Journal of Economic Perspectives—Volume 7, Number 2—Spring 1993—Pages 41–59

Employment and Unemployment in the 1930s

Robert A. Margo

The Great Depression is to economics what the Big Bang is to physics. As an event, the Depression is largely synonymous with the birth of modern macroeconomics, and it continues to haunt successive generations of economists. With respect to labor and labor markets, these facts evidently include wage rigidity, persistently high unemployment rates, and long-term joblessness.

Traditionally, aggregate time series have provided the econometric grist for distinguishing explanations of the Great Depression. Recent research on labor markets in the 1930s, however, has shifted attention from aggregate to disaggregate time series and towards microeconomic evidence. This shift in focus is motivated by two factors. First, disaggregated data provide many more degrees of freedom than the decade or so of annual observations associated with the depression, and thus may prove helpful in distinguishing macroeconomic explanations. Second, disaggregation has revealed aspects of economic behavior hidden in the time series that may be essential to their proper interpretation and, in any case, are worthy of study in their own right. Although the substantive findings of recent research are too new to judge their permanent significance, I believe that the shift towards disaggregated analysis is an important contribution.

The paper begins by reviewing the conventional statistics of the United States labor market during the Great Depression and the paradigms to explain them. It then turns to recent studies of employment and unemployment using disaggregated data of various types. The paper concludes with discussions of

■ *Robert A. Margo is Professor of Economics, Vanderbilt University, Nashville, Tennessee, and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts.*

research on other aspects of labor markets in the 1930s and on a promising source of microdata for future work. My analysis is confined to research on the United States; those interested in an international perspective on labor markets might begin with Eichengreen and Hatton's (1988) chapter in their edited volume, *Interwar Unemployment in International Perspective*, and the various country studies in that volume.

Labor Statistics Revisited

Standard unemployment and wage statistics for the 1930s appear in Table 1 (Baily, 1983, offers more detail). Two series of aggregate unemployment rates are shown, Stanley Lebergott's (1964) and Michael Darby's (1976), and an index of real hourly earnings in manufacturing compiled by the Bureau of Labor Statistics (BLS). The difference between Lebergott's and Darby's series reflects the treatment of persons with so-called "work relief" jobs. For Lebergott, persons on work relief are unemployed, while Darby counts them as employed.[1]

Between 1929 and 1933 the unemployment rate increased by over 20 percentage points, according to the Lebergott series, or by 17 percentage points, according to Darby's series. For the remainder of the decade, the unemployment rate stayed in, or hovered around, double digits. On the eve of America's entry into World War II, between 9.5 and 14.6 percent of the labor force was out of work, depending on how unemployment is measured.

In addition to high levels of unemployment, the 1930s witnessed the emergence of widespread and persistent long-term unemployment (unemployment durations longer than one year) as a serious policy problem. According to a Massachusetts state census taken in 1934, fully 63 percent of unemployed persons had been unemployed for a year or more. Similar long-term unemployment was observed in Philadelphia in 1936 and 1937 (Margo, 1991, p. 335).

[1]In counting persons on work relief as unemployed, Lebergott (1964) was effectively following census practice in 1940. Darby (1976, p. 5) challenged this practice, arguing that "[f]rom the Keynesian viewpoint, labor voluntarily employed on contracyclical... government projects should certainly be counted as employed. On the search approach to unemployment, a person who accepts a job and withdraws voluntarily from the activity of search is clearly employed." The logic of Darby's position can be debated. Although he claims that Keynesians would "certainly" count persons on work relief as employed, he identifies no Keynesians who ever held this view. The instructions to enumerators of the 1940 census specify several instances in which unemployed persons who were not actively seeking work (for example, because there was no work to be found in their occupation in their community) were still to be counted as unemployed (U.S. Bureau of the Census, 1983, section 6, p. 27). Consistency with the search approach to unemployment would, at the very least, require that such persons be separated out from the unemployed who were actively searching for work, which Darby fails to do. I am grateful to Stanley Lebergott for these points. For a detailed critique of Darby (1976), see Kesselman and Savin (1978).

Table 1
Unemployment and Real Wages in the 1930s

	Unemployment Rate		*Real Wage Index*
	Lebergott	*Darby*	*(1940 = 100)*
1929	3.2%	3.2%	69.4
1930	8.7	8.7	75.7
1931	15.9	15.3	83.2
1932	23.6	22.9	80.8
1933	24.9	20.6	79.5
1934	21.7	16.0	84.3
1935	20.1	14.2	80.4
1936	16.9	9.9	81.1
1937	14.3	9.1	85.5
1938	19.0	12.5	93.9
1939	17.2	11.3	97.3
1940	14.6	9.5	100.0

Sources: Unemployment rates: Smiley (1983, p. 488). Real wage index: average hourly earnings of production workers in manufacturing divided by the wholesale price index; hourly earnings is from U.S. Bureau of the Census (1976, series D-802, pp. 169–170; wholesale price index is from U.S. Bureau of the Census (1976, series E-40, p. 200).

Given these patterns of unemployment, the behavior of real wages has proven most puzzling. Between 1929 and 1940, annual changes in real wages and unemployment, as shown in Table 1, were positively correlated. Real wages rose by 16 percent between 1929 and 1932, while the unemployment rate ballooned from 3 to 23 percent. Real wages remained high throughout the rest of the decade, although unemployment never dipped below 9 percent, no matter how it is measured.

This information poses two central questions. Why did unemployment remain persistently high throughout the decade? How can unemployment rates in excess of 10 to 20 percent be reconciled with the behavior of real wages, which were stable or increasing?

One way of answering these questions is to devise aggregative models consistent with the time series, and I briefly review these attempts in the next section of the paper. Before doing so, however, it is important to stress that the aggregate statistics are far from perfect. No government agency in the 1930s routinely collected labor force information analogous to that provided by today's *Current Population Survey*. The unemployment rates in Table 1 are constructs, the differences between intercensal estimates of labor force participation rates and employment-to-population ratios. Because unemployment is measured as a residual, relatively small changes in the labor force or employment counts can markedly affect the estimated unemployment rate. Although

some progress has been made on measurement issues, there is little doubt that further refinements to the aggregate unemployment series would be beneficial.[2]

Stanley Lebergott (1989) has critically examined the reliability of BLS wage series from the 1930s, like the one shown in Table 1. The BLS series drew upon a fixed group of manufacturing establishments reporting for at least two successive months. Lebergott notes several biases arising from this sampling method. Workers who were laid off, he claims, were less productive and had lower wages than average. Firms that went out of business were smaller, on average, than firms that survived, and tended to have lower average wages. In addition, the BLS oversampled large firms, and Lebergott suspects that large firms were more adept at selectively laying off lower-productivity labor; more willing to de-skill (that is, reassign able employees to less-skilled jobs); and more likely to give able employees longer work periods. For example, Lebergott (1989, p. 9) cites the cases of General Electric and Westinghouse which reassigned skilled labor and foremen to factory jobs, and which gave longer hours to more able employees. Wage rates at the two companies fell 10 percent from 1929 to 1931, yet the industry average wage measured by the BLS (which, according to Lebergott, was heavily influenced by these two firms) remained constant.

A rough calculation suggests that accounting for these biases would produce an aggregate decline in nominal wages between 1929 and 1932 as much as 48 percent larger than that measured by the BLS series. Although the details of Lebergott's calculation are open to scrutiny, the research discussed elsewhere in the paper suggests that he is correct about the existence of biases in the BLS wage series.

Macroeconomic Studies

For much of the period since World War II, most economists blamed the unemployment rates in Table 1 on wage rigidity. The demand for labor was a downward sloping function of the real wage, but since nominal wages were insufficiently flexible downward (relative to producer prices), the labor market in the 1930s was persistently in disequilibrium. Labor supply exceeded labor demand, with mass unemployment the unfortunate consequence.[3]

[2]See Gene Smiley (1983) for a careful evaluation of the various attempts to date to measure aggregate unemployment in the 1930s.

[3]I recognize that many economists believe that flexible wages and prices can be destabilizing, and that this may have been the case between 1929 and 1932 (Rees, 1970, p. 308; De Long and Summers, 1986; Greenwald and Stiglitz, 1993). However, most writers in the Keynesian tradition, such as Peter Temin (1989), argue that lower wages after 1933 would have led to higher levels of employment.

The frontal attack on the conventional wisdom was Robert E. Lucas and Leonard Rapping (1969). The original Lucas-Rapping set-up continued to view current labor demand as a negative function of the current real wage. Current labor supply was a positive function of the real wage and the expected real interest rate, but a negative function of the expected future wage (Lucas and Rapping, eq. 11). If workers expect higher real wages in the future or a lower real interest rate, current labor supply would be depressed, employment would fall, unemployment rise, and real wages increase. Lucas and Rapping offer an unemployment equation, relating the unemployment rate to actual versus anticipated nominal wages, and actual versus anticipated price levels.

In a comment on Lucas and Rapping's paper, Albert Rees (1970, p. 308) argued that it was unrealistic to think that unemployed workers in the 1930s were simply waiting for times to get better. "How long," Rees asked, "[did] it take workers to revise their expectations in light of the facts? . . . [i]t is hard to imagine the long-term unemployed holding out for jobs comparable with their old jobs . . . over periods up to ten years." Lucas and Rapping (1972) conceded Rees's criticism for the period 1933 to 1941, but claimed victory for 1929 to 1933. As Ben Bernanke (1986, p. 83) pointed out, however, their victory rests largely on the belief that expected real interest rates fell between 1929 and 1933, while "*ex post*, real interest rates in 1930–33 were the highest of the century." Because nominal interest rates fell sharply between 1929 and 1933, whether expected real rates fell hinges on whether deflation—which turned out to be considerable—was unanticipated. Recent research suggests that the deflation was, at least in part, anticipated, which appears to undercut Lucas and Rapping's reply (Cecchetti, 1992).

Returning to the conventional wisdom, New Deal legislation has frequently been blamed for the persistence of high unemployment and the perverse behavior of real wages (Temin, 1989, 1990). In this regard, perhaps the most important piece of legislation was the National Industry Recovery Act (NIRA) of 1933. The National Recovery Administration (NRA), created by the NIRA, established guidelines that raised nominal wages and prices, and encouraged higher levels of employment through reductions in the length of the workweek (or "work-sharing").[4]

An influential study by Michael Weinstein (1980) econometrically analyzed the impact of the NIRA on wages. Using aggregate monthly data on hourly earnings in manufacturing, Weinstein showed that the NIRA raised nominal wages directly through its wage codes and indirectly by raising prices (Weinstein, 1980, Table 2.3, p. 53). The total impact was such that (p. 59) "[i]n the absence of the NIRA, average hourly earnings in manufacturing would

[4]The NIRA was declared unconstitutional in May 1935. Other important New Deal labor legislation includes the National Labor Relations Act (1935), which promoted unions; and the Fair Labor Standards Act (1938), which set minimum wages in certain industries and which empowered the federal government to regulate working conditions.

have been less than thirty-five cents by May 1935 instead of its actual level of almost sixty cents (assuming unemployment to have been unaltered)".[5]

It is questionable, however, whether the NIRA really had this large an impact on wages. Weinstein measured the direct effect of the codes by comparing monthly wage changes during the NIRA period (1933–35) with wage changes during the recovery phase (1921–23) of the post-World War I recession (1920–21), holding constant the level of unemployment and changes in wholesale prices. Data from the intervening years (1924–32) or after the NIRA period were excluded from his regression analysis. In addition, Weinstein's regression specification precludes the possibility that reductions in weekly hours (work-sharing), some of which occurred independently of the NIRA, had a positive effect on hourly earnings. A recent paper using data from the full sample period and allowing for the effect of worksharing found a positive but much smaller impact of the NIRA on wages (see the discussion of Bernanke's work in the next section).

Various developments in neo-Keynesian macroeconomics have filtered into the interpretation of the statistics in Table 1. Martin Baily (1983) argues that firms did not aggressively cut wages early in the 1930s because such a policy would hurt worker morale and the firm's reputation, incentives that were later reinforced by New Deal legislation.[6] Richard Jensen (1989) invokes efficiency wage theory in a provocative article. Beginning sometime after the turn of the century, large firms began to adopt personnel policies that were "designed to identify and keep the more efficient workers, and to encourage other workers to emulate them" (Jensen, 1989, p. 561; see also Jacoby, 1985). Efficiency wages were one such device, which presumably contributed to stickiness in wages.

The trend accelerated in the 1930s. According to Jensen, firms surviving the initial downturn used the opportunity to lay off their least productive workers but a portion of the initial decline in employment occurred among firms that went out of business (Lebergott, 1989). Thus, when expansion occurred, firms had their pick of workers who had been laid off. Personnel departments used past wage histories as a signal, and higher-wage workers were a better risk.[7] Those with few occupational skills, the elderly (who were

[5]According to Weinstein (1980, Table 2.3, eq. 2, p. 53), the direct effect alone (the wage codes) was responsible for increasing hourly earnings at an annual rate of 2.0 percent per month (holding constant changes in wholesale prices), or about 26 percent per year.

[6]Anthony O'Brien (1989) suggests a similar explanation of why firms did not aggressively cut wages early in the 1930s. In O'Brien's opinion, business leaders generally adhered to a belief that deep wage cuts exacerbated the recession of 1920–21, and that this social norm carried over during the early part of the Great Depression. According to O'Brien (p. 730), "[a]ny firm tempted to cut wages at the beginning of the Depression would have faced a substantial penalty in the form of diminished worker productivity" because it would have violated the social norm.

[7]Jensen (1989, p. 567) cites a study of the reemployment of factory workers in Massachusetts who lost their jobs in 1931. By 1933, only 40 percent had found regular employment; the probability was much higher, however, among workers in the top third of the wage distribution (57 percent) than in the bottom third (27 percent).

expensive to retrain) and the poorly educated faced enormous difficulties in finding work. Jensen argues that by 1935 the "reserve army" of long-term unemployed no longer exerted much downward pressure on nominal wages because employers simply did not view the long-term unemployed as substitutes for the employed at virtually any wage (see also Rees, 1970, p. 309).

A novel feature of Jensen's paper is its integration of microeconomic evidence on the characteristics of the unemployed with macroeconomic evidence on wage rigidity. Consistent with his argument, there is no evidence that wage differentials by skill or education level were wider overall in the late 1930s compared with the 1920s, even though less skilled and less educated workers were disproportionately unemployed throughout the decade (Goldin and Margo, 1992, pp. 22–23).[8] It remains an open question, however, whether use of efficiency wages and related personnel practices were as widespread as Jensen alleges, and whether such use can account empirically for the persistence of high unemployment.[9]

In brief, the macro studies have not settled the debate over the proper interpretation of the statistics in Table 1. This state of affairs has much to do with the (supreme) difficulty of building a consensus macro model of the depression economy. But it is also a consequence of the level of aggregation at which empirical work has been conducted. The problem is partly one of sample size, and partly a reflection of the inadequacies of discussing these issues using the paradigm of a representative agent. These points are illustrated in the research discussed in the next three sections.

Employment: Industry and Firm-Level Studies

In a conventional short-run aggregate production function, the labor input is defined to be total person-hours. For the postwar period, temporal variation in person-hours is overwhelmingly due to fluctuations in employment. However, for the interwar period, variations in the length of the workweek account for nearly half of the monthly variance in the labor input (Bernanke and

[8]Weinstein (1980, p. 81) argues that the NIRA compressed differences in hourly wages between skilled and unskilled labor, (although the effect was small; see pg. 82). There are two problems with Weinstein's claim. First, he compares wage differentials under the NIRA with wage differentials during the recovery phase from the post-World War I recession (1921–23). But wage differentials were compressed during World War I, and it is unreasonable to believe that they had returned to their long-run equilibrium level by 1921. Second, the data that Weinstein examines (from Beney, 1936) are relatively uninformative about movements in wage differentials in the 1930s because they refer solely to production workers in large manufacturing firms, thereby excluding most of the educated population who were not employed in manufacturing. See Goldin and Margo (1991, p. 1).

[9]See Temin (1990, pp. 302–303) and O'Brien (1989, p. 732) for critiques of the argument that efficiency wages can explain wage rigidity in the 1930s.

Powell, 1987). Declines in weekly hours were deep, prolonged, and widespread in the 1930s. The behavior of real hourly earnings, however, may have not been independent of changes in weekly hours.

This insight motivates Ben Bernanke's (1986) analysis of employment, hours, and earnings in eight pre-World War II manufacturing industries.[10] The (industry-specific) supply of labor is described by an earnings function, which gives the minimum weekly earnings required for a worker to supply a given number of hours per week; this minimum is a positive function of the reservation level of utility a worker could get in a "secondary or alternative sector" (Bernanke, 1986, p. 86). In Bernanke's formulation, the earnings function is convex in hours and also discontinuous at zero hours (the discontinuity reflects fixed costs of working or switching industries). Production depends separately on the number of workers and weekly hours, and on nonlabor inputs. Firms are not indifferent "between receiving one hour of work from eight different workers and receiving eight hours from one worker" (Bernanke, 1986, p. 91). A reduction in product demand causes the firm to cut back employment and hours per week. The reduction in hours means more leisure for workers, but less pay per week. Eventually, as weekly hours are reduced beyond a certain point, hourly earnings rise. Further reductions in hours cannot be matched one for one by reductions in weekly earnings. But when hourly earnings increase, the real wage then appears to be countercyclical.

To estimate the model, Bernanke uses monthly, industry-level data compiled by the National Industrial Conference Board (Beney, 1936; Sayre, 1940) covering the period 1923 to 1939. The specification of the earnings function (describing the supply of labor) incorporates a partial adjustment of wages to prices, while the labor demand equation incorporates partial adjustment of current demand to desired demand.[11] Except in one industry (leather), the industry demand for workers falls as real product wages rise; industry demands for weekly hours fall as the marginal cost to the firm of varying weekly hours rises; and industry labor supply is a positive function of weekly earnings and weekly hours.[12] Bernanke's estimates imply that the NIRA lowered weekly hours and raised weekly earnings and employment, although the effects were modest.[13] In six of the industries (the exceptions were shoes and lumber),

[10]The literature on work-sharing, during the 1930s and after, is voluminous; see the references cited in Bernanke (1986) for details.

[11]The inclusion of partial adjustment terms implies that, empirically, the actual supply of labor can exceed the actual demand for labor (that is, unemployment results); see Bernanke (1986, p. 95).

[12]The marginal cost to the firm of varying weekly hours is "the number of workers employed times the increment to their earnings required to get them to work the extra time" (Bernanke, 1986, p. 88).

[13]The largest percentage impacts on the level of hourly earnings (estimated as the percentage effect on weekly earnings minus the percentage effect on weekly hours) occurred in shoes, wool, and lumber, at around 10 percent. This is the basis for the claim in the previous section of the paper that Bernanke's estimate of the effects of the NIRA on wages are smaller than Weinstein's.

increased union influence after 1936 (measured with a proxy variable of days idled by strikes) raised weekly earnings by 10 percent or more.

In a related paper, Bernanke and Martin Parkinson (1991) use an expanded version of the NICB data set covering ten industries to explore the possibility that "short-run increasing returns to labor," or procyclical labor productivity, characterized co-movements in output and employment in the 1930s. Using their expanded data set, Bernanke and Parkinson (1991) estimate regressions of the change in output on the change in labor input, now defined to be total person-hours. The coefficient of the change in the labor input is the key parameter; if it exceeds unity, then short-run increasing returns to labor are present. Bernanke and Parkinson find that short-run increasing returns to labor characterized all but two of the industries under study (petroleum and leather). The estimates of the labor coefficients are essentially unchanged if the sample is restricted to just the 1930s. Further, a high degree of correlation ($r = 0.9$) appears between interwar and postwar estimates of short-run increasing returns to labor for a matched sample of industries.

Thus, the procyclical nature of labor productivity appears to be an accepted fact for both the interwar and postwar periods. One explanation of procyclical productivity, favored by real business cycle theorists, emphasizes technology shocks. Booms are periods in which technological change is unusually brisk, and labor supply increases to take advantage of the higher wages induced by temporary gains in productivity (caused by the outward shift in production functions). In Bernanke and Parkinson's view, however, the high correlation between the pre- and post-war estimates of short-run increasing returns to labor poses a serious problem for the technological shocks explanation. The high correlation implies that the "real shocks hitting individual industrial production functions in the interwar period accounted for about the same percentage of employment variation *in each industry* as genuine technological shocks hitting industrial production functions in the post-war period" (Bernanke and Parkinson, 1991, p. 451). However, technological change *per se* during the Depression was concentrated in a few industries and was modest overall (Bernstein, 1987). Further, while real shocks occurred—for example, bank failures, the New Deal, international political instability—their effects on employment, in Bernanke and Parkinson's opinion (p. 450), were felt through shifts in aggregate demand, not through shifts in industry production functions.

Other leading explanations of procyclical productivity are true increasing returns or, popular among Keynesians, the theory of labor hoarding during economic downturns. Having ruled out technology shocks, Bernanke and Parkinson attempt to distinguish between true increasing returns and labor hoarding. They devise two tests, both of which involve restrictions on excluding proxies for labor utilization from their regressions of industry output. If true increasing returns were present, the observed labor input captures all the

relevant information about variations in output over the cycle. But if labor hoarding were occurring, the rate of labor utilization, holding employment constant, should account for output variation. Their results are mixed, but are mildly in favor of labor hoarding in six of the ten industries under study.

Although Bernanke's (1986) modelling effort is of independent interest, the substantive value of his and Parkinson's empirical research is enhanced considerably by disaggregation to the industry level.[14] It is obvious from their work that industries in the 1930s did not respond identically to decreases in output demand. However, further disaggregation to the firm level can produce additional insights. Bernanke and Parkinson assume that movements in industry aggregates reflect the behavior of a representative firm. But according to Lebergott (1989), much of the initial decline in output and employment occurred among firms that exited. Firms that left, and new entrants, however, were not identical to firms that survived.

These points are well-illustrated in Timothy Bresnahan and Daniel Raff's (1991) study of the American motor vehicle industry. Their database consists of manuscript census returns of motor vehicle plants in 1929, 1931, 1933, and 1935. By linking the returns from year to year, Bresnahan and Raff have created a panel dataset, capable of identifying plants that exited, surviving plants, and new plants. Plants that exited between 1929 and 1933 had lower wages and lower labor productivity than plants that survived. Between 1933 and 1935, average wages at exiting plants and new plants were slightly higher than at surviving plants. Output per worker was still relatively greater at surviving plants than new entrants, but the gap was smaller than between 1929 and 1933.

Roughly a third of the decline in the industry's employment between 1929 and the trough in 1933 occurred in plant closures. The vast majority of these plant closures were permanent. The shakeout of inefficient firms after 1929 ameliorated the decline in average labor productivity in the industry. Although industry productivity did decline, productivity in 1933 would have been still lower if all plants had continued to operate.

During the initial recovery phase (1933–35) about 40 percent of the increase in employment occurred in new plants. Surviving plants were more likely to use mass production techniques; the same was true of new entrants. Mass production plants differed sharply from their predecessors (custom production plants) in the skill mix of their workforces and in labor relations. In the motor vehicle industry, the early years of the Depression were an "evolutionary event," permanently altering the technology of the representative firm (Bresnahan and Raff, 1991, p. 331).

While the representative firm paradigm apparently fails for motor vehicles, it may not for other industries. Some preliminary analysis of census

[14]See Whaples (1990) for an application of Bernanke's model in explaining the decline in the length of the workweek in the United States before World War I.

manuscripts for another industry, blast furnaces, is revealing on this point (Bertin, Bresnahan, and Raff, 1992). Blast furnaces were subject to increasing returns and the market for the product (molten iron) was highly localized. For this industry, reductions in output during a cyclical trough are reasonably described by a representative firm, since "localized competition prevented efficient reallocation of output across plants" (Bertin, Bresnahan, and Raff, p. 25) and therefore the compositional effects occurring in the auto industry did not happen.

These analyses of firm-level data have two important implications for studies of employment in the 1930s. First, aggregate demand shocks could very well have changed averaged technological practice through the process of exit and entry at the firm level. Thus, Bernanke and Parkinson's rejection of the technological shocks explanation of short-run increasing returns in the 1930s, which is based in part on their belief that aggregate demand shocks did not alter industry production functions, may be premature. Second, the empirical adequacy of the representative firm paradigm is apparently industry-specific, depending on industry structure, the nature of product demand, and initial (that is, pre-Depression) heterogeneity in firm sizes and costs. Such phenomena, Bertin *et al* note (p. 25), "are invisible in industry data," and can only be recovered from firm-level records, such as the census manuscripts.

Geographic Variation

Analyses of industry and firm-level data are one way to explore heterogeneity in labor utilization. Geography is another. A focus on national or even industry aggregates obscures the substantial spatial variation in bust and recovery that characterized the 1930s. Two recent studies show how spatial variation suggests new puzzles about the persistence of the Depression as well as providing additional degrees of freedom for discriminating between macroeconomic models.

State-level variation in employment is the subject of an important article by John Wallis (1989). Using data collected by the Bureau of Labor Statistics, Wallis has constructed annual indices of manufacturing and nonmanufacturing employment for states from 1930 to 1940. Wallis's indices reveal that declines in employment between 1930 and 1933 were steepest in the East North Central and Mountain states; employment actually rose in the South Atlantic states, however, once an adjustment is made for industry mix (Wallis, p. 60). The South also did comparatively well during the recovery phase of the Depression (1933–40). Wallis tests whether the southern advantage during the recovery phase might reflect lower levels of unionization and a lower proportion of employment affected by the passage of the Social Security Act (1935), but controlling for percent unionized and percent in covered

employment in a regression of employment growth does not eliminate the regional gap.[15] "What comes through clearly," according to Wallis (p. 61), "is that the [employment] effects of the Depression varied considerably throughout the nation" and that a convincing explanation of the southern difference remains an open question.

Curtis Simon and Clark Nardinelli (1992) exploit variation across cities to put forth a particular interpretation of economic downturn in the early 1930s. Specifically, they study the empirical relationship between "industrial diversity" and city-level unemployment rates before and after World War II. Industrial diversity is measured by a city-specific Herfindahl index of industry employment shares. The higher the value of the index, the greater is the concentration of employment in a small number of industries. Using data from the 1930 federal census and the 1931 Special Census of Unemployment, Simon and Nardinelli show that unemployment rates and the industrial diversity index were positively correlated across cities at the beginning of the Depression. Analysis of similar census data for the post-World War II period reveals a negative correlation between city unemployment rates and industrial diversity.

Simon and Nardinelli (1992) explain this finding as the outcome of two competing effects. In normal economic circumstances, a city with a more diverse range of industries should have a lower unemployment rate (the "portfolio" effect), because industry-specific demand shocks will not be perfectly correlated across industries and some laid-off workers will find ready employment in expanding industries (p. 385). The portfolio effect may fail, however, during a large aggregate demand shock (the early 1930s) if firms and workers are poorly informed, misperceiving the shock to be industry-specific, rather than a general reduction in demand. Firms in industrially diverse cities announce selective layoffs rather than reduce wages, because they believe that across-the-board wage cuts would cause too many workers to quit (workers in industrial diverse cities think they can easily find a job in another industry elsewhere in the same city), thus hurting production. Firms in industrially specialized cities, however, are more likely to cut wages than employment because they believe lower wages "would induce relatively fewer quits" than in industrially diverse cities (pp. 386–87).

Thus, Simon and Nardinelli (1992) conclude, wages in the early 1930s were more rigid in industrially-diverse cities, producing the positive correlation between industrial diversity and unemployment. Improvements in the quantity, quality, and timeliness of economic information, they conjecture, have caused the portfolio effect to dominate after World War II, producing the postwar negative correlation. Although one can question the historical relevance of Simon and Nardinelli's model, and the specifics of their empirical

[15]The hypothesis is that unionization and social security coverage should have reduced employment growth (by making labor more expensive to hire). The hypothesis is confirmed for unions, but not for social security; see Wallis (1989, p. 63).

analysis, their paper is successful in demonstrating the potential value of spatial data in unravelling the sources of economic downturn early in the Depression.[16]

Unemployment: The 1940 Census Sample

Postwar macroeconomics has tended to proceed as if the unemployment rates of Table 1 applied to a representative worker, with a certain percentage of that worker's time not being used. As a result, disaggregated evidence on unemployment has been slighted. Such evidence, however, can provide a richer picture of who was unemployed in the 1930s, a better understanding of the relationship between unemployment and work relief, and further insights into macroeconomic explanations of unemployment.

To date, the source that has received the most attention is the public use tape of the 1940 census, a large random sample of the population in 1940 (U.S. Bureau of the Census, 1983). The 1940 census is a remarkable historical document. It was the first American census to inquire about educational attainment, wage and salary income and weeks worked in the previous year; and the first to use the "labor force week" concept in soliciting information about labor force status. Eight labor force categories are reported, including whether persons held work relief jobs during the census week (March 24–30, 1940). For persons who were unemployed or who held a work relief job at the time of the census, the number of weeks of unemployment since the person last held a private or nonemergency government job of one month or longer was recorded. The questions on weeks worked and earnings in 1939 did not treat work relief jobs differently from other jobs. That is, earnings from and time spent on work relief are included in the totals.

I have used the 1940 census sample to study the characteristics of unemployed workers and of persons on work relief, and the relationship between work relief and various aspects of unemployment (Margo, 1988, 1991). It is clear from the public use sample that unemployed persons who were not on work relief were far from a random sample of the labor force. For example, the unemployed were typically younger, or older, than the average employed worker (unemployment followed a U-shaped pattern with respect to age); the unemployed were more often nonwhite (see also Sundstrom, 1992); and they were less educated and had fewer skills than employed persons, as measured by occupation. Such differences tended to be starkest for the long-term unemployed (those with unemployment durations longer than a year); thus, for

[16]Simon and Nardinelli (1992) provide no direct evidence that firms in industrially-diverse cities misperceived the shock in the manner described by their model; their regression specification excludes all determinants of unemployment other than the Herfindahl index, an industry mix variable, and a dummy for Western cities; and they do not estimate an equation for 1940, even though one would expect misperceptions to have vanished by that date.

example, the long-term unemployed had even less schooling than the average unemployed worker.

Although the WPA drew its workers from ranks of the unemployed, the characteristics of WPA workers did not merely replicate those of other unemployed persons. For example, single men, the foreign-born, high school graduates, urban residents, and persons living in the Northeast were underrepresented among WPA workers, compared with the rest of the unemployed. Perhaps the most salient difference, however, concerns the duration of unemployment. Among those on work relief in 1940, roughly twice as many had been without a non-relief job for a year or longer as had unemployed persons not on work relief.

The fact that the long-term unemployed were concentrated disproportionately on work relief raises an obvious question. Did the long-term unemployed find work relief jobs after being unemployed for a long time, or did they remain with the WPA for a long time? The answer appears to be mostly the latter. Among nonfarm males ages 14 to 64 on work relief in March 1940 *and* reporting 65 weeks of unemployment (that is, the first quarter of 1940 and all of 1939), close to half worked 39 weeks or more in 1939 (Margo, 1991, p. 338). Given the census conventions, they had to have been working more or less full time for the WPA.[17]

For reasons that are not fully clear, the incentives were such that a significant fraction of persons who got on work relief, stayed on.[18] One possible explanation is that some persons on work relief *preferred* the WPA, given prevailing wages, perhaps because their relief jobs were more stable than the non-relief jobs (if any) available to them. Or, as one WPA worker was quoted in E. W. Bakke (1940, pp. 421–22): "Why do we want to hold onto these [relief] jobs?... [W]e know all the time about persons... just managing to scrape along... My advice, Buddy, is better not take too much of a chance. Know a good thing when you got it." Alternatively, working for the WPA may have stigmatized individuals, making them less desirable to non-relief employers the longer they stayed on work relief (Jensen, 1989). Whatever the explanation, the continuous nature of WPA employment makes it difficult to believe that the WPA did not reduce, in the aggregate, the amount of job search done by unemployed workers in the late 1930s.

In addition to the duration of unemployment experienced by individuals, the availability of work relief may have dampened the increase in labor supply of secondary workers in households in which the household head was

[17]According to a WPA study conducted in 1939, fully 59 percent of persons with relief jobs in September 1937 held them continuously through February of 1939 (Margo, 1988, p. 34).

[18]This claim may surprise readers familiar with the statistics of work relief, because rates of assignment to, and separation from, WPA projects suggest a very high rate of turnover. Separation from project employment, however, was not the same as permanent separation from the WPA; accessions always included persons reassigned from other projects that had ended (U.S. Federal Works Agency, 1946, p. 32).

unemployed, the so-called "added worker" effect (Woytinsky, 1942). Specifically, wives of unemployed men not on work relief were much more likely to participate in the labor force than wives of men who were employed at non-relief jobs. But wives of men who worked for the WPA were far less likely to participate in the labor force than wives of otherwise employed men.[19] The relative impacts were such that, in the aggregate, no added worker effect can be observed as long as persons on work relief are counted among the unemployed (Margo, 1988, p. 348).

Although my primary goal in analyzing the 1940 census sample was to illuminate features of unemployment obscured by the aggregate time series, the results bear on several macroeconomic issues. First, the heterogenous nature of unemployment implies that a representative agent view of aggregate unemployment cannot be maintained for the late 1930s. Whether the view can be maintained for the earlier part of the Depression is not certain, but the evidence presented in Jensen (1989) and Margo (1991, p. 334) suggests that it cannot. Because the evolution of the characteristics of the unemployed over the 1930s bears on the plausibility of various macroeconomic explanations of unemployment (for example, Jensen, 1989), further research is clearly desirable.

Second, the heterogenous nature of unemployment is consistent with Lebergott's (1989) claim that aggregate BLS wage series for the 1930s are contaminated by selection bias, because the characteristics that affected the likelihood of being employed (for example, education) also affected a person's wage. Again, a clearer understanding of the magnitude and direction of bias requires further work on how the characteristics of the employed and unemployed changed as the Depression progressed.

Third, macroeconomic analyses of the persistence of high unemployment should not ignore the effects of the WPA—and, more generally, those of other federal relief policies—on the economic behavior of the unemployed. In particular, if work relief was preferred to job search by some unemployed workers, the WPA may have displaced some growth in private sector employment that would have occurred in its absence.[20] An estimate of the size of this displacement effect can be inferred from a recent paper by John Wallis and Daniel Benjamin (1989). Wallis and Benjamin estimate a model of labor supply, labor demand, and per capita relief budgets using panel data for states from 1933 to 1939. Their coefficients imply that elimination of the WPA starting in 1937 would have increased private sector employment by 2.9 percent by 1940, which

[19]The reason is not yet clear, but may have to do with eligibility requirements for work relief. Specifically, a working wife might cause family income to exceed limits set by local relief authorities, thereby causing a worker to be ineligible for employment with the WPA. T. Aldrich Finegan and I are currently investigating this issue further using the 1940 census sample (Finegan and Margo, 1992). On eligibility requirements for work relief, see Howard (1943).

[20]See Kesselman (1978) for a thorough and insightful discussion of the possible microeconomic and macroeconomic effects of work relief.

corresponds to about half of persons on work relief in that year.[21] Displacement was not one-for-one, but may not have been negligible.[22]

Other Aspects of Labor Markets

My discussion thus far has emphasized the value of disaggregated evidence in understanding certain key features of labor markets in the 1930s—the behavior of wages, employment and unemployment—because these are of greatest general interest to economists today. I would be remiss, however, if I did not mention other aspects of labor markets examined in recent work. What follows is a brief, personal selection from a much larger literature.

The Great Depression left its market on racial and gender differences. From 1890 to 1930 the incomes of black men increased slightly relative to the incomes of white men, but the trend in relative incomes reversed direction in the 1930s (Smith, 1984).[23] Migration to the North, a major avenue of economic advancement for Southern blacks, slowed appreciably. There is little doubt that if the Depression had not happened, the relative economic status of blacks would have been higher on the eve of World War II. Labor force participation by married women was hampered by "marriage bars," implicit or explicit regulations that allowed firms to dismiss single women upon marriage or prohibited the hiring of married women. Although marriage bars existed before the 1930s, their use spread during the Depression, possibly because social norms dictated that married men were more deserving of scarce jobs than married women (Goldin, 1990).

Although they have not received much attention from economists, some of the more interesting effects of the Depression were demographic or life-cycle in nature. Marriage rates fell sharply in the early 1930s, and fertility rates remained low throughout the decade (Eichengreen and Hatton, 1988, p. 47). An influential study by the sociologist Glen Elder, Jr. (1973) traced the

[21]Labor demand in Wallis and Benjamin's model is a function of the real wage, lagged employment, and various exogenous variables. Labor supply depends on the real wage, per capita relief budgets, and exogenous variables. After a three-year period the reduced form elasticity of employment with respect to relief budgets is -0.042 (calculated from Wallis and Benjamin, Table 4). Multiplying by -0.693 (69.3 percent is the WPA's share of all relief expenditures in 1937, calculated from Wallis and Benjamin's Tables 1 and 2) gives the predicted percentage change in employment (2.9 percent). Assuming that the elimination of work relief would have had no effect on aggregate labor force participation, the employed share of the labor force in 1940 would have been 87.9 percent instead of 85.4 percent ($= 1.029 \times 85.4$; from my Table 1). Since 5.1 percent of the labor force was on work relief in 1940 (see my Table 1), displacement would have equalled 49 percent ($= 2.5/5.1$). If aggregate labor force participation were higher in the absence of the WPA because of the added-worker effect (see the text and Finegan and Margo, 1992), displacement would be smaller, because the employment effect would be the same but the labor force would have been larger in size.

[22]See Kesselman and Savin (1978) for the alternate view, that work relief did not displace private employment.

[23]Other important studies containing information on the experiences of blacks in the Depression include Wolters (1970), Whatley (1983), and Wright (1986).

subsequent work and life histories of a sample of individuals growing up in Oakland, California, in the 1930s. Children from working class households whose parents suffered from prolonged unemployment during the Depression had lower educational attainment and less occupational mobility than their peers who were not so deprived. Similar findings were reported by Stephan Thernstrom (1973) in his study of occupational mobility of Boston men.

Concluding Remarks

The Great Depression was the premier macroeconomic event of the 20th century, and I am not suggesting that economists abandon macroeconomic analysis of it. I am suggesting, however, that *exclusive* focus on aggregate labor statistics runs two risks: the "facts" so derived may be artifacts in certain respects, and much of what may be informative about labor market behavior in the 1930s is rendered invisible. The people and firms whose experiences make up the aggregates deserve to be studied in their diversity, not as representative agents.

I have mentioned census microdata, such as the public use sample of the 1940 census or the manufacturing census manuscripts collected by Bresnahan and Raff, in this survey. In closing, I would like to highlight another source that could be examined in future work.

The source is the "Study of Consumer Purchases in the United States" conducted by the BLS in 1935–36. Approximately 300,000 households, chosen from a larger random sample of 700,000, supplied basic survey data on income and housing, with 20 percent furnishing additional information. The detail is staggering: labor supply and income of all family members; personal characteristics (for example, occupation, age, race); family composition; housing characteristics; and a long list of durable and non-durable consumption expenditures (the 20 percent sample). Because the purpose of the study was to provide budget weights to update the CPI, only families in "normal" economic circumstances were included (this is the basis for the reduction in sample size from 700,000 to 300,000). Thus, for example, persons whose wages were very low or who experienced long-term unemployment are unlikely to be included in the 1935–36 study. Even so, the data could prove invaluable for studies of wage determination; of labor supply decisions within families; and of the impact of unemployment on consumption expenditures. A pilot sample, drawn from the original survey forms (stored at the National Archives) and containing the responses of 6,000 urban households, is available in machine-readable format from the Inter-University Consortium for Political and Social Research at the University of Michigan (ICPSR Study 8908).

■ *I am grateful to Stanley Engerman, Claudia Goldin, Jonathan Kesselman, Stanley Lebergott, Daniel Raff, Carl Shapiro, Richard Steckel, Joseph Stiglitz, Timothy Taylor,*

Peter Temin, John Wallis, Eugene White, Gavin Wright, and seminar participants at the National Bureau of Economic Research for helpful comments.

References

Bakke, E. W., *The Unemployed Worker*. New Haven: Yale University Press, 1940.

Baily, Martin N., "The Labor Market in the 1930s." In Tobin, James, ed., *Macroeconomics, Prices and Quantities*. Washington: The Brookings Institution, 1983, 21–62.

Beney, M. Ada, *Wages, Hours, and Employment in the United States, 1914–1936*. New York: National Industrial Conference Board, 1936.

Bernanke, Ben S., "Employment, Hours, and Earnings in the Depression: An Analysis of Eight Manufacturing Industries," *American Economic Review*, March 1986, *76*:1, 82–107.

Bernanke, Ben S., and Martin L. Parkinson, "Procylical Labor Productivity and Competing Theories of the Business Cycle: Some Evidence from Interwar U.S. Manufacturing Industries," *Journal of Political Economy*, June 1991, *99*:3, 439–59.

Bernanke, Ben S., and James L. Powell, "The Cyclical Behavior of Industrial Labor Markets: A Comparison of the Pre-War and Post-War Eras." In Gordon, R. J., ed., *The American Business Cycle: Continuity and Change*. Chicago, University of Chicago Press, 1986, *583–621*.

Bernstein, Michael A., *The Great Depression: Delayed Recovery and Economic Change in America, 1929–1939*. New York: Cambridge University Press, 1987.

Bertin, Amy L., Timothy F. Bresnahan, and Daniel M. G. Raff, "Localized Competition and the Aggregation of Plant-Level Increasing Returns: Blast Furnaces, 1929–1935," mimeo, Harvard Business School, July 1992.

Bresnahan, Timothy F., and Daniel M. G. Raff, "Intra-Industry Heterogeneity and the Great Depression: The American Motor Vehicle Industry, 1929–1935," *Journal of Economic History*, June 1991, *51*, 317–31.

Cecchetti, Steven G., "Prices During the Great Depression: Was the Deflation of 1930–32 Really Unanticipated?," *American Economic Review*, March 1992, *82*, 141–56.

Darby, Michael R., "Three-and-a-Half Million U.S. Employees Have Been Mislaid: Or, an Explanation of Unemployment, 1934–1941," *Journal of Political Economy*, February 1976, *84*, 1–16.

De Long, J. Bradford, and Lawrence H. Summers, "Is Increased Price Flexibility Stabilizing?," *American Economic Review*, December 1986, *76*:5, 1031–1444.

Eichengreen, Barry, and Timothy J. Hatton, "Interwar Unemployment in International Perspective: An Overview." In Eichengreen, B., and T. Hatton, eds. *Interwar Unemployment in International Perspective*. London: Kluwer Academic Publishers, 1988, 1–59.

Elder Jr., Glen H., *Children of the Great Depression: Social Change in Life Experience*. Chicago: University of Chicago Press, 1973.

Finegan, T. Aldrich, and Robert A. Margo, "Added and Discouraged Workers in the Late 1930s: A Re-Examination," mimeo, Department of Economics, Vanderbilt University, October 1992.

Goldin, Claudia, *Understanding the Gender Gap: An Economic History of American Women*. New York: Oxford University Press, 1990.

Goldin, Claudia, and Robert A. Margo, "Skill Ratios and Wage Distributions: 1920s to 1950s," appendix to "The Great Compression: The Wage Structure in the United States at Mid-Century," National Bureau of Economic Research Working Paper No. 3817, Cambridge, August 1991.

Goldin, Claudia, and Robert A. Margo, "The Great Compression: The Wage Structure in the United States at Mid-Century," *Quarterly Journal of Economics*, February 1992, *107*, 1–34.

Greenwald, Bruce, and Joseph E. Stiglitz, "New and Old Keynesians," *Journal of Economic Perspectives*, Winter 1993, 7:1, 23–44.

Howard, Donald S., *The WPA and Federal Relief Policy*. New York: Russell Stage Foundation, 1943.

Jacoby, Sanford M., *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in American Industry, 1900–1945*. New York: Columbia University Press, 1985.

Jensen, Richard, "The Causes and Cures of Unemployment in the Great Depression,"

Journal of Interdisciplinary History, Spring 1989, *19*, 553–83.

Kesselman, Jonathan R., "Work Relief Programs in the Great Depression." In Palmer, John L., ed. *Creating Jobs: Public Employment Programs and Wage Subsidies*. Washington, D.C.: The Brookings Institution, 1978, 153–227.

Kesselman, Jonathan R., and N. E. Savin, "Three-and-a-Half Million Workers Never Were Lost," *Economic Inquiry*, April 1978, *16*, 205–25.

Lebergott, Stanley, *Manpower in Economic Growth*. New York: McGraw-Hill, 1964.

Lebergott, Stanley, "'Wage Rigidity' in the Depression: Concept or Phrase?," mimeo, Department of Economics, Wesleyan University, September 1989.

Lucas, Robert E., Jr., and Leonard A. Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy*, September/October 1969, 77, 721–54.

Lucas, Robert E., Jr., and Leonard A. Rapping, "Unemployment in the Great Depression: Is There a Full Explanation?," *Journal of Political Economy*, January/February 1972, *80*, 186–91.

Margo, Robert A., "Interwar Unemployment in the United States: Evidence from the 1940 Census Sample," In Eichengreen, B., and T. Hatton, *Interwar Unemployment in International Perspective*. London: Kluwer Academic Publishers, 1988, 325–52.

Margo, Robert A., "The Microeconomics of Depression Unemployment," *Journal of Economic History*, June 1991, *51*, 333–41.

O'Brien, Anthony Patrick, "A Behavioral Explanation for Nominal Wage Rigidity During the Great Depression," *Quarterly Journal of Economics*, November 1989, *104*, 719–35.

Rees, Albert, "On Equilibrium in Labor Markets," *Journal of Political Economy*, March/April 1970, *78*, 306–10.

Sayre, R. A., "Wages, Hours, and Employment in the United States, 1934–1939," *Conference Board Economic Record*, March 28, 1940, *2*, 115–37.

Simon, Curtis J., and Clark Nardinelli, "Does Industrial Diversity Always Reduce Unemployment? Evidence from the Great Depression and After," *Economic Inquiry*, April 1992, *30*, 384–97.

Smiley, Gene, "Recent Unemployment Rate Estimates for the 1920s and 1930s," *Journal of Economic History*, June 1983, *43*, 487–93.

Smith, James, "Race and Human Capital," *American Economic Review*, September 1984, *74*:4, 685–98.

Sundstrom, William A., "Last Hired, First Fired? Unemployment and Urban Black Workers During the Great Depression," *Journal of Economic History*, June 1992, *52*, 415–29.

Temin, Peter, *Lessons from the Great Depression*. Cambridge: M.I.T. Press, 1989.

Temin, Peter, "Socialism and Wages in the Recovery from the Great Depression in the United States and Germany," *Journal of Economic History*, June 1990, *50*, 297–307.

Thernstrom, Stephan, *The Other Bostonians*. Cambridge: Harvard University Press, 1973.

U.S. Federal Works Agency, *Final Report of the WPA Program, 1935–1943*, Washington, D.C.: U.S. Government Printing Office, 1946.

U.S. Bureau of the Census, *Historical Statistics of the United States*, Washington, D.C.: U.S. Government Printing Office, 1976.

U.S. Bureau of the Census, *Census of Population*, 1940: *Public Use Microdata Sample*, Washington, D.C.: U.S. Bureau of the Census, 1983.

Wallis, John J., "Employment in the Great Depression: New Data and Hypotheses," *Explorations in Economic History*, January 1989, *26*, 45–72.

Wallis, John J., and Daniel K. Benjamin, "Private Employment and Public Relief During the Great Depression," mimeo, Department of Economics, University of Maryland, July 1989.

Weinstein, Michael, *Recovery and Redistribution Under the NIRA*. Amsterdam, The Netherlands: North-Holland Publishing Company, 1980.

Whaples, Robert, *The Shortening of the American Work Week: An Economic and Historical Analysis of its Context, Causes, and Consequences*, Ph.d. thesis, Department of Economics, University of Pennsylvania, 1990.

Whatley, Warren, "Labor For the Picking: The New Deal in the South," *Journal of Economic History*, December 1983, *43*, 905–29.

Wolters, Raymond, *Negroes and the Great Depression: The Problem of Economic Recovery*. Westport: Greenwood Publishing Company, 1970.

Woytinsky, W. S., *Three Aspects of Labor Dynamics*. Washington, D.C.: Social Science Research Council, 1942.

Wright, Gavin, *Old South: New South*. New York: Basic Books, 1986.