Low Inflation: The Surprise of the 1990s

*Dean Croushore**

Over the past six years, inflation has been stable or declining in the United States. But over that period, economic forecasters continually predicted an upturn in inflation that never materialized. Is inflation now behaving differently than it did in prior decades? Or are the forecasters using inadequate models?

The persistent errors in the forecasts of inflation are disturbing. Forecasters had done an outstanding job of projecting inflation prior to the 1990s, as I described in my 1996 article, "Inflation Forecasts: How Good Are They?" Statistical tests discussed in that article verify that over a long period, forecasts of inflation were unbiased. So the recent, persistent overprediction of inflation is unusual.

Why should we be concerned about bad inflation forecasts? For one thing, the private sector uses inflation forecasts in a number of ways, from businesses that print catalogs showing their prices, to lenders who set interest rates

*Dean Croushore is an assistant vice president and economist in the Research Department of the Federal Reserve Bank of Philadelphia. He is also head of the department's macroeconomics section.

depending on what they think inflation will be in the future. If those inflation forecasts aren't very good, businesses and lenders will set prices incorrectly, and such a mistake may be costly to them and their customers.

In addition, inflation forecasts provide important information to policymakers who are formulating monetary policy today. If forecasts indicate a rise in inflation (if policy doesn't change), policymakers may wish to tighten monetary policy now to keep inflation from rising. Looking at forecasts helps policymakers address the problem that monetary policy actions affect the economy with a lag: actions today affect the inflation rate one to two years from now. So basing policy actions on forecasts would be desirable *if* the forecasts were accurate. Unfortunately, the persistent errors in the forecasts of inflation cast doubt on the value of using those forecasts as a basis for making policy.

Given the potential importance of inflation forecasts for policymakers and the good track record of forecasters prior to the 1990s, what explains the erroneous forecasts of the last six years? Has something changed in the way forecasters make their predictions? Or has the inflation process itself changed?

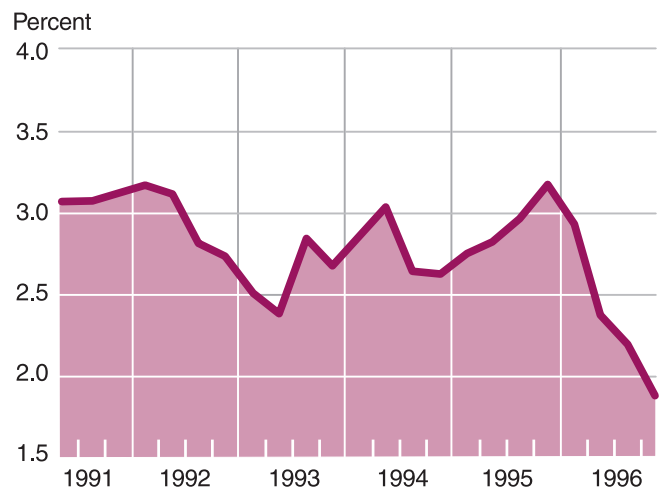
RECENT INFLATION AND FORECASTS

Inflation Has Been Surprisingly Stable . . . Perhaps the most remarkable aspect of the economic expansion that's been going on since 1991 is that the inflation rate held fairly steady for several years and has been declining recently (Figure 1).[1]

The fact that the inflation rate has not risen over the period shown in the figure is unusual, at least when compared to recent experience. Only two other economic expansions since World War II lasted at least six years (1961 to 1969 and 1983 to 1990), and in both, inflation was accelerating significantly by the sixth year. In the 1961-69 expansion, the average annual inflation rate rose from 1 percent in 1961 to about 3 percent in 1967 and to more than 5 per-

[1]The figure shows the percentage change in the Consumer Price Index for All Urban Consumers (CPI-U), which is the most commonly forecasted measure of consumer prices. It is averaged over a year to give a better picture of the trend in inflation, since inflation data vary a lot from month to month. The data points are four-quarter-ahead moving averages; for example, the data point plotted for 1991Q2 is the inflation rate in the CPI-U from 1991Q2 to 1992Q2.

FIGURE 1
The Inflation Rate
Four-Quarter-Ahead Average



cent in 1969. In the 1983-90 expansion, inflation rose from 3 percent in 1983 to about 5 percent in 1990. But since the current economic expansion began in 1991, inflation has remained remarkably steady, about 3 percent or less.

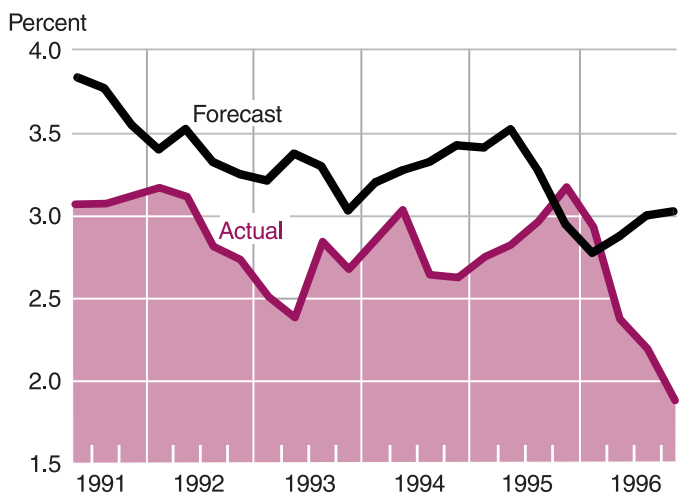
. . .But Forecasters Predicted That Inflation Would Rise. Over the past six years, forecasters have predicted an uptick in inflation, but they've been consistently wrong. The typical forecast during this expansion held that inflation would creep up by about one-half of a percentage point over the coming year, for example, from 3 to 3.5 percent.

Comparing forecasts of inflation since the current expansion began with actual inflation shows how persistently the forecasters have missed the mark (Figure 2). To interpret the figure, look at the values for the second quarter of 1991 (where the "Forecast" and "Actual" lines begin). The value shown (3.8 percent) on the "Forecast" line is the inflation forecast made in the second quarter of 1991 for the average inflation rate from that date to the second quarter of 1992.[2] The value shown (3.1 percent) on the "Actual" line is the actual inflation rate from the second quarter of 1991 to the second quarter of 1992.

The line showing the actual inflation rate is almost always below the line showing the forecast, which means forecasters persistently predicted a rise in the inflation rate that never materialized. The average forecast error (the forecast of the inflation rate minus the actual inflation rate) is about 0.4 percentage points. Actual inflation was 2.9 percent on average, while the forecast of inflation was 3.3 percent on average.[3]

Examining the quarterly pattern of the forecasts in more detail shows even more clearly that forecasters expected inflation to rise. A typical example can be seen in the forecast that was made in the second quarter of 1992 (Figure 3).

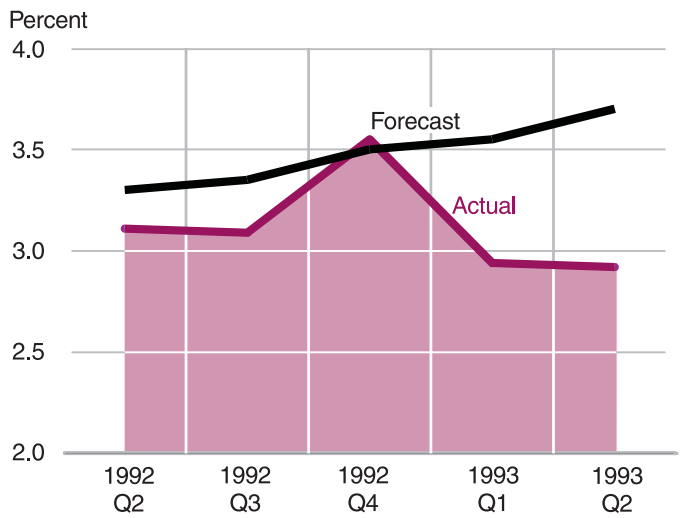
FIGURE 2
The Inflation Rate
Four-Quarter-Ahead Average



[2]The forecasts come from the *Survey of Professional Forecasters*, a quarterly survey produced by the Federal Reserve Bank of Philadelphia that covers a wide variety of macroeconomic variables, including consumer price inflation. The participants are forecasters on Wall Street, at banks, at corporations, and in consulting firms. For more details on the survey, see my 1993 article, "Introducing: The Survey of Professional Forecasters," or visit the Philadelphia Fed's Web site at 'http://www.phil.frb.org/econ/spf/spfpage.html'.

[3]The only time the forecast was below actual inflation occurred in 1995Q4 and 1996Q1, because unexpectedly high increases in food and energy prices in 1996 caused the actual inflation rate to exceed the forecast. But in every other quarter, the forecast exceeded the actual inflation rate.

FIGURE 3
The Inflation Rate
Actual vs. Forecast/Quarterly



In this diagram, the points plotted are the median forecasts for the inflation rate for the current quarter (1992Q2) and each of the next four quarters, compared with the actual inflation rate in each quarter. The figure shows that inflation, which had been 2.7 percent in the first quarter of 1992, was expected to rise gradually over time. The forecasters thought inflation would be 3.3 percent in 1992Q2, 3.35 percent in 1992Q3, 3.5 percent in 1992Q4, 3.55 percent in 1993Q1, and 3.7 percent in 1993Q2. In fact, inflation turned out to be lower than forecast in every quarter except 1992Q4.

WHY WERE THE FORECASTS SO HIGH?

We can't know for sure why forecasts for inflation have been consistently too high in the 1990s. To determine possible reasons, let's look at four different methods for forecasting inflation to see what forecasters might have missed: (1) the cost model, in which increases in the costs of producing goods translate into higher prices; (2) historical correlations, in which forecasters look at how inflation typically moves over the course of the business cycle; (3) demand-side and supply-side models, in which forecasters examine changes in aggregate demand and supply to determine their impact on inflation; and (4) the monetary model, in which inflation is directly affected by the degree of tightness of monetary policy.

The Cost Model. One method of forecasting inflation is to examine the cost of producing goods. According to this theory, if production costs begin to grow at a faster rate, firms pass on the higher costs in the form of higher prices of goods, and the inflation rate increases. The question then is, what causes production costs to rise faster? One explanation is based on an economic construct called the Phillips curve, which relates inflation to unemployment. When labor markets get tight—that is, there aren't many qualified workers available—firms increase wages more rapidly, and the cost of producing goods rises faster. This theory leads to the notion that the economy has a natural rate of unemployment, or a non-accelerating-inflation rate of unemployment (NAIRU). The Phillips curve model suggests that when the nation's unemployment rate is less than the natural rate, inflation will rise. Similarly, if unemployment rises above the natural rate, inflation will decline.

With this type of model, it's clear why fore-

casters were predicting that inflation would rise from 1994 on (though not before that). Most estimates in the early 1990s set the natural rate of unemployment at about 6 percent. The unemployment rate fell below that level in 1994. Consequently, many forecasters became alarmed about the prospects of a rise in inflation at that time. But the unemployment rate continued to decline. After averaging 6.1 percent in 1994, it fell to an average of 5.6 percent in 1995, then to 5.4 percent in 1996, and it was below 5.0 percent through most of 1997. Those forecasters who used a Phillips curve model thus produced forecasts of an increasing inflation rate.

However, as inflation continued to be benign, forecasters began to rethink their views on the natural rate of unemployment. They observed both low inflation and significant changes in the labor force, which made them realize that the natural rate of unemployment was probably less than 6 percent. So, over time, a number of forecasters have lowered their estimate of the natural rate of unemployment, and they are no longer predicting as large a rise in inflation.[4]

The fact that forecasters began to change their views about the level of the natural rate of unemployment led economists Doug Staiger, Jim Stock, and Mark Watson to investigate how the natural rate should be estimated. In their paper "The NAIRU, Unemployment, and Monetary Policy," they found that estimates of the natural rate are quite imprecise. A forecaster has no more basis for using an estimate of the natural rate of unemployment of 6 percent than one of 5.5 percent. The natural rate may be somewhere in that neighborhood, but those estimates could easily be off by a percentage point or more.

How does uncertainty about the natural rate of unemployment affect the models of the Phillips curve that forecasters use? If something happens in the economy that reduces the natural rate, but the forecasters are unaware of this event, their models will use too high a value for the natural rate. As the unemployment rate drops below their estimate of the natural rate, forecasters may think inflation will rise, but they would be wrong.

What factors might have caused the natural rate of unemployment to decline? One notable feature of the 1990s is an increased willingness by corporations to lay off workers, especially at the managerial level (for evidence, see the article by Rob Valletta). This change, in turn, has affected workers' attitudes toward their jobs and led them to reduce demands for higher wages for fear of being "downsized." Such a change in attitude is likely to have reduced the natural rate of unemployment, since workers won't demand higher wages even when the unemployment rate is very low.

Thus, one explanation of the errors in forecasting inflation, at least since 1994, may be the failure of forecasters to modify their Phillips curve models of inflation to reflect a lower natural rate of unemployment.

One other major influence on firms' production costs that may also have played a role in the 1990s in keeping production costs down is the effort to control the cost of health benefits. While benefit costs rose much more rapidly

[4]The natural rate of unemployment represents the unemployment rate when the economy is operating at full capacity. Two groups of people would still be unemployed in such a situation: (1) those who may not have the skills needed for employment (for example, people who have worked in an industry that's shrinking and who need additional training to get another job); and (2) those who are simply between jobs but unlikely to be unemployed for long. Both structural (the first case) and frictional (the second case) unemployment are necessary consequences of a growing economy in which there's always change and technical progress. The natural rate of unemployment can change whenever either structural or frictional unemployment changes. For example, the natural rate of unemployment rose in the late 1970s because the oil price shocks to the economy reduced the need for workers in industries, such as the auto industry, that depend on low oil prices.

than wages and salaries from 1988 to 1994, the situation has been reversed over the past three years, mostly because of changes in health benefits. So firms' costs haven't been rising as rapidly on the health-benefits front, thus putting less pressure on firms to raise prices. To the extent that forecasters haven't accounted for the decline in health costs, their inflation forecasts may be off the mark.

Historical Correlations. Some forecasting models are based on what many people assume to be a fact—that inflation tends to rise as expansions get longer. As we've already seen, in the long expansions of the 1960s and 1980s, inflation accelerated as time passed. As a result, many economists take it on faith that inflation rises as an economic expansion continues.

But just because our history contains two episodes in which inflation rose as the economic expansion continued doesn't mean it must always be so. Indeed, some recent empirical research challenges this notion. In a 1991 article, "The Cyclical Behavior of Prices," Tom Cooley and Lee Ohanian found many periods in U.S. history in which inflation didn't rise during expansions.[5] In fact, they found that it isn't clear whether the best forecast is for inflation to rise or to fall during expansions. This research surprised many economists who had taken the rise of inflation during expansions as fact.[6]

This line of research makes it clear that if forecasters assume that inflation will rise as an economic expansion goes on, they are ignoring important evidence to the contrary. They need to look deeper and investigate the roles of supply and demand in the overall economy.

Demand- and Supply-Side Factors. The issue of whether an expansion is primarily driven by increases in demand or supply is crucial in determining what happens to inflation. Many inflation models suggest that inflation (or the increase in the inflation rate) depends on how fast the economy is growing relative to potential growth. So if the economy's potential growth rate changes, the economy can grow faster without higher inflation. What may have happened in the 1990s is that the potential growth rate of the economy increased, but forecasters didn't recognize it. As a result, they thought the economy was growing faster than it should have, so they thought inflation would rise.

One view of the economy holds that when demand outstrips supply at existing prices, prices are bid up to a higher level until demand equals supply. Thus, if supply is stable and some factor increases demand, we'd expect inflation to rise; if some factor reduces demand, we'd expect inflation to decline. Thus, it isn't really supply or demand alone that determines inflation, but a combination of the two.

In the 1990s, as the economy has grown faster than expected, forecasters may have thought that demand for goods was outstripping sup-

[5]Examining the movement of inflation during expansions involves identifying cycles in inflation relative to its trend, which can be a tricky business. The idea is that if prices are growing 3 percent, on average, over time, we want to see if they're above or below their trend line in expansions. If the trend isn't stable over time, a number of alternative methods can be used to estimate the trend, but it isn't clear which method is best.

[6]Not all economists believe that Cooley and Ohanian's research is the final word. First, Bankim Chadha and Eswar Prasad found evidence that inflation did rise, on average, during expansions. But their results depend on this tricky business of taking out the trend, as discussed in footnote 5. John Judd and Bharat Trehan suggested that looking at simple correlations between inflation and output (as Cooley and Ohanian, as well as Chadha and Prasad, had done) is misleading. Using a simple example, they showed that such a correlation can't answer the question of whether inflation rises in expansions, because that correlation depends on the timing of the movements in output and prices, not their overall direction. Instead, a more detailed statistical analysis is needed, focusing on whether supply-side factors (what people produce) or demand-side factors (what people buy) are the dominant force in the expansion.

ply, and thus they would expect an increase in the inflation rate. For example, many forecasters thought the economy would grow at a rate of 2 to 2.5 percent in 1995, 1996, and 1997. They were about right in 1995, when the economy grew 2 percent. But in 1996, the economy grew nearly 3 percent, and it grew nearly 4 percent in 1997. So forecasters may have thought that inflation was likely to increase over the past two years, since demand was outstripping supply. In fact, inflation declined between 1995 and 1997, so supply-side factors must have come into play.

The three most important supply-side factors are growth of the labor force, productivity, and foreign competition. All have changed in ways that increased supply in the mid-1990s.

As an example of a surprising change in the growth of the labor force, consider the welfare reform bill of 1996. This bill led to an increase in the labor force that carried over into 1997: people at the lower end of the income scale entered the workforce in increased numbers. An estimate by Mary Daly suggests that the welfare reform bill caused the labor force to increase by 300,000 within the first year.

Economists had been predicting that growth in the labor force would slow down beginning in 1996; some economists even predicted that the labor force would grow by less than 1 million people in 1996 and again in 1997. In fact, the rise in the number of people in the labor force exceeded 2 million and employment grew even more—over 2.5 million—in both years. This surprising growth in the number of people working was a major factor behind the fast pace of the economy in 1996 and 1997.

Another supply-side source of economic growth is productivity. As workers become more productive, the economy's output rises. A look at the statistics is discouraging, since our national income accounts show productivity growing at a rather slow rate over the past 20 years. But is it really growing that slowly? To the casual observer, productivity appears to be exploding, especially since computer technologies are making rapid advances. So why are the statistics on productivity so bleak?

One theory is that our official statistics mismeasure productivity growth. It's hard to calculate how valuable new computers are, and it's difficult to evaluate the improvement in productivity that a new, but expensive, piece of medical equipment generates. And it's nearly impossible to evaluate productivity gains in the service sector. How much more does a lawyer do now than she did 10 years ago? How much has her ability to investigate legal questions improved in the on-line age? Does her constant availability by pager and cell phone make her more productive for her clients?

Economists like Leonard Nakamura think that these measurement issues are of primary importance. In his 1997 article, "Is the U.S. Economy Really Growing Too Slowly? Maybe We're Measuring Growth Wrong," he argues that, in fact, the U.S. economy has been doing much better than the official statistics show. According to Nakamura, the government data overstate inflation and understate productivity and economic growth. Over the last 20 years, U.S. output and productivity growth may have been as much as two percentage points higher per year than the official statistics indicate. And the mismeasurement has been increasing in the last few years. If that's right, the economy should be able to grow rapidly without causing inflation.

A final supply-side factor comes from foreign competition. In the 1990s, U.S. firms, especially in the manufacturing sector, have competed in increasingly global markets. Increased competition from foreign firms may have caused firms to set their prices lower and reduced profit margins (compared with what they would have been in the absence of that competition).

In addition, since early 1995 the dollar has been rising, on average, against other currencies. The rise in the dollar reduces the price of

imports into the United States, so U.S. firms that compete in the same markets as foreign firms must cut their prices to remain competitive. Again, the effect is likely to help reduce inflation.

All of these supply-side factors have been pushing inflation down in the 1990s, especially over the last few years. To the extent that forecasters didn't anticipate these factors, their inflation forecasts were too high.

The Monetary and Financial Model. One final method used to forecast inflation is the monetary model, which is best described by Milton Friedman's famous maxim, "Inflation is always and everywhere a monetary phenomenon." According to Friedman and other monetarists, what matters for inflation is not the growth rate of the economy but the growth rate of the money supply. Forecasters from the monetarist school look at the growth rate of the money supply as an indicator of future inflation.

These days, financial innovation has made it difficult to interpret the growth rate of the money supply itself, but other economic variables provide some indication of whether monetary policy is easy or tight.[7] One indicator of the stance of monetary policy is the difference between the interest rate on the 10-year U.S. government bond and the interest rate on three-month Treasury bills. The bigger this difference, the easier the stance of monetary policy. Since 1994 this difference has been lower than it usually is in expansions (Figure 4).

Another important indicator of monetary policy is the real federal funds rate—the nominal interest rate on overnight loans between banks minus the expected inflation rate. The higher the real federal funds rate is, the tighter monetary policy is. Again, the data show that the real federal funds rate has been a bit higher over the past few years than its average in the 1970s (Figure 5).[8] That's good, because the

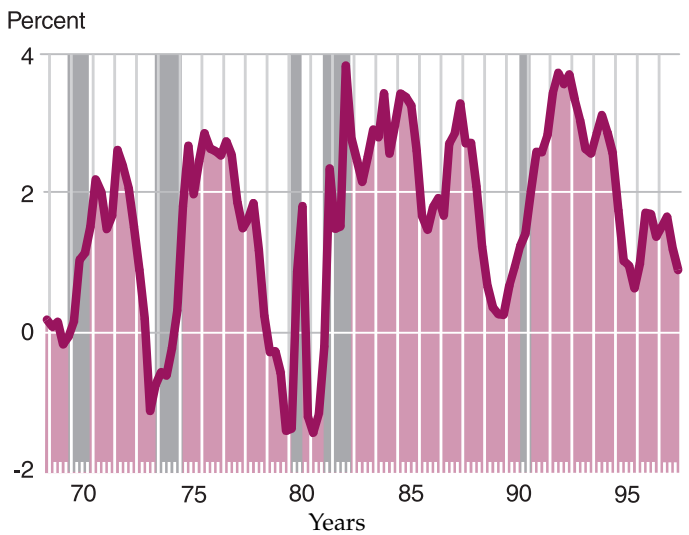
[7]For example, in 1996, the M1 measure of the money supply (which includes currency and deposits in checking accounts) *fell* 4.6 percent while the M2 measure of the money supply (which includes M1, savings accounts, and small savings deposits) *grew* 4.3 percent.

[8]The figure shows the federal funds rate minus the one-year-ahead expected inflation rate from the *Survey of Professional Forecasters.*

FIGURE 4

Interest Rate Spread

Interest Rate on 10-Year Treasury Bond Minus Interest Rate on 3-Month Treasury Bill



Source: Federal Reserve series of constant maturities, secondary markets.

1970s were a time when inflation accelerated. But the real federal funds rate in the 1990s isn't quite as high as it was in the 1980s. In the early 1980s, when the real federal funds rate was very high, inflation declined significantly. Later in the 1980s, when the real federal funds rate was about as high as it is now, inflation was fairly stable.

All these indicators point to tighter monetary policy in the 1990s than is usual in expansions. Thus monetary policy itself may be at least partly responsible for keeping inflation from rising. To the extent that forecasters misread the stance of monetary policy or relied too much on historical correlations or the Phillips curve, their forecasts called for higher inflation.

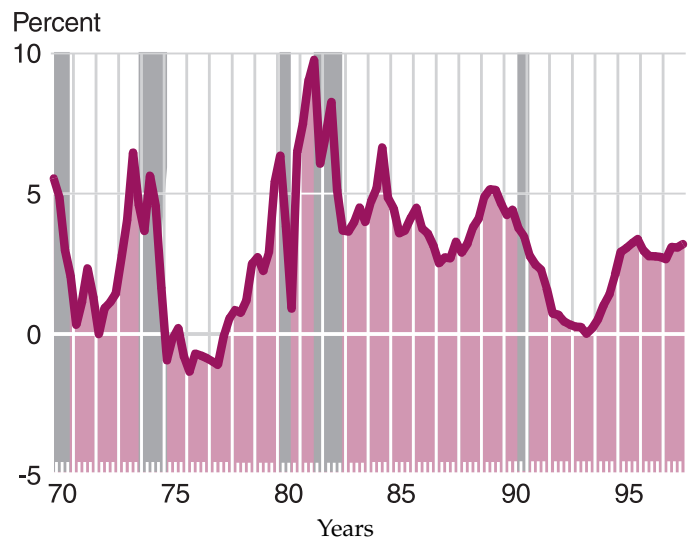
CONCLUSION

We've seen that inflation has been much lower in the 1990s than forecasters expected. Over the last six years, forecasters have continually predicted that inflation would rise, but it hasn't. It's difficult to figure out the exact source of their forecasting errors, but it's likely to be a combination of many factors. They may have based their forecasts on a cost view of inflation, using a Phillips curve, but not realized that the natural rate of unemployment was declining. They may have looked at historical correlations and assumed (incorrectly) that inflation always rises in expansions. They may have failed to take supply-side factors, such as increased growth in the labor supply, increased productivity, and foreign competition, sufficiently into account. And they may have failed to account completely for the degree of tightness of monetary policy.

With all of these considerations affecting inflation and forecasts of inflation, how should monetary policymakers react? The fact that inflation forecasts have been too high makes policymakers wary of basing decisions solely on such forecasts.

But if policymakers don't use forecasts, what do they do? They monitor the tightness of monetary policy, using a variety of financial indicators. They keep an eye on both supply and demand factors and do not assume that a rise in demand will necessarily bring higher inflation. Finally, policymakers realize that while Phillips curves and other methods of examining production costs have proven useful historically, they are difficult to use for forecasting inflation.

FIGURE 5
Real Federal Funds Rate



Source: Federal funds rate adjusted by one-year-ahead expectations of GDP price index from the *Survey of Professional Forecasters.*

REFERENCES

Chadha, Bankim, and Eswar Prasad. "Are Prices Countercyclical? Evidence from the G-7," *Journal of Monetary Economics* 34 (October 1994), pp. 239-57.

Cooley, Thomas F., and Lee E. Ohanian. "The Cyclical Behavior of Prices," *Journal of Monetary Economics* 28 (1991), pp. 25-60.

Croushore, Dean. "Introducing: The Survey of Professional Forecasters," Federal Reserve Bank of Philadelphia *Business Review*, November/December 1993.

Croushore, Dean. "Inflation Forecasts: How Good Are They?" Federal Reserve Bank of Philadelphia *Business Review,* May/June 1996.

Daly, Mary. "Labor Market Effects of Welfare Reform," Federal Reserve Bank of San Francisco *Economic Letter*, No. 97-24, August 29, 1997.

Judd, John P., and Bharat Trehan. "The Cyclical Behavior of Prices: Interpreting the Evidence," *Journal of Money, Credit and Banking* 27 (August 1995), pp. 789-97.

Nakamura, Leonard. "Is the U.S. Economy Really Growing Too Slowly? Maybe We're Measuring Growth Wrong," Federal Reserve Bank of Philadelphia *Business Review,* March/April 1997.

Staiger, Douglas, James H. Stock, and Mark W. Watson. "The NAIRU, Unemployment, and Monetary Policy," *Journal of Economic Perspectives* 11 (Winter 1997), pp. 33-49.

Valletta, Rob. "Job Loss During the 1990s," Federal Reserve Bank of San Francisco *Economic Letter* 97-05, February 21, 1997.

Trends in Metropolitan Employment Growth

*Gerald A. Carlino**

Although metropolitan areas account for only 16 percent of the total land area in the United States, they contain almost 80 percent of the nation's population and nearly 85 percent of its jobs. The United States has, on average, 24 jobs per square mile, but metropolitan areas average about 124 jobs per square mile.

The standard explanation for why firms locate in metropolitan areas is that they can lower their production costs by taking advantage of agglomeration economies—efficiency gains and cost savings that result from being close to suppliers, workers, customers, and even competitors. Although population and jobs have grown more within metropolitan areas than outside them, growth has favored smaller metropolitan areas. During the second half of the 20th century, employment has become more evenly distributed across metropolitan areas. Some observers claim that this deconcentration of people and jobs is the result of a greater preference for less urbanized living. Others say it's the result of reductions in urban agglomeration economies due to technological change and government policies, such as the building of interstate highways.

*Jerry Carlino is an economic advisor in the Research Department of the Philadelphia Fed.

The analysis presented in this article suggests a third factor: that locating new jobs in the more densely populated metropolitan areas is more expensive because these areas are nearer to using the full capacity of local resources. Adding jobs and people burdens existing support systems, leading to increases in traffic congestion, pollution, and the cost of living. These "congestion costs" are a major cause of the relatively slower growth of the largest metro areas. If the costs of congestion increase proportionately more for the larger and more dense metro areas, growth rates will be greater in the smaller and less dense metro areas.

A study undertaken at the Federal Reserve Bank of Philadelphia used a simple model to account for the postwar growth in employment in metropolitan statistical areas (MSAs). In this model, firms benefit from agglomeration economies that increase with a metro area's employment size but face congestion costs that rise more than proportionately with its density. The study found that metro areas that were less densely populated in 1951 were able to accommodate faster postwar employment growth than denser metro areas. However, growth in the densest metropolitan areas slowed less than the model predicted. This better-than-predicted growth may be the result of an ongoing process of technological change that put these denser areas technologically ahead of less dense ones. This reasoning suggests that technological change may have offset some of the effects of higher congestion costs. In fact, rather than reducing growth in the densest metro areas, as some have suggested, technological change may have promoted it.

TRADEOFF BETWEEN AGGLOMERATION ECONOMIES AND CONGESTION COSTS

Agglomeration Economies Lead to Concentration... Agglomeration economies provide a powerful incentive for the concentration of economic activity. Historically, manufacturing activity has tended to concentrate in certain areas as a means to hold down costs. Nonmanufacturing activities (such as banking, wholesale and retail trade, and services) have found it advantageous to join the cluster, supplying business services to firms or consumer services to residents. Some nonmanufacturing firms have also found it advantageous to locate near other firms in their own industry (e.g., investment banking in New York City and motion pictures in Los Angeles). Consequently, people and jobs have become concentrated in these areas.

An earlier study found that agglomeration economies for manufacturing firms tend to increase with MSA size, up to some point.[1] For example, a 1 percent increase in all inputs of production resulted in an increase of more than 1 percent in output in Peoria (1.9 percent), Cincinnati (1.4 percent), Kansas City and St. Louis (1.3 percent each), and Boston (1.2 percent). But estimates for Philadelphia, the fifth largest MSA in terms of population, indicate that the same 1 percent increase in inputs leads to an increase of only 1 percent in output.[2]

Why would a large MSA such as Philadelphia, which contained almost 5 million people in 1997, offer a smaller return to its manufacturing firms, on average, than Peoria, which contained less than 350,000 people in 1997? The answer lies in the costs to both firms and households that result from increased urban size.

...But Congestion Costs Pave the Way for Deconcentration.[3] The positive effects of agglomeration economies make up one side of the urban size ledger; the negative effects of congestion (more traffic and pollution and higher

[1]See the 1982 article by Gerald Carlino.

[2]These numbers cover the years 1957-77.

[3]Deconcentration refers to a more even distribution of employment among metropolitan areas.

housing costs) make up the other. To offset these higher congestion costs, workers must receive higher wages, and higher wages increase costs to firms. If congestion costs increase proportionately more for denser metropolitan areas, the same percentage increase in jobs raises wages more in dense locations than in less dense ones.

A 1998 study by Gerald Carlino and Satyajit Chatterjee provides a perspective on the cost of increased employment density in MSAs. They found that a 1 percent increase in employment density increased the cost of living 2.1 percent in the Jersey City MSA (the most dense MSA in the study), almost 0.3 percent in the Philadelphia MSA, and 0.04 percent in the Peoria MSA. Thus, the natural growth of the economy over the long run would lead to slower growth of the densest metro areas and faster growth in the less dense areas of the country. Called convergent growth, this process tends to equalize densities across metropolitan areas (see *Size Versus Density*).

EMPLOYMENT: A PATTERN OF DECONCENTRATION

The concentration of national employment decreased from the 1950s to the 1990s. On the basis of employment density in 1951, the top 1 percent of MSAs (or three most dense MSAs) accounted for 14 percent of total employment in the United States.[4] By 1994, the top 1 percent accounted for just 5 percent. Also in 1951, the top 10 percent of MSAs (or 30 most dense) accounted for 42 percent of total employment in the United States. By 1994, the top 10 percent accounted for only 29 percent. On the other hand, the bottom 30 percent of MSAs (or the 90 least dense) accounted for less than 4 percent of total employment in 1951. By 1994, the share of the bottom 30 percent had increased to 7 percent.

We can present the inequality of employment density among MSAs graphically by using a Lorenz curve. If employment were distributed equally across MSA land area, the Lorenz curve would be a diagonal straight line showing, for example, that any group of MSAs that contained 20 percent of total MSA land area would account for 20 percent of total MSA employment. In reality, employment is distributed unequally, resulting in the real-world Lorenz curves, which are bowed above the diagonal line (Figure 1).[5] The more unequal the employment density, the more pronounced that bowed effect will be. The figure shows Lorenz curves for 1951, 1959, 1969, 1979, 1989, and 1994. The one for 1951 is farthest from the diagonal. Over time, the Lorenz curves have moved toward more equal distribution of employment across MSAs.

Another measure, the Theil index, gauges inequality in employment density among MSAs and summarizes it in a single number. Lower values of the index are associated with less inequality. The total inequality among MSAs fell from about 1.6 in 1951 to just below 1 in 1994, a decline of about 38 percent (Figure 2).

The Theil index can be broken down to show inequality between MSAs and within MSAs. The index of inequality within MSAs is a rough

[4]The three were Jersey City, New York, and Chicago. The discussion in the remainder of this article is based on the 1998 study by Satyajit Chatterjee and Gerald Carlino. The employment data are taken from the Census Bureau's *County Business Patterns* for six years: 1951, 1959, 1969, 1979, 1989, and 1994. The official definition of a metropolitan area has changed several times since 1950; thus, this article looks at employment density (employment divided by square miles of land area) for 297 MSAs, based on constant 1983 MSA definitions. In general, MSAs are statistical constructs used to represent integrated labor-market areas that consist of counties containing a central city of at least 50,000 people along with any contiguous counties if such counties meet certain economic considerations.

[5]The Lorenz curves shown in Figure 1 plot the distribution of total *MSA* employment only, not total *U.S.* employment.

Size Versus Density

Sometimes economists have looked at the size (population or number of jobs) of metro areas when considering the benefits and costs of urbanization. A study of urban areas in France and Japan by Jonathan Eaton and Zvi Eckstein found that all cities grow at the same rate regardless of initial population size. Duncan Black and Vernon Henderson also found evidence of parallel growth for cities in the United States in that the relative size distribution of cities was unchanged during the period 1900-50.*

But the cost of urban growth may be related to the density of development rather than some measure of the size of development as in the studies by Eaton and Eckstein and Black and Henderson. Consider two cities, A and B, of equal population size, but A has twice the land area of B. In this case, B has twice the population density that A has, and many of the problems associated with increasing density (such as traffic and pollution) are likely to be greater in B, too. Thus, size alone may not be enough to gauge the costs of development. Population or employment density may be a better measure.

The ranking of MSAs based on employment size can differ markedly from their rankings based on density (see table below). For example, the Jersey City MSA ranked first in employment density in both 1951 and 1994, but it ranked 27th in level of MSA employment in 1951 and 85th in 1994. The Las Vegas MSA ranked 296th out of 297 MSAs in terms of employment density and 243rd in employment size in 1951. But by 1994, Las Vegas had moved up to rank 50th in terms of employment size, but its density, at 237th, still ranked near the bottom of the distribution.

	1951 Employment			1994 Employment		
MSA	Density	Density Rank	Levels Rank	Density	Density Rank	Levels Rank
Jersey City, NJ	4855	1	27	4636	1	85
New York, NY	2742	2	1	1969	2	3
Chicago, IL	945	3	2	1512	3	2
Bergen-Passaic, NJ	595	4	22	1344	5	35
Newark, NJ	413	5	9	651	11	21
Trenton, NJ	351	9	76	710	10	105
Philadelphia, PA	325	10	5	549	15	4
Pittsburgh, PA	193	19	7	250	38	16
Wilmington, DE	88	49	66	242	40	73
Harrisburg, PA	54	86	60	133	103	72
Ft. Meyers, FL	6	269	292	149	92	134
Las Vegas, NV	2	296	243	52	237	50

*These findings of parallel growth for cities in the United States and France appear to be evidence against convergent growth in which the initially less dense metropolitan areas grow relatively faster than the initially more dense ones. But this difference may be more apparent than real. Both studies look at population size rather than employment density as this study does. Black and Henderson's study stops in 1950, and the period of this analysis is 1951-94. Black and Henderson's notion of "parallel growth" also involves (in part) the entry of new metropolitan areas during the 1900-50 period. Because the 1983 classification of MSAs is used for all years in this article, the number of MSAs is held constant, although some locations that became MSAs in 1983 actually had not achieved MSA status in the earlier years. Thus, Black and Henderson's findings may be consistent with our notion of employment deconcentration in that their "new" urban areas were also locations that were initially less dense.

measure of suburbanization of employment, which occurs when jobs move from the MSA's central city to its adjacent suburbs. This index fell from 0.52 in 1951 to 0.37 in 1994, a decline of almost 30 percent (Figure 2).[6] Suburbanization both of people and of jobs is a widely documented pattern in the United States. But as the Lorenz curves show, a more general pattern of deconcentration of employment among MSAs is also taking place. The index for inequality between MSAs, reflecting deconcentration, fell from 1.05 in 1951 to 0.60 in 1994, a decline of almost 43 percent.

In sum, total U.S. employment has become more evenly dispersed: the most dense MSAs account for a smaller share of employment over time. This statement should not be misconstrued to mean that the largest, most dense MSAs are losing employment. Rather, the less dense MSAs are adding jobs at a faster pace.[7]

FIGURE 1
MSA Employment Becomes More Evenly Distributed

FIGURE 2
Theil Index Shows Downward Trend In Inequality Within and Across MSAs

Index/Year	1951	1959	1969	1979	1989	1994
Theil Index for Total Inequality	1.57	1.42	1.32	1.10	1.05	0.97
Theil Index for Between-MSA Inequality	1.05	0.92	0.85	0.69	0.67	0.60
Theil Index for Within-MSA Inequality	0.52	0.50	0.47	0.41	0.38	0.37

For an explanation of the Theil index, see Edward N. Wolff, *Economics of Poverty, Inequality, and Discrimination,* South-Western College Publishing, Cincinnati, 1997.

[6]Suburbanization is understated, since county-level data are used in the analysis. Most counties that contain the central city of an MSA also contain suburbs that are near the central city. This understatement is of little concern for our purposes, since deconcentration among MSAs, not suburbanization, is the main focus of this article.

[7]Only two highly dense MSAs, New York City and Jersey City, had fewer jobs in 1994 than in 1951.

ACCOUNTING FOR DECONCENTRATION

The Traditional View. Some observers believe that the faster growth of employment in the relatively less dense MSAs is a continuation of the same forces that first gave rise to suburbanization. To them, agglomeration economies have declined because of continuing innovations in production, transportation,

and communication technologies.[8] The development of the assembly line, for example, revolutionized not only how products were manufactured but also where. Because assembly lines require a horizontal flow of goods, the vertical spaces available in city factories are unsuitable. Moreover, because the price of land is less expensive outside the city, those large open spaces provide relatively cheap sites for constructing assembly-line plants.

More recent developments have also aided both suburbanization and the deconcentration of MSA employment. Dan Garnick and Vernon Renshaw point out that miniaturization and the development of lightweight materials have reduced firms' incentives to locate in the largest MSAs to lower transportation costs. Other observers have argued that the technological forces that brought about deconcentration were reinforced by certain government policies, the most important being the federal highway program.[9] The interstate highway network has connected many previously remote areas of the country with one another and with the nation's largest MSAs. Thus, some technical innovations and government policies have made the smaller and less dense MSAs more attractive for both firms and households.

Despite the speculation that deconcentration represents nothing more than a continuation of the forces that led to suburbanization, there is little independent evidence that these forces are responsible for deconcentration. In fact, another view holds that these forces are irrelevant to deconcentration.

An Alternative View. The alternative view starts with the observation that, after some point, further increases in the number of people and firms in an MSA tend to clog its roads and transportation network. In addition, the cost of transporting goods goes up and the time needed to transport them lengthens, as does the time needed to commute to work or to get to leisure activities.

Of course, the negative effects of congestion brought on by growth in an MSA are only part of the equation. The positive effects of agglomeration economies make up the other. If the net benefits of growth (agglomeration economies less congestion costs) increase proportionately less for more dense metropolitan areas, entrepreneurs will have an incentive to locate plants in less dense MSAs. Over time, growth will favor the less dense MSAs whose agglomeration economies still outweigh their congestion costs.

Evidence. What role has the disproportionate increase in congestion costs played in the deconcentration of MSA employment? To investigate this point, Satyajit Chatterjee and I developed a model of the tradeoff between agglomeration economies and congestion (see *Appendix*). Then, guided by microeconomic studies in the urban and regional economics literature, we selected values for key parameters in the model to reproduce the employment density for each of the 297 MSAs in 1951. However, by 1994, total employment in the nation's MSAs was about 2.5 times higher than in 1951. So we used our model to predict how this employment growth would be distributed across the same 297 MSAs in 1994 (Figure 3). In the figure, the solid line shows the actual distribution of MSA employment density in 1951, ordered from most to least dense, and the dashed line shows the model's prediction for job distribution in 1994. Since the dashed line—the model prediction—lies above the solid line, the model predicts a high degree of deconcentration: the less dense MSAs will attain a relatively larger share of new jobs. With the exception of the two most dense MSAs, Jersey City and New York City, the model predicts that density will increase for all other MSAs during the period 1951-94. However, employment in the less

[8]See the studies by Charles Leven (1978) and Dan Garnick and Vernon Renshaw (1980) and my 1985 article.

[9]See the studies by Charles Leven (1978) and James Coleman (1978).

dense MSAs will increase relatively more than employment in the more dense MSAs. [10]

Actual outcomes for 1994 generally confirm these predictions (Figure 4). The solid line shows actual employment densities in 1994. The data for 1994 are quite close to the predicted values (dashed line) except for the most dense MSAs and the least dense MSAs. The 18 most dense MSAs gained a larger share of employment during the period 1951-94 than predicted (the solid line in Figure 4 lies above the dashed line for these MSAs). In 1951, these 18 MSAs accounted for 34 percent of total national employment. By 1994, their share had fallen to 21 percent. But the model predicted their share would fall to 16 percent. At the other end of the employment density graph, we see that the 38 least dense MSAs gained a smaller share of employment between 1951

FIGURE 3

Employment Is Predicted To Become Less Concentrated...

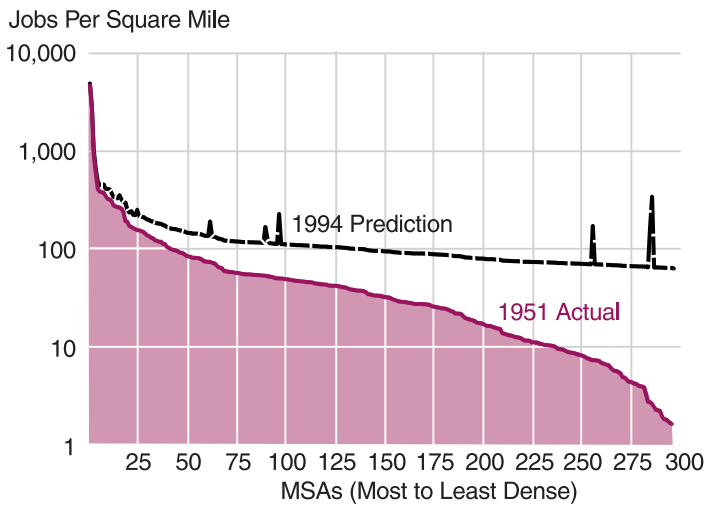
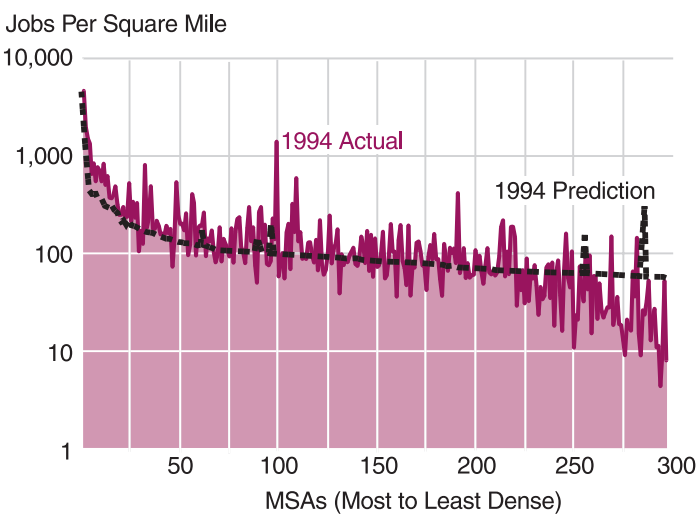


FIGURE 4

...And History Confirms Prediction



[10]The boundaries (and land area) of the MSAs discussed in this article are fixed by their 1983 definitions. In reality, as metropolitan areas grow, the boundaries of some of them spread out.

and 1994 than predicted. In 1951, these 38 MSAs accounted for 1.3 percent of total national employment. By 1994, their share had risen to 3.6 percent. But the model predicted their share would jump to 18 percent.

These discrepancies suggest that other forces, such as more rapid technological change in the densest MSAs, may have mitigated the deconcentration. Recently, some economists have argued that higher densities of people and jobs promote faster innovation and technological change and therefore growth. Economists have suggested an important link between innovation and density. They argue that the concentration of people and jobs in cities and other dense locations creates an environment in which ideas flow quickly among people. For example, the collaborative effort of many educated individuals in a common enterprise may lead to a higher sustained rate of innovation in the design of products.[11]

CONCLUSION

An examination of the data for almost 300 metropolitan areas in the United States shows a pronounced trend of deconcentration of employment from the most dense to least dense metro areas. Many economists have speculated that a decline in urban agglomeration economies accounted for the observed deconcentration of jobs in the postwar period. But our analysis suggests that growth has favored the less dense metro areas not because agglomeration economies have declined but because congestion costs associated with growth have increased faster in more dense locations.

[11]See my 1995 article for a review of this literature, and the article by Edward Glaeser, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer (1992).

APPENDIX: Modeling MSA Growth

The data in this article document a pronounced trend toward spatial deconcentration of employment. Motivated by this finding, Satyajit Chatterjee and I developed a model in which exogenous employment growth causes employment to shift in favor of less dense MSAs because congestion costs increase more rapidly for the initially more dense MSAs.[a] This general equilibrium model is described for both firms and households.

Firms. Production is subject to agglomeration economies, which are assumed to be constant (but not increasing) for those MSAs below a threshold size. Agglomeration economies are taken to increase with employment size once an MSA crosses the size threshold. There is no upper limit on agglomeration economies; beyond the threshold, they are assumed to increase in direct proportion to an MSA's employment. If agglomeration economies confer higher profits in any given MSA, firms in search of higher profits have an incentive to move to the relatively more productive MSAs. This

[a]The model discussed in this Appendix is similar in spirit to models developed by Vernon Henderson.

influx of firms increases the demand for workers and bids up local wages. The increase in local labor costs, in turn, reduces the profits of local firms. Labor costs will continue to rise until profits are once again equalized across MSAs.

Households. For workers, the increase in wages means they can increase their consumption of goods, yielding higher utility for workers in the more productive MSAs. This increase in utility attracts workers to the more productive MSAs; however, the influx of workers increases an MSA's density, and congestion costs rise. These congestion costs are assumed to increase more than proportionately with increases in MSA employment. The increased congestion costs lower real wages, and consumption and utility begin to fall. Congestion costs continue to rise, and real wages will continue to fall until worker utility is once again equal across all MSAs.

Suppose that aggregate MSA employment doubles. How would this increased employment be distributed across MSAs? Beyond the threshold size, increases in agglomeration economies are proportional to an MSA's employment size, and equal percentage increases in employment across MSAs result in equal percentage increases in agglomeration economies across MSAs. But since congestion costs increase more than proportionately with an increase in MSA density, employment growth favors the less dense MSAs. Thus, the model predicts that employment growth will be inversely related to an MSA's density. Dense MSAs also continue to grow, however, because of location-specific advantages.[b]

Calibrating the Model. The numerical specification of the model involves choosing values for four groups of parameters. These four groups are threshold size, agglomeration economies, congestion costs, and location-specific factors. We used existing studies to put bounds on the threshold, agglomeration, and congestion parameters, then selected values from within these bounds to carry out the calibration exercise. In the baseline model, we used 550,000 jobs as the threshold size after which agglomeration economies begin to increase.[c] Recall that below 550,000 jobs, agglomeration economies are taken to be constant but not increasing. Once an MSA crosses the threshold, its productivity is taken to increase 3.4 percent with each doubling of its employment size.[d] The density parameters used in the baseline model suggest that an increase in employment density of 1 percent raises the cost of living 2.1 percent in the Jersey City MSA (the most dense MSA in the study) but only 0.003 percent in the Casper, Wyoming, MSA (the least dense MSA in the study).[e] The values for the location-specific factors were chosen so that the model exactly matches the MSA distribution of employment densities in 1951 (the solid line in Figure 3). The model is then used to predict employment densities for 1994. The calibrated version of the model shows that MSAs that were less densely populated in 1951 were able to accommodate the two-and-a-half-fold increase in employment experienced during the postwar period more cheaply and thus attracted a larger share of these new jobs (the dashed lines in Figures 3 and 4).

---

[b]Location-specific factors reflect the fact that MSAs have, for example, a different mix of industries and a different quality of public infrastructure, such as roads, bridges, ports, etc.

[c]We used David Segal's study in determining the baseline and bounds for the threshold employment level.

[d]Leo Sveikauskas' study guided us in setting baseline and bounds for the agglomeration parameters.

[e]The baseline and bounds for the congestion parameters were guided by Jennifer Roback's study. See my paper with Satyajit Chatterjee for more details on the calibration exercises.

REFERENCES

Black, Duncan, and Henderson, J. Vernon. "Urban Growth," National Bureau of Economic Research Working Paper 6008 (April 1997).

Carlino, Gerald A. "Do Education and Training Lead to Faster Growth in Cities?" Federal Reserve Bank of Philadelphia *Business Review* (January/February 1995).

Carlino, Gerald A. ''Declining City Productivity and the Growth of Rural Regions: A Test of Alternative Explanations,'' *Journal of Urban Economics,* 18 (1985), pp. 11-27.

Carlino, Gerald A. ''Manufacturing Agglomeration Economies as Returns to Scale,'' Papers of the Regional Science Association 50 (1982), pp. 95-108.

Chatterjee, Satyajit and Carlino Gerald. "Aggregate Employment Growth and the Deconcentration of Metropolitan Employment," Working Paper 98-6, Federal Reserve Bank of Philadelphia (1998).

Coleman, James S. ''Social Processes and Social Policy in the Stable Metropolis,'' in C.L. Leven, ed., *The Mature Metropolis.* Lexington, Mass.: Lexington Books, 1978.

Eaton, Jonathan, and Eckstein, Zvi. ''Cities and Growth: Theory and Evidence from France and Japan,'' *Regional Science and Urban Economics,* 27 (1997), pp. 442-47.

Glaeser, Edward, Kallal, Hedi, Scheinkman, Jose, and Shleifer, Andrei. "Growth in Cities," *Journal of Political Economy,* 100 (1992), pp. 1126-53.

Garnick, Daniel, and Renshaw, Vernon. ''Competing Hypotheses on the Outlook for Cities and Regions: What the Data Reveal and Conceal," Papers of the Regional Science Association, 45 (1980), pp. 105-24.

Henderson, J.Vernon. ''General Equilibrium Modeling of Systems of Cities,'' in Edwin S. Mills, ed., *Handbook of Regional and Urban Economics,* Volume ll. New York: North Holland, 1987.

Leven, Charles L. ''The Emergence of Maturity in Metropolis,'' in C.L. Leven, ed., *The Mature Metropolis.* Lexington, Mass.: Lexington Books, 1978.

Roback, Jennifer. ''Wages, Rents and the Quality of Life,'' *Journal of Political Economy,* 90 (1982), pp. 1257-78.

Segal, David. "Are There Returns to Scale in City Size?" *Review of Economics and Statistics* 58, (1976), pp. 339-50.

Sveikauskas, Leo. "The Productivity of Cities," *Quarterly Journal of Economics,* 89 (1975), pp. 393-413.