The Estimation of Prewar Gross National Product: Methodology and New Evidence
Author(s): Nathan S. Balke and Robert J. Gordon
Reviewed work(s):
Source: *Journal of Political Economy*, Vol. 97, No. 1 (Feb., 1989), pp. 38-92
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/1831055
Accessed: 16/05/2012 09:34

The Estimation of Prewar Gross National Product: Methodology and New Evidence

Nathan S. Balke

Southern Methodist University

Robert J. Gordon

Northwestern University

The paper develops new methodology for the estimation of prewar GNP, taps previously unused data sources, and develops new estimates for the periods 1869–1908 and 1869–1928. Primary among the new data sources are direct measures of output in the transportation, communications, and construction sectors and estimates of the consumer price index. New measures of real GNP, nominal GNP, and the GNP deflator are developed. The new estimates of real GNP are as volatile on average over the business cycle as the traditional Kuznets-Kendrick series but dampen the amplitude of some cycles while raising the amplitude of others. The new estimates of the GNP deflator are distinctly less volatile than the traditional series and in fact no more volatile than those in the postwar period.

I. Introduction

Until recently, one of the least controversial stylized facts in macroeconomic history was the reduced volatility of output in the United States after World War II. Indeed, Arthur Burns devoted his entire

We are indebted to the National Science Foundation and the Sloan Foundation for financial support; to Moses Abramovitz, Brad DeLong, Ian Domowitz, Solomon Fabricant, Robert Gallman, James Heckman, Christina Romer, Mark Watson, and David Weir for stimulating correspondence and discussions; to Gabriel Sensenbrenner for help with the regressions and graphs; and to members of the Northwestern Economic History Workshop for constructive suggestions.

[*Journal of Political Economy*, 1989, vol. 97, no. 1]

1959 American Economic Association presidential address to explaining the phenomenon of a more stable postwar economy (Burns 1960). The real GNP data on which he relied had been developed in the preceding three decades primarily by Kuznets (1938, 1941, 1946, 1961), whose pioneering contributions were subsequently extended, revised, and converted into our present *National Income and Product Accounts* (NIPA) by the Department of Commerce for the period since 1929, and for the period before 1929 they were subsequently revised by Gallman (1966) and reworked to be consistent with Commerce definitions by Kendrick (1961).[1]

Innumerable studies of important macroeconomic relationships have been based on these "standard" output series. Among these are the well-known studies of money and output by Friedman and Schwartz (1963, 1982). Other examples include the research on the division of nominal GNP changes between inflation and quantity changes by Gordon (1980, 1982) and Schultze (1981, 1986). Taking for granted the reduction of output volatility after World War II, Taylor (1986) has argued that stability was achieved *despite* the greater persistence of price movements that tends to make output less stable, while DeLong and Summers (1986) claim that, on the contrary, greater postwar price persistence has contributed directly to output stability.

Recently the stylized fact of postwar stabilization has been disputed in a series of papers by Romer (1986*a,* 1986*b,* 1988, this issue), which argue that the volatility of industrial production, unemployment, and real GNP after 1947 is little different than before 1929 when revised measures are used (the high volatility of the economy in the 1930s is not disputed). Stimulated by the challenge of Romer's work to traditional output measures, this paper provides a new methodological analysis of output estimation that can be applied to any situation in which data are significantly worse prior to some "borderline" date than afterward. We apply the analysis in the development of new estimates of U.S. GNP for the period 1869–1908, but it is equally applicable to the creation of annual real GNP or price estimates prior to 1869 in the United States or in any other country prior to some

[1] This reference to Kendrick's monumental work greatly understates his contribution of compiling consistent output indexes, input indexes for both labor and capital, and indexes of labor productivity and total factor productivity for major industry subdivisions extending over the entire period between 1869 and 1955. However, for the issues of output measurement addressed in this paper, Kendrick's main contribution was relatively minor, to make the adjustments needed to convert the Kuznets output series to the concepts adopted by the Department of Commerce in the NIPA for the period since 1929. Both Commerce and Kendrick have published estimates of GNP for the period 1909–28; issues involved in choosing between these conflicting series have been treated in detail by Romer (1988) and are discussed further below.

borderline date when the coverage of available data exhibits a marked improvement.

Contribution of This Paper

Our methodological discussion compares the alternative "components" and "indicators" approaches to backcasting real GNP before a borderline date, for example, 1909 in the subsequent empirical analysis. We show that the components method makes better use of the available information, some of which is discarded by the indicators method, and we examine the circumstances in which measurement error could offset this advantage of the components method.

Our empirical section develops new estimates of real GNP by both the components and indicators methods for the period 1869–1908. These estimates go beyond the work of Kuznets, Gallman, Kendrick, and Romer, in three directions. First, unlike prior research that develops estimates of noncommodity output entirely from an assumed ratio or regression relationship to commodity output, we include direct measures of output in the construction, transportation, and communications sectors in addition to that in the commodity-producing sector. Second, both our components and indicators estimates of real GNP take advantage of the painstaking research by Hoover (1960) and Rees (1961) on prices actually paid by consumers during the period prior to World War I. Third, because no previous estimate of the GNP deflator before 1919 has made use of the Hoover-Rees consumer price index studies, we develop new estimates of the GNP deflator and the implied time series of nominal GNP to complement our new real GNP series.

Our substantive conclusions are striking. Our new real GNP series is as volatile as the standard Kuznets-Kendrick series, but our new series behaves differently over individual business cycles, with cycles of a smaller amplitude in some episodes and a greater amplitude in others. Perhaps more surprising is the finding that the aggregate price level is substantially *less* volatile than in traditional data. As a result, we conclude that the U.S. economy exhibits postwar stabilization of real output without any tendency for prices to become less volatile over the business cycle.

Plan of the Paper

We begin in Section II with a general methodological analysis of the alternative components and indicators approaches to backcasting. This is followed in Section III with a summary of the available indexes for real GNP, as well as important components of real GNP and other indicators of output movements. In this section we also discuss methods of detrending and alternative measures of volatility. Section IV evaluates the quality of alternative data series and yields a set of variables to be included in the subsequent regression analysis. Then Section V presents several alternative regression equations and the volatility of the associated backcast estimates, using both conventional price deflators and new deflators based in part on the Hoover-Rees CPI data. Section VI summarizes the results on volatility and measures their statistical significance. Section VII presents conclusions.

II. Methodological Issues

As we go backward in time, the quality of available data deteriorates.[2] Our discussion ignores the continuous nature of the deterioration and simplifies by treating the deterioration as occurring at a discrete point in time, called the "borderline" year. Before the borderline year some crucial data are missing but are available after the borderline year. Any method of estimating or "backcasting" aggregate economic activity before the borderline year must infer the level of output in the sectors lacking data.

In this paper we take the borderline year to be 1909; our task is the estimation of real GNP annually for the period 1869–1909. An important simplification is our treatment of the quality of post-1909 data as homogeneous, thus ignoring the gradual improvement that occurred between 1909 and 1929. The specific problem in developing estimates of real GNP for the period before 1909 is that data are missing for most of the non-commodity-producing sectors, particularly trade and services.

The set of data available for use in solving this inference problem can be classified as either "components" or "indicators." A component is a variable that is an actual element of real GNP, such as agricultural or manufacturing output. An indicator is a time-series variable that is correlated with real GNP in the postborderline time period when real GNP is assumed to be known. Examples of indicators include the index of industrial production, the unemployment rate, the number of building permits, and ton-miles of railroad traffic. Some variables, such as manufacturing output, can be both a component of real GNP and also an indicator, that is, correlated with aggregate real GNP after

[2] The continuous nature of the deterioration is true not just before 1929 but also since then. For instance, income data became much better after 1940 as a result of the introduction of the social security tax system and income tax withholding. The quality of both the producer and consumer price indexes, components of which underlie the NIPA deflators, has been progressive throughout the postwar period. Many other examples could be cited.

the borderline date. Other variables, such as agricultural output, may be a component of real GNP but not an indicator, having a low correlation with aggregate real GNP after the borderline date.

This classification of available data into components and indicators provides a convenient way of conceptualizing the issues involved in estimating preborderline real GNP. The components method of estimating GNP involves obtaining estimates of various components of GNP either directly or indirectly and adding them together. This is the method used by Kuznets to obtain his original GNP estimates. While he used assumed ratios to develop indirect estimates of the missing noncommodity output sectors on the basis of the known behavior of commodity output, alternatively the level of output in the missing sectors could be estimated by regression analysis, using as explanatory variables the behavior of one or more indicators in the postborderline period. The advantage of the components method is that direct information on the level of real GNP is used to the extent of the available data, and changes in economic structure in the preborderline period are automatically captured as the weights of the various components change through time. A possible disadvantage of the components method is that there may be measurement errors in the data on one or more individual components, but measurement errors can contaminate the indicators methods in the same way.

The indicators method involves estimating a regression of aggregate real GNP on a set of one or more variables (which can be components or indicators) for the postborderline estimation period and applying this relationship to the preborderline backcast period. Kuznets's "regression series" is a single-variable version of the indicators method, based on the use of commodity output as the only indicator.

Comparison of the Two Methods

In comparing the two methods, we assume that both methods use the same information set, that is, the same set of extrapolators. The difference between the two methods lies in how efficiently they use this information set. To capture the difference in the two methods in the simplest of contexts, we first examine the case in which the estimation procedures are conducted in levels. Subsequently we analyze the more complex case in which the estimation is carried out for deviations from trend. It is this second case that is relevant for our subsequent empirical analysis as well as for the previous implementation of the indicators method by Kuznets (1961) and Romer (this issue).

Let real GNP, Q_t, consist of two components,

$$Q_t = \mathbf{X}_{1t} + \mathbf{X}_{2t}, \tag{1}$$

where $\mathbf{X}_{1t}$ is a vector of components for which we have data in the preborderline period and $\mathbf{X}_{2t}$ is a vector of "residual" components for which we lack data. We have no direct means of measuring the vector $\mathbf{X}_{2t}$; therefore we must form estimates of these components. One way to accomplish this is to estimate an equation for a sample period in which the values of the residual component are known. Our estimate of $\mathbf{X}_{2t}$ would be

$$\mathbf{X}_{2t} = a_1\mathbf{X}_{1t} + a_2 Z_t + e_{2t}. \tag{2}$$

Here Z_t are any additional variables for which we have data in both the estimation and extrapolation periods, and e_{2t} is an error term. With the estimated coefficients from (2), the components estimate of real GNP is

$$Q_t^C = (1 + a_1)\mathbf{X}_{1t} + a_2 Z_t + e_{2t}. \tag{3}$$

The indicators estimate of real GNP is also based on a regression in which the same explanatory variables appear and in which the dependent variable is aggregate real GNP rather than the residual component $\mathbf{X}_{2t}$. The indicators regression is

$$Q_t^I = b_1\mathbf{X}_{1t} + b_2 Z_t + e_t^I. \tag{4}$$

Notice that the indicators and components methods in (3) and (4) are observationally equivalent. However, the components method makes use of an additional identifying restriction. This leads to more precise parameter estimates.

The discussion above applies to the case in which GNP is originally estimated in levels. However, our regression estimates below use deviations from trend rather than levels. The existing estimates of GNP (the original Kuznets components estimates as revised by Kendrick and Gallman) are generally thought to be acceptable measures of trend economic activity; only the cyclical properties of these series are at issue in the debate over the reduction in real GNP volatility after World War II. By separating each variable into trend and cyclical components and assuming the trend values to be accurate, we need to carry out our analysis only for the cyclical deviations from trend. In the analysis below, we assume that we know the trend values Z_t^*, Q_t^*, and $\mathbf{X}_{1t}^*$ and that we can calculate the trend value of $\mathbf{X}_{2t}$ ($\mathbf{X}_{2t}^* = Q_t^* - \mathbf{X}_{1t}^*$). Lowercase letters will be used to represent deviations from trend, that is, $q_t = (Q_t - Q_t^*)/Q_t^*$.

For the components method, recall that the components model for the level of real GNP is given by $Q_t^C = \mathbf{X}_{1t} + \mathbf{X}_{2t}$. The deviation from trend of the missing component vector, $\mathbf{x}_{2t}$, is estimated by

$$\mathbf{x}_{2t} = a_1\mathbf{x}_{1t} + a_2 z_t + e_{2t}. \tag{5}$$

The components model thus implies that real GNP is described by

$$Q_t^C = \mathbf{X}_{1t} + \mathbf{X}_{2t}^*(1 + a_1\mathbf{x}_{1t} + a_2 z_t + e_{2t}) \tag{6}$$

and that the components estimate of real GNP is

$$Q_t^C = \mathbf{X}_{1t} + \mathbf{X}_{2t}^*(1 + a_1\mathbf{x}_{1t} + a_2 z_t), \tag{7}$$

which is the same as (6) without the error term.

The indicators method estimates the deviation from trend of total real GNP:

$$q_t^I = b_1\mathbf{x}_{1t} + b_2 z_t + e_t^I. \tag{8}$$

On the basis of the relationship given by equation (8), the indicators model implies that real GNP is

$$Q_t^I = (\mathbf{X}_{1t}^* + \mathbf{X}_{2t}^*)(1 + b_1\mathbf{x}_{1t} + b_2 z_t + e_t^I), \tag{9}$$

and the indicators estimate of real GNP is

$$Q_t^I = (\mathbf{X}_{1t}^* + \mathbf{X}_{2t}^*)(1 + b_1\mathbf{x}_{1t} + b_2 z_t). \tag{10}$$

The indicators method, then, estimates the average deviation from trend of both $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$ and applies this average to total trend GNP.

The most illuminating way to compare the two methods is to calculate their respective errors in estimating the true deviation of real GNP from trend (q_t), which by definition is

$$q_t = \mathbf{x}_{1t}\left(\frac{\mathbf{X}_{1t}^*}{Q_t^*}\right) + \mathbf{x}_{2t}\left(\frac{\mathbf{X}_{2t}^*}{Q_t^*}\right). \tag{11}$$

This shows that the true deviation from trend is a weighted average of the deviations of the two component vectors, $\mathbf{x}_{1t}$ and $\mathbf{x}_{2t}$, with weights that shift over time in response to changes in the trend share of the two components of real GNP. The error in the components estimate is

$$q_t - q_t^C = (\mathbf{x}_{2t} - a_1\mathbf{x}_{1t} - a_2 z_t)\left(\frac{\mathbf{X}_{2t}^*}{Q_t^*}\right). \tag{12}$$

To highlight the difference in the error when the indicators method is used, we take the indicators estimate of the deviation of real GNP from trend in (8) and rewrite the estimated coefficient b_1 as the sum of the mean share of the $\mathbf{X}_1$ component during the sample period of the regression, b_{11}, plus a remaining term b_{12}:

$$q_t^I = (b_{11} + b_{12})\mathbf{x}_{1t} + b_2 z_t + e_t^I. \tag{13}$$

This allows us to write the error in the indicators estimate as

$$q_t - q_t^I = \mathbf{x}_{1t}\left(\frac{\mathbf{X}_{1t}^*}{Q_t^*} - b_{11}\right) + \mathbf{x}_{2t}\frac{\mathbf{X}_{2t}^*}{Q_t^*} - b_{12}\mathbf{x}_{1t} - b_2 z_t. \tag{14}$$

Now comparing the errors made in the two methods, we note that the error in the components method (12) contains a single component, that is, the error in estimating the deviation of the residual component of GNP ($\mathbf{x}_{2t}$), multiplied by the true trend share of that component. In contrast, the error in the indicators method has two terms. The first is the error introduced in estimating the deviation of the known element of GNP ($\mathbf{x}_{1t}$) by assuming a constant share estimated from the regression interval rather than by using the available information on the true trend share during the backcast period ($\mathbf{X}^*_{1t}/Q^*_t$). The second term is the same as in the components method, the error in estimating the residual component's share of deviations from trend GNP ($\mathbf{x}_{2t} \times \mathbf{X}^*_{2t}/Q^*_t$), except that this estimate is also contaminated by imposing a fixed trend share rather than the known variable trend shares.

As in our initial analysis for estimation conducted in levels, the components method uses additional information about the relationship between GNP and the independent variables that is not being utilized by the indicators method. However, unlike the case of estimation in levels, the two methods are not in general observationally equivalent. Only if the composition of trend GNP remains the same in both the estimation and backcast periods will the indicators and the components methods be observationally equivalent. But even in that unlikely case, the additional identifying restrictions inherent in the components method will yield more precise parameter estimates. Therefore, in the absence of measurement error, the components method yields estimates of real GNP in the backcast period that incorporate more information than the indicators method.

The Effect of Measurement Error

One of the pitfalls inherent in backcasting a real GNP estimate is the presence of measurement error in the underlying data series. Thus an examination of the sensitivity of the two methods to measurement error is important. In order to simplify the analysis, we assume that some of the components series are measured with error in the extrapolation period but are not measured with error in the estimation period. This assumption reflects the fact that, in general, the quality of almost any data series deteriorates as we move back further in time. As in the preceding section, we also assume that trend values are measured without error.

Suppose that component $\mathbf{X}_{1t}$ is measured with error. The observed component in the extrapolation period is given by

$$\mathbf{X}^0_{1t} = \mathbf{X}_{1t} + e_{mt}, \tag{15}$$

where e_{mt} is the measurement error ($e_{mt} = 0$ in the estimation period). If the extrapolation relationships are estimated in terms of deviations from trend, the estimates of deviations of GNP from trend for the two methods are

$$q_t^C = \left(\frac{\mathbf{X}_{1t}^* + a_1\mathbf{X}_{2t}^*}{Q_t^*}\right)\mathbf{x}_{1t} + \left(a_2\frac{\mathbf{X}_{2t}^*}{Q_t^*}\right)z_t + e_{2t}\left(\frac{\mathbf{X}_{2t}^*}{Q_t^*}\right) + \left(\frac{\mathbf{X}_{1t}^* + a_1\mathbf{X}_{2t}^*}{Q_t^*}\right)\frac{e_{mt}}{\mathbf{X}_{1t}^*} \tag{16}$$

and

$$q_t^I = b_1\mathbf{x}_{1t} + b_2z_t + e_t^I + b_1\frac{e_{mt}}{\mathbf{X}_{1t}^*}. \tag{17}$$

Thus the effect of measurement error depends on the parameter values and the relative trends of $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$. As long as both methods use the same set of extrapolators, as assumed here, the effect of measurement error will be of the same order of magnitude since the coefficients on the e_{mt} terms in (16) and (17) are likely to be of roughly the same size. Recall that the components method still has the advantage of being able to use information about the composition of trend GNP.

Given the superiority of the components method, the issue now becomes the choice of which component variables to include. Obviously, if none of the components were measured with error, then they should be included as extrapolators. However, when a component is measured with error, we run the risk of introducing extraneous noise into our estimate. Consequently, there is a trade-off between the information content of a component and the noise caused by measurement error.

Thus before deciding to include or exclude a component, we should carry out two steps in the research. First, we should study the source notes for each series for signs that measurement error might be present during both the sample period and the backcast period. Are the series based on primary data for each year, or for some years are they based on interpolation with proxy series? What is the coverage of these proxy series? Do the original creators of these primary or interpolated series identify sources of weakness in the data? Second, for series that show signs of measurement error, we can examine the interrelations within the sample period of the regression. Does the introduction of multiple explanatory variables lead to multicollinearity, or do multiple explanatory variables significantly improve the regression equations? If so, then the use of multiple explanatory variables may add more in terms of extra information than what is subtracted by the presence of measurement error.

III. The Basic Data Series and Their Volatility

You Can't Tell the Players without a Playbill

An essential contribution of this paper is to draw on a broader selection of data series than previous studies. Our methodological discussion in the previous section implies that it is desirable to use all the components ($\mathbf{X}_i$) and indicators (Z) variables available, particularly those that are not contaminated with measurement error. The standard annual GNP estimates of Kuznets, Gallman, and Kendrick before 1919 rely simply on commodity output data, "blown up" by assumed ratios to reflect distributive margins. In contrast, this study uses direct information not just on commodity output but also on output in the construction, transportation, and communications sectors.

Of necessity, our expanded coverage implies that our empirical work relies on a complex data set. To aid the reader in keeping track of the different authors, methods, sectoral coverage, and time interval coverage for the many time-series variables utilized in this study, tables 1 and 2 list the dramatis personae. Table 1 identifies and contrasts the alternative estimates of real GNP that form the basic stock of knowledge regarding U.S. real output behavior on which all applied macroeconomic studies have relied. Table 2, discussed subsequently, identifies the alternative annual measures of sectoral output available in the application of the methodology developed in Section II above.

Table 1 begins with the basic Kuznets GNP series on which the subsequent studies of Gallman, Kendrick, and Romer are based. The split between lines 1*a* and 1*b* identifies an important shift in Kuznets's methodology at the year 1919. Beginning in that year, the Kuznets estimates of real GNP are based on the "income-payments" approach, which sums for each year the estimated values of employee compensation, self-employment income, interest, dividends, rents, and corporate profits. Before that year, he used the components method described above, applied to a single component of GNP, commodity output. No use was made of statistical regression for most components of noncommodity GNP; instead total GNP was simply scaled up by multiplying the five major subcomponents of commodity output (consumer perishables, consumer semidurables, consumer durables, producer durables, and construction materials) by fixed ratios representing assumed distributive margins and transportation charges.

TABLE 1

Sources of Annual Real GNP Estimates, Extending Backward from 1929 to 1869, by Author, Time Interval, and Method

Author	Time Interval	Method and Nature of Contribution
1*a*. Kuznets (1961)	1919–28	Income-payments method
1*b*. Kuznets (1961)	1869–1918	Components method: blows up major subdivisions of Shaw's commodity output by assumed ratios
2. Kendrick (1961)	1889–1928	Adjusts Kuznets (1*a* and 1*b*) for different treatment of government
3. Gallman (1966)	1869–1909	Adjusts Kuznets (1*b*) for coverage in census years; no change in annual behavior between census years
4. Commerce (1986)	1909–29	Estimation method not documented
5. Kuznets (1961)	1869–1918	Indicator method, relationship established informally without statistical regression; uses single Shaw commodity output variable
6*a*. Romer (1988)	1909–18	Revises Kendrick (2) by switching to income-payments method from components method
6*b*. Romer (this issue)	1869–1908	Indicator method, relationship estimated by statistical regression; uses single Shaw commodity output variable

The assumed unitary elasticity of transportation, distribution, and construction output to changes in commodity output dominates the behavior of annual changes in the Kuznets estimates of real GNP for 1869–1918. The accuracy of the annual movements in the Kuznets series hinges totally on the elasticity assumptions and on the accuracy of the underlying estimates of commodity output.[3]

Kendrick's series, listed on line 2 of table 1, makes several adjustments, mainly involving the treatment of government spending and tax revenues, to convert the Kuznets estimates to the same conceptual basis as the current Commerce definition of GNP.[4] Because govern-

[3] A qualification is that Kuznets's estimates of construction output shift in 1915 and subsequent years from a scaled-up version of Shaw's (1947) index of construction materials to direct measures of construction output. See our subsequent discussion of table 2, lines 8 and 10.

[4] Kuznets includes just consumption and investment, not government spending, in GNP. But his augmented concept of consumption includes personal tax payments (his estimate of the value of government services to consumers), and his augmented concept of investment includes public investment. Thus Kendrick's conversion procedure involves taking Kuznets's GNP, subtracting personal tax payments and public investment, and adding government purchases of goods and services, as well as unpaid services of financial intermediaries.

ment spending was a relatively small part of GNP before World War I, the Kendrick adjustments are relatively minor in the 1869–1908 period that concerns us most in this paper, amounting to less than 1 percent of GNP, for example, in 1889. The Kendrick adjustments matter mainly for the annual volatility of GNP during and after World War I and, by the addition of government spending, have the effect of making GNP more volatile during this period.

Gallman's corrections to Kuznets's estimates are applied only for census years, primarily 1869 and 1879. He raises the level of GNP in 1869, both by revising the Shaw/Kuznets estimates of commodity output and by using a different ratio to scale up railroad construction from the output of construction materials. Gallman's estimates alter the trend of real GNP between census years but not the volatility of annual real GNP between census years. This paper shares with Romer's the adoption of the Gallman and Kendrick revisions to the original Kuznets GNP estimates for 1869–1908; this is called the "standard" GNP series in what follows.

The next series, labeled "Commerce" in line 4, is that published currently for the period 1909–28 by the Department of Commerce as part of the *National Income and Product Accounts*. However, unlike the NIPA in 1929 and later years, the series before 1929 is not currently maintained by Commerce, but rather has been inherited from work originally done in the 1950s, that is, well before the publication of Kuznets (1961) and Kendrick (1961). Commerce has changed this inherited series only by updating the base year for deflation, which currently is 1982. Two criticisms of the Commerce series have recently been made by Romer (1988). First, the use of 1982 relative prices is inappropriate in evaluating business cycles during 1909–28; in particular, the much higher relative price of government purchases in 1982 than in 1919 or 1929 causes the Commerce series to exaggerate the importance of World War I government expenditures and thus to exaggerate the volatility of GNP during and after the war. This criticism is correct as applied to the published Commerce series but can be overcome since Commerce maintains unpublished data on nominal and real expenditures for 12 components of expenditures, enough to allow recalculation of the series at the prices of any desired base year, for example, 1919 or 1929. The second and more serious criticism is that the methods used in developing the Commerce series are undocumented, in contrast to the copious documentation that underpins the Kuznets, Gallman, and Kendrick series. We follow Romer (1988) in rejecting the Commerce series from consideration as a GNP measure, simply because we do not know the sources of the substantial difference between this series and that of Kendrick.[5]

Next, line 5 lists an alternative Kuznets real GNP series for 1869–1918 based on the indicators method discussed above in Section II. This is not developed from a formal statistical regression, but rather from a "freehand regression curve" fit to a scatter plot for 1909–38 of the deviations from trend of real GNP and real commodity output (Kuznets 1961, p. 537). Detrending is achieved by drawing linear trends between midpoints of 5-year overlapping decadal averages. In the terminology of Section II, the resulting Kuznets real GNP series uses the indicators method with a single explanatory variable, the component $\mathbf{X}_{1t}$.

Romer's recent contributions, listed on lines 6*a* and 6*b*, use two different methods. For 1909–18, Romer (1988) replaces the Kuznets product-side components estimate by a series based on a little-known income-side series published by Kuznets (1961), to which she then applies the Kendrick adjustments to make her resulting 1909–18 series consistent with the existing Kendrick series for 1919–28 and the NIPA for the period since 1929. For 1869–1908, she uses the indicators method, estimating a regression of detrended real GNP (her series for 1909–18 linked to Kendrick's for 1919–28) on detrended Shaw commodity output, covering a split sample period that includes 1909–28 plus 1947–85. Detrending is achieved by running log-linear trends through the value of "standard" real GNP in specific benchmark years (which are listed and discussed further below). Her coefficient on commodity output is allowed to interact with a time trend, although the time trend is statistically insignificant, indicating support for a constant response over her entire sample period. Romer's 1869–1908 estimates amount to an update of Kuznets's indicators estimate of real GNP, using the same information set but (1) using formal regression rather than a freehand regression, (2) extending the sample period to include the postwar years but to exclude 1929–38, and (3) using a trends-through-benchmarks rather than moving average method of detrending.

[5] The use of the Kendrick series in preference to the Commerce series is an important difference between this paper and our earlier effort in this area distributed as NBER working paper no. 1999. Romer's criticisms of the Commerce series were developed after that paper was written. Readers should note also that the Commerce series is the basis for the nominal and real GNP series published in Friedman and Schwartz (1982) and Gordon (1986). In Sec. IV below we note that the cyclical behavior of real GNP in 1919–23 is still open to dispute and that there are arguments to support the Commerce version. But not knowing how the Commerce version was compiled, we join Romer in abstaining from any use of the series.

Alternative Annual Series for Components and Indicators

No primary annual real GNP data exist before 1909. Thus the estimation of real GNP poses a problem of the optimal aggregation of available information. The most important elements of the available information set are listed in table 2. The methodological discussion in Section II allows the information set to include both components of GNP (e.g., manufacturing or construction output) and indicators of economic activity that may be correlated with real GNP (e.g., the unemployment rate and stock market prices). We edit the list of series in table 2 by concentrating on estimates of output or highly correlated indicators of output in specific sectors of GNP. Indicators that may be correlated with real GNP are excluded because most are subject to difficult questions of estimation and interpretation. For instance, existing estimates of unemployment are not independent of the real GNP series listed in table 1 and are currently subject to a debate regarding estimation methodology between Romer (1986*b*) and Weir (1986). Stock market prices are a fragile indicator in the light of the ongoing debate regarding the extent to which stock market prices reflect "fundamentals," that is, underlying economic conditions relevant for the prediction of future earnings, or may rather be subject to "bubbles," that is, upward or downward movements in prices that are not justified by fundamentals.

In this light, all the annual indexes listed in table 2 are measures of or close proxies for real output in one of three major sectors of GNP. First listed is the basic Shaw (1947) commodity output series, which he originally constructed for 1869, 1879, and annually for 1889–1938 and which was extended by Kuznets to provide annual estimates for 1869–88. The Shaw series for total commodity output and the various sectoral breakdowns, for example, producer durables and construction materials, is the basis for all previous annual pre-1909 real GNP estimates.[6]

Listed next in lines 2 and 3 of table 2 are two overlapping measures of manufacturing output, Frickey's (1947) index for 1860–1914 and Fabricant's (1942) index for 1899–1939. These indexes share with

[6] Henceforth we use the phrase "commodity output" as synonymous with the sum of value added in the agriculture, mining, and manufacturing sectors. While Shaw's data on commodity output include construction materials, such materials are produced entirely within the agriculture, mining, and manufacturing sectors. Value added in the construction sector in postwar data does not correspond to Shaw's estimates of construction materials but rather consists mainly of labor and profit income earned in the construction sector. Construction output consists of the output of construction materials plus construction value added. The readily available postwar counterpart of construction output is "Structures GNP" in table 1.4 of the NIPA.

TABLE 2

Annual Sources Available on Behavior of Major Components of Real GNP before 1929, by Sector, Time Interval, and Source

Sector	Time Interval	Source*
	Commodity Output and Manufacturing Production	
1. Commodity output	1869–1938	1889–1938, Shaw (1947), extended back to 1869 by Kuznets (1961) [6C–D]
2. Manufacturing	1860–1914	Frickey (1947) [7]
3. Manufacturing	1899–1939	Fabricant (1942) [7]
4. Industrial production	1919 to date	Federal Reserve Board [7]
	Transportation and Communication	
5. Transportation and communication	1860–1914	Frickey (1947) [8]
6. Transportation	1889–1953	Kendrick (1961) [8]
7. Telephone and telegraph	1889–1953	Kendrick (1961); telephone only before 1929 [8]
	Construction	
8. Construction materials	1869–1938	1889–1938, Shaw (1947), extended back to 1869 by Kuznets (1961) [9A–B]
9. Construction output	1869–1938	Kuznets (1961) [10B]
10. Nonfarm building	1850–1939	Gottlieb (1965) [11C]
11. Construction value and volume	1915 to date	Labor-Commerce [11B]

* Number in brackets refers to series number in the Data Appendix.

Shaw's a dependence for basic data on the *Census of Manufactures*, which was conducted each decade until 1899, quinquennially for 1899–1919, and biennially for 1919–39. They share with Shaw's a dependence on partial and fragmentary data for interpolation between census years. But they differ from his series in a basic way: Shaw took nominal production at the disaggregated level and deflated by a price index, whereas Frickey and Fabricant base their indexes on a count of units produced. Below we discuss in more detail the advantages and limitations of these series. The first section of table 2 concludes on line 4 with the well-known Federal Reserve Board index of industrial production, which commences in 1919. No use is made in this study of the index since it does not extend back to 1909 and thus cannot be used as an explanatory variable in our regression equations. It is displayed here only to provide a measure of postwar volatility comparable to that of the Fabricant index during 1899–1928.

The next section of table 2 covers available annual indexes of transportation and communication. One of our most surprising discoveries is that no previous study of real GNP has made any use at all of the copious data available on transportation output, which in the 1869–1928 period primarily consisted of railroad output. Because railroads were regulated after 1890, the available data on annual output are virtually complete, and there is little doubt that data on railroads provide the most reliable single index of annual fluctuations in output over the period 1890–1929. As shown in table 2, we make use of three annual indexes. For 1869–89, Frickey's index provides a weighted index of output for railroads, street railways, canal traffic, coastal trade, telephone conversations completed, telegraph messages transmitted, and revenue from postage stamps.[7] Kendrick's annual transportation index begins in 1889 and is primarily based on railroad output through 1902, after which annual estimates for street railways are included. Line 7 summarizes Kendrick's communications index, which for annual movements is based entirely on telephone calls prior to 1929 and for both telephone and telegraph thereafter.

The final section of table 2 lists some of the available annual indexes of construction output. Shown first on line 8 is Shaw's measure of the output of construction materials, which he computed as the value of construction materials divided by individual components of the wholesale price index. Next on line 9 is Kuznets's measure of the construction component of real GNP, developed by scaling up Shaw's construction materials series, which Kuznets extended back on an

[7] All these components are measured in real terms except for postage stamp revenue. We did not bother to deflate this component since its weight is less than 5 percent of the overall index.

annual basis from 1889 to 1869. It is important to note that Kuznets shifts sources in 1915 from the scaled-up Shaw index to the Labor-Commerce series listed in line 11 and described below (see Kuznets 1961, pp. 585, 587, table R-30, nn. to cols. 5, 10).

The Gottlieb (1965) measure listed on line 10 is quite different from the pre-1915 Kuznets series. Instead of scaling up the output of construction materials, Gottlieb developed direct measures of construction output.[8] The methodology involves establishing totals of construction by decade from census data on assessed values of taxable total real estate and on the value of farms. The wealth data are converted into estimates of new building for each decade by using the relationship between wealth and new building in Ohio. Then annual movements within each decade are determined as the average of several independent interpolating series, including Gottlieb's own series for Ohio, and two different series on the value of building permits developed in earlier research by John Riggleman, Walter Isard, and Clarence Long. We choose to use Gottlieb's annual construction index rather than to rely on earlier research, for example, that of Abramovitz (1964), because Gottlieb's research was the last to be completed of the various alternatives and combines results of previous investigators with his own largely independent research on construction in Ohio.

Finally, line 11 lists the basic data source on construction activity since 1915, the volume of "construction put in place" in both current and constant dollars, as maintained jointly by the Departments of Labor and Commerce until 1959 and by Commerce alone since then. This data set is available for 18 separate categories of private and public residential and nonresidential construction, although here we use only the total. As stated above, this series is used by Kuznets beginning in 1915, and, as we shall see, it causes Kuznets's construction series to behave quite differently from Shaw's index of construction materials during the 1915–28 interval.

The Volatility of Real Output Series

We turn now to the major focus of the paper, the comparison of the prewar and postwar cyclical volatility of alternative estimates of real GNP. We require first a definition of "prewar" and "postwar" and a definition of volatility. Because the greater volatility of economic activity during the Great Depression is not at issue, the various prewar time periods displayed in table 3 terminate in 1928. For the postwar

[8] Gottlieb's only table of annual data provides a listing of his estimates in current dollars, not constant dollars. We have deflated his current-dollar series by a standard construction cost series (*Historical Statistics of the United States*, ser. N139 linked to N138 prior to 1889).

TABLE 3

STANDARD DEVIATION OF DEVIATIONS FROM TREND OF ALTERNATIVE AGGREGATE AND SECTORAL REAL OUTPUT SERIES, 1869–1986, DETRENDED BY LOG-LINEAR TRENDS THROUGH BENCHMARKS

Series and Years Covered	1869–88 (1)	1889–1908 (2)	1908–28 (3)	1869–1908 (4)	1869–1928 (5)	1954–72 (6)	1947–86 (7)	(4)/(6) (8)	(5)/(7) (9)
	Alternative Real GNP Series								
1. Standard (Commerce, 1909–28)	5.24	4.82	6.26	4.97	5.40	2.93	2.92	1.70	1.85
2. Standard (Kendrick, 1909–28)	5.24	4.95	4.57	5.03	4.89	2.93	2.92	1.72	1.67
3. Romer	4.06	3.22	3.69	3.63	3.68	2.93	2.92	1.24	1.26
4. Line 2, with alternative deflators	5.39	7.13	5.94	6.45	6.25	3.17	2.89	2.03	2.16
	Sectoral Components and Indicators								
5. Shaw commodities	6.62	6.00	4.50	6.24	5.71	4.96	4.76	1.26	1.20
6. Manufacturing production index	9.48	9.18	10.44	9.25	9.65	5.94	5.65	1.56	1.71
7. Transportation and communication	7.27	7.71	6.36	7.50	7.65	3.56	4.77	1.62	1.60
8. Shaw construction materials	10.31	10.24	14.46	12.36	13.11	...	...	...	...
9. Kuznets construction	10.35	9.62	24.15	12.03	18.05	3.92	7.41	3.07	2.44
10. Gottlieb-Commerce construction	17.85	16.74	20.18	19.28	20.69	3.92	7.41	4.92	2.80

SOURCE.—The following list cross-references the series used on each line to the series documented in the Data Appendix. Line 1 [1A–C] indicates that the source of the series computed on line 1 in this table is given in the notes to series 1A–1C in the Data Appendix. 1 [1A–C]; 2 [2A–C]; 3 [5]; 4 [4]; 5 [6A–D]; 6 [7]; 7 [8]; 8 [9A–B]; 9 [10A–B]; 10 [11A–C].

period we take a comprehensive definition, 1947–86, and also a limited definition covering 1954–72, the period after the end of the Korean War but before the beginning of oil shocks. We note in table 3 in columns 6 and 7 on any of the first three lines that our measures of postwar volatility are almost equal over the alternative 1954–72 and 1947–86 periods. Nevertheless, we maintain the distinction between the two separate periods for comparison with the prewar period since several series other than real GNP (e.g., construction and various deflators) behave quite differently in the two alternative postwar intervals.

The choice of a measure of volatility suitable for comparing the pre-1929 and postwar periods centers on two issues, alternative methods of detrending and alternative measures of the volatility of various detrended series. In this version of the paper we join Romer in adopting the method of log-linear detrending through selected benchmark years. To avoid debate, we adopt her selection of years that are, roughly speaking, "normal," which are 1873, 1884, 1891, 1900, 1910, 1924, 1947, 1955, 1962, 1972, and 1981. This choice eliminates two sources of disagreement in previous versions of our paper and Romer's, namely, methods of detrending and the choice of benchmark years.[9] Our only difference is to adopt the initial year of the investigation, 1869, as an additional benchmark year, reflecting our finding that major deviations from trend, partly or largely spurious, sometimes arise when the 1873–84 trend is extended backward to 1869. Since 1869 was a census year, it is natural to establish it as an initial benchmark. The overall effect of this extra benchmark is to reduce the volatility of GNP during 1869–72 and hence of prewar GNP overall in comparison with postwar GNP.

Finally, we share with Romer the use of a single volatility measure, standard deviations of deviations from trend.[10] Other measures, for example, standard deviations of percentage changes, are inadequate measures of cyclical volatility when trend growth rates vary over secu-

[9] This approach is a compromise. Kuznets detrended by a moving average method over a relatively short 5-year time period that implied that trend output in the late 1930s was equal to normal output and that the Great Depression (measured by the deviation of actual output from trend) had disappeared. Romer's first draft followed Kuznets by adopting a moving average detrending method, 7 years in her case (7MA), and our first draft responded to this choice by comparing results calculated with the alternative 7MA and trends-through-benchmarks method, using benchmarks different from those in the present paper, i.e., those used in previous research (Gordon 1986, app. A). Romer then responded by adopting our trends-through-benchmarks method, albeit with a different choice of benchmarks, and we have responded in turn by adopting her benchmarks, with the exception of 1869 as noted in the text.

[10] All such deviations reported in this paper are presented as percentages and are calculated from natural logarithms, e.g., $q_t = 100 \times \log(Q_t/Q_t^*)$.

lar intervals. Clearly, the reliance in both papers on standard deviations of deviations from trend raises the possibility that results could be sensitive at least in part to alternative choices of benchmark years. But there is a limit to the dimensions of sensitivity that can be tested, and so we present our results with only a single set of benchmark years and abstain from sensitivity comparisons with other sets.[11]

The first seven columns in table 3 display standard deviations of deviations from trend in three separate 20-year prewar intervals, two alternative summary prewar periods ending in, respectively, 1908 and 1928, and the two alternative postwar periods. The last two columns provide two alternative ratios of prewar to postwar volatility, with the truncated prewar and postwar periods compared in column 8 and the full prewar and postwar periods in column 9. Reading down the table, we begin with "standard" measures of real GNP, linking Kuznets (1869–88) to Kendrick (1889–1928) to NIPA (1929–86). The only difference between lines 1 and 2 is the use in the former of the undocumented Commerce series for 1909–28 in place of Kendrick's. Adopting the Kendrick series in preference to Commerce for 1909–28, we find in the right-hand two entries in line 2 our "basic standard" ratios of prewar to postwar volatility, 1.72 and 1.67. Thus our starting place is the presumption that the economy in 1869–1928 was about one and two-thirds times as cyclically volatile as the postwar economy.

Romer's two new series, when linked to NIPA in 1929, imply a radically different conclusion that the prewar economy was only about one and one-quarter times as volatile as the postwar economy, as shown on line 3. Her results thus call into question the long-standing presumption of postwar stabilization. However, as we shall show, improved estimates do not all point in the direction of lower volatility.

Since all estimates of GNP before 1909 are based on fragmentary evidence about the behavior of particular subcomponents of GNP, conclusions depend on which subcomponents are included. The importance of this choice is evident in the bottom part of table 3, which displays volatility measures for the major available indexes. The Shaw commodities output series displays a much smaller prewar/postwar volatility ratio than any of the other series listed, when compared with its closest available postwar equivalent (NIPA GNP originating in agriculture, mining, and manufacturing). The alternative manufacturing output series obtained by linking the Frickey and Fabricant

[11] Romer (this issue) shows that her prewar/postwar volatility ratios are robust to the choice of an alternative set of benchmark years, and we assume that the same would be true for our results.

indexes with the postwar Federal Reserve Board index for manufacturing yields a substantially higher prewar/postwar ratio. The ratios for transportation and communication are roughly 1.6, only a bit less than the basic ratio for standard GNP. Finally, and of considerable importance for our final results discussed below, the volatility of construction is enormously greater than that in the postwar period, whether the Kuznets or alternative Gottlieb and Commerce construction series is used (each is linked at 1929 to NIPA structures GNP). Clearly, by using evidence only on commodity output and neglecting direct evidence on output in the transportation, communication, and construction sectors, both the standard and Romer GNP estimates may be biased toward indicating too *little* prewar volatility.

IV. Data Characteristics and the Choice among Methods and Variables

Changes in Sectoral Shares over Time

Our analysis in Section II identifies as a potential advantage of the components method the ability to take into account shifts over time in the share of components in GNP. The data sources listed in table 2 allow us to assess the importance of such shifts in sectoral output shares. Table 4 divides the standard real GNP series into three components for which data exist—commodities, transportation-communication, and construction—and a residual component. From the last century as a whole, it is evident that until 1947 the share of commodity output is relatively stable in the range of 36–39 percent. Until 1918 this phenomenon is true by assumption rather than as a revealed fact about the economy since the aggregate real GNP series was little more than a scaled-up version of commodity output. After 1919 measures of real GNP and commodity output are independent.

TABLE 4

Shares of Major Sectors in Selected Benchmark Years, Constant 1929 Prices

Year	Commodities	Transportation-Communication	Construction	Residual
1873	35.7	4.3	15.8	44.1
1891	38.8	5.4	18.5	37.3
1910	38.4	8.3	14.6	38.7
1924	35.8	8.5	12.3	43.3
1947	39.4	12.0	8.8	40.0
1962	35.7	7.9	11.7	44.8
1981	33.5	9.1	3.3	51.9

Note.—Prewar and postwar series are linked as described in the Data Appendix, series 6, 8, and 11.

However, our measures of transportation-communication and construction are independent of real GNP from the beginning, and they reveal very substantial shifts in sectoral shares, with an especially marked increase in the share of transportation-communication from 4.3 percent in 1873 to 12.0 percent in 1947 and marked oscillations in the share of construction from a high of 18.5 percent in 1891 to a low of 3.3 percent in 1981. Since we do not use postwar data in our subsequent regressions that are used to backcast real GNP before 1909, only the changes in shares in the prewar period are relevant, but these are not trivial. Hence, this aspect of the historical record supports the view that components estimates of real GNP are likely to be more accurate than indicators estimates that essentially discard available information about secular changes in sectoral shares.

Data Quality and the Choice of Variables

To implement our earlier treatment of measurement error, we now review the strengths and weaknesses of the most important series, both those that have been used before and those that are used here for the first time. There has been a tendency in some recent research to simplify the universe of available data in a white and black fashion as either uniformly excellent or mediocre. The reality is far different. Most data series deteriorate as one goes back in time, and most authors, particularly Kuznets and Shaw, are aware of this and warn the reader of the limitations of their own estimates.[12] In this section we point out some of the weaknesses of the various series that are used in previous research and of the additional series that we employ. An important conclusion is that the two series likely to be most accurate in measuring year-to-year cyclical behavior have never before been used in estimating real GNP behavior. These are the relatively complete indexes of transportation and communication output provided by Frickey and Kendrick and the studies of consumer prices by Hoover and Rees.

The discussion is arranged by our rough ranking of data quality. The series discussed below are, in order, the measures of sectoral output in transportation and communication, the Hoover-Rees research on consumer prices (as contrasted to other sources of price behavior used by Kuznets and Shaw), the Shaw and Fabricant mea-

[12] Because the language used by Romer differs so drastically in tone from Shaw's own qualifications, discussed below, it is worth quoting some of her descriptions of the Shaw commodity output series: "the Shaw series appears to be quite accurate" (sec. II); "while the Shaw series is quite accurate" (sec. II); "the Shaw series is a particularly good interpolating series because it is very consistent over time" (sec. V).

sures of commodity output, the Kuznets income-side estimates, and the Gottlieb-Commerce estimates of construction output.

In any study of annual cyclical volatility, a basic distinction must be made between data sources that provide primary data on annual changes and those that provide primary data only in periodic census years while interpolating between census years on the basis of less complete data series or proxies. With two exceptions—the estimates of transportation output and consumer prices—all the other data series discussed in this section use interpolating series.

The great advantage in using data on transportation and communications is that these have been regulated industries for most of the period under study. Kendrick's index of railroad output, based on a weighted average of passenger-miles and ton-miles, is "substantially complete" after 1910 and is adjusted for "slight" undercoverage in earlier years. The Kendrick index also includes minor forms of transportation, mainly street railways, with increasing coverage in later years. These relatively complete annual transportation data are supplemented by annual data collected by Kendrick on telephone output, representing the bulk of the communications industry.[13] The transportation and communications index is extended back from 1889 to 1869 on the basis of relatively complete annual data compiled by Frickey (1947). The Frickey estimates were laboriously compiled from the leading manual of that time on railroad traffic, as supplemented by census records and annual reports of the railroads themselves. The sampling was at the rate of roughly 50 percent in each of six geographical regions, much greater than, say, the representation of Shaw's interpolators for 1889–98.[14] Frickey supplemented the annual data on railroad output, which carry more than two-thirds of the weight in his transportation-communication index, with, in descending order of importance, street railways, coastal trade, revenue from postage stamps, telephone conversations completed, and Sault Sainte Marie canal traffic.[15]

The deflators that underlie all the basic estimates by Shaw, Kuznets, and Kendrick can be contrasted with the quite different behavior of measures of consumer prices compiled by Hoover (1960) and Rees (1961). Kuznets did not make use of the Hoover-Rees indexes because they were completed well after his basic estimates were carried

[13] Kendrick indicates that the telephone output data are interpolated on the basis of an index supplied by Fabricant, but this would be relatively complete as it comes from AT&T, and in any case communications output is well under 10 percent of transportation output during the 1889–1908 period (it rose to 16 percent in 1929).

[14] See Kendrick (1961, pp. 509–10). Shaw estimates the coverage of his interpolators to be less than 25 percent for 1889–98.

[15] The weights are given in Frickey (1947, p. 112).

out. He was aware of the fact that the new indexes "show less decline from the 1870's to the 1890's, and consequently less rise from the 1890's to World War I, than the price index implicit in our estimates . . . . The long-term trends in the flow of goods to consumers in constant prices would be relatively little affected . . . . The effect on rates of growth over shorter periods is more marked" (Kuznets 1961, p. 512). As we show subsequently, the differing year-to-year behavior of the Hoover-Rees and Shaw-Kuznets price indexes is substantial and, taken together, implies quite implausible price behavior in the service and distribution sectors.

To underscore the achievement of Hoover and Rees, we begin with the fact that neither Shaw nor Kuznets used *any* direct information on prices actually paid by consumers. Shaw converted his primary data on nominal commodity output into real output by using wholesale price indexes, which, he recognized, became increasingly inadequate in coverage as he went back in time. More important for our purposes, he recognized that his increasing need to substitute price indexes of crude commodities for the missing indexes of final products imparted to his implicit deflators an excessive cyclical volatility, which, with no evidence, he called "slight" in the following passage: "Of graver import is the lack of any price series whatever for many commodities. Some of the gaps were filled by using indexes of the chief materials that enter into a commodity. . . . The use of indirect series tends to make the composite indexes fluctuate a little more than they would if based on direct series alone, for it is generally recognized that prices of materials usually fluctuate more than prices of end products. We believe, however, that the better trend representativeness more than compensates for this slight cyclical defect" (Shaw 1947, pp. 288–89). If we read closely in Shaw's notes on his price indexes, we find repeated indications that indexes for specific finished goods are lacking in earlier years, for example, no index for household appliances or luggage before 1913, for the important categories of horse-drawn passenger vehicles before 1907, nor of locomotives and railroad cars before 1910. The interpolations by Kuznets that extend the annual commodity output estimates back from 1889 to 1869 are convoluted but appear to rely on remarkably few price series (see Kuznets 1946, pp. 90–101 [nn. to "Basic Tables"]).

In contrast, the Rees contribution for 1890–1914 was to find as much evidence as possible on prices actually paid by consumers. He combined a previous study of retail prices of food by Paul Douglas with a detailed study of the prices of clothing and home furnishings from mail-order catalogs, of rents advertised in newspapers of six large cities, of a survey of gas companies of retail price schedules in force, and of other diverse sources. For our purposes it is important to note that most of the data on individual prices are collected annually, not interpolated, and that a substantial effort is made to correct for quality change in clothing and home furnishings. While the use by Rees of mail-order catalog data might suggest that his consumer price series is too sticky, that is, cyclically insensitive, on an annual basis, our own extensive data bank on mail-order catalog prices in the postwar period indicates no tendency for catalog prices to be stickier than the corresponding producer price indexes for individual commodities.[16]

Hoover's work, which covers the period 1850–80, is based on prices paid on an annual basis by one or two respondents in each of more than 40 cities in 16 states. The main defects are a limited correction for secular quality change and the stopping point of 1880, leaving a gap between 1880 and 1889 during which records of consumer prices are not based on actual retail quotations but rather are estimated by Hoover from wholesale price quotations. Interestingly, Kendrick had a high evaluation of Hoover's work, even though he did not use it in his own real GNP series. He speculated in a conference comment that Hoover might have exaggerated the limitations of her own work: "one wonders if these early statistics are really much inferior to the data underlying modern indexes" (Kendrick 1960, p. 187).

Next in reliability is the Fabricant index of manufacturing output that begins in 1899 and is linked in table 3 before that date to the Frickey manufacturing index. Romer's previous (1986*a*) criticism that prewar production indexes are too volatile applies just to the Frickey index that is based on only about 40 individual products, almost all of which are crude or intermediate products. The Fabricant (1940) index, not mentioned by Romer (1986*a*), is completely different since it is based on about 1,000 individual products (see Fabricant [1940, pp. 382–600] for a detailed list).[17] Like the Shaw index of commodity output, the Fabricant index is based on primary data in years of the *Census of Manufactures* and is interpolated between census years. The advantage of the Fabricant index is that it is based on units produced rather than deflated nominal value, thus avoiding the need to use possibly error-prone deflators, as Shaw does. For long-term comparisons an index such as Fabricant's that is based on a count of units produced is more prone to miss improvements in quality and thus may be inferior to a deflated value index such as Shaw's. But for

[16] This statement is based on more than 7,000 price observations collected from mail-order catalogs and matched to individual producer price index observations for eight-digit detailed commodities. Quality-corrected catalog prices tend to drift down relative to the producer price index observations on a secular basis with no evident cyclical characteristics.

[17] In a companion volume, Fabricant (1942) provides companion data series on employment, hours, and productivity.

cyclical comparisons, Fabricant's index is not subject to the error that may be introduced by Shaw's index by overly volatile deflators. Note that any quality change bias would cause the Fabricant index to *understate* the volatility of real output; if the average quality of units declines in recessions, then "true" constant-quality output declines more than Fabricant's index.[18]

Both the Fabricant and Shaw indexes are vulnerable to the criticism that the interpolating indexes are not as comprehensive as estimates for census years. It is well to recall Shaw's own misgivings about his annual interpolations, which are based primarily on data from individual states. He describes severe problems with the inadequate number of states and incomplete coverage within states. For 1889–98, interpolation is based on only a single state (see Shaw 1947, pp. 94, 97). Shaw's "over-all rating of the series based on state data . . . is fair for 1899–1919 and poor for 1889–99" (p. 98). In a detailed listing of his series by overall rating on three different criteria, he lists eight good, 25 fair, and 12 poor for 1899–1919, and none good, eight fair, and 32 poor for 1889–99 (p. 101). It must follow that Kuznets's backward extension of Shaw's annual estimates for 1869–88 is even more inadequate than Shaw's "poor" indexes since Kuznets used even fewer data series as interpolators for both nominal values and deflators.

The time periods shown in table 3 are not appropriate for comparing the volatility of the Shaw and Fabricant series since the figures shown for 1889–1908 represent a mixture of those of Frickey and Fabricant. The appropriate comparison is for 1899–1928, for which the Shaw series exhibits a standard deviation of deviations from trend of 4.66, roughly half the Fabricant figure of 9.20. The associated prewar/postwar ratios (using 1947–86) are 0.98 and 1.63, respectively.[19] In the light of the comparison above, we believe that the Fabricant measure provides the more accurate indicator of cyclical volatility, and we are reassured by the fact that the prewar/postwar volatility ratios of about 1.6 are almost exactly the same as those for the transportation-communication index (table 3, line 7).

We would rate the Kuznets income-side estimates of GNP for 1919–28 to be little better than Shaw's or Fabricant's and possibly

[18] Another difference is that Shaw's coverage includes only finished goods, while Fabricant included also intermediate manufactured goods. This partly accounts for the higher volatility of Fabricant's index. To control for this difference, we compare Fabricant's prewar index only with the Fed's index for the postwar period and Shaw's index only with value added in the commodity-producing sectors (agriculture, mining, and manufacturing).

[19] The prewar Shaw and Fabricant indexes are compared here with different postwar indexes (see n. 18).

TABLE 5

Alternative Measures of Economic Activity, 1919–23 (1919 = 100)

	1919	1920	1921	1922	1923
			Real GNP		
1. Commerce	100.0	92.2	85.7	98.3	109.1
2. Kendrick	100.0	98.9	96.6	102.2	115.8
3. Kuznets variant A	100.0	100.7	94.9	101.5	115.7
4. Kuznets variant B	100.0	96.0	87.9	94.7	107.3
			Labor Hours		
5. Kendrick	100.0	99.3	89.8	96.1	103.5
			Real GNP per Hour		
6. Kuznets variant A	100.0	101.4	105.7	105.6	111.8
7. Kuznets variant B	100.0	96.7	97.9	98.5	103.6

Source.—Lines 1 and 2: same as table 1; line 3: Kuznets (1946), table I-15, col. 7; line 4: line 3 plus the difference between alternative measures of flow of services; Kuznets (1946), table I-4B, col. 5, minus table I-4A, col. 5; line 5: Kendrick (1961), table A-X; line 6: line 3 divided by line 5 (%); line 7: line 4 divided by line 5 (%).

worse. Whereas at least Shaw and Fabricant had *some* data available for year-to-year interpolation, Kuznets has virtually none in a crucial sector, consumer services. This skeptical assessment may seem surprising since Kuznets's income-payments estimates have been accepted without question by previous investigators, including Kendrick and Romer, and indeed we adopt the Kendrick version as the dependent variable in our basic regressions below. However, the Kuznets data for the 1920s are relatively weak in coverage of entrepreneurial income, which is a crucial gap in view of the importance of self-employed proprietors in providing consumer services.

In assessing the Kuznets estimates, it is useful to study the behavior of alternative series during the aftermath of World War I, the five years 1919–23, as in table 5. Kuznets's variant A behaves as the Kendrick series used below in our regressions, showing a relatively small decline in 1919–21, while variant B is much closer to the Commerce series. The two Kuznets series differ in that variant A is based entirely on income estimates, calculating consumer services as a residual, while variant B is based on independent series on consumer services (Kuznets 1946, table I-4B, nn. to col. 1).

While we do not know which series is correct, the large difference between them raises a question regarding the accuracy of the Kuznets variant A series, on which the Kendrick data are based. If anything, the implied behavior of productivity using variant A, as shown on line 6, is not very plausible since productivity is shown to rise by 4 percent from 1920 to 1921, whereas it normally declines during short recessions. Kendrick records a 7.4 percent drop in labor productivity in 1907–8 and similar sharp declines in 1913–14 and 1929–30. Further, we might also ask how consumers could afford to raise their *real* expenditures on services by 13 percent, as in Kuznets's variant A over 1919–21, when over the same 2-year period hours worked in manufacturing declined by 28.3 percent and in transportation by 12.2 percent (Kendrick 1961, tables D-II, G-II). With manufacturing hours declining by nearly two-thirds as much as in the Great Contraction of 1929–33, how could the rest of the economy be entirely unaffected, as implied by the Kuznets income-payments series? These are questions worthy of further research; for his paper we include them only to suggest that the 1919–28 Kuznets income-payments estimates are not without possible flaws, and until these questions are resolved we are unwilling to follow Romer (1988) by adopting Kuznets's even cruder income-payments estimates for 1909–18 as the basis for measuring real GNP volatility.

We come last to the spliced Gottlieb-Commerce construction output series. The high volatility of construction output in the prewar period relative to the postwar period is central to the overall question of postwar stabilization. There is little question regarding the reliability of the Commerce series that begins in 1915, and in fact Kuznets switches in this year from use of scaled-up Shaw construction materials to the Commerce series. The Commerce series on "value of construction put in place" has been compiled continuously since 1915 using the same methods. It is calculated annually and so does not rely on interpolators. Thus we have some confidence that the huge difference in volatility shown in table 3, line 10, for 1908–28 versus either postwar period is based on consistent and relatively reliable data.

For the period before 1915 the Gottlieb data, as described above in Section III, are based on a complex process in which total construction is estimated over decades, and then each decade is interpolated separately. How adequate are these interpolators? One source of data is detailed county-by-county assessment records for Ohio, which have relatively complete coverage during 1870–1910 (Gottlieb 1966) and which have not been used by investigators prior to Gottlieb. The other sources are two different series on the value of building permits. Coverage is limited to 10 cities in 1868 and reaches 70 cities in 1900 and subsequent years. The geographical coverage of these cities is equivalent to or slightly better than the coverage of Shaw's interpolators for commodity output prior to 1899.[20]

Further, the Gottlieb construction series by definition includes all

[20] For details on the Long and Riggleman-Isard building permit series, see *Historical Statistics of the United States,* ser. N111 and N114.

the components that constitute a dwelling or nonresidential structure, whereas the Kuznets-Shaw series on construction materials (on which all previous real GNP estimates have been based) is very fragmentary in coverage, particularly before 1889. Kuznets took Shaw's output of construction materials in the census years 1869, 1879, and 1889 and interpolated between these years using data only on the production of nails and rails and of lumber for 1869–79 and adding cement and roofing slate in the period 1879–89. This short list of materials is a relatively flimsy basis on which to measure the volatility of construction output, and use of the Gottlieb construction output series provides additional information. Below in table 9 we report alternative results for backcast GNP that replace the Gottlieb-Commerce construction output series with the Kuznets-Shaw measure of construction materials output.

Two factors reassure us about the Gottlieb series, which is linked at 1915 to the Commerce series and used in the subsequent regression analysis. First, in the 1908–28 period dominated by the "good" Commerce data, the volatility of construction output is much higher than that of Shaw's construction materials series, 1.4 times for the Gottlieb-Commerce series (table 3, line 10 vs. line 8) and 1.7 times for the Kuznets series (line 9 vs. line 8). This suggests that the high ratio of Gottlieb to Shaw volatility before 1908 is not inconsistent with the experience based on superior data after 1915. Second, and more important, the volatility evident in the Gottlieb series before 1908 does not consist of choppy, erratic year-to-year jumps, but rather long, smooth waves in which construction rises far above trend for one decade and then falls well below trend in the next decade. In short, the Gottlieb series amply documents the famous "Kuznets building cycle," and this aspect of the Gottlieb data hinges not just on the annual interpolators but also on decadal averages that are cross-checked with assessment records and other data.

Some details are useful: The detrended and deflated Gottlieb series (linked to Commerce after 1914) is negative in every nonbenchmark year from 1874 to 1881 and from 1911 to 1923. It is positive in all years but two between 1882 and 1892 and in every year between 1899 and 1907, as well as between 1924 and 1929. During the years of "good" data, construction output displays an increase of 125 percent between 1920 and 1926. During the years covered by the Gottlieb data, construction output also more than doubles over periods of similar duration, for example, increasing 125 percent between 1878 and 1884, and 154 percent between 1898 and 1905. Taken together, the spliced Gottlieb-Commerce series records a story of continuing decade-long swings in construction activity that in frequency and amplitude equally characterize the periods 1869–1914 on the basis of the

Gottlieb data and 1915–28 on the basis of the Commerce data. The regular 20-year Kuznets cycle extends from 1869 (or earlier) through 1941 but has disappeared in the postwar period.

Implications for the Choice of Data

To carry out either the components or indicators method of backcasting real GNP before 1909, we need consistent data extending before and after 1909 on both real GNP and as many sectoral components as possible. We choose the "standard" Kuznets-Kendrick data series for 1869–1908 to establish the trend level of real GNP in benchmark years. For 1909–29 we choose Kendrick in preference to Commerce, despite the possibility raised in table 5 that the Kendrick data show too little decline in real GNP in 1919–21, in order to tilt our results toward a finding of lower prewar real GNP volatility. Not only is the Kendrick series less volatile than Commerce for 1909–28, but it yields smaller regression coefficients and hence a smoother path of real GNP before 1909 than Commerce does. Similarly, we choose to adopt the Shaw measure of commodity output as an explanatory variable in preference to the Frickey-Fabricant series, in view of the limited coverage of final goods output by Frickey before 1899. Also, by choosing Shaw, our results can be compared directly with Romer's (this issue).

For transportation and communication, we have the single-linked series already described. For construction, we use the Gottlieb-Commerce series in preference to the Shaw construction materials series (or the Kuznets construction series based on Shaw before 1915), in view of the evidence from the good construction data for 1915–28 that the Shaw series is insufficiently volatile. We cannot use the Kuznets construction series because it switches methods in 1915 in a way that would lead to a downward bias in the volatility of the backcast real GNP series.[21]

This leaves as unutilized the superior Hoover-Rees data series on prices paid by consumers. To gain perspective, we begin by creating backcast real GNP series based on the measures chosen in this section that use conventional deflators. Subsequently we examine evidence

[21] The particular way Kuznets switches methods in 1915 would lead us to substantially understate prewar volatility if we were to use the Kuznets construction series in our regressions. Because the Kuznets series is volatile after 1915, estimated regression coefficients on that construction series are relatively small and yield relatively low measures of volatility before 1909 when applied to the overly smooth Shaw-based Kuznets series. Below we provide alternative volatility measures for versions of our equations that replace the Gottlieb-Commerce construction series with Shaw construction materials.

that in fact the conventional data series are inconsistent with the Hoover-Rees evidence, and we provide alternative backcast real GNP series based on the use of alternative deflators for real GNP and for commodity output.

V. Regression Estimates and the Volatility of Backcast Real GNP

Having now displayed the volatility measures of the basic data series and discussed their limitations, we now proceed to the straightforward task of backcasting real GNP before 1909 with both the components and indicators methods. Recall that both methods carry out estimation on variables expressed as percentage deviations from trend, not on the variables expressed as levels. The components method calculates pre-1909 real GNP as the sum of the actual values of the three components (commodities, transportation-communication, and construction) plus residual GNP calculated as the trend in residual GNP times the backcast deviation of residual GNP from trend, where the backcast value is calculated from the post-1908 regression equation of residual GNP on the same three components. The indicators method calculates the backcast level of deviations of total real GNP from trend entirely from estimated coefficients of a regression in which total real GNP is the dependent variable. Because residual GNP in the components method is defined as "standard" GNP minus the three known components, by definition both the components and indicators estimates are equal to standard GNP in benchmark years and differ only in years between benchmarks.

Regression Estimates with Conventional Deflators

The regression estimates that use conventional deflators are displayed in table 6. Eight equations are shown, four each for the indicators and components methods. For each method, estimates are displayed for versions of the equations with sample periods ending in 1928 and 1938 and with both a single explanatory variable (commodity output) and all three components as explanatory variables.[22] The bottom three lines of the table show for each of the eight equations the calculated standard deviations of deviations from trend of real GNP backcast for 1869–1908.

[22] Unlike Romer, we do not include the postwar years in the sample period of our regression. The question under discussion, after all, is whether the prewar and postwar periods are different. The last thing we would want to do is to answer the question in advance by assuming that the prewar and the postwar periods exhibit the same structure.

TABLE 6

REGRESSION EQUATIONS AND BACKCAST VOLATILITY STATISTICS, STANDARD DEFLATORS, ALTERNATIVE SAMPLE PERIODS AND EXPLANATORY VARIABLES

	INDICATORS METHOD				COMPONENTS METHOD			
	1909–28		1909–38		1909–28		1909–38	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Regression Estimates							
Commodity output	.545*	.289	.868**	.374**	.249	−.187	.791**	.031
	(.224)	(.190)	(.041)	(.131)	(.365)	(.463)	(.067)	(.316)
Transportation and communication	. . .	.219	. . .	.112	. . .	.255	. . .	.074
		(.132)		(.084)		(.323)		(.204)
Construction	. . .	.178**	. . .	.161**	. . .	.159*	. . .	.142**
		(.028)		(.018)		(.068)		(.044)
$\bar{R}^2$	.111	.737	.911	.976	−.363	−.096	.720	.580
Standard error of estimate	4.364	2.374	3.966	2.062	7.102	5.790	6.440	4.978
Durbin-Watson statistic	.493	1.926	.475	1.900	.501	1.869	.467	1.836
	Backcast Standard Deviation of Deviation from Trend							
1869–88	3.92	4.75	6.25	4.10	3.60	4.30	6.20	3.85
1889–1908	3.36	5.80	5.36	5.28	3.33	5.65	5.39	5.53
1869–1908	3.63	5.48	5.78	4.84	3.45	5.18	5.78	4.87

NOTE.—Standard errors are in parentheses.
* Coefficient is significant at the 5 percent level.
** Coefficient is significant at the 1 percent level.

Looking first at the regression estimates themselves, we find that the sets of "straw-man" equations containing only the single commodity output variable are severely misspecified. In comparison of columns 3 and 4, for instance, inclusion of all three explanatory variables causes the Durbin-Watson statistic to jump from 0.5 to 1.9, causes the standard error of estimate to drop by half, and reduces the unexplained variance by 75 percent. The same improvement, both in eliminating positive serial correlation of the residuals and in improving the fit, is evident in the other pairs of equations as well.[23] This finding is not dependent on including the 1930s in the sample period; in columns 1 and 2, where the sample period ends in 1928, the reduction in unexplained variance is 73 percent in moving from column 1 to column 2.[24]

The four equations labeled "components" explain residual real GNP (i.e., total real GNP minus the three components) rather than total real GNP itself. Because the three components contain more systematic cyclical variance than the residual, which consists of the trade, service, and government sectors, it is not surprising to find that the fit of the residual equations in columns 5–8 is worse than that for total GNP in columns 1–4. In fact, in the equations ending in 1928, the adjusted R^2 is actually negative. Nevertheless, the components estimates yield roughly the same cyclical volatility of real GNP before 1909 as the indicators estimates. For the sample periods ending in the 1930s, the volatility measures in columns 7 and 8 are identical to those in columns 3 and 4. For sample periods ending in the 1920s, the components method yields volatility measures about 5 percent less than those for the indicators method.

To implement our discussion in Section II of measurement error, we have taken two steps. The first, already completed in Section IV, is a detailed assessment of the quality of individual data series. There we argued that, of the three explanatory variables included here, transportation-communication output is the most accurate, and measurement error in the construction output series is not likely to be appreciably more severe than that for commodities output over most of the

[23] The apparent paradox in cols. 7 and 8 of the R^2 and standard error moving in opposite directions is explained by the fact that the dependent variables in these two equations are different. In col. 7 the dependent variable is real GNP minus commodity output, whereas in col. 8 the dependent variable is real GNP minus all three components.

[24] These results contradict Romer's claim (this issue, sec. V) that extra explanatory variables "should not alter the prewar GNP estimates significantly. The reason for this is that commodity output is an excellent predictor of GNP." Commodity output by itself yields equations plagued by serial correlation and standard errors double those that include transportation-communication and construction as additional explanatory variables.

period between 1869 and 1899. The second step is an examination to determine whether alternative data series are highly correlated with each other or provide independent information during the sample period. It is clear from table 6 that addition of the extra explanatory variables for transportation, communication, and construction output markedly improves the fit of the equations and eliminates positive serial correlation in the residuals. Almost all this improvement is contributed by construction output, and we feel that this strong evidence that construction output adds extra information and is not highly correlated with commodity output during the sample period amply justifies including this variable in spite of the likelihood of measurement error before 1915.

If we leave aside the misspecified equations ending in 1929 that include only commodity output, that is, columns 1 and 5, the other results tell a fairly consistent story about prewar volatility. The backcast standard deviations of deviations from trend of the other equations range from 4.84 to 5.78, compared to 5.03 for the standard Kuznets-Gallman-Kendrick series. The majority of our estimates imply a greater prewar/postwar volatility ratio than the standard series, that is, a range of prewar/postwar ratios (comparing 1869–1908 with 1954–72, as in table 3) of 1.65 to 1.97, as contrasted to the ratio of 1.72 implied by the standard series. Interestingly, once transportation-communication and construction are included, the inclusion of the 1930s in the sample period is not necessary to reaffirm the traditional conclusion of greater prewar volatility. Indeed, the inclusion of the 1930s leads to lower volatility estimates for 1869–1908 in column 4 than in column 2 and in column 8 than in column 6.

The Problem of Overly Volatile Deflators

As noted above, the standard GNP estimates make no use at all of information on consumer prices. To examine the differing implications of alternative deflators, we turn to the summary in table 7 of volatility measures for price deflators. The volatility measures are standard deviations of deviations from trend, just as in table 3, and are calculated using the same benchmarks. Here, because of World War I inflation, all price measures show much greater volatility during 1909–28 than during 1869–1908, and in this subsection we focus on the volatility measures for the period ending in 1908 and the comparison with the 1954–72 postwar period.

The standard measures of price volatility include on lines 1 and 2 the standard implicit deflators for GNP and consumption, for Shaw's commodity series on line 4, and for noncommodity output (i.e., standard GNP minus Shaw commodity output) on line 7. We notice that

TABLE 7

Standard Deviation of Deviations from Trend of Alternative Aggregate and Sectoral Deflators, 1869–1986, Detrended by Log-Linear Trends through Benchmarks

Series and Years Covered	1869–88 (1)	1889–1908 (2)	1908–28 (3)	1869–1908 (4)	1869–1928 (5)	1954–72 (6)	1947–86 (7)	(4)/(6) (8)	(5)/(7) (9)
	Alternative Deflators for the Aggregate Economy								
1. Standard (Kendrick, 1909–28)	4.21	3.96	14.12	4.08	9.35	3.03	4.53	1.34	2.06
2. Standard consumption	4.07	4.26	12.16	4.49	8.52	2.29	4.73	1.96	1.80
3. Linked CPI	2.53	1.93	11.48	2.27	7.06	2.86	4.63	.79	1.52
	Sectoral Deflators								
4. Shaw commodities	4.22	4.86	17.58	4.65	11.48	2.83	9.14	1.64	1.26
5. Shaw construction materials	11.73	8.06	18.62	9.97	13.95	...	...	...	...
6. Construction price index	8.97	5.33	14.33	7.28	10.33	6.47	8.18	1.13	1.26
	Implicit Deflators for Noncommodity Output								
7. Standard	4.36	3.98	13.29	4.12	8.66	2.40	4.37	1.72	1.98
8. Alternative	6.84	6.02	14.28	6.77	9.77	3.01	4.31	2.24	2.31
	Ratio, Noncommodity Output to Consumption								
9. Standard	1.07	.93	1.09	.92	1.02	1.05	.92	.88	1.11
10. Alternative	2.70	3.11	1.24	2.98	1.41	1.05	.93	2.84	1.52

Source.—The following list cross-references the series used on each line to the series documented in the Data Appendix. Line 1 [1A–C] indicates that the source of the series computed on line 1 in this table is given in the notes to series 1A–1C in the Data Appendix. 1 [1A–C]; 2 [12A–C]; 3 [13]; 4 [6A–D]; 5 [9A–B]; 6 [11A–C]; 7: implicit deflator for standard GNP series [2A–C] minus commodity output [6A–D]; 8: implicit deflator for hybrid GNP [4] minus commodity output [6A–D]; 9: line 7 divided by line 2; 10: line 8 divided by line 3.

all these volatility measures have the same rough order of magnitude in the prewar and also in the postwar period.[25] Of particular interest is the ratio on line 9 of the noncommodity to the consumption implicit deflator, which hovers around 1.0 in both the prewar and postwar periods.

However, in none of the measures discussed in the previous paragraph is any information used on prices paid by consumers, but rather only on Shaw's selection of wholesale price indexes. In contrast, the volatility of the linked Hoover-Rees CPI series in line 3 is about half that for the standard consumption deflator over 1869–1908. Because the Hoover-Rees series is compiled from a large body of annual data on final prices paid by consumers, it represents one of the most accurate data series available for the pre-1909 period. Figure 1 displays the behavior of the standard consumption deflator and CPI for the period 1869–1928. Before 1915 the CPI shows the same general down and up movements as the standard consumption deflator yet is substantially less volatile. After 1915 the CPI rises even more than the standard deflator but with different timing, and it declines less during 1920–21.

It is doubtful that both the Hoover-Rees CPI series and the Shaw-based commodity output deflators could be accurate. To show this, we compute the implicit deflator implied for the noncommodity (i.e., transportation and distribution) sector of the economy on the assumption that both indexes were true. We begin by computing a hybrid nominal GNP series that multiplies standard real consumption by the Hoover-Rees CPI and standard real construction by our construction price index (table 7, line 6), to which are added the standard estimates of nominal and real GNP in the residual (nonconsumption, nonconstruction) sector of the economy. This hybrid deflator is likely to be more accurate than the standard GNP deflator since it uses information on prices actually paid by consumers (representing about 70 percent of GNP). We then subtract nominal and real Shaw commodity output (including construction materials) to arrive at the implicit residual deflator shown on line 8 of table 7.

The remarkable behavior of this implicit residual deflator speaks for itself. For 1869–1908 its volatility is *triple* that of the CPI, in contrast to a postwar residual/CPI volatility ratio that is close to unity. Two alternative conclusions could be drawn from this evidence. First, before 1909 the transportation and distribution sectors could have exhibited value-added prices that were much more volatile than com-

[25] The Shaw commodity series (which includes construction materials on line 4) has a higher volatility than GNP, consumption, or noncommodity output mainly because of its construction materials component (shown separately on line 5).

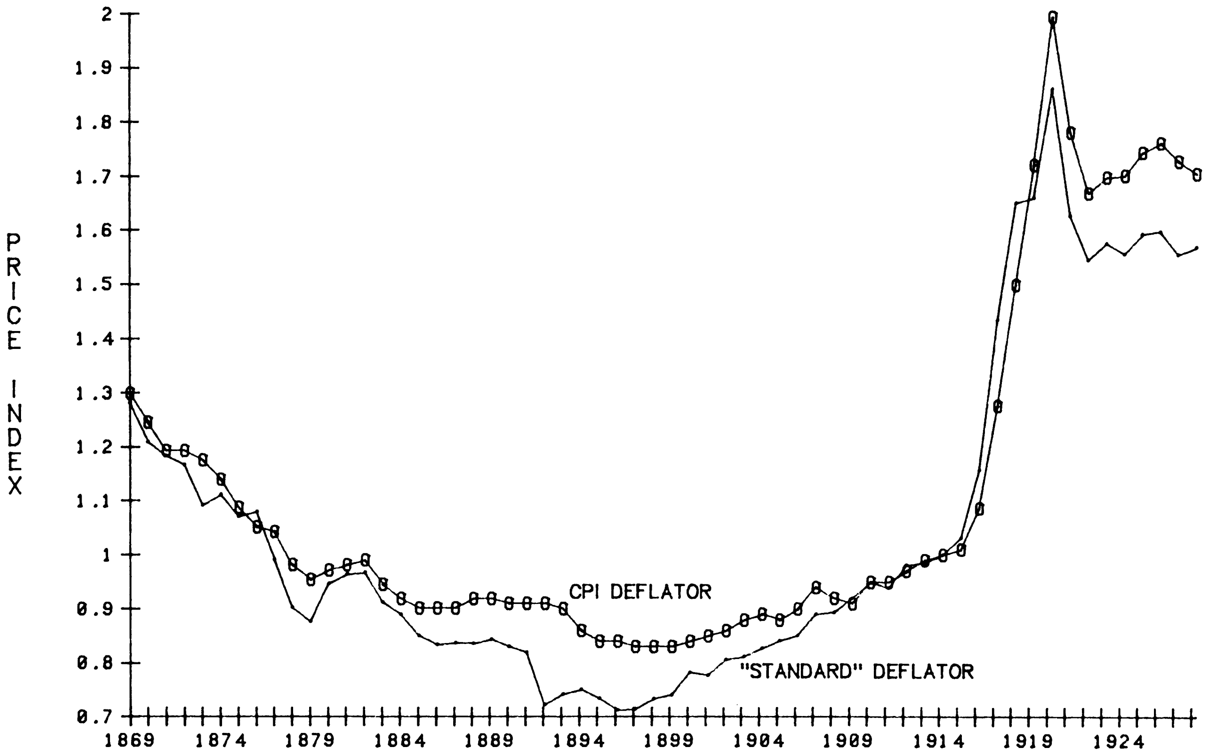

FIG. 1.—Alternative consumption deflators, 1869–1928 (1914 = 1.0)

modity prices, in contrast to the postwar period, in which both types of prices exhibit roughly the same volatility. A second and alternative view would attribute the high volatility of the implicit residual noncommodity deflator to measurement error that significantly exaggerates the volatility of commodity prices. Shaw's own misgivings about his underrepresentation of price indexes for finished goods supports this second interpretation. It seems highly implausible to us that the first explanation could be valid. Why would wholesalers and shopkeepers absorb volatile commodity price movements in order to maintain sticky consumer prices in the prewar period but not in the postwar period?

An additional piece of evidence comes from regression equations in which the annual rate of change of consumer prices is regressed on the current and one lagged change in commodity prices. The sum of coefficients indicates the degree of responsiveness of changes in consumer prices to changes in commodity prices. These sums rise from 0.39 in 1872–1908 to 0.69 in 1909–28 to 1.01 in 1949–86, which is consistent with our interpretation that measurement error in commodity prices biases down the coefficients in the prewar period.[26]

Alternative Regression Estimates for Real GNP

To test the sensitivity of regression estimates and implied backcast real GNP to the use of alternative deflators, in table 8 we provide another set of regressions for real GNP. These are identical to those of table 6 that use the full set of three explanatory variables, but here we have replaced the standard real GNP and commodity output data by the corresponding variables based on alternative deflators. Real GNP consists of nominal standard consumption deflated by the CPI instead of by the standard consumption deflator, Kuznets's nominal construction deflated by our linked construction price index, and standard nonconsumption, nonconstruction output. Commodity output consists of Shaw's nominal commodity output deflated by the CPI. Our other explanatory variables, transportation-communication output and construction output, are entered exactly as before because they are not dependent on Shaw's deflators.

Contrasting the new estimates in table 8 with the basic estimates in table 6, we first examine the coefficients. The main difference is that the coefficients on transportation-communication are larger and, in column 2, highly significant. The coefficients on commodity and construction output are of the same general order of magnitude as those in table 6. The goodness of fit of the equations as measured by R^2 is

[26] These regression equations are not shown in tabular format to save space. All sums of coefficients are significant at the 1 percent level or better.

TABLE 8

REGRESSION EQUATIONS AND BACKCAST VOLATILITY STATISTICS, ALTERNATIVE DEFLATORS, ALTERNATIVE SAMPLE PERIODS AND EXPLANATORY VARIABLES

	INDICATORS METHOD		COMPONENTS METHOD	
	1909–28 (1)	1909–38 (2)	1909–28 (3)	1909–38 (4)
	Regression Estimates			
Commodity output	.385*	.344**	−.064	−.119
	(.162)	(.074)	(.400)	(.183)
Transportation and communication	.178	.219**	.238	.363
	(.262)	(.083)	(.648)	(.206)
Construction	.178**	.149**	.106	.133**
	(.032)	(.021)	(.080)	(.052)
$\bar{R}^2$	.737	.975	−.060	.705
Standard error of estimate	3.015	2.491	7.465	6.145
Durbin-Watson statistic	1.388	1.469	1.329	1.438
	Backcast Standard Deviation of Deviation from Trend			
1869–88	4.56	4.69	4.21	4.75
1889–1908	6.09	6.11	5.95	6.34
1869–1908	5.35	5.43	5.13	5.53

NOTE.—Standard errors are in parentheses.
* Coefficient is significant at the 5 percent level.
** Coefficient is significant at the 1 percent level.

similar but standard errors of estimate are higher, reflecting the greater volatility of the dependent variable using the alternative deflators. The Durbin-Watson statistics are lower here and lie in the ambiguous range.[27]

The implied volatility of backcast GNP in the four equations of table 8 averages 5.34, as compared to 5.10 in the equivalent equations previously presented in table 6. Thus the use of the alternative deflators implies a modest increase in volatility during the backcast period, but not by enough to change substantially our overall verdict that the volatility of backcast GNP is roughly similar to that of the standard GNP series.

VI. Prewar/Postwar Volatility and Its Statistical Significance

The Range of Prewar/Postwar Volatility Ratios

Our methodological investigation compared two alternative methods of backcasting, which we called the "components" and "indicators"

[27] At the 5 percent level of significance, the upper and lower bounds of the Durbin-Watson statistic for the regression equations in table 8 are 1.21 and 1.65.

TABLE 9

SUMMARY OF EVIDENCE ON PREWAR/POSTWAR VOLATILITY RATIOS

	Volatility Ratios		Excess Prewar Volatility as a Percentage of That Indicated by Standard Series	
	1869–1908/ 1954–72	1869–1928/ 1947–86		
	(1)	(2)	(3)	(4)
1. Standard	1.72	1.67	100.0	100.0
2. Romer	1.24	1.26	33.3	38.8
3. New estimates, conventional deflators:				
a. Components to 1928	1.77	1.70	106.9	104.5
b. Indicators to 1928	1.87	1.77	120.8	114.9
c. Components to 1938	1.66	1.63	91.7	94.0
d. Indicators to 1938	1.66	1.63	91.7	94.0
4. New estimates, alternative deflators:				
a. Components to 1928	1.61	1.87	84.7	129.9
b. Indicators to 1928	1.69	1.92	95.8	137.3
c. Components to 1938	1.71	1.94	98.6	140.0
d. Indicators to 1938	1.71	1.94	98.6	140.0
5. Components method, conventional deflators, estimated through 1938:				
a. Exclude construction	2.39	2.13	193.0	168.6
b. Substitute Shaw construction materials	1.53	1.54	73.6	80.6
c. Same as 5*b*, alternative deflator	1.57	1.83	79.2	123.9

methods. Our discussion of data identified two sets of data that could be used in estimation, corresponding to "standard" and "alternative" deflators. In addition, one could have differing views regarding the relevance of Great Depression data, which calls for a comparison of results with sample periods ending in, respectively, 1928 and 1938. These different dimensions of the investigation yield a vector of implied volatility measures, and we attempt in table 9 to summarize the most important and relevant of these.

Comparisons of prewar and postwar volatility are provided in the four columns of table 9 along two dimensions. First, in columns 1 and 2, we exhibit the straightforward ratios of prewar to postwar standard deviations of deviations from trend, using as always the same set of benchmarks as Romer (this issue). Second, in columns 3 and 4, we exhibit the excess of prewar to postwar volatility, expressed as a percentage of the standard prewar/postwar volatility ratio displayed on line 1. For instance, the Romer series in column 1, line 2, has a prewar/postwar volatility ratio of 1.24, in contrast to the ratio for the standard GNP series of 1.72. This means that the excess volatility registered by the Romer measure (1.24 − 1.0, or 24 percent) is one-third of the excess volatility of the standard series (1.72 − 1.0, or 72 percent).

In subsequent lines of table 9 we report volatility measures for a variety of alternative estimated real GNP series. Lines 3*a*–3*d* summarize the results of table 6, where the alternative components and indicators backcast estimates use the standard set of deflators.[28] Here we find in columns 3 and 4 that the results cover a range of prewar/postwar excess volatility ratios between 92 and 121 percent. In this section of the table, the equations with sample periods ending in 1928 uniformly indicate a higher volatility of pre-1909 GNP than those ending in 1938. Thus the inclusion of data on the Great Depression is not a necessary condition for a finding of higher prewar volatility.

Next, on lines 4*a*–4*d* we summarize the results of table 8, in which the alternative set of CPI-based deflators is used. In column 3 the results with the alternative set of deflators are little different from those with the standard set of deflators; the range for prewar/postwar excess volatility lies between 85 and 99 percent. Here the volatility ratios with sample periods ending in 1928 are modestly smaller than those for sample periods ending in 1938, but the differences are very small. In contrast, the alternative time periods reported in column 4 indicate much higher volatility ratios when the alternative deflators are used. This occurs because the alternative deflators imply a substantially greater volatility of real GNP during and after World War I than the standard series.

Finally, we report the results of sensitivity tests that carry out the same empirical backcasting exercise with alternative sets of explanatory variables. As shown on line 5*a*, our results do not depend on the inclusion of an explanatory variable measuring construction output; when construction is excluded, the prewar variance jumps markedly. This occurs because commodities and construction output move in opposite directions in the 1875–85 period, thus reducing the volatility of backcast real GNP during this interval. Also, omission of construction increases markedly the coefficient on commodity output in the 1909–28 or 1909–38 sample period, directly implying higher pre-1909 real GNP volatility when this higher coefficient is multiplied by existing measures of commodity output. The next two lines suggest that inclusion of Shaw's series on the output of construction materials

[28] All results in lines 3 and 4 of table 9 use the regression results using the full set of three explanatory variables, in the light of the high standard errors and low Durbin-Watson statistics for the restricted regressions (e.g., in table 6, cols. 1, 3, 5, and 7) that include only commodity output.

in place of our construction output series yields somewhat smaller backcast volatility measures; however, the regression equations using the alternative Shaw measure suffer from positive serial correlation of the residuals (Durbin-Watson statistics of 0.8 or below).

Statistical Tests

At least two types of statistical tests can be applied to our results. First, because we include the Great Depression in the sample period of the equations used to develop our basic estimates of backcast real GNP, we may be concerned about the possibility of heteroscedasticity in the error terms of our regression equations. Second, we need to measure the statistical significance of differences between the prewar and postwar volatility of real GNP.

We run two tests for heteroscedasticity. The first test (Kmenta 1986, pp. 295–97) regresses squared residuals on a constant and the squared values of the explanatory variables. To test for autoregressive-conditional heteroscedasticity (ARCH), as suggested by Engle (1982), we regress the same squared residuals on a constant and lagged squared residuals. When these two tests are applied to the residuals from our indicators equation estimated to 1938 with conventional deflators (table 6, col. 4), we obtain test statistics of 0.99 and 2.15, which may be compared with the chi-squared critical values of 7.8 and 3.8, respectively, for the common and Engle-ARCH tests. This constitutes strong evidence that the hypothesis of homoscedasticity cannot be rejected.

Of more importance is the statistical significance of differences in volatility between the prewar and postwar real GNP series. We use Romer's test statistic, which relies on the Newey and West (1987) formula for calculating the standard error of the variance of deviations from trend. The test statistic has a standard normal distribution in large samples. We have calculated a substantial number of these statistics, for different prewar and postwar series. For the prewar (1869–1928) period we include the standard Kuznets-Kendrick-Gallman series, our components series based on conventional deflators, and our components series based on the alternative deflators. For the postwar (1947–86) period we include the actual data, but we also supplement the actual postwar data by estimating fitted values for the postwar period. This is necessary for symmetry since the 1869–1908 results are based on fitted values for residual GNP. We have two choices for the calculation of these fitted values. First, we can use the fitted values of a regression estimated over 1947–86 for our basic specification using the components method (table 6, col. 8). Second, we can extrapolate the prewar structure to the postwar using coefficients from the regression equation for 1909–38 (again, table 6, col. 8).

In evaluating the following test statistics, we note that any value greater than 1.96 is significant at the 5 percent level, and any value greater than 3.1 is significant at the 0.1 percent level.

	Postwar Actual	Postwar Fitted	Postwar Forecast
Prewar standard	4.42	...	...
New version, conventional deflators	4.84	3.96	3.70
New version, alternative deflators	3.11	3.32	3.02

Thus by every measure the difference between prewar and postwar GNP volatility is highly significant at better than the 1 percent level and in most cases at better than the 0.1 percent level.

The Choice of a Final Index for Real and Nominal GNP

The last stage of our investigation is the choice of one of the alternative new indexes as our central estimate of pre-1929 real and nominal GNP. We prefer the components to the indicators estimates since they use additional information on changes in the share of individual components in GNP. Yet the components estimates are no more contaminated with measurement error than the indicators estimates since each technique uses the same set of three explanatory variables. As for the sample period, we prefer the period ending in 1938 since this yields more precise parameter estimates, and yet, as shown in table 9, equations estimated through 1938 in most cases imply a lower pre-1909 volatility than equations estimated through 1928.

This leaves the choice between the two sets of estimates based on conventional deflators and the alternative deflators, that is, the Hoover-Rees CPI and linked construction cost index. Rather than choose one or the other set, we think that there are good reasons to believe that the truth lies in between these alternative sets of estimates. The conventional deflators are too volatile since they are based on an overrepresentation of crude and intermediate goods prices. Yet the alternative deflators may be too smooth when applied to the output of finished commodities. Our final index is an unweighted average of the two component indexes using the two different deflators, backcast for 1869–1908 from regression equations estimated to 1909–38. Our choice represents an example of what Granger and Newbold (1986, pp. 266–67) call "combination forecasting" using a selection of series, with weights in our case selected arbitrarily to be 0.5 on each of two series.

Our final estimates of real GNP deviations from trend for 1869–1928 are illustrated in figure 2. Here we see that our estimates do not provide a one-dimensional increase or decrease in the volatility of the standard GNP series. Instead, our estimates are less volatile than the standard series in some periods (1870–71, 1879–92, and 1913–14) and more volatile in others (1876–82, 1900–1907, and 1916–19). In this sense our estimates provide a modest revision to the economic history of the late nineteenth century. We affirm that there were prolonged depressions in the 1870s and 1890s, but we suggest that the boom of the early 1880s was less pronounced and that of the early 1900s was more pronounced than the standard data series suggests.

Our new estimates are both more and less volatile than the standard series in particular time intervals, but overall they present a similar record of substantially greater volatility before 1929 than after 1946. In this respect, we provide support for the traditional view of postwar stabilization that has been based on the standard prewar data previously developed by Kuznets, Kendrick, and Gallman, while differing with the standard data on the detailed record of particular subperiods.

Our research does not contribute any revision of the standard estimates of trend growth in real GNP between benchmark years. As shown in figure 3, when our new series for the deviation from trend is multiplied by the standard trend series, we emerge with only minor differences in the overall pattern of growth in real GNP. This aspect of our results is self-imposed to limit the scope of the study to new results on the deviation of real GNP from this trend. Data for the new real GNP series are displayed in table 10. Because we have access to improved price indexes not used in the standard GNP series, we can use our data to develop an improved estimate of the GNP deflator before 1929. This is a weighted average of the CPI, our linked construction cost estimate, and the standard implicit deflator for nonconsumption, nonconstruction output.[29] Table 10 exhibits this deflator and the associated series for nominal GNP. Since most of the weight in the new deflator is on consumption spending, the difference between our new deflator and the standard deflator is a dampened version of the difference between the CPI and standard consumption deflator illustrated in figure 1. The new GNP deflator is less cyclically volatile, it declines less from 1869 to 1896, and it increases less from 1896 to 1914 than the standard deflator.

[29] Since we have not developed new estimates of consumption spending, we use the standard Kendrick-Kuznets series on real consumption, construction spending, and the residual to weight these three price indexes in the calculation of nominal GNP. The implicit deflator is then the ratio of the resulting nominal GNP series to standard real GNP. Our new nominal GNP series is then the product of this implicit deflator and our new real GNP series.

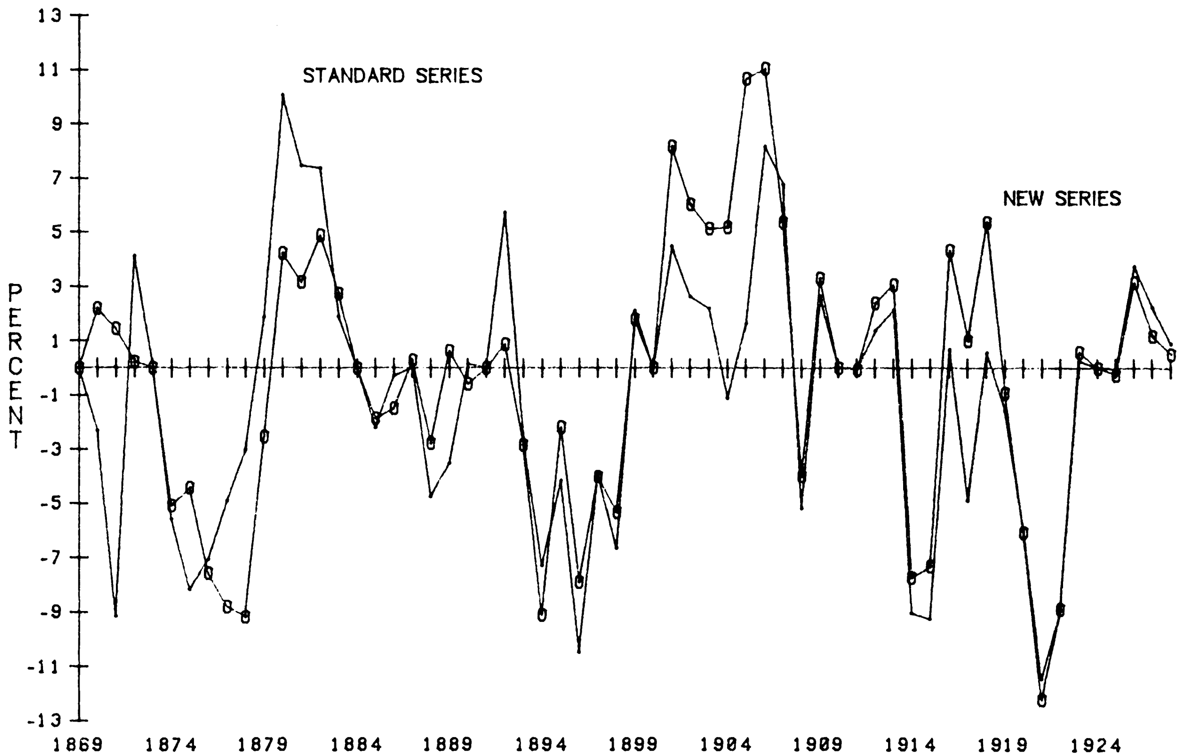

FIG. 2.—Deviations from trend, 1869–1928; standard and new components estimates

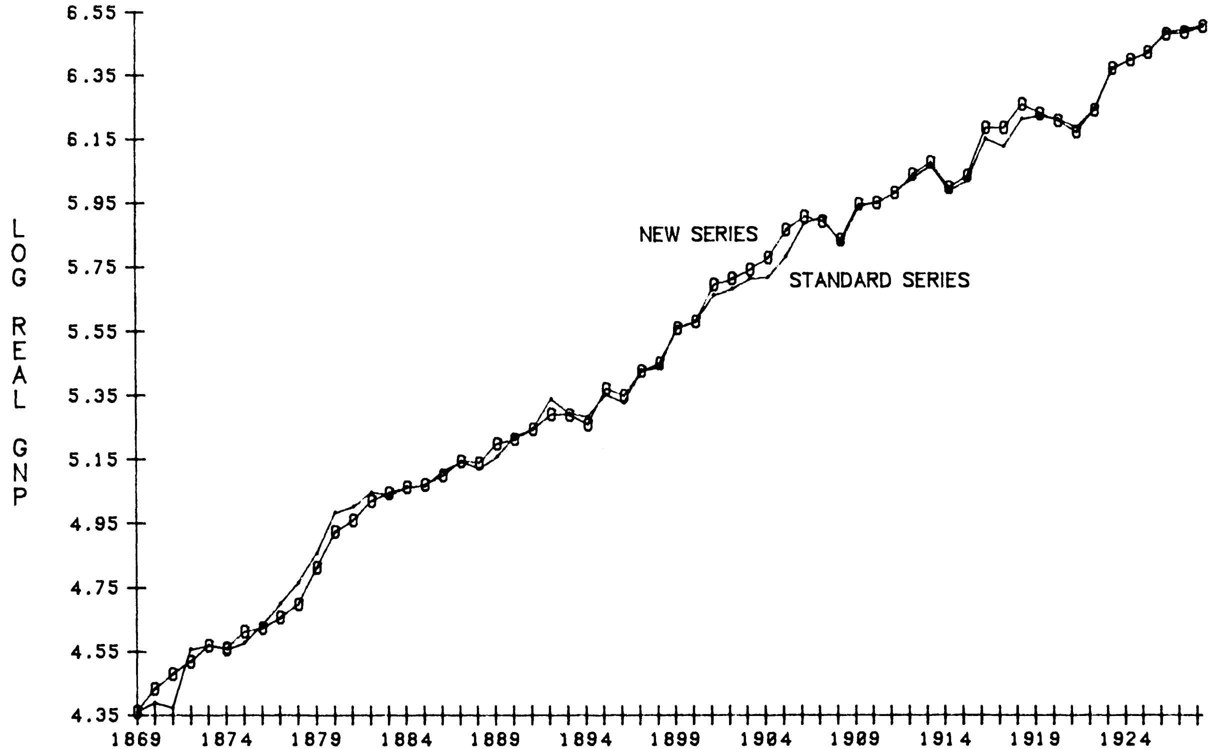

FIG. 3.—Standard and new estimates of real GNP (constant 1982 dollars)

TABLE 10

New Estimates of Nominal GNP, Real GNP in Constant 1982 Dollars, and Implicit GNP Deflator (1982 = 100), 1869–1928

Year	Nominal GNP	Real GNP (1982 Dollars)	Implicit GNP Deflator (1982 = 100)
1869	8.21	78.2	10.49
1870	8.41	84.2	9.98
1871	8.69	88.1	9.86
1872	8.81	91.7	9.60
1873	9.17	96.3	9.51
1874	8.85	95.7	9.25
1875	8.92	100.7	8.85
1876	8.68	101.9	8.51
1877	8.82	105.2	8.38
1878	8.63	109.6	7.87
1879	9.41	123.1	7.64
1880	11.06	137.6	8.03
1881	11.39	142.5	7.99
1882	12.37	151.6	8.16
1883	12.24	155.3	7.88
1884	11.92	158.1	7.53
1885	11.71	159.3	7.35
1886	12.06	164.1	7.35
1887	12.61	171.5	7.35
1888	12.75	170.7	7.47
1889	13.57	181.3	7.48
1890	13.44	183.9	7.30
1891	13.86	189.9	7.30
1892	14.33	198.8	7.21
1893	14.37	198.7	7.23
1894	13.21	192.9	6.85
1895	14.53	215.5	6.74
1896	14.25	210.6	6.76
1897	15.18	227.8	6.66
1898	15.74	233.2	6.75
1899	17.85	260.3	6.86
1900	18.58	265.4	7.00
1901	20.97	297.9	7.04
1902	21.65	303.0	7.14
1903	22.85	311.7	7.33
1904	23.93	323.5	7.39
1905	26.12	353.2	7.40
1906	28.10	367.7	7.64
1907	28.88	362.0	7.98
1908	26.72	342.2	7.81
1909	29.88	382.1	7.82
1910	31.24	383.8	8.14
1911	32.16	396.0	8.12
1912	34.85	418.9	8.32
1913	36.56	435.4	8.40
1914	34.25	402.4	8.51
1915	36.36	417.3	8.71
1916	46.02	485.0	9.49
1917	55.10	484.9	11.36
1918	69.70	522.2	13.35

TABLE 10 (*Continued*)

Year	Nominal GNP	Real GNP (1982 Dollars)	Implicit GNP Deflator (1982 = 100)
1919	77.22	507.1	15.23
1920	87.24	496.3	17.58
1921	73.27	478.8	15.30
1922	72.99	513.2	14.22
1923	85.62	585.0	14.63
1924	87.91	600.5	14.64
1925	91.49	614.1	14.90
1926	97.52	651.0	14.98
1927	96.34	654.6	14.72
1928	97.32	666.7	14.60
1929	103.90	709.6	14.64

SOURCE.—Real GNP, taken as antilog of the log of the trend value of standard GNP plus the log deviation from trend, where the deviation from trend is the average of the series listed in table 9, lines 3*c* and 4*c*. This is converted from 1929 to 1982 prices by multiplying by the ratio of NIPA 1929 real to nominal GNP. The GNP deflator is the hybrid deflator, series [4] in the Data Appendix. Nominal GNP is the product of real GNP in this table times the GNP deflator, divided by 100.

VII. Conclusion

We have developed new estimates of real GNP, nominal GNP, and the GNP deflator for the period 1869–1928. Our estimates go beyond the previous work of Kuznets, Kendrick, Gallman, and Romer by using three sets of information that have not previously been applied to the estimation of aggregate GNP prior to 1919. These include the linked Hoover-Rees–Bureau of Labor Statistics CPI, the linked Frickey-Kendrick indexes of transportation and communication output, and the Gottlieb index of construction output. None of the information in these three sets has been created for this study; instead, each set was created 20 or more years ago, and our contribution has been to extract from these sets of existing data new implications for the behavior of GNP, both real and nominal.

In applying this new information to the estimation of prewar GNP, we have made four major contributions. First, we supplement the previous discussion of the regression-based indicators approach by analyzing the components approach, which uses statistical regression to backcast the residual component of GNP for which no primary data exist, while using actual values of components for which data do exist. We show that the components approach makes better use of available data when there are secular changes in the GNP shares of the primary components. Although we judge the components method to be superior in theory, in practice the two methods yield almost identical conclusions regarding prewar GNP volatility in our sample period.

Second, the use of additional data on sectoral output substantially improves the fit of regression equations for 1909–28 or 1909–38 and eliminates the problem of serial correlation of residuals that is evident when a single commodity output variable is used. There is no systematic tendency of backcast output to be more volatile when additional explanatory variables are included in the regression equations, a surprising result that reflects the negative correlation between commodity and construction output during part of the backcast period. Instead, the use of multiple explanatory variables alters the pattern of estimated business cycles before 1909, dampening output fluctuations during 1869–76 and 1880–92 while amplifying fluctuations during 1877–79 and 1901–6.

Third, the use of the Hoover-Rees CPI series leads us to argue that the implicit deflator for the standard Shaw commodity output series is implausibly volatile, and we develop an alternative backcast real GNP series in which the consumption component of real GNP and the commodity output series are redeflated by the CPI. To avoid an excessive dependence of our results on this step, our final real GNP series is based on an average of those developed with the conventional and alternative deflators.

Fourth, our use of the CPI series allows the development of a new estimate of the GNP deflator, in contrast to the previous Kuznets deflator, which is based entirely on components of the wholesale price index and, going back in time, becomes more and more dependent on the prices of crude and intermediate goods. Our new GNP deflator is less cyclically volatile than the standard deflator and, in addition, dampens the previously estimated extent of the late nineteenth-century deflation and subsequent recovery of prices between 1896 and 1914.

Directions for Future Research

This paper has reaffirmed the standard conclusion that real GNP was more volatile before 1929 than since 1946. But it goes beyond reaffirmation to provide a new history of business cycles between 1869 and 1908, dampening some cycles while amplifying others. And it suggests strongly that the GNP deflator was substantially less volatile before 1914 than the standard version. Further research is needed on both substantive and measurement issues.

Perhaps the most important implication of our research is that the debate between Taylor and DeLong-Summers, cited in the Introduction, is based on a false premise. They disagree whether greater price stability in the postwar economy has aggravated or dampened business cycles. We show that *there has been no reduction in price flexibility in postwar business cycles* and that past estimates of the GNP deflator have been based on a fragmentary and overly volatile data base that ignores the greatly superior and long-available direct evidence on consumer prices.[30] Clearly a top priority in future research will be a reexamination of the short-run cyclical division of nominal GNP changes between price and quantity responses, as well as the persistence of price changes, since all the recent studies of this topic have used the standard deflators.

At the level of measurement, our study leaves several important questions open for further research. First, our technique forces real GNP to be equal to the standard series in benchmark years and thus contributes nothing to the estimation of secular economic growth between benchmarks. Yet our use of the new information sets on sectoral output and on consumer prices would allow us to go one step further and examine implications for trend output. Second, we have questioned the validity of the standard GNP series during and after World War I and their implausible implications for the behavior of labor productivity during the 1920–21 recession. Our use of alternative deflators shows that real GNP estimates during this period of highly volatile prices are sensitive to the choice of price indexes. Further research on the period between 1916 and 1923 is likely to yield new insights into this fascinating episode; in particular, how could manufacturing employment collapse in 1920–21 even faster than in 1929–31 without tumbling the entire economy into a Great Depression?

Data Appendix

Series with different sources during the sample period are ratio-linked together. All real series have been recalculated in 1929 dollars, and all deflators are equal to 1.0 in the base year 1929. NIPA refers to *National Income and Product Accounts* (1986; supplemented with diskettes and tapes); HS refers to *Historical Statistics of the United States* (1975); BS refers to *Business Statistics: 1982*.

1A. *GNP:* Commerce Department, 1929–86
Nominal: NIPA, table 1.1
Real: NIPA, table 1.2
Deflator: NIPA, table 7.4

1B. *GNP:* Commerce Department, 1909–28
Nominal: NIPA, table 1.25

[30] We recognize the distinction between the short-run responsiveness of the aggregate price level to nominal demand changes and the "persistence" issue, i.e., the dependence of price changes on their own lagged values. The volatility data on price indexes presented in table 7 bear directly only on the short-run responsiveness issue. Whether our new deflator exhibits as much persistence as postwar indexes is a question we defer for future research.

Real: NIPA, table 1.25
Deflator: Nominal GNP divided by real GNP

1C. *GNP:* Gallman/Kendrick adjustments of Kuznets, 1869–1908
Nominal: Net national product, Friedman and Schwartz (1982), pp. 122–29, table 4.8, col. 2. Plus capital consumption, unraveled 5-year moving average, Kuznets (1961), pp. 499–501, table R-8, col. 3; pp. 572–75, table R-29, col. 2. The series was also adjusted by adding the Kendrick adjustments of Kuznets. For 1889–1908, this adjustment is nominal GNP from Kendrick (1961), pp. 296–97, table A-IIb, minus nominal GNP from Kuznets (1961), pp. 557–58, table R-23, col. 1. For 1869–88, the adjustment is the linear interpolation of the difference between Kendrick decadal estimates (pp. 296–97, table A-IIb) and Kuznets (1961) decadal estimates (p. 520, table R-11, col. 9).
Real: Net national product, Friedman and Schwartz (1982), pp. 122–29, table 4.8, col. 3. Plus capital consumption, unraveled 5-year moving average, Kuznets (1961), pp. 499–501, table R-8, col. 6; pp. 572–75, table R-29, col. 5. The Kendrick adjustments of Kuznets for 1889–1908: Real GNP from Kendrick (1961), pp. 294–95, table A-IIa, minus real GNP from Kuznets (1961), pp. 555–56, table R-22, col. 1. For 1869–88, the adjustment is the linear interpolation of the difference between Kendrick decadal estimates (pp. 294–95, table A-IIa) and Kuznets (1961) decadal estimates (p. 521, table R-12, col. 9).
Deflator: Nominal divided by real GNP

2A. *GNP:* Standard Kendrick, 1929–86
Nominal: Same as series 1A
Real: Same as series 1A
Deflator: Same as series 1A

2B. *GNP:* Standard Kendrick, 1909–28
Nominal: Kendrick (1961), pp. 296–97, table AII-b
Real: Kendrick (1961), pp. 294–95, table AII-a
Deflator: Nominal divided by real GNP

2C. *GNP:* Standard Kendrick, 1869–1908
Nominal: Same as series 1C
Real: Same as series 1C
Deflator: Same as series 1C

3. *GNP:* Kuznets components (variant 3), 1869–1908
Nominal: Unraveled 5-year moving average, Kuznets (1961), pp. 561–62, table R-25, col. 6; pp. 557–58, table R-23, col. 1
Real: Unraveled 5-year moving average, Kuznets (1961), pp. 563–64, table R-26, col. 6; pp. 555–56, table R-22, col. 1
Deflator: Nominal divided by real GNP

4. *GNP:* With hybrid deflator, 1869–1986
Nominal: Same as series 2
Real: Series 2 minus real consumption (ser. 12) minus real construction (ser. 10) plus nominal consumption (ser. 12) divided by the CPI-based deflator (ser. 13) plus nominal construction (ser. 10) divided by the construction deflator from series 11
Deflator: Nominal divided by real GNP

5. *GNP:* Romer
Real: 1929–86, same as series 1A; 1869–1928, Romer (this issue), table 2

6A. Commodity output, 1947–86:
Nominal: GNP originating in agriculture, mining, and manufacturing, NIPA, table 6.1
Real: GNP originating in agriculture, mining, and manufacturing, NIPA, table 6.2
Deflator: Nominal divided by real output

6B. Commodity output, 1939–46:
Nominal: GNP goods, NIPA, table 1.3
Real: GNP goods, NIPA, table 1.4
Deflator: Nominal divided by real output

6C. Commodity output, 1889–1938:
Nominal: Value of finished commodities and construction materials, current dollars, Shaw (1947), pp. 62–65, table I.1; p. 69, table I.2. Multiplied by the ratio of real commodity output in 1929 (source below) to nominal commodity output in 1929 from Shaw.
Real: Finished commodity output (based on Shaw), Kuznets (1961), pp. 553–54, table R-21, col. 6
Deflator: Nominal divided by real output

6D. Commodity output, 1869–88:
Nominal: Real commodity output times Kuznets GNP deflator (ser. 3 above)
Real: Finished commodity output, Kuznets (1961), pp. 553–54, table R-21, col. 6
Deflator: Nominal divided by real output

7. Manufacturing output:
1947–86: Federal Reserve Board index of industrial production, manufacturing. 1947–60: HS, p. 667, ser. P13; 1961–86: BS, p. 3, and various issues of *Survey of Current Business*
1899–1946: Index of manufacturing, Fabricant (1940), p. 44, table 1
1869–98: Index of production for manufacture, Frickey (1947), p. 54, table 6 (this series is scaled so that in 1929 it equals the dollar value of national income from the manufacturing sector in 1929 times the ratio of nominal GNP to national income in 1929 [sources: NIPA, tables 6.1, 6.3])

8. Output from transportation and communications:
1947–86: Real GNP from transportation and communications, NIPA, table 6.2
1889–1946: A weighted average of the Kendrick (1961) transportation output index and the telephone and telegraph output index. The transportation output index is pp. 541–42, table G-II, col. 1. For 1929–46, the telephone and telegraph index is pp. 583–84, table H-III, col. 1. For 1889–1928, the telephone index is pp. 585–87, table H-IV, col. 1. The weights are the share of national income from the transportation sector and the communications sector in 1929 (NIPA, table 6.3). This average is then scaled so that in 1929 it equals the value of national income from the transportation and construction sectors multiplied by the ratio of nominal GNP to national income.
1869–88: Index of transportation and communications, Frickey (1947), p. 117, table 17

9A. Construction materials (Shaw), 1889–1938:

Nominal: Construction materials, current dollars, Shaw (1947), pp. 64–65, table I-1; p. 69, table I-2
Real: Construction materials, 1913 dollars, Shaw (1947), pp. 76–77, table I-3
Deflator: Nominal divided by real construction materials

9B. Construction materials, 1869–88:
Nominal: Nominal Kuznets construction (ser. 10B below)
Real: Real Kuznets construction (ser. 10B below)
Deflator: Nominal divided by real construction materials

10A. Kuznets construction, 1929–86:
Nominal: Structures GNP, NIPA, table 1.3
Real: Structures GNP, NIPA, table 1.4
Deflator: Nominal divided by real construction

10B. Kuznets construction, 1869–1928:
Nominal: Kuznets (1961), pp. 490–91, table R-4, col. 1, and unraveled 5-year moving average, pp. 576–87, table R-30, col. 5
Real: Kuznets (1961), pp. 492–93, table R-5, col. 1, and unraveled 5-year moving average, pp. 576–87, table R-30, col. 10
Deflator: Nominal divided by real construction

11A. Gottlieb/BLS/Commerce construction, 1929–86:
Nominal: NIPA, table 1.3
Real: NIPA, table 1.4
Deflator: Nominal divided by real construction

11B. Gottlieb/BLS/Commerce construction, 1915–28:
Nominal: Value of new construction, HS, p. 618, ser. N1
Real: Total new construction, Lipsey and Preston (1966), p. 30, ser. C22
Deflator: Nominal divided by real construction

11C. Gottlieb/BLS/Commerce construction, 1869–1914:
Nominal: Value of building, Gottlieb (1965), p. 417, table 3
Real: Nominal construction divided by the construction deflator
Deflator: Construction cost index, 1869–88, HS, p. 629, ser. N138; 1889–1914, HS, p. 629, ser. N139

12A. Consumption, 1929–86:
Nominal: NIPA, table 1.1
Real: NIPA, table 1.2
Deflator: Nominal divided by real consumption

12B. Consumption, 1889–1928:
Nominal: Kendrick (1961), pp. 296–97, table A-IIb, col. 4
Real: Kendrick (1961), pp. 294–95, table A-IIa, col. 4
Deflator: Nominal divided by real consumption

12C. Consumption, 1869–88:
Nominal: Kuznets (1961), unraveled 5-year moving average, pp. 561–62, table R25, col. 4
Real: Kuznets (1961), unraveled 5-year moving average, pp. 563–64, table R26, col. 4
Deflator: Nominal divided by real consumption

13. CPI-based deflator:
1959–86: Fixed-weight consumption deflator, NIPA, table 7.1
1915–58: Consumer price index, BLS, HS, pp. 210–11, ser. E135
1890–1914: Cost-of-living index, Rees (1961), p. 74, table 22
1869–89: Consumer price index, Hoover (1960), pp. 142–43, table 1

References

Abramovitz, Moses. *Evidences of Long Swings in Aggregate Construction since the Civil War.* New York: NBER, 1964.

Burns, Arthur F. "Progress towards Economic Stability." *A.E.R.* 50 (March 1960): 1–19.

DeLong, J. Bradford, and Summers, Lawrence H. "The Changing Cyclical Variability of Economic Activity in the United States." *The American Business Cycle: Continuity and Change,* edited by Robert J. Gordon. Chicago: Univ. Chicago Press (for NBER), 1986.

Engle, Robert F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50 (July 1982): 987–1007.

Fabricant, Solomon. *The Output of Manufacturing Industries, 1899–1937.* General Series, no. 39. New York: NBER, 1940.

———. *Employment in Manufacturing, 1899–1939: An Analysis of Its Relation to the Volume of Production.* General Series, no. 41. New York: NBER, 1942.

Frickey, Edwin. *Production in the United States, 1860–1914.* Cambridge, Mass.: Harvard Univ. Press, 1947.

Friedman, Milton, and Schwartz, Anna J. *A Monetary History of the United States, 1867–1960.* Princeton, N.J.: Princeton Univ. Press (for NBER), 1963.

———. *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867–1975.* Chicago: Univ. Chicago Press (for NBER), 1982.

Gallman, Robert E. "Gross National Product in the United States, 1834–1909." In *Output, Employment, and Productivity in the United States after 1800.* Studies in Income and Wealth, vol. 30. Conference on Research in Income and Wealth. New York: Columbia Univ. Press (for NBER), 1966.

Gordon, Robert J. "A Consistent Characterization of a Near-Century of Price Behavior." *A.E.R. Papers and Proc.* 70 (May 1980): 243–49.

———. "Price Inertia and Policy Ineffectiveness in the United States, 1890–1980." *J.P.E.* 90 (December 1982): 1087–1117.

———, ed. *The American Business Cycle: Continuity and Change.* Chicago: Univ. Chicago Press (for NBER), 1986.

Gottlieb, Manuel. "New Measures of Value of Nonfarm Building for the United States, Annually 1850–1939." *Rev. Econ. Statis.* 47 (November 1965): 412–19.

———. "Building in Ohio between 1837 and 1914." In *Output, Employment, and Productivity in the United States after 1800.* Studies in Income and Wealth, vol. 30. Conference on Research in Income and Wealth. New York: Columbia Univ. Press (for NBER), 1966.

Granger, Clive W. J., and Newbold, Paul. *Forecasting Economic Time Series.* 2d ed. Orlando, Fla.: Academic Press, 1986.

Hoover, Ethel D. "Retail Prices after 1850." In *Trends in the American Economy in the Nineteenth Century.* Studies in Income and Wealth, vol. 24. Conference on Research in Income and Wealth. Princeton, N.J.: Princeton Univ. Press (for NBER), 1960.

Kendrick, John W. "Comment." In *Trends in the American Economy in the Nineteenth Century.* Studies in Income and Wealth, vol. 24. Conference on Research in Income and Wealth. Princeton, N.J.: Princeton Univ. Press (for NBER), 1960.

———. *Productivity Trends in the United States.* General Series, no. 71. Princeton, N.J.: Princeton Univ. Press (for NBER), 1961.

Kmenta, Jan. *Elements of Econometrics.* 2d ed. New York: Macmillan, 1986.

Kuznets, Simon S. *Commodity Flow and Capital Formation.* New York: NBER, 1938.

———. *National Income and Its Composition, 1919–1938.* New York: NBER, 1941.

———. *National Product since 1869.* General Series, no. 46. New York: NBER, 1946.

———. *Capital in the American Economy: Its Formation and Financing.* Princeton, N.J.: Princeton Univ. Press (for NBER), 1961.

Lipsey, Robert E., and Preston, Doris. *Source Book of Statistics Relating to Construction.* New York: Columbia Univ. Press (for NBER), 1966.

Newey, Whitney K., and West, Kenneth D. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (May 1987): 703–8.

Rees, Albert. *Real Wages in Manufacturing, 1890–1914.* General Series, no. 70. Princeton, N.J.: Princeton Univ. Press (for NBER), 1961.

Romer, Christina D. "Is the Stabilization of the Postwar Economy a Figment of the Data?" *A.E.R.* 76 (June 1986): 314–34. (*a*)

———. "Spurious Volatility in Historical Unemployment Data." *J.P.E.* 94 (February 1986): 1–37. (*b*)

———. "World War I and the Postwar Depression: A Reinterpretation Based on Alternative Estimates of GNP." *J. Monetary Econ.* 22 (July 1988): 91–115.

———. "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869–1908." *J.P.E.*, this issue.

Schultze, Charles L. "Some Macro Foundations for Micro Theory." *Brookings Papers Econ. Activity,* no. 2 (1981), pp. 521–76.

———. *Other Times, Other Places: Macroeconomic Lessons from U.S. and European History.* Washington: Brookings Inst., 1986.

Shaw, William H. *The Value of Commodity Output since 1869.* New York: NBER, 1947.

Taylor, John B. "Improvements in Macroeconomic Stability: The Role of Wages and Prices." In *The American Business Cycle: Continuity and Change,* edited by Robert J. Gordon. Chicago: Univ. Chicago Press (for NBER), 1986.

Weir, David R. "Unemployment Volatility, 1890–1984: A Sensitivity Analysis." Manuscript. New Haven, Conn.: Yale Univ., 1986.