13 Empirical Evidence on Aggregate Supply Models and Business Cycles

Chapter 13 Contents

A. Topics and Tools

The literature testing the sources of business-cycle fluctuations and the nature of aggregate supply is voluminous. This chapter describes a few selected studies on a few of the major topics. In particular, the literature examining variants of the real business cycle model is enormous and only a tiny fragment is presented here.

Many basic tools of time-series econometrics are used in the studies reviewed here. You don't need to understand all the details of the statistical models to read this summary, but a good background in econometrics would be helpful for reading the source papers.

B. Basic Empirical Facts of Business Cycles

Some aspects of business cycles are subject to heated dispute, but many patterns are unambiguous regardless of the country or time period one examines. In addition to Stock and Watson (1999), which focuses on the United States, you may wish to examine the overview of evidence for the United States, Europe, and Japan presented in Chapter 14 of the macroeconomics text by Burda and Wyplosz (1998). Much of the discussion presented here is based on the results reported by Burda and Wyplosz.

Length and magnitude of cycles

Burda and Wyplosz present evidence in their Table 14.1 on the length and severity of business-cycle fluctuations in six countries from 1970 to 1994. The average peak-to-peak length of the business cycle varies from 5.5 years in the United States to just over 9 years in Japan. This is somewhat longer than the typical business cycle before 1970. Within the post-1970 sample there is considerable variation in the length of cycles:

from one as short as 6 quarters in the United States to 12-year cycles recorded in France and Germany.

The average percentage deviation of real GDP at the peak or trough from its level at the midpoint of the cycle is between 2% and 3.5% in these six countries. Thus, although business cycles still receive a lot of attention, recent cycles have been far less severe than the 30% decline in output that occurred in the United States from 1929 to 1933.[1]

Behavior of GDP components over the cycle

All of the components of private spending tend to be procyclical. Consumption is strongly correlated with income over the cycle, but tends to be quite a lot smoother. Since economic theory tells us that households would like to smooth their consumption, this is not surprising.

Investment is the most volatile of the components of expenditures. Investment bears a disproportionate share of the decline in recessions and experiences the strongest relative expansion in booms. Inventory investment is especially strongly cyclical, though its magnitude is small relative to the total economy.

Government purchases of goods and services are not strongly correlated with income over the cycle. Government budgets are typically set well in advance of actual expenditures, implying that decisions are made before the state of the economy is known. This may seem surprising because much of the government budget takes the form of "entitlements." Entitlement programs stipulate rules for eligibility for such programs as Social Security, welfare, and unemployment insurance. Anyone who qualifies is given benefits and the total cost to the government is not known in advance. Some of these outlays, such as welfare and unemployment insurance, are very cyclically sensitive, so we might expect that the government would spend more in recessions, making government spending countercyclical. However, recall that these entitlement programs are not government purchases of goods and services; they are transfer payments. Thus, they are not included in the government-spending variable that is added in as a component of expenditure in the GDP accounts.[2]

---

[1] Opinions differ on the causes of this apparent reduction in business-cycle severity. Some authors have argued that the responsiveness of policy authorities to business cycle conditions has effectively smoothed the cycle. Others claim that the macroeconomy has been subject to smaller and less frequent shocks. Christina Romer (1986) has demonstrated that much of the apparent reduction may be an illusion created by the method used to construct the prewar data. This evidence is reviewed in Romer (1999).

[2] Transfer payments are treated as "negative taxes" in the national-income accounts. They enter households' disposable income, but they are not part of the breakdown of GDP by expenditures because they are gifts rather than purchases of goods and services.

Imports are strongly procyclical; exports are less strongly so. Part of the increase in private spending associated with a business-cycle expansion is usually spent on foreign goods, which drives the cyclical behavior of imports. It is less obvious why exports should be procyclical, though those who believe that cycles are caused by aggregate-demand fluctuations would argue that causality may run from export demand to GDP.

Cyclical behavior of other variables

Romer's Table 5.3 on page 191 summarizes one way of characterizing the cyclical behavior of variables: the average change from peak to trough during a recession. The changes in employment and unemployment are as expected; the former is strongly procyclical and the latter countercyclical. Average weekly hours in manufacturing are also procyclical, but even with the strong movement of employment and average hours in the same direction as output, labor input still declines proportionally less in recessions than output does. This makes average labor productivity procyclical, which is a principal argument used in support of real-business-cycle models.

The cyclical behavior of prices and inflation is highly controversial. Prior to 1973, both conventional (Keynesian) wisdom and the bulk of the empirical data indicated that inflation was procyclical. The Phillips curve, which we shall study shortly, suggested a negative relationship between unemployment and inflation, which provided evidence that inflation was higher in booms and lower in recessions.

However, the "stagflation" that occurred in the middle 1970s after the OPEC oil embargo ushered in a new pattern of cyclical behavior of prices and inflation. Since then, inflation has often tended to be higher in recession periods than in expansions in many countries. As shown in Romer's Table 5.3, the overall correlation during the 1947–2004 period is slightly positive (inflation declines in recessions on average), supporting a procyclical inflation rate. However, the right-hand column shows that inflation has actually declined in only four of ten postwar recessions.

The cyclicality of inflation is important because it is a prediction on which some theories of business cycles disagree. Thus, clear evidence of either procyclical or countercyclical inflation might allow one set of theories to be rejected in favor of the other. However, the conclusion that inflation is procyclical is quite sensitive to the time period, country, and method chosen for the analysis. Because the evidence is unclear, it has merely served to expand the focus of the debate to include the proper interpretations of empirical observations as well as the competing theoretical models themselves.

Romer's table suggests that real wages are slightly procyclical, which is consistent with the conventional view. However, some microeconomic studies have found evidence of countercyclical real wages for some samples. Again, the cyclicality of wages is an important point that might allow us to discriminate among theories. While the bulk of the evidence supports an acyclical or a weakly procyclical real wage, there is sufficient disagreement to allow competing theories to claim validation.

On average, nominal interest rates have tended to fall in U.S. recessions, as has the real money stock. This finding for the nominal interest rate is robust across other economies, though the real money stock is less cyclical in some economies than in the United States.

C. Real vs. Keynesian Interpretations of Cycles

The most active question of investigation in recent empirical business-cycle analysis has been the relative importance of aggregate demand shocks and technology shocks as a source of fluctuations. The motivation for this question is the issue of whether Keynesian or real-business-cycle models provide the more relevant description of cycles. Because of the widely different implications of the two theories for macroeconomic policy, the answer to this question matters a great deal. This section lays out the basic empirical cases for RBC and Keynesian models; subsequent sections describe research strategies that have been used to assess the relative importance of the two models.

The basic case for RBC models

As discussed in Chapter 7, the empirical case for the real-business-cycle theory was initially presented in terms of calibrated simulations rather than econometric models or statistical tests. Based on estimates of fundamental behavioral parameters from external (non-macroeconomic) sources, analysts simulate the behavior of the RBC model under alternative patterns of shocks. They then compare the properties of the resulting simulated business cycles with those of actual cycles. To the extent that the simulations mimic the cyclical properties of actual time series, success is claimed for the RBC model.

There have been many dozens of empirical applications of RBC models in the last 25 years. We consider here the study that is generally recognized as the earliest published paper in the RBC literature, Kydland and Prescott (1982). These authors shared the Nobel Prize in 2004 for this work and the literature that followed. It is representative of the empirical successes and shortcomings of RBC models. Kydland and Prescott built a dynamic RBC model in which investment projects require several periods of construction time before they become productive. They imposed estimates of some parameters from microeconomic studies or from general economy-wide observations such as labor's share of output. Other parameters were chosen by exploration of alternative possibilities and examination of the implied results.

Kydland and Prescott report three sets of simulated results and compare them with actual, quarterly U.S. data from 1950 to 1979. Table 1 shows some of the results they report for autocorrelations of real output, standard deviations of variables, and correlation of other variables with real output.[3]

Table 1. Kydland and Prescott's empirical results

	Autocorrelations of output			
Lag	Model		Actual	
1	0.71		0.84	
2	0.45		0.57	
3	0.28		0.27	
4	0.19		–0.01	
5	0.02		–0.20	
6	–0.13		–0.30	
	Standard deviations		Correlations with output	
Variable	Model	Actual	Model	Actual
Real output	1.8	1.8		
Consumption	1.3	0.6	0.74	0.94
Investment	6.5	5.1	0.80	0.71
Hours worked	1.1	2.0	0.93	0.85
Productivity	0.9	1.0	0.90	0.10

The results reported in Table 1 demonstrate that a suitably calibrated RBC model is capable of generating cyclical fluctuations that capture key features of the U.S. economy. The top section shows a realistic degree of persistence in real output fluctuations.[4] The second section shows that the pattern of variability and covariation with output in the RBC model are broadly similar to real values. The biggest deviation from reality in Table 1 is the behavior of hours worked and productivity. Hours worked do not vary as much in the model as they do in real life, while productivity is far too strongly correlated with real output.

A second paper summarizing the basic empirical case for the RBC models is Plosser (1989), which we read earlier in the course. Plosser's method of generating shocks was quite different, though no less controversial than Kydland and Prescott's.

[3] The statistics in Table 1 are for the cyclical component of the series, with detrending based on the Hodrick-Prescott filter.

[4] It should be noted, however, that considerable persistence was "built into" the model, A major component of the productivity shock followed a first-order autoregressive process with parameter 0.95, as in Romer's Chapter 5 model.

Rather than generating repeated random shocks imposing a pattern of strong autocorrelation, Plosser used estimated Solow residuals for the U.S. economy as his shocks.

Plosser's Table 1 is typical of the results of RBC models. This is reproduced below as our Table 2. Although the literature is large and the details of the results vary from study to study, those presented in Table 2, like those of Kydland and Prescott discussed above, are representative of typical RBC model outcomes.

Table 2. Plosser's simulation results.

Variable	Mean	Standard deviation	Autocorrelations ρ_1	ρ_2	ρ_3	Corr w/ output	Corr w/ actual
Actual U.S. Annual Data							
$\Delta\log(Y)$	1.55	2.71	0.13	–0.17	–0.16	1.00	1.00
$\Delta\log(C)$	1.56	1.27	0.39	0.08	0.05	0.78	1.00
$\Delta\log(I)$	2.59	6.09	0.14	–0.28	–0.19	0.92	1.00
$\Delta\log(L)$	–0.09	2.18	0.17	–0.32	–0.24	0.81	1.00
$\Delta\log(w)$	0.98	1.80	0.44	–0.16	–0.08	0.59	1.00
Simulated Predictions from Plosser's RBC Model							
$\Delta\log(Y)$	1.56	2.48	0.30	0.18	0.14	1.00	0.87
$\Delta\log(C)$	1.65	1.68	0.55	0.44	0.37	0.96	0.76
$\Delta\log(I)$	1.37	4.65	0.14	0.00	–0.02	0.97	0.72
$\Delta\log(L)$	–0.08	0.89	0.07	–0.09	–0.12	0.87	0.52
$\Delta\log(w)$	1.64	1.76	0.51	0.40	0.33	0.97	0.65

How similar are the simulated predictions from Plosser's model to the actual U.S. values? Is the glass half full or half empty? Most of the basic qualitative characteristics of business cycles seem to be captured by the simulation results. Most growth rates are positively autocorrelated at the first order, meaning that a high value of the variable in period t tends to be associated with a high value in period $t + 1$. The relative sizes of the means and variances of the growth rates are pretty similar. All are strongly positively correlated with the actual values they are attempting to simulate and all variables are procyclical, as they are in the actual data.

However, there are also some significant differences between the top and bottom halves of Table 2. These discrepancies are common in basic RBC models and mirror those pointed out above in Kydland and Prescott's results, showing aspects of the macroeconomy that the models are not very successful in capturing. For example, look at the behavior of $\Delta\log(L)$, the growth rate of employment. The standard deviation in the simulations is less than half of the actual standard deviation, indicating that the RBC model predicts much less variation in employment growth over the business cycle.

Moreover, aside from being too small, the movements in employment growth predicted by the model are not always in the right direction and at the right time, which is indicated by the correlation coefficient of only 0.52 between actual and predicted. This result means that the RBC assumption that labor markets clear continuously and that fluctuations in employment result from intertemporal substitution in labor supply does not do a very good job of explaining actual employment fluctuations.

A second discrepancy is in the cyclical behavior of real wages. The RBC model predicts a correlation between wage growth and output growth of 0.97—almost perfect correlation. The actual data show a correlation of 0.59, which implies that there is a much weaker association than the RBC model predicts. Another difficulty is the mean and standard deviation of investment growth, both of which are much smaller than the actual values. Finally, consumption seems to vary too much in the RBC models and too closely with output, perhaps as a result of the assumption that changes in output growth are due to permanent technology shocks that have a predicted MPC of one.

The bottom line on these simulations is that while they produce business cycles that resemble the broad outlines of actual cycles, the congruency is not sufficient to accept the current versions of these models as a full explanation of economic fluctuations. The response of RBC modelers to these shortcomings is to suggest ways that their models can be improved to try to capture the effects they are missing. Plosser does this in the final section of his paper entitled "The Real Business Cycle Research Agenda." While much progress has been made in the decade since Plosser's article was published, the deficiencies of RBC models have proven to be difficult to resolve.

Productivity shocks, wages, and labor input

The most discordant aspect of the business cycle for RBC models seems to be the labor market. The RBC model associates (marginal) productivity with real wages (which are only barely procyclical) and assumes that workers are always on their labor supply curves (there is no involuntary unemployment). Thus, if employment is strongly cyclical but productivity does not move as strongly with output, then labor supply must be extremely sensitive to real wage movements. This contradicts a broad consensus of empirical evidence suggesting that aggregate labor supply is quite inelastic with respect to wages. Much of the subsequent refinement of RBC models has focused on developing alternative models of the labor market that attempt to resolve this inconsistency.[5]

One frequently cited model is that of Christiano and Eichenbaum (1993), which includes two variations on the standard RBC framework. First, as in the model in Romer's Chapter 5, Christiano and Eichenbaum incorporate government spending shocks in addition to productivity shocks. Second, they include indivisibilities in labor

[5] A readable summary of some of this early work is in Hansen and Wright (1992).

supply, so that workers cannot vary their hours of work continuously. With these two changes, Christiano and Eichenbaum's model exhibits realistically large variability in labor input relative to the amount of variability in labor productivity. However, the correlation between labor input and productivity (real wages) is still (counterfactually) strongly positive in this model.

Galí (1999) shows that while the Christiano and Eichenbaum model is able to generate realistic *unconditional* variances and (in some cases) correlations among the variables, the *conditional* variances and correlations are still unrealistic. He uses a vector autoregression technique (discussed in a subsequent section) to decompose the overall variation and correlations into those arising from productivity shocks and those arising from other kinds of shocks (government-spending shocks in the Christiano and Eichenbaum model). He finds that the estimated effects of productivity shocks in the U.S. and other economies do not correspond to those predicted by the Christiano and Eichenbaum model.

Among Galí's results is the controversial finding that the immediate effect of a productivity shock on labor input is *negative*. This effect is strongly inconsistent with the RBC model, but Galí shows that it is consistent with a new Keynesian model with price stickiness.[6] Attempts to identify productivity shocks statistically from vector autoregressions require strong assumptions and are always, therefore, subject to criticism from those who take issue with the particular set of assumptions used. However, Shea (1998) finds a similar result using data on patents and R&D expenditures to measure productivity shocks directly. Although one can argue that changes in patents and R&D are at best imperfect measures of productivity shocks, the consistency of this result with Galí's econometric evidence has established a serious channel of empirical doubt on the RBC framework. This has been an active area of recent empirical research on which no consensus yet exists. For example, Francis and Ramey (2005) show that Galí's result can be explained by RBC models that include measures to reduce the speed of adjustment of expenditures, such as "habit-formation" models of consumption and adjustment costs to investment.

Microeconomic evidence on intertemporal substitution

When most people contemplate their own labor-supply decisions, they do not seem to attribute a very large role to considerations of intertemporal substitution. Apart from extreme examples such as a million-dollar-per-hour wage, intertemporal wage differences and the level of real interest rates do not seem too important in deciding how many hours to work.

---

[6] The new-Keynesian story is that output does not respond immediately to a productivity shock because it is determined by aggregate demand. So when productivity increases, firms are able to lay off workers and still produce the amount of output demanded.

However, it is very difficult to measure these effects empirically in order to know whether people's *actual* labor-supply behavior (as opposed to their *self-perceived* behavior) responds to these intertemporal tradeoffs. Aggregate data confound the effects of permanent and temporary wage changes and are averaged across huge numbers of people. The individual data that we have from standard sources such as the Current Population Survey and the Panel Study on Income Dynamics also fail to provide data where we can document the effects of differentials between the current wage and the expected future wage. This absence of clear empirical evidence has left the argument unsettled between those who argue that intertemporal substitution is important and those who follow the conventional interpretation and believe that these factors have at most minor effects on labor supply.

An interesting piece of evidence from a unique data source is analyzed by Camerer et al. (1997). They argue that the nature of cabdrivers' pay system allows one to observe with great clarity the effects of temporary wage changes on a day-by-day basis.

Cabdrivers can (and do) vary their work hours quite flexibly on any given working day, so they have a good opportunity to perform intertemporal substitution in response to wage differentials. Such wage differentials also arise with great frequency because the driver's hourly earnings depend mainly on how easy or difficult it is to find fare-paying passengers. Fluctuations in the weather, the day of the week, and the timing of special events cause the wage to vary a lot from day to day, so we can observe how much drivers substitute intertemporally in response to these changes.

Camerer et al. look at these responses in detail and find that drivers, especially relatively inexperienced drivers, actually behave in a completely opposite way. On days when passengers are plentiful and hence hourly earnings are high, drivers tend to work *shorter* hours than on less profitable days. For some samples, the elasticity of labor supply with respect to daily wage changes is near –1, which is consistent with a daily income target under which drivers work until they earn a certain target amount of money and then quit for the day.

Such income-target behavior could be consistent with utility maximization over a longer time horizon if utility is a highly convex function of income. However, it seems implausible that this behavior should occur for one-day changes in the wage. The authors estimate that even with a zero supply elasticity—which corresponds to working a fixed number of hours regardless of the wage—drivers would earn on average 5 percent more than they do by following their estimated behavior pattern. It is hard to imagine that the increase in income of 5 percent coupled with a more even distribution of work and leisure across days would not increase utility. Thus, Camerer and his coauthors conclude that the evidence casts serious doubt on whether the conventional intertemporal utility-maximization model is realistic one for this group of workers.

However, Farber (2005) finds opposite results using Camerer's own data. He uses a different econometric methodology and comes to the conclusion that the large income effects that Camerer et al. observe are not present. He finds that a driver's stopping decision depends mainly on the number of hours he or she has driven during the shift and that there is little effect of earnings.

Direct evidence that supply shocks cause cycles

The plausibility of supply (or technology) shocks as a potential explanation for business cycles gained prominence as a result of the recessions that followed the Arab oil embargo of 1973 and the subsequent rises in oil prices at the end of that decade. James Hamilton (1983) went back to earlier recessions in the postwar period and discovered that all but one recession since World War II had been immediately preceded by a significant rise in the price of oil.

Hamilton's Figure 1 shows the coincidence of oil-price rises and recessions from 1948 through 1975. This finding was quite remarkable because no one had thought about the rise in oil prices as a cause of recessions in these earlier periods and in fact no economists even realized until Hamilton's study that price increases had occurred at those times.

However, Hamilton is careful to point out (top of page 230) that the proximity in time of increases in the oil price and recessions could be explained by any of three hypotheses. One is random coincidence; a second is that another variable may influence both oil prices and output; the third is that oil-price increases may cause recessions. Traditional statistical analysis is ideally suited to testing the first hypothesis. Tests of the statistical significance of a correlation coefficient are designed to determine how likely it is that the observed coefficient would have occurred by chance if the variables are truly unrelated. Hamilton found that the coincidence was too strong and regular to be explained by random chance.

Statistical testing for correlation between two variables is quite easy; determining the direction(s) of causality between them is much more difficult. All that statistics can demonstrate is that two things consistently happen at the same time. Other considerations, such as economic theory, must usually be used to assess what causes the events to occur together. However, there is a technique called ***Granger causality*** that helps us to assess whether past values of one variable help to predict current and future values of another. Although it can be dangerous to infer the direction of causality between two variables from the timing of their movements, this is the closest that macroeconomists have come to an acceptable test of causality.[7]

[7] Rational expectations provides a good example of why this inference can be incorrect. Suppose that the Federal Reserve pre-announces a sudden increase in the money supply and that

Tests of Granger causality are usually done in the way that Hamilton describes on pages 231 and 232. His equation (1) is a regression equation that attempts to predict the current value of z based on its own past values and past values of x. Granger causality tests the null hypothesis that all of the b coefficients on the lagged x values are zero. If this were the case, then x would not improve the prediction of z_t relative to the prediction that can be made solely on the basis of past values of z. If this null hypothesis is true, then we says that x does not Granger cause z. If the null hypothesis is rejected—meaning that we conclude that lagged x values *do* help to predict z_t—then x does Granger cause z.

The statistical test that is used for Granger causality is an F test, which examines whether the closeness of the prediction of z_t in Hamilton's equation (1) is significantly worse when the x_{t-i} terms are omitted (the b_i coefficients are constrained to zero). If the estimated b_i coefficients are all exactly zero (which would happen with near-zero probability), then the calculated F statistic would be zero and we would accept the null hypothesis. The extent to which the F statistic is larger than zero measures the degree to which including the x_{t-i} terms (with nonzero coefficients) improves the prediction of z_t. If the F statistic if large enough to exceed the critical value for the test at some given significance level (often 5%), then we reject the null hypothesis and conclude that x Granger causes z.

In order to investigate the second of his three hypotheses—that some third variable is causing changes in both oil prices and GDP—Hamilton's Table 2 presents Granger causality tests of whether six macroeconomic variables affect oil prices and vice versa. The F statistic is shown in the left column of each section of the table and the associated p value in the right section. The F statistic measures the degree to which the prediction of the z variable is improved *in the present sample* by including the x values. The p value (for "probability value") is the probability that this much improvement in prediction would occur if the lagged x values actually had no effect and only random chance led us to see an improved prediction. Using the conventional significance level of 5%, we reject the null hypothesis of no Granger causality if the p value is smaller than 0.05.

Based on the tests in his Table 2, Hamilton was unable to find any evidence that a third variable had caused oil prices, but he did find strong evidence that oil prices caused output and unemployment in his sample. Thus, he rejected his second hypothesis in favor of the third; he concluded that oil prices were a major cause of business cycles in the postwar period.

---

interest rates fall as a result. The change in interest rates would precede the actual change in money supply even though the causality ran the other direction.

Hamilton's statistical sample ends in the 1972. The October 1996 issue of *Journal of Monetary Economics* published an update of Hamilton's study on oil prices and recession by Mark A. Hooker, along with a response from Hamilton and a further reply from Hooker.[8] When he extended Hamilton's study through 1994 using quarterly data, Hooker found that oil prices had a much weaker Granger causal effect on real variables and for the 1973–94 sub-sample there was no causal effect at all.

Hooker considers three hypotheses for why the empirical structure of the oil-price-macroeconomy relationship changed in the 1970s. The first is that the overall behavior of macroeconomic variables was simply different after 1973. Second, he considers whether oil prices became endogenous after 1973, responding to changes in the U.S. macroeconomy—in other words, whether macroeconomic variables Granger caused oil prices in the later sample. Finally, he questions whether increases in oil prices (as occurred regularly in the earlier sample) have effects on the macroeconomy that are asymmetric to those of oil-price declines such as those of the 1980s.

Hooker finds considerable evidence that the macroeconomy was structurally different after 1973, which corroborates a sizable empirical literature indicating that major macroeconomic relationships changed at that time. He does not, however, find evidence to support the second hypothesis: oil prices are still not Granger caused by other variables after 1973. Nor does asymmetry between the effects of increases and decreases in oil price seem to be important.

Based on his estimates, Hooker concludes that the extension of the sample period into the 1990s significantly weakens the case for oil prices as a regular source of business-cycle fluctuations. He does not question, however, the importance of the 1973 OPEC embargo and price increase as a trigger for the ensuing recession. Moreover, it would be a remarkable coincidence if the events in the oil market in 1973 were not at least partially responsible for the structural changes in the macroeconomy that Hooker finds.

A further skeptical view is offered by Barsky and Kilian (2004), who argue that the evidence for oil-price effects arising from political events is weak even in the 1970s, where the consensus opinion has emphasized the importance of oil shocks.

The basic case for Keynesian models

The empirical case for Keynesian models evolved in the thirty years following World War II as economists utilized the newly developed data on national income and product to estimate key relationships of the Keynesian system such as the consumption function, investment function, demand function for money, and the Phillips

[8] See Hooker (1996), Hamilton (1996), and Hooker (1996).

curve. These estimated equations were often combined into macroeconometric simulation models and simulated to estimate the properties of the macroeconomy.[9]

This model-building strategy reached its apex around 1970. After 1970, instability in several equations of the model (notably the Phillips curve and the money-demand equation) diminished the performance of the models. Theoretical criticism (such as Friedman's famous presidential lecture on the Phillips curve and Lucas's microeconomic approach to modeling) also caused economists to question the validity of this approach. Despite these difficulties, most macroeconomists retain a belief in the short-run importance of demand shocks, following a mechanism similar to the one described in the early Keynesian macro models.

Among the first macroeconometric models was the Klein-Goldberger model.[10] Given the computational technology available to them, the successful estimation and simulation of even their simple macroeconomic model was a considerable feat. The model itself was a stylized version of the *IS/LM* framework, with a strong emphasis on spending as the determinant of output and on the determination of fiscal policy multipliers.

Klein and Goldberger's equations, like those of later macro models, seemed to fit the data closely.[11] The experiments that they and others performed with their models predicted realistic effects for spending shocks and for monetary and fiscal policy changes. The result of this apparently successful modeling activity was a growing consensus that aggregate demand has a strong effect on real variables such as output and interest rates, and that demand shocks are the dominant cause of business cycles.

A second influential study that argues for strong, short-run real effects of monetary shocks is Friedman and Schwartz (1963). Friedman and Schwartz's book is a monumental contribution to monetary economics in several dimensions. They used records of banks and government agencies to construct measures of the money supply and reserves going back into the 19th century. Based on these measures and on a richly detailed analysis of contemporary writings, they meticulously document numerous episodes where disturbances emanating from the monetary sector were followed by recessions in real macroeconomic activity.

---

[9] A collection of papers describing many of the more prominent macro models of the day is Klein and Burmeister (1974).

[10] See Klein and Goldberger (1955).

[11] In retrospect, we now recognize that the goodness of fit was often overstated due to the "spurious regression" phenomenon. Any time that highly trended time series are regressed on one another, they are likely to appear correlated even if the underlying changes in the series are unrelated.

The detailed analysis of movements in monetary measures over nearly a century of highly varied U.S. history and comparison of these movements with the contemporaneous changes in economic variables lead Friedman and Schwartz to some broad conclusions about the effects of money on the economy. In their final chapter, they summarize their findings as follows:

> 1. Changes in the behavior of the money stock have been closely associated with changes in economic activity, money income, and prices.
> 2. The interrelation between monetary and economic change has been highly stable.
> 3. Monetary changes have often had an independent origin; they have not been simply a reflection of changes in economic activity.
> 4. In monetary matters, appearances are deceiving; the important relationships are often precisely the reverse of those that strike the eye.[12]

Both the Keynesian tradition embodied in macroeconometric models and the monetarist school represented by Friedman and Schwartz predicted that shocks to the money supply, and to aggregate demand more broadly, would have strong effects on real variables, at least in the short run. It is this broad consensus that the more avid proponents of the RBC approach have attacked.

Direct evidence that demand shocks cause cycles

There are many empirical approaches that have found significant effects of money and other aggregate-demand variables on real output. We shall examine two recent studies here. The first is a survey by Christiano, Eichenbaum, and Evans (1999) (hereafter CEE), which describes and updates the literature using vector autoregressions (VARs) to examine the effects of monetary policy.

The use of VARs has become widespread since they were introduced by Christopher Sims (1980).[13] We will not elaborate the details of the VAR approach here, except to introduce the most basic idea. A VAR is a set of regressions in which the current value of *each* of a set of variables is regressed on lagged values of *all* of the variables. In specifying the VAR, one must decide which variables should be in the set and how many lags should be included. For example, a VAR involving two variables, output and money, with two lags, would look like equation (1).

[12] Friedman and Schwartz, p. 676.

[13] A good introduction to the use of VARs is in Chapter 5 of Enders (1995).

$$Y_t = \alpha_Y + \beta_{YY1}Y_{t-1} + \beta_{YY2}Y_{t-2} + \beta_{YM1}M_{t-1} + \beta_{YM2}M_{t-2} + \eta_{Yt},$$
$$M_t = \alpha_M + \beta_{MY1}Y_{t-1} + \beta_{MY2}Y_{t-2} + \beta_{MM1}M_{t-1} + \beta_{MM2}M_{t-2} + \eta_{Mt}. \tag{1}$$

Because there are no current values of either variable on the right-hand side of (1), the usual difficulties of simultaneous equations do not affect the estimation of the α and β parameters.

The results of a VAR are usually presented in one or both of two forms: *impulse response functions* (how one variable affects all others) and *variance decompositions* (how one variable is affected by all others). Impulse response functions are estimates of the dynamic effects of a shock to one variable on the future time paths of all of the variables of the system. Variance decompositions measure how much of the variation in one variable is caused by past shocks (at various distances in the past) to itself and the other variables in the system.

Although the estimation of the parameters is straightforward in a VAR, one must make often-controversial *identifying assumptions* in order to identify the shocks associated with each variable and calculate impulse responses or variance decompositions. For example, in system (1), it is tempting to call the error term η_{Mt} a monetary shock and η_{Yt} an output shock. If the two error terms were independent of one another, this would be reasonable. However, these error terms are almost always correlated in practice, indicating that unexpected changes in current output and current money are related.

If one is willing to make the identifying assumption that output does not affect money within the current period, then one can interpret the relationship between current money and output as money causing output. In this case, η_{Mt} is a pure monetary shock because output shocks are assumed not to affect money immediately and η_{Yt} is a combination of an output shock and the effects on current output of the money shock.

However, one could alternatively assume that money does not affect output within a single period, which means that the relationship between the current values is output causing money. In this case, η_{Yt} is a pure output shock and η_{Mt} combines the monetary shock with the contemporaneous effect of output on money.

The two alternative identifying assumptions lead to two different sets of "money shocks," so the implied effects of monetary policy can be quite different. The large literature that uses VARs to estimate the effects of monetary shocks on the economy reflects a wide variation in the choice of which variable is to represent monetary policy and in the associated identifying assumptions. CEE survey this literature in detail. We shall examine only one set of results that they generate.

Current Federal Reserve policy is formulated in terms of a target level for the federal funds interest rate. The VAR system we examine uses this rate as the monetary policy variable. Figure 1 shows their estimated responses of output over time to an increase in the federal funds rate (a contractionary monetary policy shock). The dotted band indicates standard errors around the estimated effect.

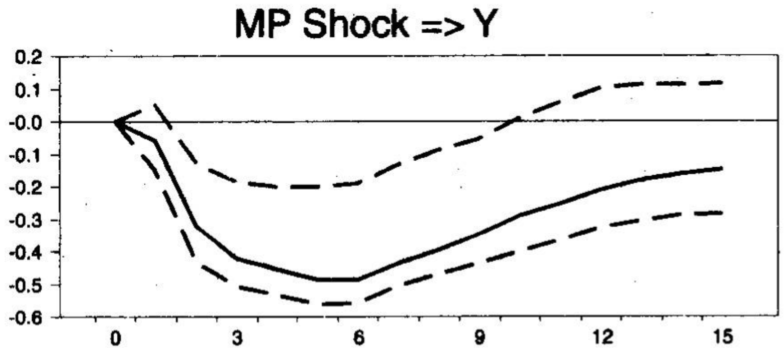

Figure 1. Christiano, Eichenbaum, and Evans's effect of funds rate on output.

Figure 1 shows that output decreases relatively quickly after a monetary contraction, but that the effect begins to die out after about 6 quarters. By the end of the period shown (15 quarters), the output effects of the monetary shock have largely died away. This is very much consistent with the aggregate demand effects predicted by modern Keynesian models.

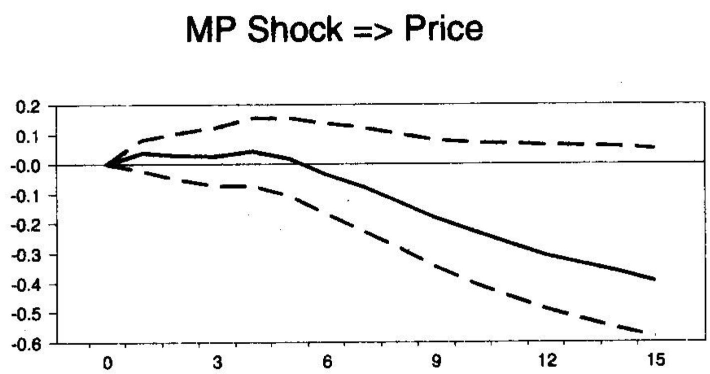

Figure 2. CEE's estimated dynamic effect of funds rate on prices

The effect on prices shown in Figure 2 is also consistent with Keynesian models. The general price level does not begin to react until 5 to 6 quarters after the monetary contraction—about the time that the output effects reach their maximum. Prices then begin to fall relative to their expected path in response to the low level of real output.

The evidence from CEE can be viewed as a modern version of econometric evidence that monetary policy (i.e., aggregate demand) affects real output. However, the uncertainty about the appropriate identifying assumption makes any VAR result open to question. An alternative strategy for identifying monetary shocks was used by Christina and David Romer.

Romer and Romer (1989) identified monetary shocks by what they called the "narrative approach." They examined the detailed minutes of Federal Open Market Committee (FOMC) meetings in the postwar era and identified six dates (now commonly called "Romer dates") on which the Federal Reserve intentionally pursued contractionary monetary policy. They then examined the behavior of industrial production and unemployment after each of the six Romer dates relative to a forecast based on the variable's own history up to that date.

They found that industrial production was substantially lower and unemployment substantially higher in the months after Romer dates than one would have predicted. They attribute this result to the real effects of monetary contraction on output and unemployment. This provides further evidence, independent of econometric methods of identifying monetary shocks, that aggregate demand shocks have real effects in the short run.

D. Are Output Shocks Permanent?

Both RBC advocates and most modern Keynesians agree that aggregate demand shocks should not have significant *permanent* effects on real output, whereas supply shocks may have permanent effects on the output growth path. This suggests an empirical research strategy of testing whether output shocks have temporary or permanent effects on the output growth path. An empirical finding that most output disturbances are temporary deviations from a fixed growth path could be interpreted as supporting demand-based explanations, while a finding that output shocks tend to have permanent effects on the path would indicate that such shocks are more likely supply/productivity disturbances.

Nelson and Plosser's test for reversion to trend

Nelson and Plosser (1982) examined U. S. historical data using the Dickey-Fuller test to assess whether output is better characterized by a trend-stationary process (with temporary cyclical fluctuations) or a difference-stationary process (with fluctuations in output not reverting to a fixed trend). The Dickey-Fuller test examines whether large positive (negative) deviations from the predicted trend tend to cause negative (positive) future changes that would bring the series back to the trend. The null hypothesis of the test is that the series *does not* revert to a fixed trend, *i.e.*, that it is difference stationary.

Only if there is sufficient evidence to be 95% certain of trend reversion will this null be rejected in favor of the alternative hypothesis of trend stationarity.[14]

Nelson and Plosser find insufficient evidence of reversion to trend to reject the difference-stationarity hypothesis. Thus, they conclude that output fluctuations in the United States have tended to be permanent, supporting the real-business-cycle theory. This result sparked a large literature testing for a unit root (*i.e.*, difference stationarity) in output, using the Dickey-Fuller test and other tests. The Dickey-Fuller test and related tests have been criticized for having low power in discriminating between difference-stationary and trend-stationary processes, so some economists have sought alternative methods of testing the permanence of output fluctuations.

Studies not based on unit root tests

We shall examine two such tests, one by Campbell and Mankiw (1987) and one by Cochrane (1988). Campbell and Mankiw addressed the question of whether current changes in output lead to changes of in optimal long-run forecasts of future output levels. They estimate a variety of statistical forecasting models for output using quarterly data from 1947 through 1985 then calculate how the implied forecasts for real output many quarters ahead would respond to a unit change in current output.

For the vast majority of these models (13 out of 15 specifications), a one-unit increase in current output raises the optimal 10-year and 20-year forecast by *more than one unit*. This implies that output fluctuations are not only permanent, but that they tend to be followed by future permanent fluctuations in the same direction. Like Nelson and Plosser, Campbell and Mankiw argue that these results could be interpreted as supporting the real-business-cycle theory. However, they also suggest that alternative models of permanent output shocks, including possible permanent effects of demand-based shocks, could also be consistent with these results.

Rather than testing an "either/or" model of temporary vs. permanent changes in output, Cochrane examines a model in which output is subject to both permanent and transitory shocks. He then estimates the relative importance of the two kinds of shocks using the behavior of "long differences," which are the changes in output between two points in time that are far apart. For a pure random walk, in which all changes are permanent, the variance of $x_t + x_{t-k}$ increases linearly with the lag length k. In contrast, for a purely stationary series (or one with stationary deviations from a fixed trend), the variance of the k-difference levels off for large k.

Cochrane's results, based on a sample of U.S. annual data starting in 1869, suggests that temporary shocks account for about two-thirds of the variation in GNP, with the permanent, random-walk component accounting for one-third. In contrast to the

[14] The common technical term for this hypothesis is the "unit-root hypothesis." If a time-series process such as real output has a "unit root," then it is difference stationary.

studies finding strong support for permanent shocks, this study suggests that aggregate-demand fluctuations may be an important source of fluctuations. However, because technology shocks in the RBC model can be either permanent or transitory, an important transitory component cannot be interpreted as evidence against it.

E. Interpreting the Cyclicality of Prices and Inflation

The simple, textbook model of aggregate supply and aggregate demand suggests that the cyclical behavior of prices could help distinguish between supply- and demand-induced fluctuations. Shocks to aggregate supply should cause prices and output to move in opposite directions, while demand shocks should induce positively correlated movements.

If there were a clear procyclicality or countercyclicality of prices and inflation, this could give a decisive indication of the predominant source of cycles. However, in practice, the cyclical behavior of prices varies over time and depending on how one specifies the variables.

Table 3. Stock and Watson's cross correlations between output and prices/inflation.

	Lead or lag relative to real GDP												
	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6
GDP def.	.12	.12	−.02	−.18	−.33	−.46	−.54	−.60	−.61	−.59	−.52	−.42	−.30
GDP infl.	.45	.55	.61	.58	.48	.32	.15	−.01	−.14	−.25	−.34	−.41	−.47
CPI level	.34	.24	.12	−.04	−.21	−.38	−.51	−.62	−.68	−.67	−.59	−.48	−.34
CPI infl.	.34	.47	.58	.64	.62	.52	.35	.14	−.08	−.27	−.40	−.48	−.51

Table 3 shows the cross correlations between U.S. postwar quarterly price and inflation variables and real output, calculated by Stock and Watson based on cyclical components of the series extracted with the bandpass filter.[15] Stock and Watson's evidence suggests a striking difference in cyclical behavior between the price level and the inflation rate, regardless of whether prices are measured by the GDP deflator or the consumer price index. The price level seems to be strongly countercyclical, with a lead of approximately two quarters relative to GDP. However, the inflation rate seems to be procyclical with a lag of 3 to 4 quarters. Moreover, the evidence of countercyclical

[15] The bandpass filter uses spectral analysis do decompose output fluctuations into regular components having various frequencies, then filters out the high-frequency (less than 6 quarter wave length) and low frequency (more the 32 quarters), leaving only the cyclical components of business-cycle frequency. See Baxter and King (1999) for details.

price level is strongest for the period since 1970. Prior to this date, the price level seemed to be procyclical.

It is difficult to interpret this apparent contradiction between the behavior of prices and inflation. Macroeconomic theories rarely distinguish dynamics in sufficient detail to predict different behavior of prices and their rate of change. Moreover, the empirical detrending of the data has no direct correspondence to the dynamic structure of theoretical models. Using the basic results, RBC proponents cite the apparent countercyclical behavior of the price level in recent cycles as support for supply-induced cycles. Keynesians claim that the procyclicality of inflation supports demand-based fluctuations.

Ball and Mankiw (1994) present a Keynesian explanation of the countercyclical price level since 1973. First, they accept that some business cycles (those of the mid-1970s and early 1990s) may have supply-side origins, due to changes in oil prices. However, they argue that the largest cyclical shock in the period, the recession of the early 1980s, was due to contractionary monetary policy. They reconcile the apparent countercyclical behavior of prices in this period to the detrending of the data.

Suppose that the successful disinflation engineered by the Fed in the early 1980s reduced the trend rate of inflation. The actual path of prices might look like the bold line in Figure 3. The period where the underlying inflation rate declines is the marked recession period at the cusp of the actual price line. Notice that conventional detrending shows the price level as being considerably *above* its trend during the recession, which would show up as countercyclical price behavior. Although most studies use more sophisticated detrending methods such as the Hodrick-Prescott filter or the band-pass filter, they are still likely to come to the same conclusion, even if the disinflation is caused entirely by aggregate demand (monetary) shocks.

The example of Figure 3 shows that aggregate-demand shocks that have permanent effects on the inflation rate can have countercyclical effects on detrended prices, but procyclical effects on inflation—inflation falls during recessions. This helps reconcile the apparent countercyclicality of post-1970 prices with demand-driven business cycles, and with the evidence of procyclical prices from earlier historical periods.

Given the ambiguity of evidence and the possibility of alternative interpretations, it appears unlikely that the cyclical behavior of prices and inflation can resolve the relative importance of supply shocks and demand shocks in business cycles.

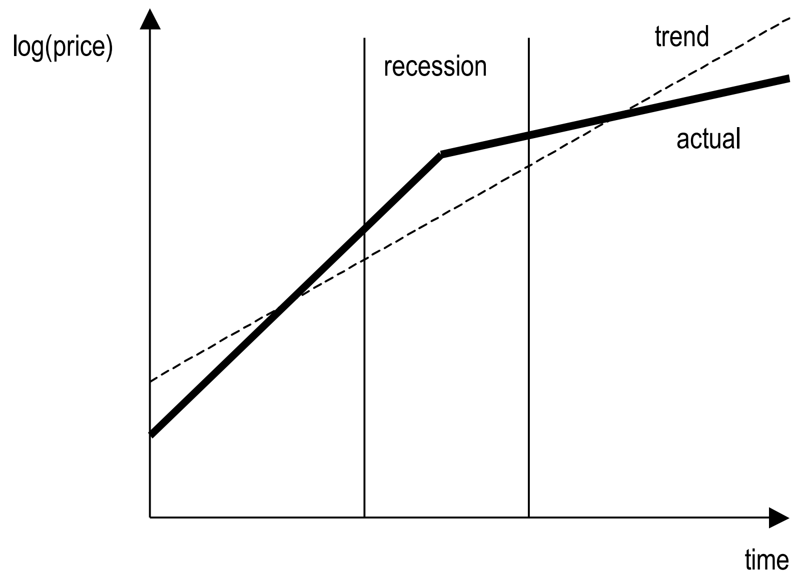

Figure 3. Ball and Mankiw's interpretation of countercyclical prices.

F. Interpreting Procyclical Labor Productivity

Most studies find that labor productivity (output divided by labor input) is procyclical. Since the RBC model explains booms as positive shocks to productivity, this evidence provides strong empirical support. Some variants of the Keynesian framework, on the other hand, suggest that labor productivity should be countercyclical. If employment rises in a boom, with no change in the production function, then the marginal (and average) product of labor should decline, making labor productivity (and the real wage) countercyclical. Once again, however, there are several possible explanations that can reconcile the procyclical behavior of productivity with a Keynesian explanation of business cycles.

Explaining procyclical productivity without technology shocks

As noted above, the procyclical behavior of multifactor productivity is a principal underpinning of the real-business-cycle model. On first examination, this seems to require a supply-based explanation of business cycles such as technology shocks. If technology is fixed and business cycles are caused by demand shifts, then the decline in employment should *increase* the marginal and average product of labor in a recession, not *decrease* it. In other words, productivity (and presumably real wages) should be countercyclical.

We examine in the following sections several alternative arguments that have been advanced to explain why total factor productivity and labor productivity might be procyclical. These alternatives do not rely on changes in the production function or in technology. Instead they focus on the nature of the production function itself and on whether the measured relationship between inputs and outputs always represents the actual production function. The specific explanations that we shall consider are:

- **Increasing returns to scale.** If returns to scale are increasing, then expansion of output (in booms) lowers costs and leads to greater efficiency. The Solow residual in would rise with output because inputs do not need to increase as much as output when there are increasing returns.

- **Labor hoarding.** If firms keep workers on the payroll in recessions even though they are not producing at their normal rate, then measured labor input will not fall with output, even though actual work effort does decline. This will show up as a reduction in measured productivity.

- **Variations in the rate of utilization of the capital stock**. If the capital stock is used more intensively in booms than in recessions, then capital input is also mismeasured. Because changes in measured capital input do not change with changes in utilization, but output does change, the Solow residual will pick up these procyclical variations in utilization.

Increasing returns to scale and procyclical productivity

We discussed issues related to returns to scale quite extensively in our analysis of modern growth models (see Chapter 5). When we examined the theoretical rationale for constant, increasing, or decreasing returns in that context, we often used the "replication principle" as a reference point. According to this idea, each plant or firm should expand to the most efficient (lowest average cost) scale of operation. The size of the market then determines how many such plants or firms can be supported, with each plant operating at minimum average cost. Returns to scale are constant because expansion or contraction of total industry output simply means increasing or reducing the number of plants in operation with no change in average cost.

There are several circumstances in which the replication principle may not hold. One is the case of ***external economies***, where returns to scale are constant within the firm, but expansion of the industry leads to cost reductions. In our growth models, we introduced publicly accessible "knowledge capital" as a means of motivating external economies and increasing returns to scale at an aggregate level.

Internal economies of scale are also possible, but are generally not consistent with markets being perfectly competitive. Consider the usual **U**-shaped long-run average cost curve that economists believe characterizes most production processes. The downward-sloping part of this curve has increasing returns to scale because expansion of output leads to lower average cost. If competitive firms are producing on this part of the LRAC curve, market forces will lead them to combine operations and achieve lower average cost by exploiting these economies of scale.

There are two situations where firms would not expand or merge to achieve the efficient scale of output. One is when the total market demand is not large enough to absorb the entire output of a single firm of optimal size. This is the case of ***natural monopoly***, where costs are minimized by concentrating production in a single firm, which generally still produces in a region of falling costs (increasing returns).

The second situation where increasing returns may be an equilibrium is under imperfect competition. Recall the model of ***monopolistic competition*** from your introductory economics course. In this model, firms' products are sufficiently differentiated from one another that each firm has some monopoly power for its portion of the market. However, in the long run, free entry drives economic profit to zero. Because firms' demand curves slope downward (rather than being horizontal under perfect competition), marginal revenue is less than price. This means that a profit-maximizing firm that sets marginal revenue equal to marginal cost ends up producing where price exceeds marginal cost. This happens on the downward-sloping part of the average cost curve. Zero profit in the long run implies that the firm will produce where the downward-sloping demand curve is tangent to the long-run average cost curve.

Thus, there is a close connection between increasing returns to scale and imperfect competition. Competitive industries may have increasing returns only due to external economies, but internal increasing returns can be a long-run equilibrium only when firms have some monopoly power.

Two papers by Robert Hall provide some empirical support for the presence of increasing returns and imperfect competition. Hall (1986) used business-cycle fluctuations in output to examine the resulting changes in cost. He found that the changes in cost were proportionally much smaller than the changes in output, suggesting increasing returns. Because economic profit in these industries did not seem to be large, Hall concluded that they were characterized by a market structure similar to monopolistic competition, with each firm having considerable short-run market power, but with entry keeping economic profit at zero in the long run.

Hall (1990) tests an "invariance property" of the Solow residual: if the Solow residual measures only exogenous changes in technological capability, then it should be uncorrelated with variables that have only demand-side effects. He finds evidence that the Solow residual is correlated with, among other indicators, government military spending. This rejects the interpretation of the Solow residuals as a pure technology variable. Among the explanations he considers for this correlation, Hall builds a case for increasing returns and imperfect competition as being most consistent with the industry-level data he examines.

Evidence on labor hoarding and job hoarding

De Long and Waldmann (1997) use a cross-country sample to shed light on the sources of procyclical variation in productivity. They pose the question of their analysis on page 33: "Some believe that procyclical productivity is ingrained in the technology of production. But a standard view of procyclical productivity sees it as a consequence, not a cause, of changes in activity. Labor productivity falls when output falls because firms retain more workers than required to produce low current output. They do this to avoid the costs of laying workers off now and hiring replacements in the future when activity recovers."

The fundamental problem faced by anyone trying to distinguish between these two hypotheses is identifying shocks to technology and to aggregate demand. De Long and Waldmann do this by making a set of "identifying assumptions." The essential assumption they use is that "there are no technological or other supply-side shocks that are (a) specific to a single country, yet (b) affect a broad range of industries within manufacturing." (p. 34) In simpler words, they are assuming that shocks that are common across industries within a country are demand shocks, not technological shocks. In contrast, shocks that are common to an industry across countries are assumed to be technology shocks.

This identifying assumption (if true) allows them to interpret any within-country but cross-industry shock as being a demand shock. If these shocks seem to cause procyclical changes in productivity, then there must be factors other than exogenous technological shocks that are causing productivity to move with output. De Long and Waldmann's study estimates a large set of regressions to measure the effect of national, cross-industry changes in output (demand shocks) on labor productivity, correcting for the effects of industry-specific, international productivity changes (technology shocks).

They conclude that the effects of these demand shocks are significant and positive, and thus aggregate-demand factors play an important role (alongside technology shocks) in explaining procyclical labor productivity. Three explanations are advanced for why demand shocks might affect productivity in this way. De Long and Waldmann rule out the first—increasing returns to scale—on the basis of implausible differences across countries in the response of productivity to demand. These differences support the varying importance of the other two explanations: ***labor hoarding*** by firms and ***job hoarding*** by workers.

Labor hoarding refers to the practice of firms keeping redundant workers on the payroll during bad times, so that they will retain their trained workforce for use when desired output picks up again. If such workers are laid off, some will find other jobs and be unavailable to the firm when it wants to increase labor input. New workers will require training and may turn out to be less reliable than the proven workers already on the payroll. Thus, the firm may save "turnover costs" by hoarding labor.

However, notice that the tendency of firms to hoard labor may be related to the aggregate unemployment rate in the economy. If the unemployment rate is very high, then laid-off workers are unlikely to find alternative employment, so the firm may not need to hoard labor. When unemployment is low, labor hoarding will be more beneficial since laid-off workers are more likely to have accepted other jobs before the firm tries to recall them.

Job hoarding refers to workers' practice of pursuing legislative and contractual restrictions on the freedom of firms to lay off workers. The effect of job hoarding on productivity is likely to be related to the unemployment rate as well, but in the opposite direction as labor hoarding by firms. When unemployment is low, some of any firm's workers will find alternative employment opportunities and quit voluntarily. If this particular firm wishes to reduce its overall level of employment, this will allow it to do so. However, there will be few quits when unemployment is high, so the firm will not be able to reduce its workforce and measured labor productivity will decline.

Thus, examining the effect on the labor-productivity/aggregate-demand relationship of changes in the unemployment rate may allow one to identify whether labor hoarding by firms or job hoarding by workers is more common. Performing this task, De Long and Waldmann found that labor hoarding was the apparent cause in the United States, but job hoarding was more prominent in Europe.

This is consistent with the general characteristics of labor markets in the two regions. American labor markets tend to be more "flexible," with layoffs being more common, unions being relatively weak, and employment-protection legislation almost nonexistent. Continental Europe features much stronger unions, which have pushed through laws restricting the ability of firms to lay off workers and have negotiated labor contracts that include similar restrictions. Given these differences, we would expect labor hoarding to be more common in the United States and job hoarding to be the norm in Europe, which is exactly the result that De Long and Waldmann found.

Evidence on the workweek of capital

The U.S. Bureau of Labor Statistics and its foreign counterparts have long maintained detailed statistics on the hours that wage-earning employees work. Because wage earners are paid by the hour, it is relatively easy for firms to report their hours of work quite accurately. Moreover, the Department of Labor is very interested in the well-being of workers, which gives it a good incentive to monitor changes in average hours worked. Because these data are available, it has become quite standard to measure labor input as "person-hours" rather than as number of employees, which corrects for the hours-worked dimension of variations in labor utilization.

There are no comparable data on the hours of utilization of capital. One reason is that capital is rarely paid based on the hours it is used. That means that firms do not automatically collect information comparable to labor hours for their capital stock. Another reason is that there is no government "Department of Capital" that worries about the well-being of machines in the way that the Labor Department is concerned with workers.

As a result, econometricians who estimate productivity are usually forced to assume that the entire installed capital stock is used at a constant rate of utilization in each period. In the terms used by Matthew Shapiro (1993), we can write the production function as

$$Y = F(SK, N, L, E, M) = F(Z, N, L, E, M), \qquad (2)$$

where Y is gross output, K is the stock of installed capital, L is labor input measured in person-hours, E is energy input, and M is input of materials. The factor S measures the capital utilization rate as its average weekly work hours during the period. $Z = SK$ is the actual flow of capital input: the average workweek of capital times the installed stock.

Shapiro points out that mismeasuring Z by K (i.e., assuming that S is always constant) can lead to serious errors in the estimation of productivity shocks using Solow residuals:

> Consider the production function of equation (2) with fixed utilization. The standard Solow total factor productivity residual, ε, is given by $\varepsilon = \Delta y - \Delta x$, where Δy is the log change in gross output and $\Delta x = \alpha_K \Delta k + \alpha_N \Delta n + \alpha_L \Delta l + \alpha_E \Delta e + \alpha_M \Delta m$, is the share-weighted log change in the inputs. [Lower-case variables represent the logs of the corresponding capital-letter variables.] … Robert Solow shows that under the assumptions of constant returns to scale, perfect competition, and correct measurement of the factors and shares, the residual, ε, equals the rate of technological change, ε^*. Observed Solow residuals are highly procyclical. An obvious source of this procyclicality is unaccounted variation in the inputs. Production might rise because factor utilization increases. If the increase in factor utilization is not reflected in total factor input, measured ε will be spuriously procyclical. [Shapiro (1993)]

Shapiro goes on to show that an adjusted Solow residual that measures technological change in a way that corrects for variations in capital use is

$$\tilde{\varepsilon} \equiv \varepsilon - \alpha_K \Delta u, \tag{3}$$

where Δu is the growth rate of utilization. He uses data from several sources on the workweek of capital to calculate adjusted Solow residuals from equation (3) and examines their cyclical behavior. The residuals corrected for capital utilization do not move procyclically, which suggests that the procyclical movement in the traditional Solow residual ε is due *not* to exogenous shocks to technology as the RBC models assume but instead to changes in the utilization rate of capital.

G. Direct Estimation of the Sources of Cycles

Given the evidence cited above that both supply and demand shocks play an important role in output fluctuations, many authors have attempted to carry Cochrane's mission forward by not only estimating the relative importance of the two kinds of shocks, but also to estimate the shocks themselves. This allows us to estimate the sources not only of business cycles in general, but also of individual cyclical fluctuations such as the recessions of the 1970s and 1980s.

Blanchard and Quah

One of the earliest and most cited decomposition of supply and demand shocks was undertaken by Blanchard and Quah (1989). They estimated a model analyzing both supply shocks having permanent output effects and demand shocks with temporary output effects. Blanchard and Quah use a vector autoregression methodology to analyze the joint behavior of output and unemployment, employing the absence of long-run output effects to identify demand shocks.

Blanchard and Quah present estimates the effects of a demand shock and a supply shock on output and unemployment. Figure 4 (their Figure 1) shows the effect of a positive demand shock in their model. Output rises quickly, peaking two quarters after the demand shock, then falling gradually back to zero over a period of five to six years. The effects on unemployment are similar, but slightly slower and in the opposite direction.

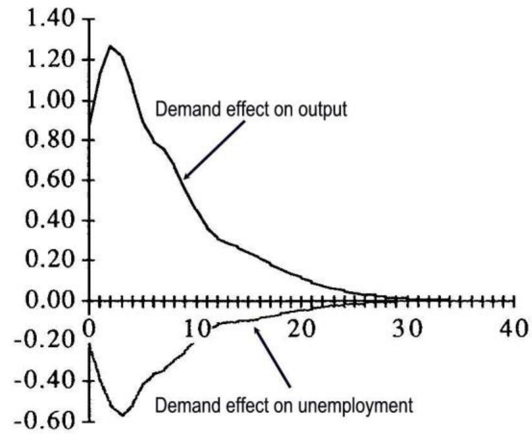

Figure 4. Blanchard and Quah's estimated effects of demand shock.

Figure 5 (their Figure 2) shows the effect of a positive (beneficial) supply shock in Blanchard and Quah's model. Output increases gradually, peaking after about two years and settling back to a somewhat smaller permanent effect. Perhaps surprisingly, the initial impact on unemployment is an increase (perhaps reflecting sectoral reallocation of resources in response to the changing conditions on the supply side), followed by a decline and no long-run effect.

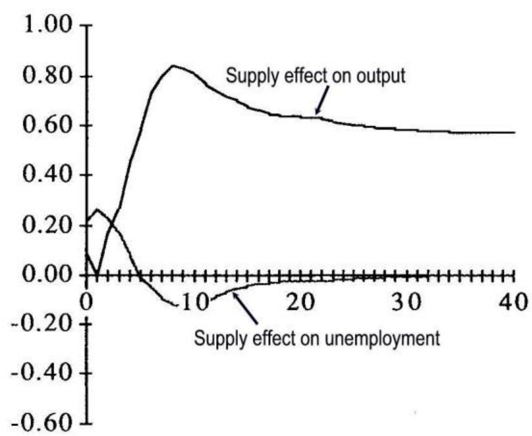

Figure 5. Blanchard and Quah's estimated effects of positive supply shock.

Blanchard and Quah also present actual historical time paths representing the effects of supply shocks and demand shocks on output. Figure 6 and Figure 7 (Figures 7 and 8 from the original) show the effects of supply and demand shocks. Figure 6 shows that if there were no demand shocks, output would grow steadily but not totally smoothly through the sample period. In particular, the recessions of the mid-1970s and the early 1980s seem to have a significant supply component.

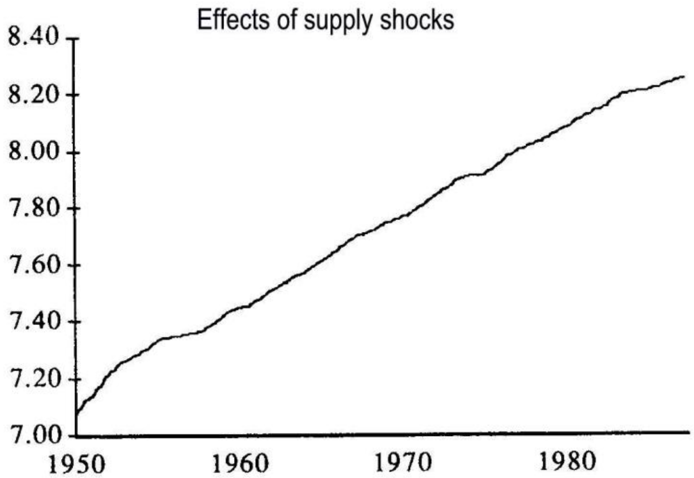

Figure 6. Blanchard and Quah's estimates of output path with no demand shocks.

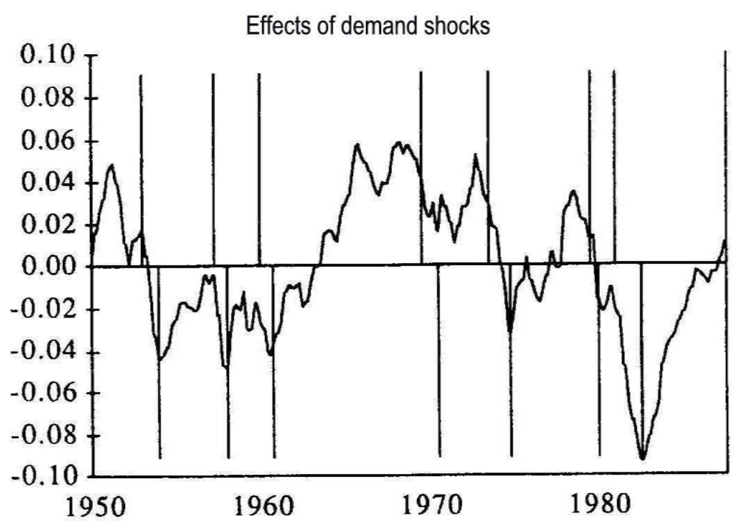

Figure 7. Blanchard and Quah's estimated effects of demand shocks.

Figure 7 shows that Blanchard and Quah's model attributes a great deal of cyclical fluctuation to demand shocks.[16] In particular, the boom of the late 1960s (during the Vietnam War) and the recession of the early 1980s (the Volcker disinflation) both seem to be largely demand driven.

In the analysis of their results, Blanchard and Quah point out that the oil-chock recessions of the mid- and late-1970s are a combination of supply and demand effects:

[16] The vertical bars in Figure 7 show official business-cycle turning points. Peaks are shown by lines above the curve; troughs by lines extending downward.

"The recession of 1974–75 is … explained by an initial string of negative supply disturbances, and then of negative demand disturbances. Similarly, the 1979–80 recession is first dominated by a large negative supply disturbance in the second quarter of 1979, and then a large negative demand disturbance a year later…. [W]e find these estimated sequences of demand and supply disturbances consistent with less formal descriptions of these episodes."[17]

Finally, Blanchard and Quah calculate the fractions of the total variation in output that are attributable to demand and supply shocks over various horizons. Not surprisingly, the short-run fluctuations in output are dominated by demand effects, with supply effects being more important in the long run. For their baseline model, 99 percent of output fluctuations attributable to shocks in the most recent three quarters is due to demand shocks. Over a ten-year horizon, that fraction drops to 40 percent.

Jordi Galí (1992) estimates a VAR model of the postwar U.S. economy with identification restrictions that attempt to identify four kinds of shocks: money supply shocks, money demand shocks, spending (IS) shocks, and aggregate supply shocks. He finds that supply shocks and spending shocks have the greatest contribution to postwar business cycles. Table 4 (taken from Galí's Table IV), shows the shares of the unpredictable variation at various horizons that is attributable to each kind of shock. He finds that the supply shock dominates at all horizons, but that the spending shock is very important in short-run fluctuations. Money supply shocks have a modest impact at business-cycle frequencies.

Table 4. Galí's estimated decomposition of variation in unpredictable output changes.

Quarters ahead	Kind of shock			
	Aggregate supply	Money supply	Money demand	Spending
1	0.69	0.00	0.00	0.31
5	0.67	0.12	0.02	0.19
10	0.72	0.14	0.04	0.10
20	0.83	0.09	0.02	0.06

Galí also presents a detailed breakdown of the sources of the six recessions in his sample. (his Table V) shows that the six postwar recessions show great variation in their causes. In particular, the two largest contractions, in the mid-1970s and the early 1980s, show remarkable differences that correspond closely with economists' intuitive understanding of these cycles.

[17] Blanchard and Quah, pp. 664–665.

Table 5. Galí's estimates of causes of historical business cycles.

	Kind of shock			
Recession period	Aggregate supply	Money supply	Money demand	Spending
1957III–1958II	0.62	0.06	–0.01	0.33
1960II–1961I	1.16	0.16	–0.18	–0.14
1969IV–1970IV	0.38	0.45	0.14	0.03
1973IV–1975I	0.78	–0.02	0.00	0.24
1980I–1980III	0.11	0.29	0.08	0.52
1981III–1982IV	0.06	0.38	0.26	0.30

The recession of the mid-1970s followed the first and largest oil shock. Galí's model attributes nearly all of this shock to supply factors, with a minor contribution from spending (which may be associated with the winding down of Vietnam military expenditures). In contrast, the large recession that began in 1981 is usually attributed to contractionary monetary policy. Galí's model attributes over 60 percent of this recession to monetary factors and only 6 percent to supply shocks.

In a recent study, Hartley and Whitt (2003) use data from six countries to estimate a model that includes both temporary and permanent shocks to aggregate demand and aggregate supply. Table 6 (from their Table 7) shows their estimated percentages of the variation in GDP at business-cycle frequencies.

Table 6. Hartley and Whitt's estimates of relative importance of supply and demand shocks.

Nature of Shock		Country					
		Germany	France	U.K.	Netherlands	Italy	U.S.
Supply	Permanent	18.6	13.2	6.2	8.1	4.2	24.6
	Transitory	5.3	10.1	2.7	18.3	0.5	4.9
Demand	Permanent	76.1	23.2	91.2	73.6	32.0	67.7
	Transitory	0.0	53.5	0.0	0.0	63.2	2.7

The results vary considerably by country, but in each case Hartley and Whitt attribute much more variation to demand shocks than to those in supply. Demand shocks tend to be transitory (reversed) in France and Italy, but permanent in the other four countries. Another study of the G-7 countries by Canova and de Nicoló (2003) also finds demand shocks to be dominant in most countries.

H. Empirical Tests of the Lucas Aggregate Supply Model

International evidence on the slope of the AS curve

The Lucas model implies that the response of real output to changes in aggregate demand should depend on $V_r / (V_r + V_p)$, the share of the variability of individual nominal prices that is associated with relative-price changes. We know little about how V_r, the variability of relative prices, differs across countries or over time. However, we can observe large international differences in V_p, the variability of inflation. Some countries have inflation rates that vary over the course of a few years by hundreds of percent, while the variation in others is a couple of percentage points. Given the magnitude of V_r, output should respond less to aggregate demand in countries with higher inflation variability than in countries with stable inflation. Therefore, countries with more variable inflation should have steeper *AS* curves.[18]

To test this hypothesis empirically, one must be able to identify aggregate-demand shocks and to measure their effects on real output. Lucas (1973) used changes in nominal GDP to represent aggregate demand shifts.[19] He estimated the response of real output to changes in nominal GDP growth and found considerable evidence in support of his hypothesis. In countries such as Paraguay and Argentina, output hardly responded at all to nominal GDP changes. In the United States and other stable-inflation countries, real output was much more responsive.

Subsequent research has extended Lucas's sample to more recent time periods and additional countries. In general, the results have proved to be robust. Volatile inflation seems to be associated with steep short-run aggregate supply curves. However, other models share this prediction of the Lucas model. So while this empirical result supports the Lucas model in general, it may not succeed in demonstrating its superiority overall competing models. For example, higher variability of monetary shocks may lead to shorter labor contracts or to greater indexation of contracts, either of which will reduce the elasticity of aggregate supply in new Keynesian wage-contract models.[20]

---

[18] Remember that price is on the vertical axis and output on the horizontal axis of the *AS* curve.

[19] Also see the erratum published in *American Economic Review* 66(5), December 1976, 985.

[20] See Joanna Gray (1978) or Romer's Problem 6.12.

Ball, Mankiw, and Romer (1988) extended the Lucas cross-country analysis to attempt to discriminate between the new classical and new Keynesian models.[21] In Lucas's model, the slope of the short-run aggregate-supply curve depends *only* on the relative variances of relative prices and aggregate inflation. Changes in the level of actual (or expected) inflation do not affect the slope because even very high inflation will not have real effects in the Lucas model as long as people anticipate it correctly.

However, in the standard new Keynesian model with price stickiness induced by menu costs, *either* more variable inflation *or* a higher expected level of inflation will induce more frequent price adjustment, which implies a less elastic aggregate-supply curve. Thus, a test of whether the level of inflation (in addition to the variance) helps predict supply elasticity can provide evidence on the relative merits of the two theories. If only inflation variance affects supply elasticity, then the Lucas model is supported; if both variance and level of inflation affect elasticity, then the price-stickiness model seems more appropriate.

Following Lucas, Ball, Mankiw, and Romer (BMR) estimate the elasticity of the short-run aggregate-supply curve from equation (4)

$$y_{i,t} = \alpha_i + \tau_i x_{i,t} + \lambda_i y_{i,t-1} + \gamma_i t + \varepsilon_{i,t}, \tag{4}$$

where i indexes countries, t is time, y is the log of real GNP, x is the log of nominal GNP, and ε is an error term. The nominal GNP variable x is the measure of aggregate demand, so τ_i measures the elasticity of short-run aggregate supply for country i—the share of an increase in nominal demand that is satisfied through increased output.

BMR estimate equation (4) for 43 countries using all available postwar data. Their τ_i estimates range from a high of 0.85 for Denmark to (insignificantly) negative values for a few high-inflation countries. As the next step in their analysis, they estimate the relationship between their estimated τ_i values and the level and variability of inflation. Their results are reported in their Table 5 and by Romer in equation (6.101). They find that the level of inflation has a significant effect on the elasticity of short-run aggregate supply and that the effect is negative (as predicted by the new Keynesian model) except when inflation gets very high.[22]

Moreover, including the level of inflation in the equation for τ makes the effect of the variability of inflation statistically insignificant. While the new Keynesian model predicts a negative effect of inflation variability on elasticity, it is less central to the

---

[21] This study is discussed in section 6.10 of Romer.

[22] As shown by Romer's Figure 6.16, the relationship between τ_i and inflation appears to be nonlinear. BMR estimates a quadratic relationship and, since parabolas eventually turn, the slope of the relationship is likely to be positive for some values of inflation and negative for others. They find that the relationship is negative for inflation rates under 34%.

Keynesian model than to the Lucas model. Thus, the strength of the effect on supply elasticity of the level of inflation and the weakness of the variability of inflation lends empirical support to the new Keynesian framework in comparison to Lucas's new classical model.

Evidence from hyperinflations

The Lucas model has strong implications for the Phillips curve tradeoff between unemployment and inflation. Fully anticipated changes in inflation should have no real effects, so there should be no tradeoff at all when expectations are correct. We could test this hypothesis if we could identify changes in inflation that were correctly anticipated. However, the absence of good measures of inflationary expectations makes this task difficult.

Thomas Sargent (1982) argued that reforms of monetary and fiscal institutions at the end of periods of hyperinflation were cases of correctly anticipated changes in inflation. He looked at inflations immediately following World War I in Germany, Austria, Hungary, and Poland. The first three of these countries were defeated in the war and, under the Treaty of Versailles, required to pay enormous reparations to the victorious Allies.[23] The reparations burden overwhelmed these countries' public finances, leading them to inflation as a source of government revenue. While Poland owed no reparations, it was a newly formed country with small specie reserves. Inflation in Poland was used to pay for a defensive war against Russia, which invaded shortly after the end of World War I.

In each of these countries, prices rose by a factor of more than 10,000 in the early 1920s. In Germany, prices went up by a factor of 100,000,000,000 (100 billion) between 1920 and 1924. A similar inflation in the United States would put the price of one copy of Romer's textbook at perhaps one hundred times the current dollar wealth of Bill Gates—about 6 trillion dollars!

Each of these four inflations ended when the government finally got control of its expenditures and reformed the institutions of the central bank. Fiscal reform came through a combination of renegotiation of treaty terms, improved tax collection, and in Poland's case the end of its costly wars. Monetary reform made the central banks more independent of the government's fiscal demands, so that inflationary finance was not as tempting.

Sargent argued that these reductions in inflation were likely to be correctly anticipated. The reforms were well publicized and economically credible. Both economists

[23] One of John Maynard Keynes's earliest publications is *The Economic Consequences of the Peace*, written in 1919, in which he predicted that the huge reparations demanded would be unpayable and prevent recovery of the European economies.

and the general public looked at the changes as likely to end the extremely high inflation. Under these circumstances, then, the Lucas model predicts that disinflation should be possible with no associated recession. The expectations-adjusted Phillips curve itself should shift downward, rather than disinflation moving the economy down along the curve to high unemployment rates.

Indeed, Sargent finds that output did not decline following these disinflations. Indeed, the end of the disruption associated with inflation often produced an increase in real activity. He concludes from these episodes that, consistent with the Lucas model, correctly anticipated changes in inflation have no real effects of the kind predicted by the Phillips curve.

However, Sargent's disinflations are extreme examples. He selected these extreme cases because correct anticipations are plausible in situations such as the extensive reforms to the policy process that followed them. Does this conclusion apply to more modest inflations? The absence of good measures of inflation expectations makes drawing clear conclusions difficult. In most recent cases, reducing inflation has involved a substantial unemployment cost. How much of this could have been mitigated by more accurate expectations is an open question.[24]

Tests of policy ineffectiveness

As noted above, the most remarkable implication of the Lucas model is the prediction that correctly anticipated changes in money growth should have no real effects. The first direct test of the policy-ineffectiveness proposition was a pair of papers by Robert Barro.[25] The principal difficulty in testing this theory is measuring agents' expectations about the money supply. If we are to estimate the effects of anticipated vs. unanticipated changes in the money supply, we must be able to distinguish between them.

Barro developed a measure of expectations about monetary policy using a natural empirical extension of the rational-expectations hypothesis. He estimated a regression equation in which the dependent variable was *DM*, the growth rate of the money supply, and the explanatory variables were variables that agents in the economy could observe in the previous year and that might logically be used to forecast central-bank monetary policy. The predicted values (fitted values) from this regression equation (which he called *DM*) were then assumed to correspond to the anticipated part of monetary policy; the residuals (*DMR*) were taken to be the unexpected component.

---

[24] A seminal work in the analysis of the unemployment costs of disinflation is Laurence Ball (1994).

[25] See Barro (1977) and Barro (1978).

Barro included both the anticipated and unanticipated components of money growth in a regression equation with the unemployment rate as the dependent variable. His results supported the Lucas proposition that the *DM* variables had no effect on unemployment while the *DMR* terms had positive effects. In a companion paper, he found supporting evidence for Romer's equations (6.91) and (6.92). The anticipated component of money growth seemed to affect prices proportionally and did not affect real output.

Despite the initial support from Barro's econometric studies, the tide of evidence eventually turned against Lucas's model. Other economists repeated Barro's study for different countries, different time periods, and most importantly with different sets of variables appearing in the agents' prediction equation. Barro's results turned out to be quite fragile; the bulk of the statistical evidence now suggests that even anticipated changes in the money supply have real effects, though they may be smaller than those of unanticipated changes.[26]

I. Direct Evidence of Price and Wage Stickiness

Pitfalls in testing for price and wage stickiness

Given the vast quantity of data collected by the federal government on prices and wages, it might seem an easy task to assess whether they are sticky or flexible. These data, however, are published at an aggregate level, which makes them useless for the analysis of nominal stickiness in individual prices. Consider the change in a typical price index from one month to the next. Under price stickiness we would expect that many, perhaps most, individual prices would be the same in the two months. However, for a few goods, their occasional price jumps will occur between the two monthly observations. Averaging together the many prices that do not change with the few that change a lot will yield an aggregate price index that changes gradually, even if there is considerable underlying stickiness in individual prices.

In order to assess price stickiness, one must have time-series data on actual transaction prices for a specific and unchanging commodity. No published data of this kind exist, so economists looking to study price stickiness have had to be creative to find

[26] For an example of a detailed study that comes to the opposite conclusion, see Frederic S. Mishkin (1982). For an extensive survey of the results in this literature, see Table 13.1 of Goodhart (1989).

situations in which such prices could be observed over time. A few ingenious examples, such as cover prices of magazines and prices published in catalogs, have been uncovered, but data on actual transaction prices remain scarce.[27]

Moreover, even if one has data on the prices of actual transactions, it is not always trivial to assess whether such prices are indeed sticky. Even perfectly flexible prices would remain unchanged for long periods if there were no change in the underlying demand and supply for the good. Price stickiness means that prices change less frequently than they would if they responded fully to market forces. Thus, in order to establish the existence of price stickiness, one must be able to show that competitive market prices would have changed more frequently than prices actually did.

Economists often use the aggregate rate of inflation as a benchmark for movements in equilibrium prices. One can also examine the response of prices to discrete events that have a predictable (and large) effect on the equilibrium price. Solow response to such events could be evidence of price stickiness.

Data on transaction prices

One of the earliest studies to examine stickiness of transaction prices was performed by Dennis Carlton (1986).[28] He used data on transaction prices collected by George Stigler and James Kindahl in the 1950s and 1960s. The transactions in the data set were (wholesale) purchases of materials by manufacturing companies from suppliers.

Carlton finds that many prices change infrequently, often staying unchanged for several years. Even though his sample period of 1957–66 was one of relatively low inflation, so equilibrium prices might have been fairly stable, he infers from this evidence that firms are holding prices steady in the face of changes in equilibrium prices. Among the observations that lead him to this conclusion is that prices of identical goods for different buyers change at different times, which would not happen in a competitive market with a homogeneous price that was constantly at the market-clearing level.

Carlton's study uses wholesale prices for purchases by large manufacturing firms. The behavior of prices in these transactions could be quite different than those in retail market. For example, because manufacturers and their suppliers tend to have long-

---

[27] A recent collection of papers, which are discussed below, has used the actual firm-level data collected by the agencies that calculate price indexes to examine price stickiness.

[28] It is worth noting that Carlton is a specialist in industrial organization, not macroeconomics, so although his study is widely cited in the macroeconomics literature, the emphasis of his analysis is on the implications of the price data for market structure and other industrial organization issues.

term relationships, it would be easy for them to vary non-price terms of their transactions rather than price. Thus, a supplier might delay delivery rather than raising price when a surge in demand raises the equilibrium price of the good.

Two studies have examined data on retail prices. Stephen Cecchetti (1986) looked for rigidity in the cover prices of 38 magazines over the period from 1953 to 1979. Because this sample includes both low-inflation and high-inflation periods, it allows him to examine the effects of changes in the underlying inflation rate on the frequency of price changes. Since we expect the equilibrium price to change with inflation, we expect prices to be very flexible during times of high inflation.

For every year in his sample, even those with high inflation, Cecchetti finds that at least half of the magazines in his sample do not change prices. However, the number of magazines changing prices is very sensitive to the inflation rate, as shown in Figure 8. During the high-inflation period of the 1970s, the number of price changes is much higher than during the low-inflation period before about 1966.[29]

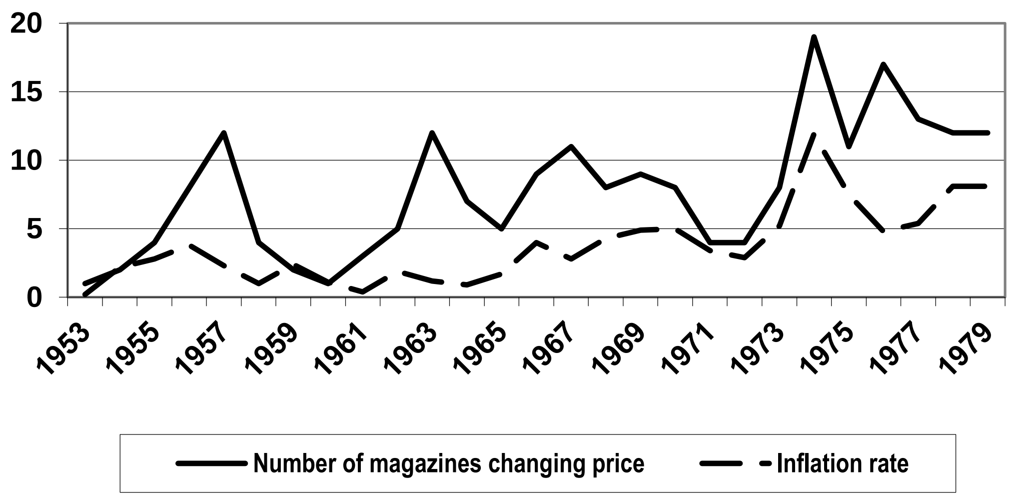

Figure 8. Cecchetti's relationship between inflation and the number of magazines whose prices change.

Cecchetti also examines the magnitude of the average price change and the amount of inflation that had occurred since the last time a magazine had changed its price. The average price change in each year is between 15 and 30 percent, with the

[29] The period of 1972 and 1973 appear to be an anomaly because inflation was high, yet few price changes occurred. This was the period of Nixon's wage and price controls that attempted to prevent firms from raising prices. General inflation continued because many firms were successful in either exploiting loopholes in the price controls or hiding their price increases. Because the cover price of magazines was easily observed, it would have been difficult for publishers to circumvent these price controls without being caught, so there were probably fewer price changes for this highly visible good in these years than would otherwise have occurred.

amount of inflation since the last price change being of similar magnitude. Thus, firms seem to have been holding prices constant for two years, three years, or longer, then changing by a relatively large amount to keep up with the considerable price inflation that had occurred since the last change. Cecchetti compares this model to the predictions of an *Ss* model of the kind used by Caplin and Spulber.[30] He finds the evidence broadly consistent with the Ss model, but not with a simple version having constant menu costs.

Anil Kashyap (1995) constructs and examines a data set of retail transaction prices using mail-order catalogs. He uses prices from the regular, semi-annual catalogs of three prominent mail-order sellers of outdoor equipment. To assure the consistency of the prices over time, he uses only "well-established, popular-selling items that have undergone minimal quality changes" (Kashyap 1995, 248). Prices of selected items are collected over the sub-periods for which they are available from 1953 to 1987. Few items are available in unchanged form for the entire 25-year sample, but all of the selected goods have observable samples of at least 12 years.

Kashyap's sample begins in a period of low inflation, includes the high-inflation years from 1974 to 1982, and ends as inflation returned to low levels in the mid and late 1980s. His results conform strongly to Cecchetti's: "prices are adjusted infrequently, by differing amounts, and although prices are more likely to change during periods of high overall inflation, the synchronization of changes across goods is generally low" (Kashyap 1995, 261).

To summarize, both Cecchetti and Kashyap find considerable evidence that patterns of price changes do not conform to those one would expect under perfect competition with perfectly flexible prices. Both find considerable evidence for price stickiness.

Studies using micro-data underlying price indexes

Each of the studies discussed above focuses on a small and opportunistic selection of goods on which the authors have access to transaction-price data. Bils and Klenow (2004) look at a much broader selection of goods and services: the commodities comprising the Consumer Price Index.

Bils and Klenow find wide variation across commodities in the frequency of price changes. For example, only 1.2 percent of coin-operated laundries change their price from one month to the next, whereas 79 percent of gas stations do (Bils and Klenow 2004, 951). Figure 9 reproduces Figure 1 from the Bils and Klenow paper. It shows the distribution of the percentage of outlets changing prices across the 350 product

[30] See Romer's section 7.5.

categories they study.[31] Figure 9 shows the wide variation across products arrayed between the laundries on the left end and the gas stations on the right.

Averaged across all goods, Bils and Klenow report that about 20% of firms change their prices in the average month and that the average interval between measured price changes is between four and five months. This is a higher frequency of price adjustment than those reported in the selective studies above. Moreover, Bils and Klenow find that the pattern of price changes for individual goods does not correspond to the exponential pattern predicted by the Taylor model that underlies Romer's fixed-price model. They interpret these results as casting serious doubt on such models with strict, time-dependent pricing rules.

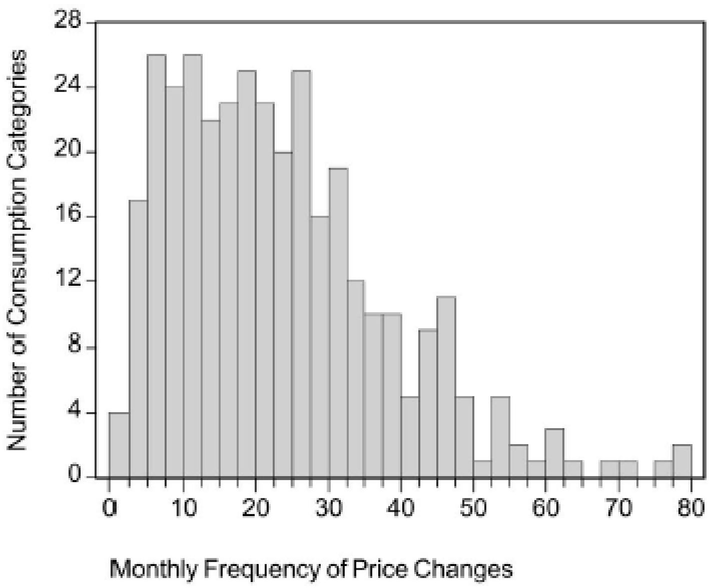

Figure 9. Distribution of price-change frequencies across goods

Good ideas travel fast in academia. Following on Bils and Klenow's research for the United States, the European Central Bank has initiated a series of similar studies using data from major countries in the Euro-area. Studies currently available as working papers from the ECB Web site include Álvarez and Hernando (2004) for Spain; Aucremanne and Dhyne (2004) for Belgium; Baudry et al. (2004) and Loupias and Ricart (2004) for France; Dias, Dias, and Neves (2004) for Portugal; Fabiani, Gattulli,

[31] This figure can be a little hard to interpret. The left-most bar indicates that for 4 of the 350 products, between 0 and 2.5% of outlets changed their price in an average month. The next bar shows that 2.5 to 5% of outlets changed price in an average month in 17 product categories, etc.

and Sabbatini (2004) for Italy; and Jonker, Folkertsma, and Blijenberg (2004) for the Netherlands.[32]

In another important paper using the BLS-CPI data set, Klenow and Kryvtsov (2008) compare actual price changes to the predictions of several models. Not surprisingly, they find that none of the models captures all aspects of price-adjustment behavior. They are most optimistic about newer variants of state-dependent pricing models that may come closest to the properties of their data.

J. Why Are Prices Sticky?

The evidence discussed in the previous section presents a case for the existence of significant price stickiness. We now examine a small collection of papers that investigate why firms do not change prices more frequently. Is it menu costs or something more complex?

Survey evidence on the causes of price stickiness

Sometimes the most obvious way to find out why someone does something is to ask him. This might seem obvious to anyone except an economist, but interview surveys are a methodology that economists rarely employ. There is a long-standing tradition in economics of testing theories based on what people do rather than on what they say. During the 1990s, a research team headed by Alan Blinder followed the unorthodox strategy of asking business executives to rate the importance of a dozen theoretical reasons why prices might be sticky. The results are summarized in Blinder (1994) and presented in more detail in Blinder et al. (1998).

Blinder asked firms to respond to 12 possible theoretical reasons why prices might adjust slowly in response to a change in demand or costs. He provides the following table of descriptions of the 12 theories.[33] For each theory, Blinder read the executive being interviewed a short description of why it might lead to price stickiness and asked whether this was an important factor leading to price stickiness for that firm. He then categorized the results on a four-point scale where 4 = very important, 3 = moderately important, 2 = of minor importance, and 1 = totally unimportant. The theories marked with asterisks in Table 7 are the ones that Blinder thinks were the most prominent among macroeconomists of the 1980s.

[32] See http://www.ecb.int/home/html/index.en.html.

[33] From Table 4.3 of Blinder (1994).

The mean scores are shown in Table 8. The bold lines in the table separate adjacent scores whose difference is statistically significant. Among the academically popular theories, only coordination failure falls in the top group. Menu costs of price adjustment rank sixth among the theories, with an average response just below "of minor significance."

However, we should not necessarily reject the importance of nominal rigidities for two reasons. First, nominal rigidities may manifest themselves in implicit or explicit contracts, which seem more important to respondents. Second, theory (Ball and Romer (1990)) tells us that the presence of real rigidities may strongly enhance the effect of nominal rigidities. Real rigidities through coordination failure is the theory that resonated most positively with decision-makers.

Table 7. Blinder's 12 theories of price stickiness.

Theory Number	and Name	Brief Description
B1	Nominal contracts	Prices are fixed by contracts
B2	Implicit contracts	Firms tacitly agree to stabilize prices, perhaps out of "fairness" to customers (Okun 1981)
*B3	Judging quality by price	Firms fear customers will mistake price cuts for reductions in quality (Allen 1988)
B4	Pricing points	Certain prices (like $9.99) have special psychological significance (Kashyap 1992)
*B5	Procyclical elasticity	Demand curves become less elastic as they shift in (Bils 1989; Shapiro 1988)
B6	Cost-based pricing with lags	Price rises are delayed until costs rise (Gordon 1981; Blanchard 1983)
*B7	Constant marginal cost	Marginal cost is flat and markups are constant (Hall 1986)
*B8	Costs of price adjustment	Firms incur costs of changing prices (Rotemberg 1982; Mankiw 1985)
B9	Hierarchical delays	Bureaucratic delays slow down decisions
*B10	Coordination failure	Firms hold back on price changes, waiting for other firms to go first (Ball and Romer 1991)
B11	Inventories	Firms vary inventory stocks instead of prices (Blinder 1982)
B12	Delivery lags, service, etc.	Firms prefer to vary other elements of the "vector," such as delivery lags, service, or product quality (Carlton 1990)

Table 8. Mean scores for theories in Blinder's survey.

Rank	Theory	Mean Score
1	*Coordination failure	2.77
2	Cost-based pricing with lags	2.66
3	Delivery lags, service, etc.	2.58
4	Implicit contracts	2.40
5	Nominal contracts	2.11
6	*Costs of price adjustment	1.89
7	*Procyclical elasticity	1.85
8	Pricing points	1.76
9	*Constant marginal cost	1.57
10	Inventories	1.56
11	Hierarchical delays	1.41
12	*Judging quality by price	1.33

Blinder's survey thus yields mixed evidence for macroeconomists seeking to understand price rigidity. While most price-setters rejected the direct importance of menu costs in motivating price stickiness, they strongly supported the importance of coordination failures and real rigidities, which can make even small nominal stickiness much more important.

Following on Blinder's work in the United States, a group of European economists have undertaken surveys of price setters in countries using the euro. In addition to asking about the motivations for price stickiness, these surveys also queried firms about the process of reviewing prices to determine whether and how to change them. This included information on how often price reviews occur and whether the review and resetting of prices was done on a standard timetable or depended on the state of the market. They also asked about the information on which pricing decisions are based. The detailed results have been published in book form in Fabiani, Loupias, Martins et al. (2007); the second chapter of this book, Fabiani, Loupias, Druant et al. (2007), presents an overview of the results.

Table 9. Reasons for price stickiness in European countries

Theories	AT	BE	FR	GE	IT	LU	NL	PT	SP	**Euro Area**
Implicit contracts	3.0	2.5	2.2	—	—	2.7	2.7	3.1	2.6	**2.7**
Explicit contracts	3.0	2.4	2.7	2.4	3.0	2.8	2.5	2.6	2.3	**2.6**
Cost-based pricing	2.6	2.4	2.5	—	—	2.7	—	2.7	—	**2.6**
Coord. failure	2.3	2.2	3.0	2.2	3.0	2.1	2.2	2.8	2.4	**2.4**
Judge qual. by price	1.9	1.9	—	—	—	2.2	2.4	2.3	1.8	**2.1**
Temporary shocks	1.5	1.8	2.1	1.9	2.0	1.7	2.4	2.5	1.8	**2.0**
Δ non-price factors	1.7	1.7	—	—	—	1.9	1.9	—	1.3	**1.7**
Menu costs	1.5	1.5	1.4	1.4	2.0	1.8	1.7	1.9	1.4	**1.6**
Costly information	1.6	1.6	—	—	—	1.8	—	1.7	1.3	**1.6**
Pricing thresholds	1.3	1.7	1.6	—	1.0	1.8	1.8	1.8	1.5	**1.6**

Source: Fabiani, Loupias, Druant et al. (2007), Table 2.8. Results are average responses on the following scale: 1 = unimportant, 2 = of minor importance, 3 = important, 4 = very important. Countries are Austria, Belgium, France, Germany, Italy, Luxembourg, Netherlands, Portugal, and Spain.

As shown in Table 9, the most prominent reasons for price stickiness in most European countries are implicit and explicit pricing contracts, along with cost-based pricing and coordination failures. These are four of the five highest-rated causes in Blinder's U.S. survey, so the evidence is largely consistent. While menu costs and pricing points/thresholds have some degree of importance in the United States, neither is at all important in European countries. Price as a signal of quality is more prominent in the European sample (especially in Luxembourg, Nethlands, and Portugal) than in the U.S. Many of the other causes queried in the surveys do not match up directly. Although not in the U.S. questionnaire, the idea of sticky (costly) information does not get support from the Euro Area survey.

Evidence on menu costs from supermarket data

The advent of computerized cash registers in grocery stores has enabled merchants to track sales with great precision. It also has provided another source of data on transaction prices for economists. Levy et al. (1997) analyzed the behavior of five supermarket chains and constructed a test for the impact of menu costs on the frequency of price changes.

Levy et al. collected data from the stores on the frequency and magnitude of price changes. They also monitored store employees and calculated the actual cost of each of the many steps involved in changing a price. While menu costs may seem trivial, especially under a computerized system where individual items do not bear price tags, Levy found the process to be both complicated and costly.

Figure 10 shows the steps identified by Levy. For each step, they identified the several-dozen individual tasks that were involved and measured how many seconds of labor time were required to perform each task. Adding up the costs led to the conclusion that each store in the average chain spent about $100,000 on menu costs each year, or about 52 cents per price change and about 0.7 percent of sales.

Although this seems like a very small magnitude, it is much larger than the magnitudes (less than 0.1 percent) reported in earlier research. Moreover, Levy found evidence that menu costs do cause price stickiness. Three of the chains reported the frequency of cost increases that were not passed along to customers because of the cost of changing prices. For these chains, 22 percent, 34 percent, and 30 percent of the desired price changes were not implemented because of menu costs. This suggests that even small menu costs can cause substantial price stickiness.

Another feature of their data allowed Levy's team to estimate the effects of menu costs on pricing decisions. One of the five chains operated in a state that required "item-pricing"—each individual item had to have a separate price tag. For this chain, the cost per price change was significantly higher at $1.33. If menu costs are important, then the chain subject to item-pricing should change its prices less frequently than the other chains. Whereas the four chains not subject to item-pricing changed an average of 15.66 percent of their prices in a typical week, the item-pricing chain changed only 6.31 percent.

Thus, Levy et al. present evidence that, at least for supermarkets, menu costs can be larger than earlier estimates suggested. Moreover, menu costs seem to lead to a significant degree of price stickiness.

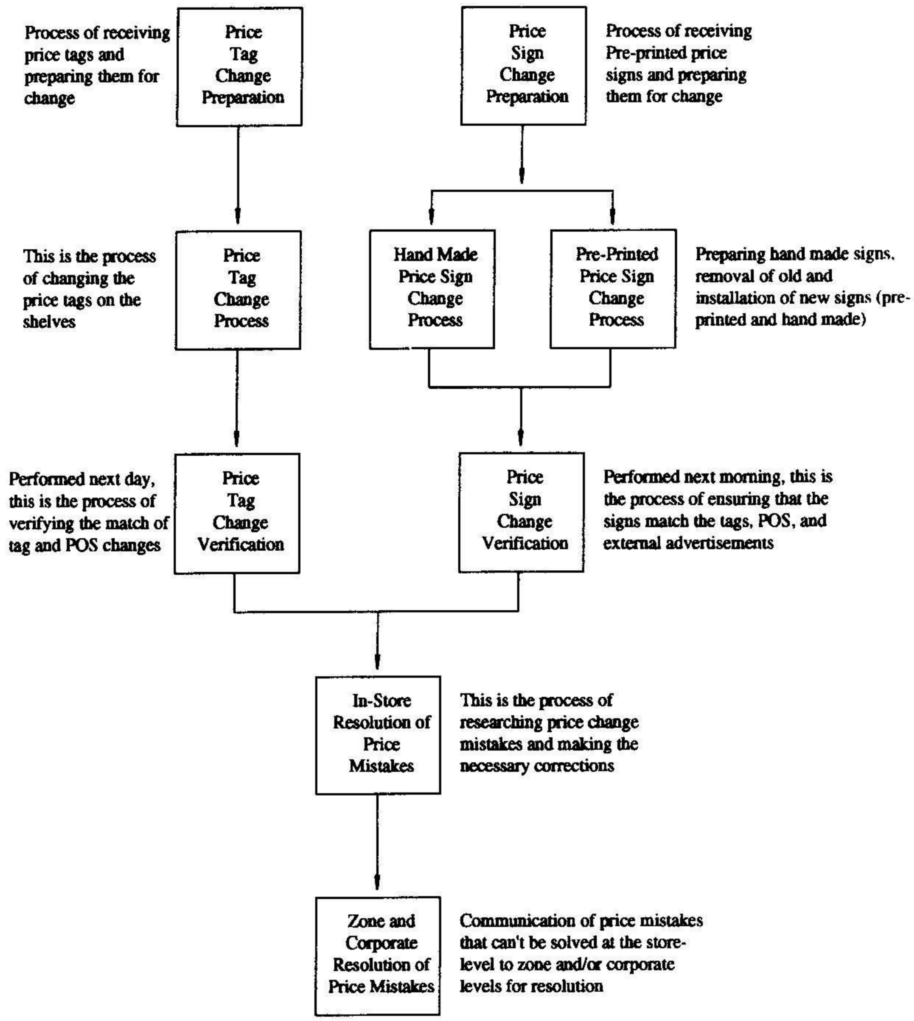

Figure 10. Steps involved in changing supermarket prices.

Levy and Young use internal company documents to examine the causes of this amazing run of price stickiness. They find that Coca-Cola debated long and hard about raising its price. Among the major reasons why the price change took so long were customer relations, the relationship between Coke and its bottlers, and the technology of automated vending machines.

The "Five-Cent Coke" was a major advertising theme for Coca-Cola for decades. Many within the company felt that increasing the price would betray the trust of their loyal customers. The company's national price advertising also forced local bottlers (who actually set prices) to maintain uniform national pricing, which Coca-Cola strongly desired.

Beginning in 1936, Coca-Cola began marketing their bottles in vending machines. These primitive machines worked on a single nickel; technology to accept multiple coins or to make change would not be available until decades later. By the 1950s, a large part of Coke sales were through vending machines, which raised an obvious problem: the only feasible increase in price was to double the price to a dime. Apparently, the Coca-Cola top brass convinced high-ranking friends in the Eisenhower Administration to propose the creation of a 7½-cent coin, expressly to provide an intermediate price for Coke machines. Fortunately, wiser heads prevailed in Washington!

Levy and Young conclude that Coca-Cola was a classic case of price stickiness based on nominal rigidity. The combination of factors involving implicit contracts with customers, advertising and affiliate issues, and technological indivisibilities acted much like menu costs.

Administrative costs of price adjustment

The ubiquitous use of the term "menu costs" invokes a particular image of nominal adjustment costs: the costs of physical reprinting of menus or perhaps manually changing price tags on goods. One of the reasons for skepticism about menu-cost theories is that, apart from fancy restaurants with gold-leaf printing on their menus, these costs simply cannot be all that high. It doesn't cost much to hire a minimum-wage employee to change shelf stickers or even to stamp new price tags on existing stock.

However, as noted in Levy et al. (1997) and reported here in Figure 10, the process of changing prices involves many steps, including research and decision-making by highly paid executives and the public-relations task of explaining (or perhaps negotiating) price increases with customers. Zbaracki et al. (2004) examine the process of changing prices in one large industrial firm. They find that the physical "menu" costs are a tiny fraction of the total costs that the firm incurs when it changes prices. Managerial costs (information collection, decision making, and internal communication) are six times as large as menu costs and "customer costs" associated with communicating and negotiating price changes with buyers exceed menu costs by a factor of twenty.

This evidence places adjustment costs in quite a different light. It raises the possibility that adjustment costs may be quantitatively more significant than the often-trivialized physical menu costs. Managerial and customer-related adjustment costs may also be convex—increasing in the amount by which prices are changed—rather than a simple lump sum that is incurred any time prices change at all. Decisions to make large price changes require more weighty consideration by management and much greater effort in communication with customers than small changes that are likely to have small effects.

Another possibility, which is important in considering the possibility of "inflation stickiness" or information stickiness that is raised by Mankiw and Reis (2002), is that managerial and customer costs could be associated with changes in inflation rather than just changes in the level of prices. For example, if a firm has done extensive research to decide that it should increase its price by 2% per year and has informed its customers of this decision then the only costs associated with annual price changes of 2% would be the (small) menu costs. As long as they keep this rule setting their rate of price inflation unchanged, managerial and customer costs are low or nil. However, should the firm decide to change the rate of price increase to 3%, this would require another decision by managers and a new round of communication and negotiation with customers. Thus, managerial and customer costs can lead to rigidity of inflation as easily as to rigidity in prices.

K. Further Evidence on Recent Models

Mankiw and Reis (2002), in a paper described in Romer's section 7.7, argue that slowness in the spread of information leading to "sticky expectations" can lead to dynamic behavior of output and inflation that avoids some of the empirical inconsistencies of the sticky-price model. They test this model in Mankiw, Reis, and Wolfers (2003) by looking at the degree of disagreement among the respondents to surveys of inflation expectations.

They find that disagreement about future inflation increases when inflation is high, when inflation changes suddenly, or when rates of price change among commodities vary strongly. When they relate actual survey expectations to the predictions of an econometric model based on various vintages of past data, they find evidence consistent with agents updating their expectations roughly once per year. This is broadly consistent with the calibration used by Mankiw and Reis (2002) and with other empirical evidence and they interpret the result as consistent with the sticky-information model. Klenow and Willis (2007) find further evidence in support of a sticky-information variant using data from the disaggregated CPI sample.

Another recent strand of macroeconomic modeling is the development of theoretical models with coordination failures. Because coordination failures can occur in a variety of contexts, empirical testing of this concept tends to be idiosyncratic, looking for events or situations that might represent coordination failure and examining their specifics to assess whether this is a reasonable interpretation.

Cooper and John (1988) argue that coordination failures occur when there are strong strategic complementarities between the actions of agents. Cooper and Haltiwanger (1996) find several situations that they believe reflect strategic complementarities in the macroeconomy. They identify several symptoms that should reflect the possibility of complementarities. First, there should be ***positive co-movement*** of variables between different decision-makers. When one agent decides to do something, others are likely to do the same thing.[34] The second symptom is ***temporal agglomeration***, where discrete decisions of many agents are coordinated in time. Synchronizing decisions can help to avoid the real rigidities associated with overlapping contracts in the Fischer and Taylor models. Third, strategic complementarities should lead to ***magnification and propagation*** of macroeconomic effects of shocks relative to what would have occurred in their absence. Of course, one can measure magnification only if one can estimate what would have happened in the absence of complementarities.

Cooper and Haltiwanger's evidence comes from a variety of observations. They note that movements in major variables (output, employment, and prices) tend to be positively correlated across industries, which could be a sign of complementarities between them. This is true both with respect to movements at business-cycle frequencies and for seasonal changes.

They find a number of circumstances in which discrete decisions of firms, such as investment in new factories or replacement of machines, are synchronized. Among other examples, they point to temporal agglomeration occurring around the introduction of the new model year of automobiles, which was altered via policy intervention in the 1930s.

Finally, they present evidence supporting the existence of magnification mechanisms in the macroeconomy through technological externalities and other means. Cooper and Haltiwanger conclude from their eclectic set of examples that situations of strategic complementarity exist in the macroeconomy. This raises the possibility that coordination failures could arise and that explicit government coordination could improve resource allocation.

L. Works Cited in Text

Álvarez, Luis J., and Ignacio Hernando. 2004. Price Setting Behaviour in Spain: Stylized Facts Using Consumer Price Micro Data. Frankfurt: European Central Bank.

Aucremanne, Luc, and Emmanuel Dhyne. 2004. How Frequently Do Prices Change? Evidence Based on the Micro Data Underlying the Belgian CPI. Frankfurt: European Central Bank.

Ball, Laurence. 1994. What Determines the Sacrifice Ratio? In *Monetary Policy*, edited by N. G. Mankiw. Chicago and London: University of Chicago Press.

[34] Of course, one must be careful about this condition because there are many reasons other than strategic complementarities that might induce agents to make similar decisions. Any exogenous shock to which all respond in a similar way would qualify.

Ball, Laurence, and N. Gregory Mankiw. 1994. A Sticky-Price Manifesto. *Carnegie Rochester Conference Series on Public Policy* 41:127-151.
Ball, Laurence, N. Gregory Mankiw, and David Romer. 1988. The New Keynesian Economics and the Output-Inflation Trade-off. *Brookings Papers on Economic Activity* 1988 (1):1-65.
Ball, Laurence, and David Romer. 1990. Real Rigidities and the Non-neutrality of Money. *Review of Economic Studies* 57 (2):183-203.
Barro, Robert J. 1977. Unanticipated Money Growth and Unemployment in the United States. *American Economic Review* 67 (1):101-115.
———. 1978. Unanticipated Money, Output, and the Price Level in the United States. *Journal of Political Economy* 86 (4):549-580.
Barsky, Robert B., and Lutz Kilian. 2004. Oil and the Macroeconomy Since the 1970s. *Journal of Economic Perspectives* 18 (4):115-134.
Baudry, Laurent, Hervé Le Bihan, Patrick Sevestre, and Sylvie Tarrieu. 2004. Price Rigidity: Evidence from the French CPI Micro-Data. Frankfurt: European Central Bank.
Baxter, Marianne, and Robert G. King. 1999. Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *Review of Economics and Statistics* 81 (4):575-593.
Bils, Mark, and Peter J. Klenow. 2004. Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy* 112 (5):947-985.
Blanchard, Olivier Jean, and Danny Quah. 1989. The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review* 79 (4):655-673.
Blinder, Alan S. 1994. On Sticky Prices: Academic Theories Meet the Real World. In *Monetary policy*, edited by N. G. Mankiw. Chicago and London: University of Chicago Press.
Blinder, Alan S., Elie R. D. Canetti, Devid E. Lebow, and Jeremy B. Rudd. 1998. *Asking about Prices: A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.
Burda, Michael, and Charles Wyplosz. 1998. *Macroeconomics: A European Text*. 2nd ed. Oxford: Oxford University Press.
Camerer, Colin, Linda Babcock, George Loewenstein, and Richard H. Thaler. 1997. Labor Supply of New York City Cabdrivers: One Day at a Time. *Quarterly Journal of Economics* 112 (2):407-41.
Campbell, John Y., and N. Gregory Mankiw. 1987. Are Output Fluctuations Transitory? *Quarterly Journal of Economics* 102 (4):857-880.
Canova, Fabio, and Gianni de Nicoló. 2003. On the Sources of Business Cycles in the G-7. *Journal of International Economics* 59 (1):77-100.
Carlton, Dennis W. 1986. The Rigidity of Prices. *American Economic Review* 76 (4):637-658.
Cecchetti, Stephen G. 1986. The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines. *Journal of Econometrics* 31 (3):255-274.
Christiano, Lawrence J., and Martin Eichenbaum. 1993. Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations. *American Economic Review* 83 (3):430-450.
Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. Monetary Policy Shocks: What Have We Learned and to What End? In *Handbook*

of Macroeconomics, edited by J. B. Taylor and M. Woodford. Amsterdam: Elsevier Science.
Cochrane, John H. 1988. How Big Is the Random Walk in GNP? *Journal of Political Economy* 96 (5):893-920.
Cooper, Russell, and John Haltiwanger. 1996. Evidence on Macroeconomic Complementarities. *Review of Economcs and Statistics* 78 (1):78-93.
Cooper, Russell, and Andrew John. 1988. Coordinating Coordination Failures in Keynesian Models. *Quarterly Journal of Economics* 103 (3):441-63.
De Long, J. Bradford, and Robert Waldmann. 1997. Interpreting Procyclical Productivity: Evidence from a Cross-Nation Cross-Industry Panel. *Federal Reserve Bank of San Francisco Economic Review* 1997 (1):33-52.
Dias, Mónica, Daniel Dias, and Pedro D. Neves. 2004. Stylized Features of Price Setting Behaviour in Portugal: 1992-2001. Frankfurt: European Central Bank.
Enders, Walter. 1995. *Applied Time Series Analysis*. New York: Wiley.
Fabiani, Silvia, Angela Gattulli, and Roberto Sabbatini. 2004. The Pricing Behaviour of Italian Firms: New Survey Evidence of Price Stickiness. Frankfurt: European Central Bank.
Fabiani, Silvia, Claire Loupias, Martine Druant, Ignacio Hernando, Claudia Kwapil, Bettina Landau, Fernando Martins, Thomas Mathä, Roberto Sabbatini, Harald Stahl, and Ad Stokman. 2007. Summary of Results for the Euro Area. In *Pricing Decisions in the Euro Area*, edited by S. Fabiani, C. Loupias, F. Martins and R. Sabbatini. Oxford: Oxford University Press.
Fabiani, Silvia, Claire Loupias, Fernando Martins, and Roberto Sabbatini, eds. 2007. *Pricing Decisions in the Euro Area*. Oxford: Oxford University Press.
Farber, Henry S. 2005. Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers. *Journal of Political Economy* 113 (1):46-82.
Francis, Neville, and Valerie A. Ramey. 2005. Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited. *Journal of Monetary Economics* 52 (8):1379-99.
Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States, 1867-1960*. Princeton, N.J.: Princeton University Press.
Galí, Jordi. 1992. How Well Does the IS-LM Model Fit Postwar U.S. Data? *Quarterly Journal of Economics* 107 (2):709-738.
———. 1999. Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review* 89 (1):249-271.
Goodhart, C. A. E. 1989. *Money, Information and Uncertainty*. 2nd ed. Cambridge, Mass.: MIT Press.
Gray, JoAnna. 1978. On Indexation and Contract Length. *Journal of Political Economy* 86 (1):1-18.
Hall, Robert E. 1986. Market Structure and Macroeconomic Fluctuations. *Brookings Papers on Economic Activity* 0 (2):285-322.
———. 1990. Invariance Properties of Solow's Productivity Residual. In *Growth, Productivity, Unemployment*, edited by P. Diamond. Cambridge, Mass.: MIT Press.
Hamilton, James D. 1983. Oil and the Macroeconomy since World War II. *Journal of Political Economy* 91 (2):228-48.

———. 1996. This Is What Happened to the Oil Price-Macroeconomy Relationship. *Journal of Monetary Economics* 38 (2):215-20.
Hansen, Gary D., and Randall Wright. 1992. The Labor Market in Real Business Cycle Theory. *Federal Reserve Bank of Minneapolis Quarterly Review* 16 (2):2-12.
Hartley, Peter R., and Joseph A. Whitt, Jr. 2003. Macroeconomic Fluctuations: Demand or Supply, Permanent or Temporary? *European Economic Review* 47 (1):61-94.
Hooker, Mark A. 1996. This Is What Happened to the Oil Price-Macroeconomy Relationship: Reply. *Journal of Monetary Economics* 38 (2):221-22.
———. 1996. What Happened to the Oil Price-Macroeconomy Relationship? *Journal of Monetary Economics* 38 (2):195-213.
Jonker, Nicole, Carsten Folkertsma, and Harry Blijenberg. 2004. An Empirical Analysis of Price Setting Behaviour in the Netherlands in the Period 1998-2003 Using Micro Data. Frankfurt: European Central Bank.
Kashyap, Anil K. 1995. Sticky Prices: New Evidence from Retail Catalogs. *Quarterly Journal of Economics* 110 (1):245-74.
Klein, Lawrence R., and Edwin Burmeister, eds. 1974. *Econometric Model Performance: Comparative Simulation Studies of the U.S. Economy*. Philadelphia: University of Pennsylvania Press.
Klein, Lawrence R., and Arthur S. Goldberger. 1955. *An Econometric Model of the United States, 1929-1952*. Edited by J. Tinbergen, P. J. Verdoorn and H. J. Witteveen. Vol. 9, *Contributions to Economic Analysis*. Amsterdam: North-Holland.
Klenow, Peter J., and Oleksiy Kryvtsov. 2008. State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation? *Quarterly Journal of Economics* 123 (3):863-904.
Klenow, Peter J., and Jonathan L. Willis. 2007. Sticky Information and Sticky Prices. *Journal of Monetary Economics* 54S:S79-S99.
Kydland, Finn E., and Edward C. Prescott. 1982. Time to Build and Aggregate Fluctuations. *Econometrica* 50 (6):1345-70.
Levy, Daniel, Mark Bergen, Shantanu Dutta, and R. Venable. 1997. The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains. *Quarterly Journal of Economics* 112 (3):791-825.
Loupias, Claire, and Roland Ricart. 2004. Price Setting in France: New Evidence from Survey Data. Frankfurt: European Central Bank.
Lucas, Robert E., Jr. 1973. Some International Evidence on Output-Inflation Tradeoffs. *American Economic Review* 63 (3):326-334.
Mankiw, N. Gregory, and Ricardo Reis. 2002. Sticky Information and Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics* 117 (4):1295-1328.
Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. 2003. Disagreement about Inflation Expectations. *NBER Macroeconomics Annual* 18:209-248.
Mishkin, Frederic S. 1982. Does Anticipated Monetary Policy Matter? An Econometric Investigation. *Journal of Political Economy* 90 (1):22-51.
Nelson, Charles R., and Charles I. Plosser. 1982. Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications. *Journal of Monetary Economics* 10 (2):139-162.

Plosser, Charles I. 1989. Understanding Real Business Cycles. *Journal of Economic Perspectives* 3 (3):51-77.
Romer, Christina D. 1986. Is the Stabilization of the Postwar Economy a Figment of the Data? *American Economic Review* 76 (3):314-34.
———. 1999. Changes in Business Cycles: Evidence and Explanations. *Journal of Economic Perspectives* 13 (2):23-44.
Romer, Christina D., and David H. Romer. 1989. Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. In *NBER Macroeconomics Annual 1989*, edited by O. J. Blanchard and S. Fischer. Cambridge, Mass.: MIT Press.
Sargent, Thomas J. 1982. The Ends of Four Big Inflations. In *Inflation: Causes and Effects*, edited by R. E. Hall. Chicago: University of Chicago Press.
Shapiro, Matthew D. 1993. Cyclical Productivity and the Workweek of Capital. *American Economic Review* 83 (2):229-33.
Shea, John. 1998. What Do Technology Shocks Do? In *NBER Macroeconomics Annual 1998*, edited by B. Bernanke and J. J. Rotemberg. Cambridge, Mass.: MIT Press.
Sims, Christopher A. 1980. Macroeconomics and Reality. *Econometrica* 48 (1):1-48.
Stock, James H., and Mark W. Watson. 1999. Business Cycle Fluctuations in U.S. Macroeconomic Time Series. In *Handbook of Macroeconomics*, edited by J. B. Taylor and M. Woodford. Amsterdam: Elsevier Science.
Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen. 2004. Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets. *Review of Economics and Statistics* 86 (2):514-533.