity.

Okun’s approach also seems vulnerable to the same criticism often directed at Lucas-type new-classical models: Why should a firm be afraid to lose customers when raising prices in response to higher nominal demand if information on higher nominal demand is instantly available? What prevents all firms from indexing to nominal demand and advertising price specials on items priced lower than would be warranted by the indexation formula? As we have seen already, the elementary theory of monopoly pricing behavior by itself sug-
gests little for price flexibility. Everything depends on the response of marginal cost to aggregate demand shocks.

B. The Independence of Costs and Demand

Nevertheless, there is a deeper insight in Okun’s distinction between cost and demand. Firms raise price in response to an upward shift in the marginal cost schedule not just because it is optimal in a textbook model, but because they will go bankrupt if cost rises sufficiently in relation to price. There is no such economic necessity of raising price in response to an increase in demand when cost is fixed, and for a monopolist such a price increase is not even optimal with a constant-elasticity demand curve and a flat MC schedule. When OPEC raises the price of oil sharply in relation to the level of nominal aggregate demand, everybody understands why the local service station raises the price of gasoline at the pump, but they do not understand why an increase in aggregate demand requires any such response of the gasoline price if the costs of service station inputs are fixed.

This distinction hinges on the possibility that shifts in marginal cost can be independent of shifts in aggregate demand. Our historical study of price adjustment in Parts II and III stressed the theoretical and empirical importance of supply shifts. Ball and Romer’s (1989b) study of staggering emphasizes the critical role of idiosyncratic firm-specific shocks. Giuseppe Bertola and Ricardo Caballero (1990) emphasize the role of idiosyncratic uncertainty in explaining infrequent price adjustment at the micro level. New-classical Mark I macroeconomics was built on Lucas’ distinction between local and aggregate shocks. Okun’s emphasis on cost-based pricing leads us to broaden Lucas’ two-way distinction between local and aggregate demand shocks and suggest a four-way distinction between local and aggregate cost shocks.\(^{42}\)

This four-way distinction creates two complementary sets of reasons why firms may rationally expect marginal cost to move differently from marginal revenue. First, marginal revenue may move with aggregate nominal demand but marginal costs may not. This would occur if a firm believes that its costs depend not just on nominal demand but on local supply factors (e.g., harvests, strikes, price changes for imported materials). Second, in a situation with nominal aggregate demand fixed, a firm might face a local shift in demand (e.g., a decline in beer drinking in response to drunk-driving laws) that reduces marginal revenue, while marginal cost is fixed, tied to aggregate nominal demand. More generally, any set of covariances among the four shock concepts is possible.

C. The Role of the Input-Output Table

To be a credible explanation of real price rigidity, the distinction among local and aggregate cost and demand shocks must be embedded in a world with many heterogeneous firms interacting within a complex input-output table. With only two firms, each supplying the other, information would be cheap enough to permit both firms to disentangle the local versus aggregate component of their costs. But with thousands of firms buying thousands of components, containing ingredients from many other firms, the typical firm has no idea of the identity of its full set of suppliers when all the indirect links within the input-output table are considered. Because the informational problem of trying to anticipate the effect of a currently perceived nominal demand change on the weighted-average cost of all these suppliers is difficult to formulate and probably impossible to

\(^{42}\) I have previously (1981, pp. 520ff.) suggested a distinction between aggregate and local components of both cost and demand with explicit reference to Lucas’ original two-way classification.
solve, since (as Bresnahan emphasizes) thousands of elasticities are involved, the sensible firm just waits by the mailbox for news of cost increases and then, Okun-like, passes them on as price increases.

The input-output table approach provides a critical contribution not just to understanding real price rigidity, but also nominal rigidity. The standard accusation against all theories of real rigidity, made often above, is that they are consistent with nominal flexibility achieved through indexation to nominal demand. Yet the input-output table approach emphasizes the high fraction of a firm's costs that is attributable to suppliers of unknown identity, with some unknown fraction producing in foreign countries under differing aggregate demand conditions. This environment would give pause to any firm considering nominal demand indexation of the product price, because the failure of all suppliers to adopt similar indexation could lead to bankruptcy when nominal demand declines. Thus the input-output approach borrows one element that is basic to Keynesian economics, the coordination failure that arises from the lack of private incentives to solve a social problem, with another element inherited from Lucas, the distinction between aggregate and idiosyncratic shocks.

One criticism of the input-output approach claims that with perfect information about aggregate variables, the only equilibrium of the economy would be for immediate adjustment of all prices to nominal shocks. Yet this ignores the fundamental assumption that marginal cost and marginal revenue are imperfectly correlated with aggregate demand. Under these conditions each firm would be unwilling to index price to nominal GNP both because marginal cost may not move with nominal GNP even if marginal revenue were to do so, and vice versa.

A good reason for every domestic firm to refuse to index its product price to domestic nominal demand would occur to any economist from, say, Belgium or Chile. Because we know that purchasing power parity (PPP) fails and that real exchange rates are volatile, why would any firm adopt indexation of its price to domestic Belgian or Chilean nominal GDP, which would disconnect its price from the large share of its costs that are imported? The input-output approach, by stressing the independence of marginal cost and aggregate demand, provides an understanding of the lack of indexation to domestic nominal GNP and thus the critical link that converts a theory of real rigidity into a theory of nominal rigidity.43

A firm's viability depends on the relation of price to cost, not price to nominal GNP. Aggregate macroeconomic stability is a public good subject to a free-rider problem. No individual firm has an incentive to take the risk posed by nominal GNP indexation, which would take away from the firm the essential control required of the relation of price to cost. In this sense, the input-output explanation of nominal rigidity requires capital markets that are imperfect enough to penalize the profit volatility that would result if a firm tried to index its prices to nominal demand without being sure in advance that its suppliers would do likewise.

There is another sense in which the input-output table explains nominal rigidity. It creates a technological environment for staggered price setting, similar to but more complex than Taylor's staggered wage setting. Today's product price is based on costs set at many different dates in the past as product components weave their way through the input-output table. This may appear to violate the maxim that prices should be based on replacement cost. But there are too many links in the input-out-
put table for the producer even to guess what the replacement cost may be. The automobile firm may receive a notice from the headlight maker of a price increase, but no warning of a price-increase notice that is already in the mail from the filament maker to the headlight maker or from the copper maker to the filament maker. Blanchard (1987b) uses the term *cumulation hypotheses* to describe the role of the input-output table in translating prompt price adjustment at the individual level to gradual price adjustment at the aggregate level. He provides suggestive supporting evidence that in disaggregated data prices adjust faster than in aggregate data. The automobile, headlight, filament, and copper maker may all respond to cost increases within a day, but months can separate the effects of a change in the price of copper from the ultimate change in the price of automobiles.

The input-output table approach dominates menu costs in explaining why the price of strawberries is more volatile than the price of automobiles (because strawberries are not physically transformed from farm to market). It explains the different rate-of-change adjustment coefficient ($\alpha$) across industries by two auxiliary assumptions. First, auction markets are distinct for customer markets and are limited mainly to crude and intermediate goods. Thus products like strawberries and plastics that appear relatively early in the input-output chain have relatively flexible auction-like prices. But what is it that creates more price rigidity for more complex products later in the chain? Partly it is the law of large numbers that cancels out idiosyncratic supply shocks for final products incorporating large numbers of different purchased materials. But there also may be a role for wage rigidity, as the prices of products embodying relatively large amounts of embodied labor, like automobiles, tend to be more rigid than that of products embodying large amounts of embodied land, like wheat or strawberries.

Thus the input-output approach is complementary to theories of rigidity in the labor market.

A safe compromise place to end the discussion of product-market rigidities is to admit that the input-output approach is complementary as well to at least one of the new-Keynesian approaches based on nominal rigidities. The input-output approach needs some additional element to explain why we do not observe in the real world extremely frequent small price changes every day as firms react to each tiny cost change as it arrives in the mail through the input-output table (while Carlton documents some such small changes, long intervals of complete rigidity are common as well).

The core element that needs to be added to the input-output approach is a cost to making price changes every day that causes rational managers to concentrate price-setting decisions at discrete intervals. There is no need to force this sort of nominal rigidity into a single semantic category; the core factor for some firms may best be described as staggered time-dependent rules and for others as state-dependent rules based on menu-type costs. Undoubtedly these categories overlap because many firms face both time-dependent and state-dependent costs. Some firms that routinely hold price-setting meetings once a week or month to save on managerial time costs may decide at those periodic meetings to leave some or all prices unchanged when the difference between the current and optimal price does not yet exceed the perceived cost of printing new menus and catalogs.

### VII. The Search for Structure: Labor Market Behavior

#### A. The Relation of Wage and Price Behavior

The dominant new-Keynesian view is that nominal rigidities originate in the product market, not the labor market. The
path of wages, to use the words of Mankiw, is “completely indeterminant and completely irrelevant” (1988, p. 446). Yet surely this goes too far. Mankiw follows earlier writing, notably by Barro (1977b) and Robert Hall (1980), who have argued that wage rigidity is irrelevant for employment determination. In the context of a long-term or even lifetime job, there is no reason for the wage in a given time period to be equal to the marginal product of labor. The wage can be an installment payment on a lifetime contract.

However, the claim that sticky wages are irrelevant to allocation calls for prices to be perfectly flexible, as is required for perfect market clearing, while wages are sticky. We have already observed in Section II.C that capital markets are likely to impose a tax on the resulting profit variability. Further, the monopolist example shows that prices will not be perfectly flexible unless all elements of marginal cost are perfectly flexible. This brings us back to indexation: The sticky wages that are installment payments for lifetime jobs must be fully indexed to nominal demand for Barro, Hall, and Mankiw to be correct that sticky wages are irrelevant to allocation.

Just as it is implausible for wages to be sticky while prices are perfectly flexible, so is the reverse, for wages to be perfectly flexible while prices are sticky. Yet this is just what is assumed in much of the menu-cost literature reviewed in Part V. When menu costs lead rational firms to avoid price changes and meet demand through changes in output, corresponding fluctuations in labor input are required. In menu-cost models the real wage adjusts to make workers willing to change the amount they work; that is, the nominal wage rate is perfectly flexible. In the Blanchard-Kiyotaki canonical model of monopolistic competition, the representative agents set their relative price to minimize the marginal disutility of work; that is, they slide along their voluntary labor supply curve. As Rotemberg has noted, “both of these approaches have the very un-Keynesian implication that in recessions workers are close to indifferent between working and not working.”

For new-Keynesian models to avoid inconsistency, their distinction between small menu costs of price changes and large social costs of output changes must apply equally in the labor and product markets. The same costs of adjustment that inhibit price changes must apply equally to wages, which are just another price. Sticky prices cause changes in nominal aggregate demand to be transmitted directly to shifts in the demand curves facing not just individual firms, but also individual workers. The Barro-Grossman spillover model discussed in Part IV.B achieves the desired symmetric treatment, in which sticky wages and prices cause both firms and workers to face constraints on the amount they can sell.

Blanchard and Fischer (1989, p. 427) state that the key issue in new-Keynesian economics is explaining why “labor and output supply functions [are] relatively flat.” They intend this phrase to mean real rigidities. Yet their choice of words is unfortunate, because it ignores the distinction between aggregate and individual supply curves, as well as between notional and effective supply functions. The labor supply function of an individual head of household may be vertical, but any mechanism that rigidifies the real wage will cause the individual to be pushed off of this notional supply function. Actual behavior traces shifts in the effective labor demand schedule and tells us nothing about the shapes of notional functions. In our interpretation the key issue is the explanation of wage and price rigidity, not the explanation of

44 Blanchard has written to me that what he means by flat labor supply is “the set of real wages and employment traced out as the marginal product of labor shifts” and not “the competitive labor supply curve.” The issue here is the possibly misleading choice of words, not any substantive difference between my interpretation and that of Blanchard and Fischer.
why labor and output supply curves are flat. And at a deeper level, as argued above in Section IV.B, the really central element is the coordination failure that underlies wage and price rigidity.

B. Early Theories of Real Rigidity: Search Models and Implicit Contracts

Just as theories of price stickiness can usefully be divided between theories of nominal versus real rigidity, so can theories of wage stickiness. Reflecting the chronological development of the field, we begin with models of real wage rigidity and then turn to models of nominal wage rigidity.

Widely recognized as the first attempt to build structural models of labor market behavior as the outcome of maximizing behavior was the new microeconomics of the famous volume edited by Phelps (1970). Most of the papers in the Phelps volume, including Phelps' own desert island parable, yielded market-clearing conclusions and as such should be regarded as part of the development of new-classical rather than new-Keynesian ideas. In the parable, workers are on isolated islands and react to a wage cut by boarding rafts to sample wage offers on other islands. Variations in employment during business cycles are due solely to the voluntary response of workers to changes in the expected real wage. The parable ignores the prompt availability of aggregate information, fails to explain layoffs and "no help wanted" signs, and yields the counterfactual implication that voluntary quits vary countercyclically (see Okun 1981, ch. 2).

The main contribution of the new microeconomics volume was not to business cycle theory but rather to explain why the natural unemployment rate is greater than zero, due chiefly to the work of Dale Mortensen (1970a, 1970b). In a world of costly information and heterogeneous jobs and workers, workers sample from an array of job offers and firms sample from an array of workers. Unemployment is a voluntary activity, but all voluntary unemployment is not socially beneficial, and government unemployment benefits tend to stretch out the interval between searches, imposing a social cost through the taxes levied on some to support the extended search interval of others. The new microeconomics volume also contained the important Phelps-Winter (1970) theory of customer markets, based also on the assumption of imperfect information. We have seen that this was later picked up and developed by Okun in a new-Keynesian rather than new-classical setting.

While the new microeconomics was explicitly classical in approach, the next wave of contributions under the heading of implicit contract theory was the first to develop what some initially thought was a microeconomic explanation for Keynesian wage stickiness. In the simultaneously written and independent contributions by Costas Azariadis (1975), Martin Baily (1974), and Donald F. Gordon (1974), employees were assumed to be relatively more risk averse than their employers, mainly because of self-selection of individuals to become entrepreneurs. Firms maximized profits by minimizing the variability of income to their workers, who disliked variability, in effect providing a compensation package that consisted partly of pecuniary wage payments and partly of insurance services.

It was soon recognized that this approach provides no satisfactory explanation of Keynesian unemployment; it justifies only a fixed-income contract (i.e., tenure) rather than the fixed-wage variable-employment

---

45 Thus I concur with Barsky and Solon (1989, pp. 29–30), who find that procyclical real wage behavior at the individual level in micro data is consistent with noncyclical behavior in aggregate data. This pattern reflects a cyclical variation in the "employment opportunities" (read "constraints") that face both "stayers" at firms who face changing opportunities to work overtime, and "switchers" who face cyclicality in opportunities for across-firm career advancement.
contracts actually observed. Variable employment is explained only by a gratuitous element patched onto the theory, government side-payments during periods of unemployment. Even when these variable employment fixed-wage contracts are generated by the theory, they have the un-Keynesian implication that workers are equally happy when employed and unemployed. Further, workers are shown to care about stability in real income, not nominal income, so implicit contract theory has no explanation for the failure of workers to insist on full indexation of wage contracts.

C. Labor Unions

The effects of labor unions have been extensively analyzed by labor economists. Bargaining models have been developed in which firms and unions, which in turn act on behalf of their member workers, bargain over wages and employment. Some models characterize the employment decision as a unilateral decision of management, as it is in many contracts. These models that are concerned only with wage setting are sometimes called the right-to-manage model and fall between two extremes. At one extreme firms are all powerful and are able to pay the minimum wage possible, that is, the competitive wage. Because firms have complete control of both employment and the wage, this subclass of models does not warrant the label bargaining model at all; the efficiency wage models discussed in Section VII.E fall into this class. At the other extreme is the union-monopoly model dating back to Dunlop (1944); here too there is no bargaining, because firms set employment and unions set the wage.

A more general model is developed by Ian McDonald and Robert Solow (1981), who show that a bilateral monopoly between a firm and a union can lead to relatively large employment fluctuations and relatively small real-wage fluctuations, thus contributing a source of real-wage rigidity. In an extension, McDonald and Solow (1985) examine the impact of business-cycle fluctuations on a labor market segmented into a union primary sector and a competitive secondary sector. Reflecting the small real-wage fluctuations in the union sector, they show that either permanent or temporary changes in real aggregate demand widen sector wage differentials in recession and cause greater fluctuations in primary sector than secondary sector employment.

Yet the formal theory of unions does not provide a general explanation of Keynesian wage rigidity. If union members care about stability of employment, it is difficult to understand why they are willing to tolerate a wage rate that is set for a substantial interval, while the decision on the amount of employment is left to the firm. Obviously if the wage rate is predetermined as part of a union contract, this rigidifies marginal cost and hence prices, and nominal demand fluctuations are transmitted to output and employment. But, overall, the union literature leaves open the question why the wage rather than the level of employment is set by contracts, and why the wage rate is not indexed to nominal demand so as to stabilize employment.

Another problem is raised by the empirical evidence of Table 3 in Part III. Unions became important in the U.S. only after the mid-1930s, yet the estimated rate-of-change (α) and level (γ) effects for the U.S. are the same during 1873–1914 and 1954–87. Some factor other than unions must account for the price stickiness evident in

46 More technically, as pointed out to me by Blanchard, it is the marginal utility of consumption that is equalized between the employed and unemployed. The ranking of utility depends on the form of the utility function.

47 This point is made by Blanchard and Fischer (1989, p. 453).
U.S. data for the nineteenth century. The main contribution of unionization may have been the particular American phenomenon of the three-year staggered wage contract, which has doubtless contributed to the higher inertia parameter for both wages and prices. Yet here too there are problems, because inertia in price change seems to have increased more than for wage change (Table 3), and, further, inertia increased substantially for price change in the postwar period in the U.K., France, and Germany (Table 4), nations in which the three-year contract is not prevalent.

D. Insider- Outsider Theory

Another body of work that deals with the existence and persistence of unemployment is the insider-outsider theory. The insiders are experienced incumbent employees whose jobs are protected by a variety of labor turnover costs which make it costly for firms to replace them. The outsiders, who are either unemployed or work in the casual or secondary labor market, have no such protection. Assar Lindbeck and Dennis Snower (1986, 1988) argue that, owing to these turnover costs, the insiders gain market power, which they use to their own advantage, without necessarily taking fully into account the interests of the outsiders. Further, the insiders often can influence the turnover costs themselves by agreeing to cooperate among themselves but not with outsiders should the latter attempt to gain employment by underbidding the insider wage. This structure causes unemployment for the outsiders, who cannot find jobs even though they would be willing to work for less than the prevailing insider wage.

Although the insider-outsider theory contributes to our understanding of union behavior, it is not primarily a contribution to the union literature. A wide variety of labor turnover costs may well be significant even in the absence of unions, for example, hiring, training, negotiation, litigation, and firing costs, as well as costs that can be directly imposed by the insiders when they shirk or fail to cooperate in the presence of underbidding outsider entrants. Nevertheless, the insider-outsider theory does suggest a rationale for unionization by showing how unions can organize and coordinate insiders' rent-seeking activities.

The insider-outsider theory sheds light on a variety of labor market phenomena, such as the persistence of unemployment, differences in variability of employment across industries and countries, labor market segmentation, the duration and composition of unemployment, and the interindustry wage structure. The theory has been applied to the puzzle of persistently high unemployment in Europe in the 1980s by Blanchard and Lawrence Summers (1986) and Lindbeck and Snower (1988) and has become one of several explanations of the hysteresis hypothesis (see Section II.D), in which the rate of unemployment depends on the history of actual unemployment rather than, as in Friedman's original version, being "ground out" by the microeconomic structure of the economy. The insider-outsider approach explains the emergence of high unemployment in the 1980s as an indirect consequence of the oil shocks of the 1970s, which created a temporary adverse reduction in labor demand and caused the insider work force to contract. When labor demand recovered the remaining insiders set wages to maximize their own welfare, thereby discouraging employment and making the high unemployment persist.48 The best evidence in support of this approach is the work of Layard and Nickell (1987) which shows that the demand pressure variable entering the Phillips-curve wage equation is not total unemployment, but rather total unemploy-

48 To this point the discussion of the insider-outsider model is largely based on several paragraphs of text kindly contributed by Assar Lindbeck and Dennis Snower.
ment minus the long-term unemployed. However, to the extent that it explains the persistence of high European unemployment by high insider real wages, it is subject to the criticism (R. Gordon 1988) that high unemployment was immune to the moderation of real wage growth and the disappearance of the European wage gap in the 1980s.

E. Efficiency Wage Theory

If any development in the microeconomics of labor markets could be called the “rage of the 80s,” it is efficiency wage theory, based on the hypothesis that worker productivity depends on the level of the real wage. When there is such a link between the wage rate and worker efficiency, firms may rationally pay a real wage rate that exceeds the market-clearing level. Firms may refuse to reduce the wage to hire members of a pool of unemployed workers who may be available at a lower wage, fearful that a reduction in real wages for existing workers may reduce productivity by more than the gain in lower wages. The appearance of an excess supply of labor in such a setting can be shown to be consistent with maximizing behavior of both firms and workers. There is substantial overlap between the insider-outsider and efficiency wage models, as they both focus on barriers to underbidding by unemployed outsiders. While the insider-outsider approach emphasizes the market power of incumbent workers, the efficiency wage approach stresses the choice problem of firms that have imperfect information about the productivity of their employees.49

The reasons for the response of productivity to the real wage vary across models and include effort, reduced shirking, lower turnover and training costs, the ability of high-wage firms to screen and obtain a higher-quality labor force, and improved morale and loyalty.50 Virtually all the literature with implications for macroeconomics dates from the 1980s. Although most surveys trace the germ of the idea back three decades to early work on less-developed countries that posited a linkage among wages, nutrition, and health (e.g., Leibenstein 1957), the terms efficiency wages and efficiency earnings appear in Alfred Marshall’s Principles (1920, pp. 456-69). Another precursor of the idea is the negative relationship between wages and quit rates embedded in Phelps’ (1970) desert island parable and other early models in the new microeconomic literature. Efficiency wage theory provides a rare common meeting ground for mainstream and radical economists, because the far left in U.S. economics has taken the lead in developing theories of dual labor markets and for setting out policy proposals for higher minimum wages based on the assumed validity of the efficiency wage approach.51

The basic efficiency wage result is obtained in a simple model with identical, perfectly competitive firms and a production function in which labor input is multiplied by an efficiency factor $e$ that depends on the real wage. Because the elasticity of $e$ with respect to the real wage declines as the real wage increases, the first-order conditions require the firm to choose an optimal real wage rate ($w^*$) at which this

49 An excellent comparison of the two approaches is provided by Lindbeck and Snower (1988, ch. 3).

50 Two surveys of the literature that identify those authors and papers who have studied particular channels of efficiency wage effects are Katz (1986) and Weiss (1990).

51 On dual labor markets, see especially Peter Doeringer and Michael Piore (1971) and David Gordon, Richard Edwards, and Michael Reich (1982). More recent evidence by mainstream economists is provided by William T. Dickens and Kevin Lang (1985). For policy proposals based on efficiency wage assumptions, see Samuel Bowles, David Gordon, and Thomas Weisskopf (1983). Their policy proposal to raise the minimum wage assumes implicitly that the current wage is below the optimum efficiency wage, whereas all the work in the new-Keynesian tradition examines the implications of assuming that the actual wage is already at the optimum efficiency wage level.
elasticity is unity. Workers are hired up to the point where their marginal product equals the optimal wage \((w^*)\). The intuition of the unit elasticity result is that firms forgo efficiency gains that yield more than they cost when they pay below \(w^*\), while a wage above \(w^*\) would cost more than it yields in efficiency gains. Stated another way, effective labor cost is minimized at \(w^*\).

Because \(w^*\) is completely fixed by whatever factors of taste and technology that determine the \(e\) function, the firm’s reaction to any change in its relative price (i.e., a demand shock) is to cut employment while maintaining the wage rate at \(w^*\). Firms have no incentive to cut the actual wage, because this would actually increase their wage bill per unit of output. The extreme result of a fixed real wage in this model stems from the assumption that a worker’s efficiency depends on the absolute level of the real wage rather than on the real wage relative to something else, whether some measure of economy-wide real earnings or real wages in a perceived peer group or comparison group. A variant of this approach, in which effort depends on the relative real wage and on the unemployment rate (a high value of which raises effort by increasing the cost of job loss), allows the real wage to regain some flexibility and to depend inversely on the unemployment rate (Carl Shapiro and Stiglitz 1984; Summers 1988).

Several criticisms of the efficiency wage approach have been offered.\(^{52}\) One line is to propose that job applicants should “buy” high-wage jobs either by offering lump-sum payments or performance bonds to employers, or by offering to work in low-wage apprentice status for an initial period. Efficiency wage proponents point out, however, that unemployed workers lack sufficient wealth and are risk averse, and that the same monitoring problems that generate the efficiency wage result also make it unlikely that banks or other financing sources will come forth to provide finance for the initial lump-sum payments or performance bonds. This defense does not rule out low-wage apprenticeships, which are in fact observed, but these can be interpreted alternatively as a means of sharing the cost of training rather than as the “sale” of a job by a firm. A second criticism is that the efficiency wage model is dominated by direct payments to workers in proportion to their efficiency, whether through piece-rate contracts or through “tournaments” that pay workers according to their ranking by performance. The defense against this criticism is similar to the first. Piece rates and tournaments are subject to information problems. Workers involved in joint production do not often have a unique claim to a “piece,” while payments to a team invite shirking by some members of the team. Tournaments are also difficult to implement; there are rarely many workers in a firm doing exactly the same tasks and no way to rank across tasks.

Overall, the efficiency wage approach seems to be an essential ingredient in explaining numerous aspects of microeconomic labor market behavior, including segmented labor markets, persistent wage differentials for similar workers that are not equalizing differences, queues for high-paid jobs, and procyclical fluctuations of the quit rate.\(^{53}\) Variations on the model can explain why firms sometimes dismiss workers instead of cutting their wage. However, as a theoretical underpinning of the new-Keynesian paradigm, it suffers from the same defect as all models of real rigidities. If workers gear their effort to the real wage, they have the incentive to increase their effort when the real wage increases.

\(^{52}\) This paragraph summarizes Weiss (1990, pp. 6–10).

\(^{53}\) The most controversial item on this list is persistent wage differentials, as argued by Katz and Summers in a series of papers, including Katz (1986) and Katz and Summers (1989). For a sample of a dissenting view, see Robert Topel’s comment which appears after the latter paper.
there appears to be no barrier to full wage indexation that allows firms simultaneously to maintain worker effort through maintenance of the optimal real wage \( w^* \), while changing the nominal wage in tandem with the nominal price in order to achieve macroeconomic self-correction. Further, the efficiency wage theory has little to say about the sources of variations in wage and price responsiveness over time and across countries that were identified in Part III.

This negative verdict applies only if a new-Keynesian explanation of nominal wage and price rigidity is erected on the sole base of the efficiency wage theory. However, once the input-output approach and the independence of local and aggregate costs and demand are accepted as the underlying reason why actual economies do not index to nominal demand, the way is open to accept the efficiency wage approach as another source of cost rigidity within the input-output table, of potentially equal importance with the uncertain evolution of the prices of purchased materials. Once again, we find that the new-Keynesian approach is most convincing when sources of real and nominal rigidity are combined rather than when either one or the other is proposed as the sole explanation.\(^{54}\)

F. Nominal Rigidities: Wage Contract Models

Theories of wage stickiness can be based on real rigidities, as in the approaches outlined above, or on nominal rigidities. The most influential work that rationalizes nominal rigidities in new-Keynesian labor market analysis is the staggered-contract approach of Fischer (1977a) and Taylor (1980). The classification of contract rigidities as nominal is subject to the preceding criticism—that the negotiation costs that rationalize the existence of contracts do not rule out fully indexed contracts. The costly negotiations set the real wage, while the nominal wage is costlessly indexed, preferably to nominal GNP.

The Fischer-Taylor contract literature is set up entirely in nominal terms and does not discuss the option of full nominal demand indexation, so we will discuss it on those terms, as a source of nominal rigidity. In Fischer’s version the wage for half the workers is set for two periods at the beginning of period \( t \) and for the other half at \( t + 1 \). The wage set for the first group can respond to any change in the money supply in the first period but not in the second. The greater flexibility of nominal money than of nominal wages is an assumption rather than a result and leads to real effects of perceived monetary disturbances that cannot occur within the new-classical framework. Fischer’s version assumes no barriers to price flexibility and market clearing in product markets. The unemployment his model generates during a period when money has declined but wages have not declined is classical (because of an excessive real wage), not Keynesian.

In the setting of an \( n \)-period rather than two-period contract model, Taylor (1980) makes the nominal wage fixed over the life of the contract (at a level that depends on the expected price and expected output) and setting the price as a simple markup over the average wage rate. A monetary disturbance falls fully on output during the period until the next contract renegotiation. Then wages can adjust quite rapidly, because the dependence of the negotiated wage on expected future output creates a strong feedback loop between unemployment and wage behavior. Nevertheless, because prices depend on wages set in any previous contract still in force, the duration of the real output response to a nominal monetary shock can last for much longer than the length of the contract, the same

\(^{54}\) In this important conclusion we fully endorse the basic message of Akerlof and Yellen (1985) and Ball and Romer (1987).
result as was subsequently derived by Blanchard (1983) for the product market (see Section V.E).

Taylor’s approach is sufficiently plausible and important to take seriously. Yet it is subject to at least two criticisms. First, the assumptions of staggering and of fixed contract length are arbitrary. In some places (especially Japan) contract expiration dates among firms in the union sector are nearly simultaneous. If contract length depends on a balancing of negotiation costs and the allocative costs of infrequent adjustment, one would expect changes in contract length in response to the variability of either local or aggregate shocks. A more general statement of this first criticism is that the existence of nominal wage contracts is not explained from the first principles of microeconomics. Models of optimal contracting do not produce the nominal stickiness generated by the Taylor-type contracting models. The Taylor approach needs to be supplemented by an extension to wage setting of recent work on staggered price setting by Ball and Romer (1989b) and others, as reviewed above.

The second problem is that, once the contract expires, the adjustment of the wage in response to expected future output is not complete but is bounded in Taylor’s model by an arbitrary Phillips-type adjustment coefficient. As argued by Blanchard (1987a), Taylor’s results require this adjustment coefficient to be relatively “small” and, if a cyclical response of the markup of prices over wages is allowed, that must be “small” as well. While Blanchard’s point suggests that Taylor’s wage adjustment may be too slow, I would argue the opposite. In particular, Taylor (1983) has claimed that it is possible for the monetary authorities to engineer a disinflation with no output loss. However, this result depends heavily on Taylor’s assumption that the effect of real demand on wage-setting decisions works through expected future real demand rather than past and current real demand. For wage setters to use a model to calculate the implications of their current wage-contract decisions on future real demand requires not only a universal belief that the announced disinflationary path of nominal demand will be maintained on target, but also a universal ability to forecast the response of actual prices to the path of nominal demand, as is required for future real demand to be predicted.

Where then do staggered wage contracts fit in? We have seen that full price flexibility for a monopolist requires full flexibility of marginal cost, and staggered contracts eliminate that full flexibility in the absence of instantaneous nominal GNP indexation. Barro, Hall, and Mankiw have argued that it is possible for firms to adjust their prices in proportion to a change in nominal aggregate demand if wages do not adjust. But this is not profit maximizing. In almost any model of monopolistic price setting, an incomplete adjustment of wages implies less than full adjustment of profit-maximizing prices. In this sense, recent new-Keynesian theorists have gone overboard in shifting the emphasis from the labor to the product market.

VIII. Conclusion

We have stressed throughout the need for new-Keynesian theory to address the most important elements of variability in the adjustment of prices along three dimensions, the inertia effect ($\lambda$), the rate-of-change effect ($\alpha$), and the level effect ($\gamma$). The industrial organization literature contributes ample evidence of differences in rate-of-change effects across industries. It also shows that intervals of fixed prices lasting months or years in some industries can coexist with frequent small price changes in other industries. It stresses that some industries are competitive, some are monopolistic, and some industries combine monopolistic cartels with competitive price wars. Within industries, all firms do not
exhibit the same price behavior, and given firms do not even charge the same prices to all customers. Heterogeneity is rampant, with hundreds of products common in many industries, and many products combine labor with hundreds or thousands of purchased components.

The time-series evidence shows a wide variety of price adjustment patterns across time and countries. The inertia (λ) effect has become more prevalent since World War II in every country but Japan. The rate-of-change (α) and level (γ) effects were remarkably similar before World War I and after World War II in most countries, but exhibited sharp divergences in between. The rate-of-change effect increased sharply during and after World War I, while the level effect virtually disappeared during the interwar period in the U.S., U.K., and Germany. The inertia-prone postwar U.S. is at one extreme, and Japan throughout most of its history is at the other extreme of relative flexibility.

A convenient image for understanding the desirable direction of new-Keynesian theory is a small 1 × 1 box set next to a gigantic n × n matrix, where n is measured in the thousands, if not the millions. The small box represents the identical representative agents of both new-classical models and the canonical monopolistic competition model, with their i subscripts, and the practice of treating the macro economy as identical to the representative agent with the subscripts removed. The gigantic matrix represents the real world, full of heterogenous firms enmeshed in a web of intricate supplier-demander relationships. This n × n matrix suggests two main themes of the theoretical review in this paper.

First, the key to introducing theories of real rigidity as a source of nominal price stickiness is to find a good reason why we do not observe nominal GNP indexation. That reason is simple, and is at the heart of all good microeconomics. Individual firms maximize profit by setting their own marginal cost equal to their own marginal revenue. They have no reason whatsoever to care about nominal GNP unless it provides useful information to supplement what they can learn from observing their “local” cost and demand. There are many reasons for firms to expect their nominal marginal costs and local demand to contain idiosyncratic elements that cause them to evolve independently from nominal demand. The most straightforward argument, which is enough to make the case, is that firms in a small open economy know that their costs are determined outside the national boundaries within which domestic nominal demand applies. This principle generalizes to firms in large open economies, because we know that even under flexible exchange rates purchasing power parity does not hold over long periods, so costs of imports and domestically produced import substitutes can evolve independently of domestic aggregate demand.

The independence of cost and demand, and the input-output table approach, represent two separate components in the required (but as yet missing) new-Keynesian analysis that can come to grips with the industry, cross-time, and cross-country facts summarized here. The idea of independent cost and demand shocks seems crucial to come to grips with the time-series evidence. Just as Lucas argued (1973) that Argentina had a more vertical Phillips curve because agents knew that aggregate demand shocks dominated local shocks, so we can argue in parallel that the increase in the rate-of-change coefficient (α) in Table 4 for the U.S., U.K., and Japan during 1915–22 reflected a recognition by price setters that the increasing importance of aggregate disturbances created a greater than usual correlation between changes in marginal costs and changes in aggregate demand. Similarly, the increase in persistence (the λ parameter) observed almost everywhere after World War II reflects a
widespread belief that government full-employment policies and the end of the gold standard created an upward drift in prices, leading to the expectation that marginal costs would have an upward drift and would no longer be a stationary process.

The input-output component is complementary to the independent shocks idea, and helps to explain why firms do not simply assume that marginal costs will move in parallel with aggregate nominal demand: Most firms do not know the identity of their suppliers, their suppliers' suppliers, and so on, because the input-output table is so broad and so deep. The input-output component of the proposed explanation is required to grapple with the industry evidence. Prices of corn and wheat on auction markets exhibit sharp daily swings, subject to administered limits. Prices of strawberries exhibit frequent sharp weekly swings. Prices of many crude materials exhibit frequent changes, both small and large. Yet prices of newspapers and many finished goods can remain unchanged for more than a year. A unified explanation that explains the degree of volatility of fixity for every product may be impossible to achieve, but the basic idea that crude materials are relatively volatile and finished goods relatively fixed seems compatible with the input-output approach which stresses the number of steps and number of purchased components that are mixed together with labor input in each final good. The input-output approach also leaves open a role for a theory of real wage rigidity, once it is admitted that nominal GNP indexation is unlikely. The input-output approach emphasizes the time lags in transmitting news of cost and demand changes back and forth within the input-output table. However, to explain why prices do not change by small amounts every day, this approach needs to be supplemented with a plausible mixture of time-dependent and state-dependent costs of daily price changes.

Once the independence of local costs, local demand, and aggregate demand is admitted as the fundamental explanation for the lack of nominal demand indexation, the way is open to take seriously new-Keynesian research on real rigidities in the labor market. Work on union behavior and on nominal contracting in the labor market does not appear promising, in light of the similarity of the $\alpha$ and $\gamma$ coefficients in most countries before World War I and after World War II. However, the efficiency wage model has strong persuasive power as to why firms resist real wage cuts, and the independence of shocks and input-output table explanations contribute the needed supplementary explanation of why real wage rigidity becomes translated into nominal wage rigidity. The other most promising development in the labor market literature is the insider-outsider approach, if only because the disenfranchisement of outsiders holds up the best available ray of hope that we have for understanding why the Phillips-curve level ($\gamma$) effect disappeared in the U.S., U.K., and German interwar periods, and perhaps in some European countries in the 1980s.

Our perspective that emphasizes independent shocks and the input-output approach reinforces the view that coordination failures are the essence of macroeconomic inefficiency in new Keynesian models. Should the government attempt to intervene to provide the missing coordination of microeconomic wage and price decisions, or should its activities be limited to the traditional Keynesian use of monetary and fiscal policy to manipulate aggregate demand directly? Clearly, traditional forms of internalization through tax and subsidy policy are infeasible in light of pervasive heterogeneity among products and decision makings in the millions; to go in this direction would mean slipping into the quagmire from which Eastern Europe is trying to emerge. Even mandatory indexation to domestic nominal demand may be suboptimal in many countries
where an important component of nominal marginal cost is set in foreign currencies and responds more to foreign than to domestic aggregate demand. This shifts the ultimate weapon for fighting business cycles back to the traditional instrument, aggregate demand policy, but not in the form of any old-fashioned Keynesian bias in favor of fiscal policy. If prices respond slowly to fluctuations in nominal GNP growth, then the optimal objective of stabilization policy should be to stabilize the growth rate of nominal GNP growth. Whether and how this can be achieved is beyond the scope of this paper.

Some commentators (particularly Blanchard 1987a) have lamented that, far from being a set of facts looking for a theory, the new-Keynesian paradigm suffers from too many unrelated theoretical explanations. Yet the essential features emphasized here, the independence of shocks, and the input-output table, embody a core set of realistic microeconomic elements: A technology of transactions, heterogeneity of goods and factor inputs, imperfect competition, imperfect information, and imperfect capital markets. Unlike time-dependent or place-dependent factors like unions, these essential features are timeless and placeless. They lead us to expect that the degree of price flexibility in the early nineteenth century would not be much greater than today, except insofar as the $n \times n$ matrix was smaller, with fewer steps from primary producer to final consumer, and indeed we find a basic similarity within each country in the $\alpha$ and $\gamma$ parameters before World War I and after World War II.

Recognition of the universality of these imperfections in economic life is overdue—perhaps a campaign can be started to change economic language so that these features will be considered the norm, rather than some aberrant or exotic flower. Rather than thinking of basic aspects of transaction and capital-market technology as imperfections, perhaps we could all start recognizing that these features are part of the way that markets function.

But these suggestions represent only the beginning of a needed research program. At the truly micro-micro level of relations between individual firms and customers, imperfections go far beyond anything that the independence of shocks, input-output, or efficiency wage approaches can explain by themselves. The evidence presented by Carlton that firms charge different prices to different customers for the same product, and apply nonprice allocation rules differently across customers, opens up a whole new dimension of heterogeneity that future theorists will need to consider. The ultimate merger of the new empirical industrial organization and the new-Keynesian macroeconomics (it is hoped not by leveraged buyout) seems a long way off, but it is a worthy goal to support.

APPENDIX

A

THE VARIETY OF HISTORICAL EXPERIENCE: REGRESSION METHODOLOGY AND ESTIMATES

Specification of Regression Equations

The aim is to estimate the three parameters $\lambda$, $\alpha$, and $\gamma$ in equation (9) in the text, which is repeated here for convenience:

$$p_t = \lambda p_{t-1} + \alpha x_t + \gamma q_t + z_t. \quad (9)$$

There may be some concern regarding the close resemblance of (9) to an identity obtained by rearranging (3):

$$p_t = \hat{x}_t - \hat{q}_t + \hat{q}_{t-1}. \quad (a)$$

Comparing (9) and (a), the former includes $p_{t-1}$ and excludes $\hat{q}_{t-1}$. Because inertia may be absent in some historical eras ($\lambda = 0$), the difference between (9) and (a) boils down to the exclusion of $\hat{q}_{t-1}$. Thus if (9) is a true structural equation, the identity (a) provides the value of the missing variable, $\hat{q}_{t-1}$. This argument is more transparent when (9) is transformed to include $\hat{q}_{t-1}$ but to exclude $\hat{q}_t$:

$$p_t = [1/(1 + \gamma)][\lambda p_{t-1} + (\alpha + \gamma)\hat{x}_t + \gamma \hat{q}_{t-1} + z_t]. \quad (9')$$

If (9) is a structural relation, so is (9'). Given the values of the right-hand variables in (9'), two of which

---

are predetermined \( (p_{t-1} \text{ and } Q_{1-t}) \) and one of which is endogenous \((\hat{x}_t)\), the role of the identity \((a)\) is to determine output as \( Q_t = Q_{1-t} + \hat{x}_t - p_t \). In short, the identity shows how output must change, given the structural price equation \((9')\). This just restates the basic point about Keynesian economics: If the current price is predetermined by an equation like \((9')\), then the current output level \(Q_t\) is determined as a residual.

The main estimation problem is not the fact that there is an identity linking some of the variables in \((9')\), but rather the endogeneity of \(\hat{x}_t\), which we have discussed above in the context of policy feedback. Because the essence of the problem is policy feedback, there can be no escape by replacing nominal GNP by the money supply, or by using money as an instrument for nominal GNP. And alternative versions with real GNP or unemployment are also subject to bias if policy feedback is not complete, as illustrated in Table 1. Our solution, which is to bracket the \(\alpha\) parameter by estimating alternative versions of \((9')\) with \(\hat{x}_t\) and \(\hat{q}_t\) as alternative explanatory variables, seems to be the best alternative.

To do this, we provide a pair of estimates for each dependent variable (price change, nominal wage change, and real wage change for the U.S., and price change for the U.K., France, Germany, and Japan). The specification for the first member of each pair is \((9')\). The specification for the second member of each pair is the transformation of \((9')\) that results when identity \((a)\) is used to replace \(\hat{x}_t\) by \(\hat{q}_t\):

\[
p_t = \left[1/(1 - \alpha)\right] \{ [1/(1 - \alpha)] [p_{t-1} + (\alpha + \gamma) \hat{q}_t + \gamma Q_{t-1} + z_t] \}.
\]

\((9')\)

The values of the three parameters \(\lambda, \alpha, \) and \(\gamma\) can be easily unscrambled. If in \((9')\) \(a_1\) is the estimated coefficient on \(\hat{x}_t\), \(a_2\) is the estimated coefficient on \(p_{t-1}\), and \(a_3\) is the estimated coefficient on \(Q_{t-1}\), then the parameters resulting from the estimation of \((9')\) are \(\gamma = a_2/(1 - a_3)\), \(\alpha = a_1 - \gamma (1 - a_3)\), and \(\lambda = a_3(1 + \gamma)\). If in \((9')_1\) \(b_1\) is the estimated coefficient on \(\hat{q}_t\), \(b_2\) is the estimated coefficient on \(p_{t-1}\), and \(b_3\) is the estimated coefficient on \(Q_{t-1}\), then the parameters resulting from the estimation of \((9')_1\) are \(\alpha = b_1 - b_2(1 + b_1 - b_3)\), \(\gamma = b_3(1 - \alpha)\), and \(\lambda = b_3(1 - \alpha)\).

Data, Detrending, and Parameter Shifts

Postwar data are taken from standard U.S. and OECD sources, and data prior to World War II are based on Nathan Balke and R. Gordon (1989) for the U.S., Charles Feinstein (1972) for the U.K., Kazushi Ohkawa and Mihohei Shinohara (1979) for Japan, and national sources as summarized by Angus Maddison (1982) for France and Germany. Data for Japan, the U.K., and the U.S. measure nominal GNP, real GNP, and the GNP deflator. Data for Germany and France prior to World War II measure real GNP, the CPI, and a hybrid concept of nominal GNP equal to the CPI times real GNP. The nominal wage equations for the U.S. are based on Rees’ data on average hourly earnings in manufacturing linked in 1980 to the BLS index of average hourly earnings in manufacturing. The real wage is this nominal wage series divided by the GNP deflator. Data sources are given in Appendix B.

The use of output data in estimating \((9')\) and \((9')_1\) requires a detrending procedure to define the \(\hat{x}_t\), \(\hat{q}_t\), and \(Q_t\) variables. Significant variations in population and productivity growth over the past century prevent the use of a single trend and require the choice of benchmark years to separate multiple piecewise log-linear output trends. The choice of the wrong benchmark years would introduce measurement error into all three of these variables. To avoid the possible criticism that benchmark years might have been selected to support or refute a particular hypothesis, all are copied from previous research directed at other issues. The main control for supply shocks is a set of dummy variables to proxy the effects of government intervention both in the form of price controls (as during World War II) and intervention to raise prices and wages, as during the National Recovery Act period in the U.S. Great Depression. Also for the U.S. we include a variable to measure the effect on aggregate inflation of changes in the relative prices of food and energy. The specific values of the supply-shock dummy variables are given in Appendix B.

The key issue of changing cyclical responsiveness can be addressed by two alternative methods. One obvious way of providing information on parameter shifts would be to estimate separate versions of \((9')\) and \((9')_1\) for each major subperiod within the available

---

1 For the postwar U.S., benchmark years are taken from my macroeconomics textbook (1990) and for the other four countries from Gordon (1988); for the pre-World War II period, U.S. benchmarks are taken from Christina Romer (1989), for France, Germany, and the U.K. from Solomos Solomou (1987), and for Japan from Gordon (1983). Inconsistency may result from the use of benchmark years originally selected by varying criteria—peak output in some cases, average output in others, and the level of output consistent with a particular unemployment rate in still others. For the European countries, where the benchmark years before World War I are all peaks (thus eliminating any positive values of \(Q_t\)), the resulting \(Q_t\) series is adjusted by subtracting its (negative) mean and converting the mean to zero. This results in a mix of positive and negative values. No such adjustments are carried out in the interwar of postwar periods.

2 The larger number of such supply-shock variables for the U.S. than for other countries may indicate that supply shocks have been more important in the U.S., or they may simply indicate that I am more familiar with the history of the U.S. than of the other countries. However, the extra attention given to the U.S. is largely due to the inclusion of wartime data for the U.S. but not for the other countries, where the years of World War II and its aftermath are excluded for all four of the other countries, while World War I and its aftermath are excluded for France and Germany.
TABLE A

Equations explaining annual changes in the GNP deflator, the nominal wage rate, and the real wage rate in the U.S., 1873–1987

<table>
<thead>
<tr>
<th>Variable</th>
<th>Price (1)</th>
<th>Nominal Wage (2)</th>
<th>Real Wage (3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged inflation (p_{t-1})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic effect</td>
<td>0.23*</td>
<td>0.38*</td>
<td>0.32*</td>
<td>0.55*</td>
<td>0.09</td>
<td>0.18**</td>
</tr>
<tr>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915–22</td>
<td>—</td>
<td>0.38**</td>
<td>—</td>
<td>0.31</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1954–87</td>
<td>0.45*</td>
<td>0.66*</td>
<td>0.15</td>
<td>0.37**</td>
<td>−0.31**</td>
<td>−0.29**</td>
</tr>
<tr>
<td>Excess nominal GNP growth (\dot{x}_t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic effect</td>
<td>0.35*</td>
<td>—</td>
<td>0.44*</td>
<td>—</td>
<td>0.09**</td>
<td>—</td>
</tr>
<tr>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915–22</td>
<td>0.40*</td>
<td>—</td>
<td>0.37*</td>
<td>—</td>
<td>−0.04</td>
<td>—</td>
</tr>
<tr>
<td>Excess real GNP growth (\dot{q}_t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic effect</td>
<td>—</td>
<td>0.32*</td>
<td>—</td>
<td>0.51*</td>
<td>—</td>
<td>0.19*</td>
</tr>
<tr>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915–22</td>
<td>—</td>
<td>0.73*</td>
<td>—</td>
<td>0.89*</td>
<td>—</td>
<td>0.16</td>
</tr>
<tr>
<td>Detrended log output (\dot{Q}_{t-1})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic effect</td>
<td>0.22*</td>
<td>0.29*</td>
<td>0.30*</td>
<td>0.43*</td>
<td>0.08</td>
<td>0.14**</td>
</tr>
<tr>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1930–53</td>
<td>−0.16*</td>
<td>−0.21**</td>
<td>−0.31*</td>
<td>−0.41*</td>
<td>−0.16**</td>
<td>−0.20*</td>
</tr>
<tr>
<td>Supply-shock variables (z_t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>World War I controls</td>
<td>−4.97*</td>
<td>−6.98**</td>
<td>2.00</td>
<td>−1.85</td>
<td>6.99*</td>
<td>5.13**</td>
</tr>
<tr>
<td>NRA</td>
<td>7.21*</td>
<td>9.52**</td>
<td>13.12*</td>
<td>17.53*</td>
<td>5.91**</td>
<td>8.01*</td>
</tr>
<tr>
<td>World War II controls</td>
<td>−16.68*</td>
<td>−20.19*</td>
<td>−2.13</td>
<td>−8.87**</td>
<td>14.54*</td>
<td>11.32*</td>
</tr>
<tr>
<td>Nixon controls</td>
<td>−2.94</td>
<td>−3.60</td>
<td>−2.47</td>
<td>−3.81</td>
<td>0.46</td>
<td>−0.21</td>
</tr>
<tr>
<td>Food-energy effect</td>
<td>0.63</td>
<td>0.97</td>
<td>0.39</td>
<td>0.81</td>
<td>−0.24</td>
<td>−0.16</td>
</tr>
<tr>
<td>R²</td>
<td>0.85</td>
<td>0.52</td>
<td>0.83</td>
<td>0.58</td>
<td>0.36</td>
<td>0.43</td>
</tr>
<tr>
<td>S.E.E.</td>
<td>2.02</td>
<td>3.59</td>
<td>2.45</td>
<td>3.90</td>
<td>2.53</td>
<td>2.38</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.10</td>
<td>2.09</td>
<td>2.08</td>
<td>2.11</td>
<td>2.23</td>
<td>2.15</td>
</tr>
</tbody>
</table>

Notes: Supply-shock variables are defined in Appendix B. * indicates statistically significant at 1 percent level, ** at 5 percent level.

Regression Results

Table A addresses the issue of changing cyclical responsiveness of prices, nominal wage rates, and real wage rates in the U.S.\(^3\) Six columns of results are shown for the entire 1873–1987 sample period, with equations for price, nominal wage, and real wage changes presented in pairs. The first member of each pair uses specification (9') in which nominal GNP

\(^3\) Here the wage data are adjusted for the trend in productivity growth (using piecewise linear trends between benchmarks), so that the dependent variable in the columns labeled Nominal Wage is actually the change in trend unit labor cost, and in the columns labeled Real Wage is actually the change in labor’s income share adjusted for cyclical fluctuations in productivity.
### Table B

Equations Explaining the Annual Inflation Rate, Five Countries, 1873–1986

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged inflation ((p_{t-1}))</td>
<td>Basic effect</td>
<td>0.18*</td>
<td>0.22*</td>
<td>-0.15**</td>
</tr>
<tr>
<td></td>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1923–38</td>
<td>—</td>
<td>—</td>
<td>0.27*</td>
</tr>
<tr>
<td></td>
<td>1960–86</td>
<td>0.24*</td>
<td>0.72*</td>
<td>0.59*</td>
</tr>
<tr>
<td>Excess nominal GNP growth ((\Delta \lambda_t))</td>
<td>Basic effect</td>
<td>0.58*</td>
<td>—</td>
<td>0.61*</td>
</tr>
<tr>
<td></td>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1915–22</td>
<td>0.12**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>1923–38</td>
<td>-0.26**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>1960–86</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Excess real GNP growth ((\Delta \lambda_t))</td>
<td>Basic effect</td>
<td>—</td>
<td>0.30*</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1915–22</td>
<td>—</td>
<td>1.33*</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>1923–38</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>1960–86</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Detrended log output ((Q_{t-1}))</td>
<td>Basic effect</td>
<td>0.26*</td>
<td>0.36*</td>
<td>0.36*</td>
</tr>
<tr>
<td></td>
<td>Extra effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1915–22</td>
<td>—</td>
<td>0.83*</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>1923–38</td>
<td>-0.18*</td>
<td>-0.27*</td>
<td>—</td>
</tr>
<tr>
<td>Supply-shock variables ((z_t))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.K. World War I</td>
<td>-10.16*</td>
<td>-23.48*</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>U.K. 1972–73 controls</td>
<td>-6.37*</td>
<td>-10.24*</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>France Poincare</td>
<td>—</td>
<td>—</td>
<td>10.48*</td>
<td>26.69*</td>
</tr>
<tr>
<td>France Popular Front</td>
<td>—</td>
<td>—</td>
<td>16.76*</td>
<td>44.37*</td>
</tr>
<tr>
<td>Hitler controls</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Japan oil shock</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.96</td>
<td>0.85</td>
<td>0.94</td>
<td>0.82</td>
</tr>
<tr>
<td>S.E.E.</td>
<td>1.28</td>
<td>2.46</td>
<td>1.51</td>
<td>2.65</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>1.99</td>
<td>2.03</td>
<td>1.88</td>
<td>2.35</td>
</tr>
</tbody>
</table>

**Notes:** Sample period for U.K. begins in 1958, and for Japan begins in 1888. Supply-shock variables are defined in Appendix B.

* Indicates statistically significant at 1 percent level, ** at 5 percent level.

---

**Appendix B: Data Appendix**

**United States**

GNP, deflator, and food-energy effect 1929–87:
Output and prices from National Income and Product Accounts, Tables 1.1 and 7.4, U.S. Department of Commerce. Food-energy effect (1959–87 only) is the difference between the growth rates of the fixed-weight consumption deflator and the fixed-weight deflator for consumption expenditures net of food.
and energy, from National Income and Product Accounts, Table 7.1. 1869–1929: Balke and Gordon (1989, Table 10).


Dummy variables World War I: 1918 = 1.0, 1919–20 = 0.5. NRA: 1933–34 = 0.5, 1935–36 = −0.5. World War II: 1943–44 = 0.5, 1946–47 = −0.5. Nixon: 1972–73 = 0.5, 1974 = −0.3, 1975 = −0.7.


Dummy variables France Poincare: 1926 = 1.0. France Popular Front: 1936–38 = 0.33. Hitler controls: 1937–38 = 0.5.


Output trend The output trend is calculated as a log-linear trend between the following benchmark years: 1856, 1865, 1873, 1882, 1889, 1907, 1913, 1920, 1940, 1951, 1961, 1972, 1979, and 1987.


Output trend The output trend is calculated as a log-linear trend between the following benchmark years: 1885, 1890, 1903, 1914, 1919, 1929, 1933, 1953, 1961, 1972, 1979, and 1987.

Dummy variable Japan Oil Shock: 1974 = 1.0.

References


———. “Output Fluctuations and Gradual Price Ad-


PHelps, EDMUND S. Microeconomic foundations of


