NBER WORKING PAPER SERIES

THE PREWAR BUSINESS CYCLE RECONSIDERED: NEW ESTIMATES OF GROSS NATIONAL PRODUCT, 1869-1918

Christina Romer

Working Paper No. 1969

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1986

Working Paper #1969
July 1986

The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1918

ABSTRACT

This paper shows that the existing estimates of prewar gross national product exaggerate the size of cyclical fluctuations. The source of the exaggeration is that the original Kuznets estimates are based on the assumption that GNP moves one-for-one with commodity output valued at producer prices. New estimates of GNP for 1869-1918 are derived using the estimated aggregate relationship between GNP and commodity output for the interwar and postwar eras. The new estimates of GNP indicate that the business cycle is only slightly more severe in the pre-World War I era than in the post-World War II era.

Christina D. Romer
Woodrow Wilson School of
Public and International Affairs
Princeton University
Princeton, NJ 08544

INTRODUCTION

The existing estimates of gross national product for the seventy years before World War II have perhaps done more to shape economists' perceptions of prewar business cycles than any other macroeconomic series. The historical GNP data have been analyzed in great detail and are constantly cited in research on prewar fluctuations in economic activity. Hence, much of what economists believe about prewar cyclical fluctuations is derived directly from the cyclical behavior of prewar GNP. As a result, the accuracy of the prewar estimates of GNP is one of the main determinants of the accuracy of our views are about the prewar cycle.

The importance of the prewar estimates of GNP has, if anything, increased in recent years. Economists interested in a variety of cyclical relationships have recently included the historical GNP series in their analyses. Much work has been done, for example, on the changing severity of cycles over time. Studies by Baily (1978) and DeLong and Summers (1984) use prewar GNP movements to argue that cycles have become less severe in the postwar era. Much current research also concerns the changing cyclical relationship of prices and output (see, for example, Schultze, 1981, and Gordon, 1982) and money and output (see, for example, Friedman and Schwartz, 1982). In all these studies the existing historical estimates of GNP play a key role in the analysis.

While the short-term cyclical behavior of the prewar GNP series has been analyzed in detail, the accuracy of the existing prewar estimates of GNP for

such cyclical analyses has rarely been discussed. The standard estimates of GNP before 1929 are still those derived by Simon Kuznets. Several economists have amended the Kuznets series, most notably Robert Gallman and John Kendrick, but these amendments have not changed the cyclical behavior of the Kuznets series. Most of the alterations of Kuznets's series concentrate on correcting the long-term trend of the prewar GNP series or making the Kuznets series conceptually similar to the modern Department of Commerce GNP series. Few economists have questioned whether the Kuznets estimates reflect cyclical movements accurately.

There is reason to believe, however, that the representation of cyclical movements in the Kuznets estimates should be critically examined. Kuznets clearly never intended for his estimates of GNP to be used for the analysis of short-term cyclical movements. While he is very careful about determining long-run movements of various series, but he is less careful about cyclical movements. Annual movements in various components of GNP are often derived by simply interpolating by means of another series, without taking into account the actual cyclical relationship between the series being created and the interpolating series. As a result, it is very likely that the existing GNP estimates are not adequate for the detailed cyclical analysis to which they are being put today.

Kuznets himself believed that his annual series did not represent cyclical movements accurately. He states in Capital in the American Economy that the version of his prewar estimates of GNP that are most frequently used today "would not be acceptable measures of the amplitude of short-term changes" (Kuznets, 1961, p. 546). He repeatedly warns readers not to use his data for short-term cyclical comparisons and urges them to use the data in five-year moving average form.

This paper examines Kuznets's pre-1918 GNP series and finds that Kuznets's caveats concerning the accuracy of the annual estimates are justified. There is evidence that Kuznets's methods accentuate the size of cyclical movements in the prewar GNP series. The basic problem with the Kuznets estimates is that they are derived almost entirely from data on commodity output at producer prices. Total output at consumer prices, including the value added in transportation and distribution, is essentially assumed to move proportionately with commodity output at producer prices. This, however, may not be true. A variety of interwar and postwar evidence shows that gross national product moves less over the cycle than does commodity output. If the same is true in the prewar era, then Kuznets's series will exaggerate cyclical fluctuations.

Because the existing estimates of GNP may not be accurate for the short-term cyclical analysis to which most economists wish to put them, I derive what I believe is a more accurate prewar GNP series for 1869-1918. I create a new series by explicitly modeling the fact that GNP does not move as much over the cycle as does commodity output. Data from the interwar and postwar eras are used to estimate the relationship between GNP and commodity output. In estimating this relationship, the sensitivity of cyclical movements in GNP to cyclical movements in commodity output is allowed to change over time. I then use the resulting time-varying sensitivity estimate to convert pre-1918 data on commodity output into estimates of GNP for this period. In this derivation I do not alter the Kuznets/Gallman estimates of long-run trend GNP. Rather, I change only the representation of short-term cyclical fluctuations.

The cyclical properties of the new estimates of GNP that I derive are very different from those of the standard GNP series. The pre-1918 business cycle appears to be much less severe in the new data than in the existing Kuznets series. As a result, the decline in the severity of the business cycle between

the pre-1918 and the postwar eras is also much smaller than is typically believed. In fact, the average amplitude of the cycle for 1869-1918 using the new data is only slightly larger than the average amplitude for 1947-1983.

This reworking of the historical estimates of GNP is organized as follows. Section I describes the Kuznets estimates of GNP. Section II discusses the possible problems with these data and provides empirical evidence that the Kuznets estimates are excessively volatile. Section III suggests a new method for deriving annual estimates of prewar GNP and presents the new series. Section IV examines the basic volatility characteristics of the new GNP series and analyzes the degree of stabilization between the pre-World War I and post-World War II eras.

SECTION I

KUZNETS'S ESTIMATES OF PREWAR GROSS NATIONAL PRODUCT

The first step in forming new estimates of prewar gross national product is to analyze the existing historical estimates. It is important to discover whether there is reason to believe that the existing GNP data do not measure short-term cyclical fluctuations correctly. In this section I describe the methods that Kuznets uses to estimate real prewar GNP. I also discuss briefly the revisions of Kuznets's estimates suggested by Gallman and Kendrick.

Kuznets provides estimates of gross national product and its components for 1869-1938. For the most part his estimates after 1928 have been superseded by independent estimates by the Department of Commerce. For the period before 1928 Kuznets derives several versions of GNP estimates and uses different methods for various subperiods.

For the period after 1919, the Kuznets estimates are derived using the income-payments approach. These estimates are described and presented in Kuznets's book National Income and its Components 1919-1938 (1941). While the accuracy of the GNP data after 1919 has not been thoroughly established, there is reason to believe that this series is relatively accurate. For the period after 1919 Kuznets has ample income data for most sectors and uses methods very similar to those used by the Commerce Department after 1929. Therefore, it is

likely that the Kuznets data for 1919-1928 are not substantially worse than the Commerce Department data after 1929.

For the period 1869-1918 Kuznets does not have the income data necessary to derive income-side estimates of GNP. As a result, he must use different methods to construct data for this period. Kuznets actually forms two annual series for 1869-1918 which he identifies as the components series and the regression series. Both these series are derived by what can best be described as the product-side approach. Both use data on commodity output to estimate GNP. The two series differ in how they convert this base data into estimates of GNP and at how fine a level of disaggregation the conversion is made.

The pre-1918 series are described in two sources. National Product since 1869 (1946) describes the derivation of decadal averages of GNP. These decadal averages form the long-term trend for both the components series and the regression series. The annual components and regression series are described in Capital in the American Economy (1961). Only some of the actual data are published. The components series for 1869-1918 is only published in five-year moving average form. The annual regression series is published in Capital in the American Economy only for the period 1889-1918.

Because the GNP estimates before 1918 are based on much less data than are the later estimates and are derived using methods that are quite different from those used after 1919, it is reasonable to be more suspicious of the earlier estimates than of the later series. For this reason, this study focuses on the quality of the GNP data before 1918. It analyzes both the components series and the regression series for the period 1869-1918 to see if Kuznets's methods may have yielded GNP estimates that misrepresent cyclical movements.

1. Components Series

The Kuznets components series for 1869-1918 is by far the more important of the two series to analyze. It is the one that is most commonly used in modern econometric analyses. This fact may seem odd given that it has only been published by Kuznets in moving-average form. The reason for its widespread use is that it forms the basis of the widely used Kendrick estimates of GNP (see Kendrick, 1961). Kendrick adjusts the Kuznets components series for 1889-1918 to be conceptually similar to the post-1929 Department of Commerce series. While Kendrick's adjustments are quite important for classifying aggregate GNP into its various components and for determining the level of output in wartime, his correction factors do not change the cyclical properties of Kuznets's components series. This fact can be seen by comparing the cyclical movements in the Kendrick GNP series to those in the Kuznets components series. This comparison shows that cyclical movements are very similar in the two series.[1]

The same is true of other modifications of Kuznets's series. Friedman and Schwartz present estimates of real net national product back to 1869 in their book Monetary Trends in the United States and the United Kingdom (1982). This series is a combination of the Kuznets components series and revisions to some individual components of GNP prepared by Robert Gallman (see Gallman, 1966). While the trend of this series is quite different from that of the Kuznets components series, the cyclical movements are nearly identical.[2] Gallman's revisions are primarily revisions of decadal averages, not of annual movements. Again, therefore, the accuracy of the representation of cyclical movements in the Kuznets components series is the key determinant of the accuracy of the cyclical properties of this important revision of Kuznets's data.

The base data underlying the Kuznets components series come from a study by William H. Shaw entitled Value of Commodity Output since 1869 (1947). Shaw uses

data from the Census of Manufactures to derive comprehensive benchmark estimates of commodity output for various census years starting in 1869. He then forms annual estimates of commodity output for 1889-1938 by interpolating between benchmark observations by various annual series. The annual data come from a plethora of state reports and industry publications. Shaw presents commodity output data for a variety of major and minor subgroups valued in current and in 1913 prices.

Kuznets extends the annual Shaw series on commodity output to cover the earlier period 1869-1888. He uses annual series similar to those used by Shaw to interpolate between Shaw's census year benchmarks for 1869, 1879, and 1889. Kuznets also transforms Shaw's data on real commodity output to correspond to a 1929 base year rather than a 1913 base year. While Kuznets's contribution to the derivation of the basic commodity output data is substantial, in what follows I refer to the Shaw/Kuznets series on real finished commodity output simply as the Shaw series. I do this primarily to distinguish the commodity output data from the Kuznets GNP series.[3]

It is important to be very clear about what the Shaw series does and does not cover. The Shaw series appears to be a very good compilation of data on the value of finished commodity output. This series provides a good measure of the gross value of finished commodity output valued at producer prices. However, this series does not include the value added to a good after it reaches its final physical state. Because goods are valued at producer prices, the value added associated with transportation and distribution is excluded. Furthermore, this series also excludes many components of GNP not represented by commodity output such as the flow of services to consumers and the value of new construction.

Kuznets's derivation of the components GNP series centers on converting the commodity output data into estimates of GNP. In this transformation Kuznets uses very little data other than the Shaw series. For the components series Kuznets makes the transformation from commodity output to GNP at a fairly disaggregate level. He derives estimates of the flow of goods to consumers and gross capital formation. For the flow of goods to consumers he estimates the flow of four categories of goods: perishables, semidurables, durables, and services. For gross capital formation he estimates construction, the flow of producer durables, the change in total inventories, and the change in net claims held by foreigners. For all the components of GNP Kuznets first derives estimates of real quantities and then converts them to nominal estimates by means of a price index for that component. For this reason, I only discuss the procedures he uses to derive estimates of real GNP.[4]

Though he is working at a disaggregate level, Kuznets's methods are quite similar for three categories. For the flow of perishable, semidurable and durable goods, Kuznets's problem is to convert the Shaw series on commodity output valued in 1929 producer prices to the value of the flow of goods to consumers at the cost to them in 1929 dollars. To do this, Kuznets begins by taking the ratio of the average flow of a category of goods to consumers to average commodity output in that category for overlapping decades. He then forms a series of the linear trend of this ratio. For the most part, the decadal averages of the flow of goods to consumers are formed by scaling up the decadal averages of commodity output. The scale factors are determined by an analysis of the trends in distributive margins and transportation charges. This analysis is described in National Product Since 1869.

To form the annual series on the flow of a category of goods to consumers in 1929 prices, Kuznets multiplies the trend ratio of the flow of goods to

consumers to commodity output in a given category by commodity output in that category. The key assumption used in the derivation of the annual series is that the relationship between the flow of goods to consumers and commodity output does not vary over the cycle. This is the same as assuming that the components of the flow of goods to consumers not included in commodity output valued in producer prices, primarily the value added in transportation and distribution, move one-for-one with commodity output.

Kuznets's methods for other sectors use similar assumptions. Estimates of the flow of gross producer durables and total construction form the main components of Kuznets's series on capital formation. Both these series are estimated in ways completely analogous to those for the flow of goods to consumers. For producer durables, the Shaw series on the output of producer durables at producer prices is multiplied by the trend ratio of the flow of producer durables to ultimate users to commodity output. For construction, Kuznets forms the trend ratio of construction output to the output of construction materials by linearly interpolating between decadal averages. This trend ratio is then multiplied by the annual Shaw series on the output of construction materials in 1929 prices to form an annual series on total construction output. Again, the key assumption used by Kuznets is that the relationship between commodity output at producer prices and total product does not vary over the cycle.

Whereas for most pieces of GNP Kuznets uses simple linear interpolation by commodity output, he uses a different method for measuring components of GNP not directly involving commodities such as the flow of services to consumers or the net change in inventories. For these series he estimates the actual sensitivity of the component in question to commodity output or to the flow of goods to consumers in a period for which he has good data on both items. He then uses

the estimated sensitivity to transform pre-1918 data on commodity output or consumption into estimates of the needed components.

Kuznets's method for estimating the flow of services provides an example of this procedure. Kuznets has reasonably good data on the flow of both goods and services to consumers for 1919-1941. He calculates the trend level of the two series for this period by connecting overlapping decadal averages. He then regresses the deviation of the flow of services from trend on the deviation of the flow of commodities to consumers from trend. The parameter estimates are then used to create fitted values for 1869-1918 where independent estimates of services do not exist. In creating the fitted values, Kuznets uses his pre-1918 estimates of the flow of goods to consumers. As discussed before, these estimates are themselves created using only a very simple filter of commodity output valued in producer prices.

There are two possible problems with this procedure. The first is that the sensitivity estimate that is used to convert data on consumer expenditures on goods into estimates of consumer expenditures on services is much higher than expected. While the coefficient that Kuznets uses is substantially less than one, it is much larger than the sensitivity estimate for other sample periods. For example, in the postwar era the relationship between consumer expenditures on goods and consumer expenditures on services is actually negative.[5] As a result, the Kuznets prewar estimates of the flow of services may move too much over the cycle.

A second problem is that Kuznets's procedure involves estimating the relationship between true data on consumer expenditures on services and true data on consumer expenditures on goods in the interwar era. He then uses the sensitivity estimate to form prewar estimates of services. The possible flaw in this procedure is that the base data on the flow of goods to consumers may not

be uniformly good between the pre-1918 era and the 1919-1941 period. From the description of the methods used to form the pre-1918 consumption estimates it is clear that these estimates may be excessively volatile because they are formed using the assumption that the flow of goods to consumers moves one-for-one with commodity output. If this is true, then using the coefficient based on accurate consumption data will pass on too much of the movement in the flawed data on the flow of goods to consumers to the estimates of the flow of services. The resulting estimates of the flow of services to consumers could thus be excessively volatile.

2. Regression Series

The second GNP series that Kuznets creates for the period 1869-1918 is called the regression series. The regression series is derived by estimating the aggregate relationship between real commodity output and real GNP for the period 1909-1938. This relationship is then used to convert Shaw's pre-1918 data on commodity output into estimates of GNP. This method attempts to take into account at the aggregate level the fact that GNP does not vary as much over the cycle as does commodity output valued in producer prices.

The actual derivation of the regression series is not complicated. Kuznets first forms estimates of trend GNP and trend commodity output. To do this he calculates averages for overlapping decades and uses these to represent trend GNP or commodity output for the midpoint of the decade. Annual trend values are derived by linear interpolation between these midpoints.[6] GNP and commodity output for 1909-1938 are then expressed as percentage deviations from trend. Kuznets then fits "a freehand regression curve" to a scatter plot of the deviations of real commodity output and real GNP from trend (Kuznets, 1961, p. 537). This curve is used to form fitted values for the pre-1918 era.

Though Kuznets asserts that the regression curve cannot be expressed as a simple mathematical function, it is actually very close to a simple linear regression. One can replicate Kuznets's pre-1918 regression series very accurately using the results of a simple linear regression of the percentage deviations of GNP from trend on the percentage deviations of the Shaw series on total commodity output from trend and a constant. The actual parameter estimates are:

$$gnp_t = \underset{(0.005)}{-0.001} + \underset{(0.046)}{0.896}\ co_t + e_t$$

where small letters denote percentage deviations from trend.[7] The fitted values from applying this relationship to the pre-1918 era differ from Kuznets's regression estimates usually by less than one-half of one percent and only for two years by more than one percent.

From this representation of the derivation of the regression series it is clear that it has much in common with the components series. Though the two GNP series are derived at very different levels of aggregation and using somewhat different assumptions, both series suggest that the cyclical movements in GNP move very closely with those in commodity output. The components series for the most part assumes that GNP moves one-for-one with commodity output. The regression series uses the estimated sensitivity of GNP to commodity output over the period 1909-1938 to estimate GNP for 1869-1918. The assumption derived from this procedure is that cyclical movements in GNP are 90 percent as large as those in commodity output.

While the basic procedures underlying the components and regression series are quite similar, there are obvious differences between the two series. The regression series is derived at a much more aggregate level than is the components series. Also, the regression methodology removes the random movements in

GNP that are not perfectly correlated with commodity output. Finally, the regression series is derived using the assumption that the deviations from trend of GNP move 0.9-for-one with the deviations of commodity output from trend rather than one-for-one as is assumed in the derivation of the components series. The comparison of volatility properties of the two series given in Section IV shows that the effect of these differences is exactly what one would predict. The regression series is consistently less volatile than the components series. However, the magnitude of the difference in volatility is small and the prewar regression series is still substantially more volatile than the actual postwar GNP series.

SECTION II

EXCESS VOLATILITY IN THE KUZNETS GNP ESTIMATES

The description of the Kuznets components and regression estimates of GNP points out a key similarity between the two series: both data sets are derived by postulating that the deviations of GNP from trend move very closely with the deviations of commodity output from trend. This section analyzes the accuracy of this key assumption underlying both Kuznets GNP series for 1869-1918.

1. Possible Sources of Excess Volatility

At an abstract level there is reason to suspect that both of Kuznets's procedures yield estimates of GNP that overstate the size of cyclical movements. The key problem is that commodity output represents a very cyclically sensitive portion of GNP. Data on commodity output valued at producer prices do not include the pieces of GNP that are typically considered to be the less cyclically sensitive components of total national output such as trade, transportation, and services. As a result, assuming that the total GNP has cyclical movements identical (or nearly so) to those in commodity output yields a GNP series that overstates the size of cyclical fluctuations.

That commodity output is an unrepresentative fraction of GNP is for the most part a simple empirical regularity in postwar data. However, this regularity is plausible at a theoretical level. First, in the case of consumer

expenditures on services it is likely that the demand for many services is invariant to the state of the cycle. Except under extreme circumstances one would not expect expenditures on haircuts or physicians' services to fluctuate dramatically. The same is not true of expenditures on commodities (especially durable goods) where demand may not have a strong time-specific character. Hence, one would not expect expenditures on services to vary closely with commodity output.

Fixed costs may provide an explanation of why the trade and transportation component of consumer expenditures on goods does not move one-for-one with commodity output valued at producer prices. It is certainly possible that fixed costs for trade and transportation firms are relatively larger compared to variable costs than they are for manufacturing firms. This could be due, for example, to the fact that retail firms are typically quite small. As a result, overhead expenses may be very large relative to labor costs.

If this is true, then one would expect the value added in transportation and distribution to be less cyclically volatile than commodity output. In a cyclical downturn, a retail distributor or a transportation firm may be unable to cut costs proportionately with the decline in volume simply because variable cost is a small fraction of total cost. In this case, the distributive margin on each good will rise. As a result, the final gross value to consumers of the commodity in question will not have fallen by as much as the value of the commodity to producers. Hence consumer expenditures on commodities will tend to be less cyclically volatile than commodity output valued at producer prices.

2. Evidence of Excess Volatility

While it is easy to construct stories about why GNP should not move one-for-one with commodity output valued at producer prices, such stories do not

provide evidence that both the Kuznets GNP series are in fact excessively volatile. However, estimates of the actual sensitivity of the deviations of GNP from trend to the deviations of commodity output from trend from periods where good data are available on both quantities does suggest that Kuznets's assumptions of a one-to-one relationship or a .9-to-one relationship are erroneous.

To examine the validity of Kuznets's assumptions I regress the deviations of aggregate GNP from trend on the deviations of aggregate commodity output from trend for various time periods. That is, I estimate the equation

$$(1) \qquad gnp_t - \overline{gnp}_t = \gamma(co_t - \overline{co}_t) + \varepsilon_t$$

where gnp_t is the logarithm of aggregate real gross national product, co_t is the logarithm of aggregate real commodity output, and bars over a variable denote trend values. As is discussed in detail in Section III, trend values are calculated using piecewise linear trends between years that are identified as times of full employment (rather than times of exceptionally low unemployment). This method yields estimates of trend values that correspond to normal output, not to peak or potential output.

The earliest time period that can be examined begins in 1909. For 1909-1918 Kuznets provides very rough income-side estimates of GNP.[8] While the estimates surely have a great deal of random measurement error, they should be free of the cyclical exaggeration present in the product-side estimates. Hence, they should provide a consistent estimate of the sensitivity of cyclical movements in GNP to the cyclical movements in commodity output.

The relationship between GNP and commodity output can be examined in the postwar era as well. Data on commodity output valued at producer prices are available from the Federal Reserve Board in the series titled the gross value of

final products.[9] Since this series is only available after 1954, I also examine the relationship between GNP and the FRB value-added index of final goods production over the entire postwar era. The similarity in the results using both the gross value and value-added series for the same sample period suggests that this approximation is very close.

The coefficient estimates from the regression of the cyclical movements in GNP on the cyclical movements in commodity output for various time periods are shown in Table 1. For most sample periods over the interval 1909-1983 the coefficient estimate is substantially below 0.9 or 1. It generally ranges between 0.6 and 0.7. The coefficient estimates for the postwar era are somewhat smaller than those for the interwar era.

The one exception to this pattern involves any sample period that includes the Great Depression. When the 1930s are included in any regression, even a pooled sample covering both the interwar and postwar eras, the coefficient estimate is approximately 0.9 or higher. This result explains why Kuznets in deriving his regression estimates of GNP found that GNP and commodity output had very similar cyclical properties. He specifically estimated the relationship between the two series over a period that was dominated by the Great Depression.

There can be little doubt that the results from the Depression are anomalous. For no other sample period is the coefficient estimate anywhere near as high. As Table 1 shows, the relationship between GNP and commodity output is very stable in periods outside the Depression. In both the periods 1909-1918 and 1919-1928 the coefficient estimate is approximately 0.65. In contrast, the coefficient estimate for 1929-1938 is 0.97. These results are completely robust to minor changes in the definition of decades.

That the Great Depression should show a much higher sensitivity of movements in GNP to movements in commodity output than do other periods is not

TABLE 1

Coefficient Estimates

	Independent Variable	Sample Period	Coefficient γ	Standard Error of γ
Interwar	CO-Shaw[a]	1909-1938	.929	.045
	CO-Shaw	1909-1928	.649	.106
	CO-Shaw	1909-1929	.667	.090
	CO-Shaw	1909-1930	.667	.089
	CO-Shaw	1919-1928	.647	.111
	CO-Shaw	1919-1929	.670	.087
	CO-Shaw	1919-1930	.670	.089
	CO-Shaw	1919-1938	.939	.046
	CO-Shaw	1909-1918	.655	.235
	CO-Shaw	1909-1919	.700	.239
	CO-Shaw	1909-1920	.549	.238
	CO-Shaw	1929-1938	.971	.050
	CO-Shaw	1930-1938	.983	.050
Postwar	CO-FRB[b]	1955-1983	.505	.031
	CO-FRB	1955-1962	.602	.071
	CO-FRB	1963-1972	.432	.034
	CO-FRB	1973-1983	.585	.058
	IP-FRB[c]	1947-1983	.541	.034
	IP-FRB	1947-1962	.597	.089
	IP-FRB	1963-1972	.478	.027
	IP-FRB	1973-1983	.578	.040
Interwar and Postwar	CO-Shaw and CO-FRB	1909-1938 and 1955-1983	.866	.036
	CO-Shaw and CO-FRB	1909-1928 and 1955-1983	.565	.048
	CO-Shaw and IP-FRB	1909-1938 and 1947-1983	.867	.034
	CO-Shaw and IP-FRB	1909-1928 and 1947-1983	.584	.045

[a]Shaw's series on the value of commodity output in 1929 dollars.

[b]Federal Reserve Board series on the gross value of final products and construction materials in constant dollars.

[c]Federal Reserve Board value-added index of the production of final products and construction materials.

surprising. Most theories of consumer and producer behavior suggest that agents will respond much differently to a particularly long and severe depression than they will to a recession that is expected to be mild or brief. For example, while consumers will typically smooth their expenditures on services, in a very severe depression they are likely to reduce those expenditures. This could be due to revisions in estimates of permanent income or to increasingly binding liquidity constraints. Similarly, while distributive margins typically rise in a recession, they may actually fall in a long and severe depression. This could be caused by bankruptcies among transportation and retail establishments that allow remaining firms to have sufficient volume to cover fixed costs. Both of these responses to severe economic decline can explain why GNP moves much more closely with commodity output during the 1930s than during any other period.

The fact that the sensitivity of movements in GNP to movements in commodity output for most sample periods excluding the Great Depression is substantially below that assumed by Kuznets suggests that both the Kuznets components and regression series are excessively volatile. This is especially true given how anomalous the decline of the 1930s is in comparison to the economic fluctuations of the late 1800s and early 1900s. Judging from the cyclical behavior of commodity output, the percentage decline in production during the Great Depression was nearly three times as large as the percentage decline in the worst depression of the 1869-1918 period.[10] Hence it is much more reasonable to believe that the relationship between GNP and commodity output is more like that exhibited in both the interwar and postwar eras than during the 1930s. In this case, GNP would be far less volatile than the Kuznets estimates suggest.

3. The Decline in the Ratio of Commodity Output to GNP

While the available evidence suggests that the Kuznets series are

excessively volatile, conclusive proof is impossible to find. Most importantly one must worry that GNP moved more closely over the cycle in the period 1869-1918 than it did in the interwar and postwar eras. It is possible that structural changes in the economy caused the sensitivity of GNP to commodity output to have been larger in the prewar era than in the postwar era. If this is true, then Kuznets's estimates of prewar GNP may not overstate cycles as much as the previous evidence suggests.

The prime structural change that might have caused the relationship between GNP and commodity output to change over time is the relative decline in the commodity sector of the economy. If commodity output represented a larger fraction of GNP in the pre-1918 period than in later periods, then it is possible that GNP should move more closely with commodity output in the earlier period than it does today.

Two observations on the decline of the commodity sector suggest that this change is not a major cause for concern. The first is that the decline in the size of the commodity sector is a modern occurrence. Furthermore, the decline in the size of the commodity sector within the modern period has been quite small. The second observation is that the change in the relative size of the commodity sector between the interwar and postwar eras has little effect on the sensitivity of cyclical fluctuations in GNP to fluctuations in commodity output.

Evidence on the changing share of commodity output in GNP is of differing quality. For the period after 1909 we possess independent estimates of both real GNP and real commodity output valued in producer prices. Hence one can simply examine the ratio of these two series. The ratio of commodity output to GNP in selected years beginning in 1910 is given in Table 2. The years shown are the same benchmark years used to construct the piecewise linear trends discussed earlier. The ratio shows a definite downward trend over time in the

TABLE 2

Ratio of Commodity Output to GNP

Year	Ratio[a]
1910	.458
1918	.480
1924	.419
1937	.397
1955	.414
1962	.416
1972	.379
1981	.353

[a]The ratio compares the value of commodity output in constant dollars to gross national product, also valued in constant dollars.

share of commodity output in GNP. However, it is crucial to observe that the magnitude of the decline is quite small. Over the period 1909-1983 the ratio only fell from approximately 0.46 to 0.35. This suggests that even if this same downward trend were present from 1869 on, the ratio of commodity output to GNP in the prewar era would not be dramatically larger than it is in the postwar era.

Fragments of prewar evidence suggest that the decline in the ratio of commodity output to GNP in the pre-1918 era was less pronounced than the modern behavior of this ratio suggests. In deriving the components series on GNP, Kuznets assumed that the ratio of commodity output to GNP in various sectors from 1869-1918 was the same as the ratio of the two quantities over the period 1919-1928. Kuznets based this assumption on the behavior of the shares of total income accruing to workers in the transportation and distribution sectors in census years in the prewar era. He felt that while the income studies showed "a mild downward trend" in the ratio of commodity output to GNP, the trend was both slight enough and variable enough that it was best to assume that the ratio of commodity output to GNP was essentially flat over the period 1869-1929 (Kuznets, 1946, p. 67).

A study by Harold Barger on the role of distribution in the American economy confirms Kuznets's belief to a large degree. Barger finds that as the cost of distribution rose slowly, the cost of transporting goods from producers to distributors fell steadily. As a result, "the ratio of distribution cost plus freight charges to producers' value scarcely rose at all" between the Civil War and the Great Depression (Barger, 1955, p. 63). Barger does suggest that there may nevertheless have been a slight fall in the ratio of commodity output to GNP over time simply because more goods began to enter the distribution process rather than remain on the farm or within the household (Barger, 1955,

p.64). However the implied change in the ratio of real commodity output valued in producer prices to real GNP between 1869 and 1929 is still very small.

The Barger and Kuznets results are important because they suggest that there is little reason to suspect that the relationship between GNP and commodity output changed significantly between the prewar and interwar eras. Since the ratio of commodity output to GNP declined only very slightly (if at all) between these two periods, it is hard to believe that the sensitivity of GNP to commodity output was much different between the two periods.

One further piece of evidence on the excess volatility of the Kuznets GNP series is the effect of changes in the ratio of commodity output to GNP on the sensitivity of cyclical movements in GNP to movements in commodity output. As mentioned earlier, the ratio of commodity output to GNP fell somewhat between 1909-1983. The coefficient estimates shown in Table 1 suggest, however, that this fall did not cause dramatic changes in the measured sensitivity of GNP to commodity output. For example, the coefficient estimate for 1909-1918 is 0.66 while the estimate for 1973-1983 is 0.59. This implies that even if the ratio of commodity output to GNP were noticeably larger in the prewar era than it is in the interwar and postwar eras, the estimated sensitivity of GNP to commodity output would be substantially lower than 0.9 or 1. Hence, Kuznets's GNP estimates would still be excessively volatile.

SECTION III

NEW ESTIMATES OF PREWAR GROSS NATIONAL PRODUCT

The preceding descriptions and analysis of the two Kuznets series suggest that the existing estimates of prewar GNP do not measure cycles correctly. Rather, both the Kuznets components and regression series exaggerate the size of cyclical fluctuations in the period 1869-1918. Given that there are problems with the Kuznets estimates, it is only natural to attempt to create better estimates of prewar GNP. Since none of the analysis has challenged the accuracy of the trend values of GNP, the derivation of new estimates presented in this section will concentrate on improving only the representation of cyclical movements.

1. Overview of Methods

In deriving new estimates of prewar GNP the most obvious problem is a lack of data. Obviously, modern researchers are forced to estimate prewar GNP precisely because we do not have data on many of the components of total output. Kuznets is quite convincing that we possess little data on the components of GNP other than comprehensive figures on the value of commodity output in producer prices. Data on distributive margins, the change in inventories, and consumer expenditures on services are very scarce and subject to exceedingly large margins of error. While it is possible that new and extensive research could

provide adequate direct estimates of these components of GNP, for now it is necessary to assume that the main information we possess about movements in GNP is provided by the available data on commodity output.

Given the limitations on data, it is thus sensible to view the creation of new estimates of GNP as finding the best method for transforming data on commodity output into estimates of GNP. If one views the problem this way, it is clear that one should use some sort of regression technique. Simply imposing a one-to-one relationship between the two series will not yield the best estimates of GNP. As Friedman shows in his article on linear interpolation, using some estimate of the relationship between the series to be created and the interpolating series (perhaps derived from a fragment of data or from a different sample period) will, in most circumstances, yield a more accurate series than simply imposing a one-to-one relationship (Friedman, 1962).

The method that I use to derive new estimates of prewar GNP corresponds very closely to Kuznets's aggregate regression technique. I assume that the deviations of aggregate real GNP from trend are correlated with the deviations of aggregate real commodity output from trend. The actual sensitivity of cyclical movements in GNP to cyclical movements in commodity output is estimated from a sample period that includes both the interwar and postwar eras, but excludes the Great Depression.

The main improvement that I make in Kuznets's procedure is to allow the measured sensitivity of aggregate GNP to aggregate commodity output to change over time. As shown in Section II, there is some evidence that aggregate GNP is less sensitive to movements in aggregate commodity output in the postwar era than in the interwar era. This is probably due to the modest decline that has occurred in the ratio of commodity output to GNP. Since the same trend may extend to the pre-1918 period as well, it is useful to allow the estimated

sensitivity to decline over time as well.[11] By doing so, one can derive an estimate of the sensitivity of GNP to commodity output for the period 1869-1918 that represents this time period more accurately.

The decision to create a new aggregate regression series is a choice that should be defended. The obvious alternative procedure is to employ regression techniques at a disaggregate level. For example, one could estimate the relationship between GNP and commodity output for each sector and use these relationships to create new prewar GNP data. The obvious benefit of this approach is that it allows the sensitivity of GNP to commodity output to vary across sectors. Furthermore, to the extent that the relative size of sectors changes over time, it allows the aggregate sensitivity of GNP to commodity output to vary over time as well.

The main reason that I choose not to follow the disaggregate approach is that by using a time-varying sensitivity estimate, the aggregate regression methodology preserves the main benefit of the disaggregate approach. The key argument in favor of the disaggregate approach is that it allows GNP to move more closely with commodity output in the pre-1918 era when the noncommodity components of GNP were slightly smaller. The modified aggregate regression approach that I employ also allows the estimated sensitivity of GNP to commodity output to be higher in the prewar era. However, it does so in a way that is much more straightforward and much less time consuming than does the disaggregate approach.

2. Specifics of the Deviation

The basic procedure that I use to create new GNP data is thus straightforward. In some period when I have good aggregate data on both GNP and commodity output I estimate the relationship between the deviations from trend

The choice of sample period is also influenced by the sensitivity analysis of Section II. The regressions reported in Table 1 show that the relationship between GNP and commodity output during the Great Depression was much different than the relationship during any other time period. Since the Great Depression appears to be a downturn of unprecedented magnitude, it is best not to let the experience of the 1930s determine the relationship between GNP and commodity output in the prewar era. Hence, I exclude the Great Depression from the sample period over which I estimate equation (2).

Trend Values. The calculation of trend values for GNP and commodity output is a third complicating factor in the derivation of new estimates of GNP. To estimate equation (2) and form new estimates of GNP for 1869-1918, one must specify trend values for both GNP and commodity output for the entire period 1869-1983. The methods that can be used are limited by the fact that the procedures for forming trend values must be applicable to the pre-1918 period for which we do not possess estimates of GNP that represent cycles accurately. The method I use for calculating trend values involves interpolating linearly between benchmark estimates of the logarithms of GNP and commodity output. Since the benchmark values chosen are generally only 8 to 12 years apart, this method allows the trend of each series to change frequently over the period in question.

In using piecewise-linear trends, the key step is deciding which years to use as benchmarks. Often, researchers choose to connect peak years and thus form an estimate of potential rather than trend GNP or commodity output (see, for example, Gordon, 1982). For this study, I specifically choose years that correspond only to trend output or unemployment at the natural rate rather than to peak output. This was a necessary change because the existing Kuznets estimates of prewar GNP accentuate the size of cyclical fluctuations. As a

result, the only years for which the pre-1918 GNP estimates are accurate are years when the economy is neither below or above trend.

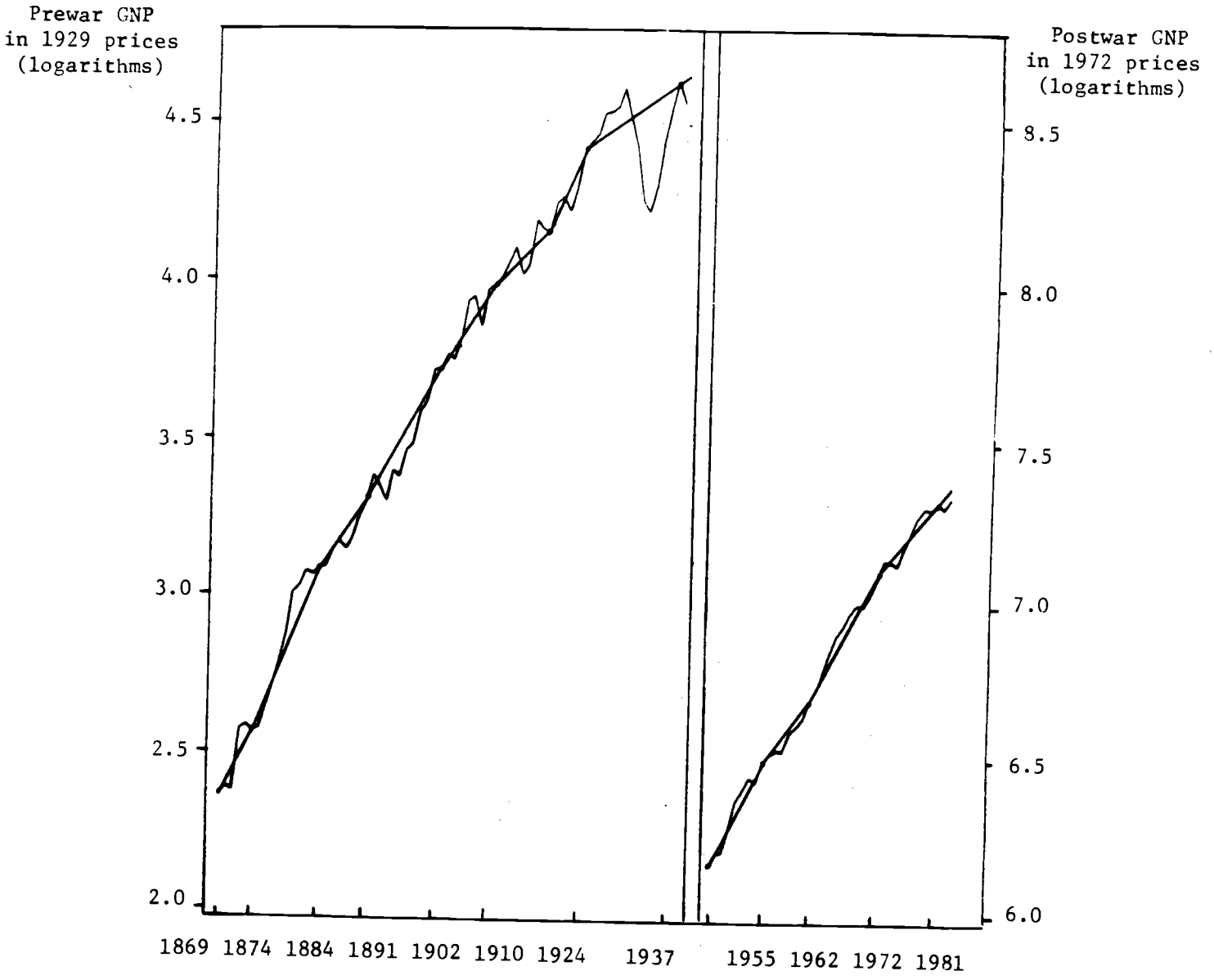
Deciding during which years the economy was on trend involves an admittedly arbitrary and imperfect procedure. In choosing benchmark dates I use a mixture of examining a plot of the data in logarithms and a qualitative knowledge of which prewar and postwar years are typically considered to correspond to periods of boom and recession. From the plot of the data I try to choose years that correspond to points of mid-expansion in the business cycle. For the postwar era I also use data on the unemployment rate to confirm that the years chosen do correspond to conventional estimates of full, rather than over-full, employment.

The actual years chosen as benchmark estimates for both GNP and commodity output are: 1869, 1874, 1884, 1891, 1902, 1910, 1918, 1924, 1937, 1947, 1955, 1962, 1972 and 1981.[13] Figures 1 and 2 show the resulting trend values of GNP and commodity output and the existing underlying actual data. From these graphs, it should be clear that the years chosen as benchmarks do correspond reasonably well to points of mid-expansion in the cycle.[14] Another fact that should be apparent from Figures 1 and 2 is that even relatively large changes in the choice of benchmark years would not alter the trend values substantially.[15]

Data. The final issue to consider in deriving new estimates of prewar GNP involves which data to use. For estimating the relationship between GNP and commodity output in the interwar and postwar eras, the choice of data is fairly obvious. For real GNP the Kuznets income-side estimates are used for 1909-1928. The Commerce Department GNP series is used for 1955-1983. For real commodity output, the Shaw series is used for 1909-1928. The Federal Reserve Board series on the gross value of final products (amended to include the gross value of construction materials) is used for 1955-1983. All data are expressed in terms of constant dollars.

Figure 1

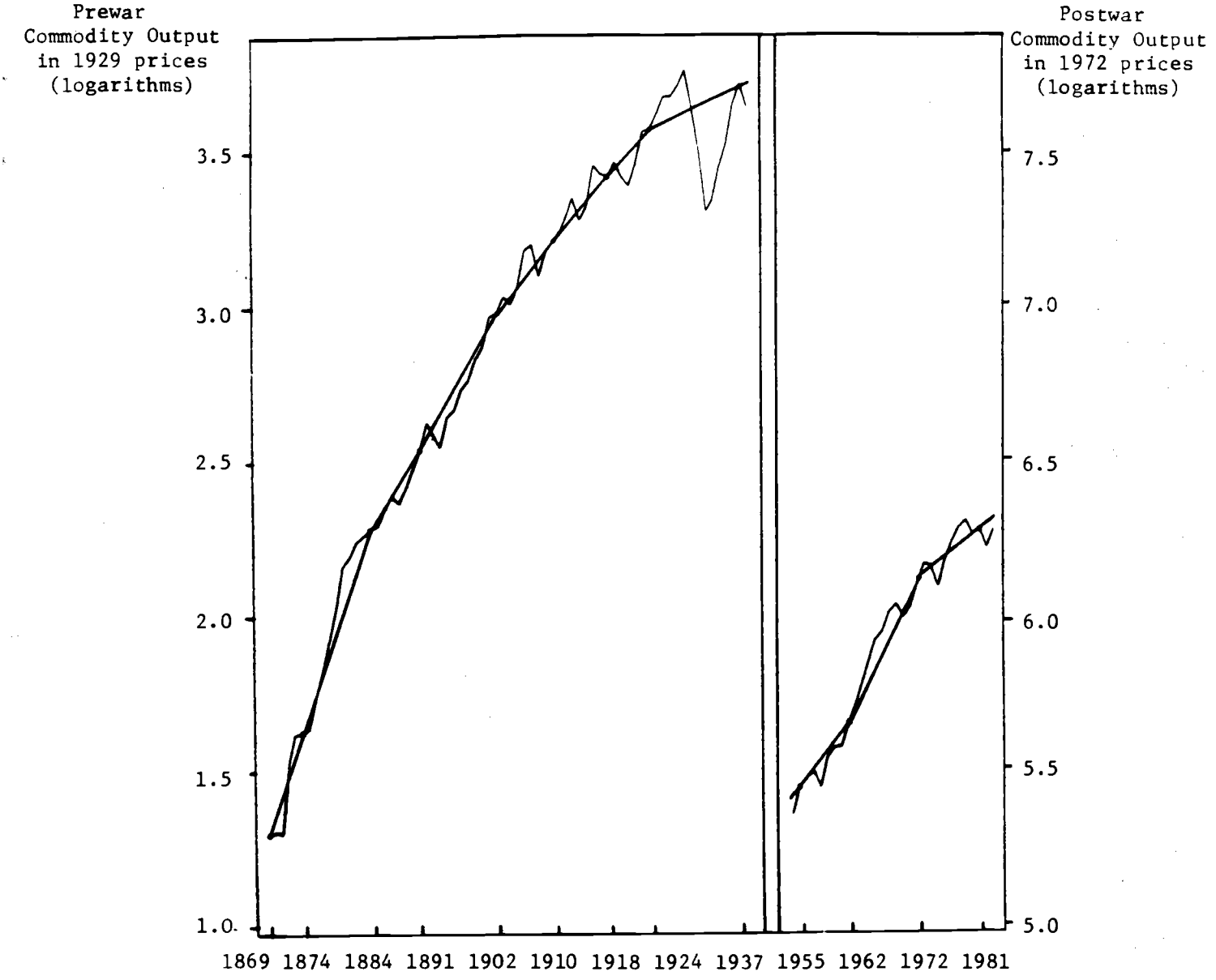
Gross National Product and Trend Gross National Product
1869-1938 and 1947-1983



Notes and Sources: For 1869-1918 the Kuznets/Gallman components series on GNP in 1929 prices is used. The net national product component of this series is available in Friedman and Schwartz (1982, Table 4.8, pp.121-128). The capital consumption component is from unpublished work by Kuznets. For 1919-1938 the Kuznets income-side estimates, Variant III (also in 1929 prices) are used. This series is published in Capital in the American Economy (1961, Table R-2, p. 487). For 1947-1983 the GNP data are the standard Commerce Department values in 1972 prices. The data are available with current revisions in the Economic Report of the President for 1985. Trend values are calculated using the procedures described in the text.

Figure 2

Commodity Output and Trend Commodity Output
1869-1938 and 1954-1983



Notes and Sources: For 1869-1938 the Shaw/Kuznets series on commodity output in 1929 prices is used. This series is published in Capital in the American Economy (1961, Table R-21, pp. 553-554). For 1954-1983 the Federal Reserve Board series on the gross value of final products is used. I amend this series to include the gross value of construction supplies. The gross value data for 1954-1967 (on a 1963 base) are available in Industrial Production, 1971 edition. The data for 1967-1975 (on a 1972 base) are available in Industrial Production, 1976 edition. The data for 1976-1983 are available from monthly Federal Reserve Board releases. The 1963 and 1972 base year series are combined using a ratio splice based in 1967. Trend values are calculated using the procedures described in the text.

In forming new estimates of prewar GNP, the choice of data is less straightforward. For commodity output the Shaw series is the obvious series to use. However, for calculating trend GNP there are some alternatives. One can either use the original Kuznets series or one of the revisions of this series to derive benchmark estimates. I choose to use Gallman's amended version of the Kuznets components series. I do this because the Gallman series appears to correct some of the suspected anomalies in the trend of the Kuznets series.[16] Again, both commodity output and GNP are expressed in constant (1929) dollars.

3. Results

Parameter Estimates. The ultimate result of following the procedures described is the derivation of new estimates of real GNP for 1869-1918. An important intermediate step is the estimation of the time-varying relationship between the deviations from trend of aggregate real GNP and aggregate real commodity output. When equation (2) is estimated over the period 1909-1928 and 1955-1983 the parameter estimates are:

$$(2') \qquad gnp_t - \overline{gnp}_t = \underset{(.094)}{(.667} - \underset{(.0020)}{.0025 \cdot trend_t})(co_t - \overline{co}_t)$$
$$S.E. = 0.017$$

where standard errors are in parentheses.[17]

The coefficient estimates are perfectly sensible. They suggest that the deviations from trend of GNP are substantially smaller than those of commodity output and that the sensitivity of GNP to commodity output has declined over time. However, the change in the coefficient over time is reasonably small. According to the estimates, the time-varying coefficient measuring the sensitivity of GNP to commodity output fell from 0.664 in 1909 to 0.483 in 1983.

The presence of a nontrivial trend in the coefficient implies that the coefficient one should use in converting prewar data on commodity output into estimates of GNP is bigger than any of the coefficients one finds using an interwar or postwar sample. Indeed, projecting the simple linear trend back in time, the coefficient ranges between 0.763 in 1869 and 0.642 in 1918. By using these time-varying estimates of the coefficient, one is able to capture the very plausible decline in the sensitivity of GNP to commodity output.

New Estimates of Prewar GNP. The new estimates of prewar gross national product that result from following the procedures described are given in Table 3. The new estimates cover the period 1869-1918 and represent real GNP in constant (1929) dollars. I create new estimates of GNP for 1909-1918 despite the fact that rough income-side estimates of GNP are available for this decade. I do this because the data for the 1910s have sufficient measurement error that using the new methods is likely to provide point estimates of GNP that are more accurate than the original series. The best evidence for this proposition is that Kuznets also chose to use his regression procedure to create estimates of GNP for 1909-1918 rather than use the income-side series.

The new estimates of prewar GNP are graphed in Figure 3. For comparison, the Kuznets/Gallman components series is also presented in Figure 3. Two characteristics of the new series are apparent from the figure. The first is that the new estimates of prewar GNP are identical to the Kuznets/Gallman estimates in benchmark years. This is to be expected given the procedures used to derive the new series. The second is that cyclical movements in the new series are noticeably smaller than cyclical movements in the Kuznets/Gallman series. This difference between the two series is analyzed in detail in Section IV.

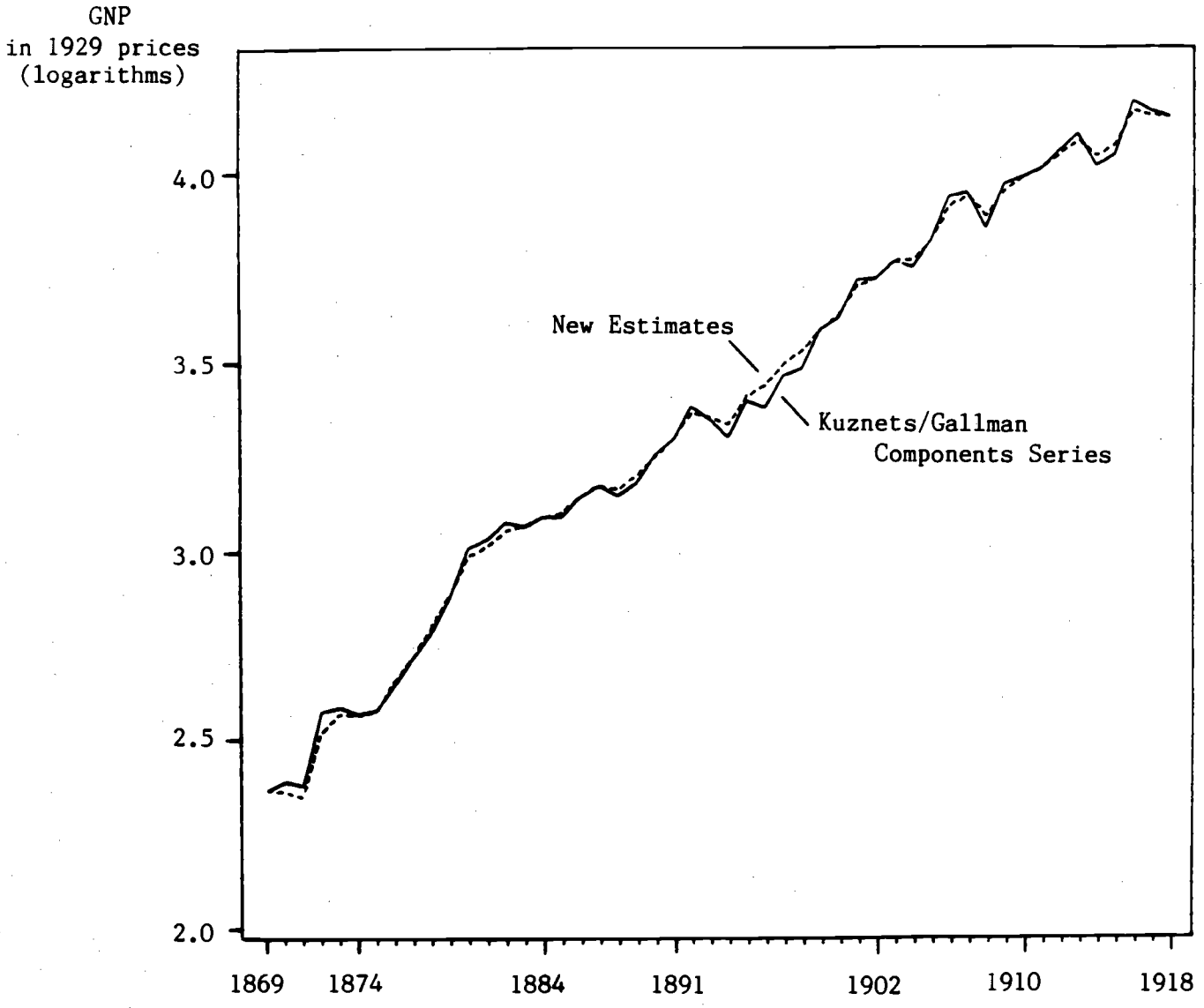
TABLE 3

New Estimates of Gross National Product
1869-1918

Year	GNP (In Billions of 1929 Dollars)	Year	GNP (In Billions of 1929 Dollars)
1869	10.654	1894	28.095
1870	10.593	1895	30.395
1871	10.470	1896	31.186
1872	12.398	1897	33.041
1873	13.046	1898	34.160
1874	13.020	1899	36.180
1875	13.177	1900	37.621
1876	14.241	1901	40.662
1877	15.177	1902	41.416
1878	16.440	1903	43.574
1879	17.817	1904	43.462
1880	19.806	1905	45.819
1881	20.298	1906	50.174
1882	21.154	1907	51.584
1883	21.507	1908	48.853
1884	21.995	1909	52.199
1885	22.214	1910	54.263
1886	23.190	1911	55.427
1887	23.906	1912	57.572
1888	23.676	1913	59.824
1889	24.555	1914	57.363
1890	25.736	1915	59.086
1891	27.057	1916	64.690
1892	28.995	1917	63.830
1893	28.550	1918	63.640

Figure 3

New and Old Estimates of Gross National Product
1869-1918



Notes and Sources: The net national product component of the Kuznets/Gallman series is available in Friedman and Schwartz (1982, Table 4.8, pp. 121-128). The capital consumption component is from unpublished work by Kuznets. The derivation of the new estimates is described in the text.

For anyone interested in using the new estimates, certain facts are relevant. One is that the current estimates are in constant dollars. A nominal series can be formed by multiplying the new estimates by the Kuznets implicit price deflator series given in Friedman and Schwartz (1982, pp. 122-129). Another is that the new GNP series is designed to be consistent with the original Kuznets series after 1918. While the new estimates incorporate Gallman's revisions to the trend of the Kuznets series before 1909, they are consistent with the Kuznets data for the 1920s and 1930s. Thus, the new series can easily be extended through 1938 using the Kuznets estimates of GNP in 1929 dollars (Variant III) given in Capital in the American Economy (1961, Table R-2, p. 487).

While the new estimates are consistent with the Kuznets series after 1918, they are not consistent with the modern Department of Commerce estimates. In the derivation of the new estimates I abstract from the work of Kendrick and the Department of Commerce that amends the Kuznets data to fit modern Department of Commerce procedures. I do this for two reasons. The main one is that the amended base data using Commerce Department concepts begin in 1889. Since the goal to this paper is to form new data back to 1869, following Kendrick would have required a tremendous amount of research before one could even begin to analyze problems with cyclical representations. While the derivation of Kendrick-like corrections for 1869-1888 is certainly a useful endeavor, it was clearly out of the range of this study.

The second reason is that for cyclical analyses, the Kendrick/Commerce Department corrections are of essentially no importance. These corrections primarily involve the treatment of military expenditures. They also involve separating government expenditures into a separate category rather than including them in the flow of goods to consumers and capital formation categories as

Kuznets does. There is nothing in the corrections that significantly changes the cyclical properties of the underlying Kuznets components series. As a result, they can be ignored for the purpose of deriving new GNP estimates that measure cycles more accurately.

SECTION IV

VOLATILITY PROPERTIES OF THE NEW GNP SERIES

Having derived new estimates of GNP, it is natural to ask how these estimates change one's perception of the prewar economy. Since the new estimates are designed to measure short-term cyclical fluctuations more accurately than do the original Kuznets estimates, the most interesting application is to analyze the business cycle properties of the new GNP data. This section considers the basic volatility properties of the prewar GNP series. Various measures of the severity of cycles are used to compare Kuznets's GNP estimates, my new estimates, and postwar data on actual GNP.

The cyclical properties of the Kuznets components series are analyzed in detail in existing studies. Several studies use the existing prewar GNP data to argue that business cycles are more severe and the economy is in general more volatile in the prewar era. This is certainly the case in two recent articles, one by Martin Neil Baily (1978) and one by J. Bradford DeLong and Lawrence Summers (1984). While neither of these studies uses Kuznets's data directly, both use amended versions of prewar GNP data that derive their cyclical movements from the Kuznets components series.[18]

Both of these studies conclude that there has indeed been a dramatic stabilization of the postwar economy. Even when the Great Depression is excluded, cyclical fluctuations in prewar GNP are nearly twice as large as

fluctuations in postwar GNP. However, the discussion in Sections I and II should make one skeptical of these results. That discussion shows that there is reason to believe that both the Kuznets components series and the regression series exaggerate cyclical movements in GNP. Given that the new estimates of GNP are derived in such a way that they should represent cyclical movements more accurately, it is interesting to see if these estimates also show that cycles were more severe in the prewar era.

1. Measures of Volatility

Table 4 shows three common measures of volatility for the various GNP series. I analyze the Kuznets/Gallman components series and the Kuznets regression series on real GNP, my new prewar estimates of real GNP, and the standard Commerce Department series on postwar GNP in 1972 dollars. For the three prewar series I examine the volatility properties over the period 1869-1918. For the postwar series I use the sample period 1947-1983.

The first measure to consider is the mean cyclical amplitude of the detrended GNP series. This is obviously a measure of the severity of cycles in the two periods. The various GNP series are detrended using the piecewise linear trends discussed in Section III. Peaks and troughs are defined by the actual turning points in detrended commodity output.[19] Peaks and troughs defined this way are somewhat different from NBER reference dates which are derived using undetrended data.[20] Cycles with declines in GNP of less than two percent are excluded from calculations of the mean. The results, however, are very similar when only cycles of less than one percent are excluded from the calculation of the mean.

The results are quite striking. A comparison of the average amplitude of the prewar components series and the actual postwar series shows the usual

TABLE 4

Measures of Volatility

Series	Mean Cyclical Amplitude of Detrended Series[a]	Standard Deviation of Percentage Changes	Standard Deviations of Deviations from Trend
GNP - Kuznets/Gallman, Components (1869-1918)	.101	.058	.054
GNP - Romer (1869-1918)	.074	.040	.039
GNP - Kuznets, Regression (1869-1918)	.088	.052	.050
GNP - Commerce Department (1947-1983)	.053	.027	.028
GNP - Constructed Postwar[b] (1955-1983)	.050	.026	.024
Commodity Output - Shaw (1869-1918)	.104	.057	.055
Commodity Output - FRB (1955-1983)	.098	.046	.051
Industrial Production - FRB[c] (1947-1983)	.101	.049	.048

[a]Series are detrended using the piecewise linear trends described in the text. Cycles of less than 2 percent are excluded from the calculation of the mean.

[b]The constructed postwar regression series is formed using the same methods and time-varying coefficient that is used to form the new estimates of prewar GNP. This series can only be constructed for the period 1955-1983 because the commodity output data are only available after 1954. For comparison, the measures of volatility for actual GNP for 1955-1983 are: Mean Cyclical Amplitude - .054; Standard Deviation of Percentage Changes - .025; and Standard Deviation of Deviations from Trend - .028.

[c]The industrial production series used is a weighted combination of the FRB final products index and the construction materials index. To facilitate comparison with the gross value series, the measures of volatility for the industrial production series for 1955-1983 are: Mean Cyclical Amplitude - 0.104; Standard Deviation of Percentage Changes - .047; and Standard Deviation of Deviations from Trend - .049.

stylized fact. There appears to be a dramatic decline in the severity of cycles between the prewar and postwar eras. The average amplitude of the prewar Kuznets/Gallman components series is approximately 64 percent larger than that of the postwar GNP series. This damping of business cycle fluctuations over time is also apparent when one compares the Kuznets regression series to the postwar GNP data. However, the degree of stabilization is somewhat smaller when the regression series is used in place of the components series. The average cyclical amplitude of the regression series is approximately 14 percent smaller than the average amplitude of the components series. Thus, some of the apparent volatility of the prewar economy is simply an artifact of the problem that modern researchers have continued to use the components series despite Kuznets's belief that the regression series provides the better measure of pre-1918 cyclical fluctuations.

The new estimates of prewar GNP are much less volatile than either the Kuznets components or regression series. The average amplitude for the new GNP series for 1869-1918 is 31 percent smaller than that of the Kuznets/Gallman components series and 17 percent smaller than that of the Kuznets regression series. The average cyclical amplitude of the new prewar GNP series is also larger than that for the postwar Commerce Department series. The cyclical amplitude of the new prewar series is 33 percent larger than that of the postwar GNP series. From these results it is clear that the measured severity of cyclical movements in the new prewar GNP series is approximately midway between that of the prewar Kuznets/Gallman components series and the postwar Commerce Department series.

The results using different measures of volatility are roughly similar to those using the mean cyclical amplitude. I consider both the standard deviation of the percentage changes in GNP and the standard deviation of deviations from

trend. These two measures are more indexes of general volatility than of cyclical severity. Using these measures, the prewar Kuznets/Gallman components series is again roughly twice as volatile as the actual postwar GNP series. The Kuznets regression series is only slightly less volatile than the components series.[21] The new estimates of real prewar GNP are substantially less volatile than either the prewar components or regression series. The standard deviations of both percentage changes and deviations from trend for the new GNP series are approximately halfway between those for the prewar Kuznets/Gallman components series and those for the postwar GNP series.

One issue that arises in examining the standard deviations of the new GNP series involves the fact that the new GNP estimates are derived as the forecast values of a regression. As a result, one would expect the new prewar estimates to have a lower standard deviation than a true GNP series would have just simply because the variance of the residual has been suppressed. One way to gauge the importance of this effect is to construct a GNP series for the postwar era using the same methods as those used to create the new prewar estimates. Since this constructed postwar series will also be the fitted values of a regression, it too should have standard deviations that are biased downward. By comparing the standard deviations of the actual and constructed postwar series, one can estimate the magnitude of the possible bias.

The volatility properties of the constructed postwar GNP series are given in Table 4. A comparison of the standard deviations of both percentage changes and deviations from trend for the actual and constructed postwar series suggests that the bias resulting from using fitted values is very small. The standard deviation of percentage changes for 1955-1983 is only 4 percent smaller for the constructed postwar estimates than for the actual data. While the variance of the residual could certainly be larger in the prewar era, the simple postwar

comparison suggests that the standard deviation estimates of the new prewar GNP series are not very biased. As a result, these measures of volatility probably do provide a good indication of the amount of stabilization that has occurred over time.

2. Interpretation of Results

The comparisons of the various measures of volatility suggest that economists may have greatly overstated the degree to which the U.S. economy has stabilized over time. While the Kuznets/Gallman components series for 1869-1918 is by most measures approximately 69 percent more volatile than the Commerce Department GNP series for 1947-1983, the new GNP series for 1869-1918 is on average only 35 percent more volatile than the postwar series. This suggests that almost exactly half of the often-noted stabilization of gross national product over time disappears when the new estimates of prewar GNP are used in the calculation.

While it is crucial to note that much of the observed-stabilization of GNP is due to flaws in the original Kuznets data, it is also important to stress that even the new estimates of prewar GNP reveal some stabilization between the prewar and postwar eras. The business cycle before 1918 continues to look somewhat more severe than the cycle after 1947 even when more accurate estimates of prewar GNP are examined. However, the change over time now appears to have been mild rather than dramatic.

It is useful to note that some of the remaining evidence of stabilization may be due to the fact that in deriving the new estimates of prewar GNP, I have deliberately erred on the side of creating a series that is too volatile rather than too smooth. This tendency is most obvious in the use of a time-varying estimate of the relationship between GNP and commodity output. Despite the

evidence of Kuznets and Barger that the ratio of commodity output to GNP was roughly stable between 1869 and 1929, I allow the sensitivity estimate used to create new data to be higher in the prewar era than in the interwar and postwar eras.

Other steps in the derivation have also tended to overstate the volatility of the new pre-1918 GNP series. For example, I have not taken into account the possible measurement error in the prewar commodity output data. If early commodity output is measured with error as Weir suggests, then using a sensitivity estimate derived using predominantly good postwar data will overstate cyclical movements in prewar GNP. As a result, a comparison of the new prewar estimates and the postwar GNP series will, if anything, overstate the amount of stabilization of the business cycle that has occurred over time.

Given that the new prewar estimates of GNP reveal significantly less stabilization over time it is useful to try to understand the source of the results. The key fact to consider in this regard is that commodity output has stabilized only very slightly over time. Table 4 reports various measures of volatility for both the prewar Shaw series and the postwar Federal Reserve Board data on real commodity output. These measures show that commodity output was only 12 percent more volatile on average in the prewar era than in the postwar era.

Given the slight stabilization in commodity output, one might reasonably expect only a slight stabilization in GNP. Indeed, if the relationship between GNP and commodity output had been stable between 1869 and 1983, GNP would also have decreased in volatility by 12 percent. The new estimates of prewar GNP show a 35 percent stabilization precisely because the relationship between GNP and commodity output is allowed to have declined slightly over time as the noncommodity sectors of the economy have expanded. The key point is that most

of the stabilization shown by the new series stems not from the underlying commodity output data, but from assumptions and estimates about the relationship between commodity output and GNP.

The same is true to a much larger degree for the original Kuznets estimates. The dramatic stabilization shown by the Kuznets components and regression series is due almost entirely to Kuznets's assumption that GNP moved nearly one-for-one with commodity output from 1869-1918, while GNP only moved about 0.6-for-one with commodity output in the postwar era. Since the analysis of Section III suggests that this assumption is not accurate, neither are the estimates of the degree of stabilization that can be derived from the Kuznets series.

CONCLUSION

The comparisons of Section IV show that perceptions of the prewar business cycle are very different when the new GNP series is used in place of Kuznets's prewar estimates. Using the new GNP series business cycle fluctuations from 1869-1918 appear to be much less severe than is conventionally thought. As a result, one of the most often cited generalizations about business cycles, the dramatic damping of economic fluctuations, disappears when the new estimates of prewar GNP are considered. In its place is the much more modest stylized fact that the severity of business cycles has declined only slightly over time.

Given that the new prewar data presented in this paper offer such different conclusions about the prewar economy, it is important to evaluate the quality of the new series and to stress its limitations. In presenting the derivation of the new series in Section III, I stress the assumptions being made to derive the new estimates. The key assumption is that prewar GNP moves significantly less than one-for-one with commodity output. This assumption is justified on the basis of the estimated relationship between GNP and commodity output over the period 1909-1928 and 1955-1983. Even when one allows the measured sensitivity of GNP to commodity output to decline over time, one's best estimate is that cyclical movements in prewar GNP are only about 60-70 percent as large as those in commodity output. As a result, it seems likely that the key assumption underlying the new estimates is correct.

In addition to the evidence concerning the validity of my assumptions, a further piece of evidence about the cyclical accuracy of the new estimates of GNP is provided by my research on the historical industrial production and unemployment data (see Romer, 1986a, 1986b). In these studies, a variety of historical evidence is used to show that the assumptions underlying the industrial production and unemployment data are incorrect and that the true data should be substantially less volatile. If one believes that various output and employment series should show roughly similar cyclical movements, then the fact that the existing Kuznets series is more volatile than the new estimates of unemployment or industrial production provides evidence that the Kuznets series is excessively volatile. By the same reasoning, the fact that my new estimates of GNP show cyclical movements of approximately the magnitude the previous studies suggest is accurate for industrial production and unemployment provides evidence that the new GNP series is reasonably reliable.

While the analysis above stresses the accuracy of my new estimates of GNP for cyclical analysis, it is important to realize that the new series is still subject to many limitations. The most important of these is the fact that the new estimates of GNP are still derived entirely from data on commodity output. Since the necessary data on most noncommodity components of GNP do not exist, this limitation is essentially inevitable. However, this fact means that any movements in GNP not perfectly correlated with commodity output will be missed by the new GNP series as they are by the existing Kuznets series.

Another limitation of the new estimates is that no adjustments have been made to the trend level of GNP. The new estimates have also not been derived to be completely consistent in trend values with the Kendrick or Department of Commerce estimates. Therefore, my new estimates of GNP should not be used for analyzing long-run trends or in any application where actual levels are

important. The new estimates have been derived to represent cyclical movements more accurately than do existing estimates of GNP. It is thus in cyclical analyses that the new estimates of real gross national product will be both most accurate and most useful.

FOOTNOTES

*I would like to thank Robert Barro, J. Bradford DeLong, Rudiger Dornbusch, David Romer, Paul Romer, Peter Temin, and two anonymous referees for extremely helpful comments and suggestions. I am also grateful to the Social Science Research Council for financial support.

[1]Various measures of cyclical volatility are very similar for the Kuznets and Kendrick series. The standard deviation of percentage changes for the original Kuznets components series for 1889-1918 is 0.064. The same measure for the Kuznets/Kendrick series for 1889-1918 is 0.059. The standard deviation of deviations from trend for the Kuznets series for 1889-1918 is 0.052, while that for the Kuznets/Kendrick series is 0.047. Hence, the Kuznets components series is the main determinant of the cyclical movements in the Kuznets/Kendrick estimates of GNP.

[2]The standard deviation of percentage changes for the Kuznets components series for 1869-1918 is 0.063 while that for the Kuznets/Gallman series is 0.058. The standard deviation of deviations from trend for the Kuznets series for 1869-1918 is 0.055, while that for the Kuznets/Gallman series is 0.054.

[3]The Shaw/Kuznets series on commodity output in 1929 prices for 1869-1938 is published in Capital in the American Economy, Table R-21, pp. 553-554.

[4]Kuznets also derives three statistical variants of his components estimates of GNP. The only difference between the three variants is how the trend level of the flow of services to consumers is measured. Variant III derives the trend level of services by interpolating the trend from the Commerce Department series after 1929. Since this series is considered to be the variant most consistent in levels with the postwar GNP series, this variant of the Kuznets components series is used in the all of the following analysis.

[5]To estimate the sensitivity of the real flow of services to consumers to the real flow of commodities to consumers, Kuznets regresses the deviations from trend of the two series on one another. Trends are calculated by connecting overlapping decadal averages. Kuznets's results for 1919-1941 are:

$$SER_t - \overline{SER}_t = \underset{(.134)}{.304}(CF_t - \overline{CF}_t) + e_t$$

$$R^2 = .20$$

where SER is the level of the real flow of services to consumers, CF is the level of the real flow of commodities to consumers, and bars over variables denote trend values. (See Kuznets, 1961, notes to Table R-28, p. 571, for details.) The same regression for the period 1952-1976 yields estimates:

$$SER_t - \overline{SER}_t = \underset{(.139)}{-.154}(CF_t - \overline{CF}_t) + e_t$$

$$R^2 = .05$$

where data on SER and CF are from the National Income and Product Accounts and are measured in 1972 dollars.

[6]An example may help to clarify Kuznets's procedure. Kuznets calculates averages for decades that overlap by five years, say, for 1899-1908, 1904-1913, and 1908-1917. He then uses these decadal averages to represent the midpoint of the decade. In this example the midpoints would be 1903.5, 1908.5, and 1913.5. He then interpolates linearly between these values to form an annual trend series.

[7]Trend values of GNP and commodity output are calculated in the same way Kuznets calculates trends: by connecting midpoints of overlapping decadal averages.

[8]Kuznets's income-side estimates of GNP for 1909-1918 are described in detail in Capital in the American Economy, (1961, pp. 546-552).

[9]The gross value series is described in the Federal Reserve Board publications Industrial Production for 1971 and 1977. I amend both the gross-value and value-added indexes of final goods production to include the output of construction materials. I do this because Shaw's measure of commodity output includes construction materials production.

[10]The detrended Shaw series on real commodity output declined by 49.3 percent between 1929 and 1932. The decline in commodity output in the worst depression in the period 1869-1918, 1880-1885, was 17 percent. The Shaw series is detrended using a piecewise linear trend.

[11]The evidence from Kuznets and Barger suggests that the trend is almost certainly milder in the period 1869-1918 than in the postwar era. As a result, this procedure will overstate the sensitivity of GNP to commodity output in the prewar era.

[12]The FRB gross value series actually begins in 1954. But, since 1955 was chosen as a benchmark year, it is better to use the data beginning in 1955.

[13]To calculate trend values for years not between two benchmarks, the trend from the nearest two benchmarks is continued either forward or backward.

[14]The two exceptions to this pattern are 1918 which corresponds to a mild trough in production and 1937 which corresponds to a mild boom. 1918 is included because it seemed likely that the apparent decline in production is due to the unusually high level of wartime production in 1916 and 1917. As a result,

1918 appears to be a year when output may have been essentially on trend. 1937 is included because it appears to only be a boom in relation to the severe depression of the early 1930s. Given the relatively low level of output in 1937 it seems plausible that production in that year only reflects trend output rather than potential output.

[15]I have tested the effects of the choice of benchmark years on the estimated sensitivity of GNP to commodity output over the interwar period. In addition to constructing trend values using the years 1910, 1918, 1924, and 1937 as benchmark values, I construct alternative trend values using the years 1911, 1925, and 1936 as benchmark values. The alternative trend values are different from the original trend values in that GNP and commodity output are now substantially below trend during World War I and very much above trend in the late 1920s. Despite these changes in trend values, the estimated sensitivity of GNP to commodity output for the period 1909-1928 does not change substantially. The coefficient estimate using the original trend values is 0.649. The estimate using the alternative trend values is 0.733.

[16]The actual data that I use are presented in Friedman and Schwartz (1982, Table 4.8, pp. 122-129). Friedman and Schwartz compile a series on GNP that is essentially identical to Kuznets's components series except that it includes Gallman's unpublished adjustment of trend levels for certain components of GNP for the period before 1909. Friedman and Schwartz actually only present data on net national product. To form a series on real GNP I add Kuznets's unpublished data on capital consumption for 1869-1918.

[17]In estimating equation (2) the simple linear trend jumps at 1955 to take account of the missing observations. For reference, the trend series is equal to one in 1909.

[18]Baily uses the Kendrick estimates of GNP reported in Historical Statistics. DeLong and Summers use Gordon's quarterly data set, which, depending on the version used, derives its annual movements from either the Kendrick/Commerce Department series or the Kuznets/Gallman series presented in Friedman and Schwartz (See Gordon, 1982).

[19]The commodity output series is used to determine peaks and troughs because it is the one series that is nearly uniformly good over the prewar and postwar eras. In nearly all cases the cycles in GNP correspond perfectly with those in commodity output.

[20]The actual dates of prewar business cycles using my method of dating are, peak to trough, 1869-1871, 1873-1875, 1880-1885, 1887-1888, 1892-1894, 1895-1896, 1901-1902, 1903-1904, 1906-1908, and 1913-1914. The dates for postwar business cycles are 1947-1949, 1953-1954, 1956-1958, 1959-1961, 1966-1967, 1968-1971, 1973-1975, and 1979-1982.

[21]Some of the similarities between the components and the regression series may be due to inconsistencies between the two series. The components series includes Gallman's corrections to the trend of GNP, the regression series does not. One of Gallman's main corrections is to raise the 1869 level of GNP substantially. As a result, percentage changes and deviations from trend may appear quite large in the early 1870s in the uncorrected regression series. However, even when the early 1870s are excluded the Kuznets regression series has volatility properties very similar to those of the components series.

REFERENCES

Baily, Martin Neil. "Stabilization Policy and Private Economic Behavior." Brookings Papers on Economic Activity (1978:1): 11-50.

Barger, Harold. Distribution's Place in the American Economy since 1869. National Bureau of Economic Research No. 58, General Series. Princeton University Press, 1955.

DeLong, J. Bradford and Lawrence H. Summers. "The Changing Cyclical Variability of Economic Activity in the United States." National Bureau of Economic Research, Working Paper No. 1450, September 1984.

Friedman, Milton. "The Interpolation of Time Series by Related Series." American Statistical Association Journal 57 (December 1962): 729-757.

Friedman, Milton and Anna J. Schwartz. Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975. National Bureau of Economic Research. Chicago: University of Chicago Press, 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, Volume 30. New York: National Bureau of Economic Research, 1966, 3-76.

Gordon, Robert J. "Price Inertia and Policy Ineffectiveness in the United States, 1890-1980". Journal of Political Economy 90 (December 1982): 1087-1117.

Kendrick, John W. Productivity Trends in the United States. National Bureau of Economic Research, No. 71, General Series. Princeton: Princeton University Press, 1961.

Kuznets, Simon S. Capital in the American Economy: Its Formation and Financing. National Bureau of Economic Research. Princeton: Princeton University Press, 1961.

_____. National Income and Its Composition, 1919-1938. New York: National Bureau of Economic Research, 1941.

_____. National Product since 1869. New York: National Bureau of Economic Research, 1946.

Romer, Christina. "Spurious Volatility in Historical Unemployment Data." Journal of Political Economy, 94 (February 1986): 1-37. (a)

_____. "Is the Stabilization of the Postwar Economy a Figment of the Data?" American Economic Review 76 (June 1986), forthcoming. (b)

Schultze, Charles L. "Some Macro Foundations of Micro Theory." Brookings Papers on Economic Activity (1981:2): 521-576.

Shaw, William H. Value of Commodity Output since 1869. New York: National Bureau of Economic Research, 1947.

U.S. Board of Governors of the Federal Reserve System. Industrial Production. 1971 and 1976 Editions. Washington, D. C., 1972 and 1977.

U.S. Bureau of Economic Analysis. The National Income and Product Accounts of the United States, 1929-76. Washington, D. C., 1981.

U.S. President. Economic Report of the President, 1984. Washington, D. C., 1985.

Weir, David. "Stabilization Regained: A Reappraisal of the U.S. Macroeconomic Record, 1890-1980." Mimeograph, Yale University, 1985.