A NEW MONTHLY INDEX OF INDUSTRIAL PRODUCTION, 1884-1940

DATA APPENDIX

Jeffrey A. Miron

Christina D. Romer

November 1989

DATA APPENDIX

This appendix describes the sources, availability, and necessary adjustments for the thirteen individual production series that are used in the derivation of the Miron-Romer index. It also describes the procedures used to combine the individual series into an aggregate index of industrial production.

I. Component Series

As described in the text, the individual series that we include in the new index come from a variety of primary sources. Many of the series that we use, however, have previously been collected by researchers of the National Bureau of Economic Research. These series are available on a tape entitled "National Bureau of Economic Research: Macroeconomic Time Series for the United States, United Kingdom, Germany, and France" that is stored at the Inter-University Consortium for Political and Social Research (ICPSR 7644).

In cases where the series we use is available on the NBER tape, we use the information in the following way. We first compare the data on the tape to the original source documents. Whenever there is a discrepancy between the tape and the original source, we update the series to be consistent with the primary source. We do this because there appear to be numerous coding errors in the tape. The only time when we do not follow this rule is when there is an obvious typographical error in the original source.

After checking and updating the data on the tape, we then make any further adjustments that are necessary to the series. In some cases there are no adjustments to make. In other cases there may be a change in the units reported or the in the coverage of the series that must be dealt with.

In what follows we describe the coverage and nature of each series and the appropriateness of using a given series as a measure of production. We then discuss the primary sources of the series that we use and all of our transformations of the data on the NBER tape into the series that are ultimately used to form the new aggregate index. The name listed for each series is the mnemonic under which the series is stored in a RATS data file that is available on diskette from the authors.

A. Pig Iron Capacity

PICAPM
Gross Tons (2240 Pounds) Per Day, Monthly Average
Available 1884:6 - 1940:12

1. Basic Facts about the Series

This series shows the daily operating capacity of the average number of pig iron furnaces in blast in a given month. Since blast furnaces are not turned on and off frequently and are generally run at full capacity, this series should be closely correlated with the actual production of pig iron. Indeed, after

1902 when production data become available the correlation between production and capacity is nearly perfect. Furthermore, since pig iron is the primary input to the production of steel, this series should also provide a good proxy for basic steel production, though there may be a lag between the production of the two goods.

2. Original Source

The base data are from Iron Age, the major trade publication of the iron and steel industry, and are based on a comprehensive survey of iron and steel producers. The data available from Iron Age show the daily operating capacity of furnaces in blast on the first of the month. These data are usually broken down by the fuel used in the furnace: charcoal, anthracite, and coke. Since for some purposes the data on the capacity of charcoal furnaces are not reported, we restrict our data collection to the capacity of anthracite and coke furnaces. After 1930 the only data available are for the capacity of coke furnaces. Since this appears to be due to the fact that anthracite went out of use, we do not view this as a change in the coverage of the series and do not make any adjustment for it.

3. Transformations

This series is not available from the NBER tape and so was entered by hand. Whenever a capacity number for a given month was revised from one issue of Iron Age to another, we use the latest version as the true number.

To use the series we have as a proxy for monthly production, it is necessary to average the first day of the month observations to yield average daily capacity for the month. The formula that we use to do this is:

$$\text{Monthly average} = [1/(2*\text{Days}_t)] [(\text{Days}_t + 1)\text{Cap}_t + (\text{Days}_t - 1)\text{Cap}_{t+1}]$$

where Days is the number of days in the month and Cap is the daily capacity on the first of the month.

There are missing observations for 1918:2 and 1921:2 that cannot be eliminated because the data just are not available. To deal with this, we use the following procedure. We take the first of the month capacity observations for January and March and distribute the average of these two numbers over two months, using a formula analogous to that above. The actual formulas are:

January: 44/59*Capacity(Jan 1) + 15/59*Capacity(March 1)

February: 29/118*Capacity(Jan 1) + 89/118*Capacity(March 1).

B. Anthracite Coal Shipments

ACSHIP
1000s of Long Tons (2240 Pounds) Shipped Over the Month
Available 1884:1 - 1940:12, with 1924:1-1926:12 missing

1. Basic Facts about the Series

This series shows the shipments of anthracite coal over a given month. This series should be highly correlated with actual production because inventory holdings at the mines tended to be fairly small

in the period covered by our index.

2. Original Sources

The base data are from The Coal Trade (later called Saward's Annual), a major industry publication. The data, at least after 1905, appear to have been originally collected by the Anthracite Bureau of Information and are based on the reports of carrier companies. For some unknown reason no data were collected for the period 1924:1-1926:12. Beginning in 1933, the data are reported in net, rather than long tons.

3. Transformations

This series is available for 1884:1-1922:12 from the NBER tape (variable 212 of dataset 01A). We have checked the data on the tape against the original source. For the period 1884-1922 there are few discrepancies between the tape and the original source. In some years the original data change from one issue of The Coal Trade to another, presumably due to data revisions. As is appropriate, the NBER uses the later numbers. In 1903:12, 1913:11, 1915:4,8 and 1919:9 the numbers from The Coal Trade and the tape differ. However, it appears that there may be typos in the original source. We deduce this from the fact that the monthly numbers given in The Coal Trade do not add up to the annual figure while the NBER numbers do. This suggests that the NBER had additional information and that their numbers should be used.

For 1918-1922 we have data from both The Coal Trade and Saward's Annual. In 1918:1, 1920:3, 1922:9-10, and 1922:12 the numbers from the two sources differ slightly and the tape accords with The Coal Trade. We update the data on the tape to be consistent with Saward's Annual because after 1922 this is the only original source available. We also convert the data after 1933 from net tons to gross tons so that it is in the same units as the earlier data.

While the NBER codes the data for 1922:4-8 as missing because of the strike, we record coal shipments as zero in these months. A footnote in Saward's Annual states: "No shipments in April, May, June, July, and August, 1922, on account of strike." Furthermore, for this period we have actual production data and it is essentially equal to zero.

C. Crude Petroleum Production, Appalachian Region

APPAPETR
1000s of Barrels Per Day
Available 1884:1 - 1940:12

1. Basic Facts about the Series

This series shows the average daily production of crude petroleum in the Appalachian region in a given month. This region is composed of New York, Pennsylvania, Kentucky, West Virginia, and parts of Ohio. For the period 1884-1890, data are only available for New York and Pennsylvania. This represents only a slight change in coverage because New York and Pennsylvania account for over 90% of Appalachian production in this period. Both petroleum series are on a daily average basis, where daily average has been calculated as monthly production divided by the number of days in the month.

For the late 1800s Appalachian production represents nearly 100% of total U.S. petroleum

production (based on a comparison of available annual numbers). However, by World War I, Appalachian production accounts for only about 10% of total production. This suggests that the Appalachian series becomes a decidedly less good proxy for total production over time.

2. Original Sources

The base data are from Mineral Resources and Minerals Yearbook. These figures were compiled by the U.S. Geological Survey from monthly reports of pipe-line and other companies. They show the quantity of petroleum transported from producing properties. Beginning in 1919 there appears to have been some adjustment for stocks.

3. Transformations

This series is available from the NBER tape. It is presented as two series, variable 243 and variable 241 of dataset 01A. Variable 243 begins in 1894:1 and covers production in the entire Appalachian region and variable 241 extends back to 1884:1 and covers production only in New York and Pennsylvania.

Mineral Resources provides data on Appalachian production for 1890:1 - 1893:12 as well as the later data that appear on the tape. There is no evidence of any change in coverage or alteration of procedures. Therefore, we simply continue the tape series for Appalachian production with this series.

We have checked the data against the original source. Minor discrepancies in the New York and Pennsylvania series exist in 1884:2 and 1884:7. These discrepancies appear to be due to rounding errors. The data on the tape are updated to be consistent with the original source. Discrepancies in the Appalachian series exist in 1921:2 and 1924:1 and these are corrected to be consistent with the original source. In addition, there are many discrepancies between the tape and the original sources in 1935-36 and 1937 and 1939. The discrepancies are always very small. The likely source of the discrepancies for 1935-36 is that Minerals Yearbook only gives data by state and hence it is impossible to aggregate exactly in the way that is done earlier for the Appalachian Region. The discrepancies for 1937 and 1939 are most likely due to data revisions that we cannot document. Evidence in favor of this is that the numbers from Minerals Yearbook say subject to revision. Furthermore, one can see that there are substantial revisions for 1937 and 1939 because the annual numbers given for those years in 1938 and 1940 are quite different from the preliminary numbers. This is not true of the numbers for 1938 and 1940 for which our numbers are identical to the NBER tape. Because it seems likely that the tape has revised data, we accept the tape's numbers for 1935-1940.

While NY and PA production account for almost all of Appalachian production prior to 1890, it is sensible to connect the two series with a simple ratio splice to ensure that there is no jump in the combined series. The year chosen for the splice is 1890:1 and the splice was done in levels. The procedure resulted in multiplying the NY and PA series for 1884:1 - 1889:12 by (70.0/68.0).

D. Sugar Meltings at Four Ports

SUGM4
1000s of Long Tons (2240 Pounds) Melted over the Month
Available 1890:1 - 1937:12

1. Basic Facts about the Series

       This series shows the amount of raw sugar being made into refined sugar at the four main ports of sugar import. As such, it should provide an excellent proxy for the actual production of refined cane sugar.

2. Original Source

       The base data are from Willett and Gray's Weekly Statistical Sugar Trade Journal, the major trade publication of the sugar industry. For most of the late 19th and early 20th centuries the Weekly Statistical Sugar Trade Journal tracks only the meltings in four Atlantic ports: New York, Boston, Philadelphia, and Baltimore. Following World War I, the Journal also tracks meltings in eight or more ports. Because the four-port series continues through 1937, we opt to use this consistent series for the entire sample period. However, while for the early period it is clear that 4 ports covered most of the meltings, by 1940, they appear to account for only about 50 percent of all meltings. This means that our procedure may cause us to miss some of the expansion in the sugar industry in the interwar period.

3. Transformations

       This series is available from the NBER tape (variable 392 of dataset 01A) for the period 1890:1-1921:12.

       We expand the NBER series using monthly data from the original source to cover the period 1922:1-1937:12. The monthly data stop in 1937:12. Though there are weekly data on meltings at the Atlantic ports for 1938-1940, it is difficult to use these because no observations exist for the last week of each year. Neither carrying forward the previous week's growth rates or adopting the first week of January's growth rate seem to be appropriate. In particular, using either of these procedures for the eight-port series (for which both monthly and weekly data exist) yields a reading for December that is higher than the actual monthly observation.

       We have checked the data on the NBER tape against the original sources. Minor discrepancies exist for 1897:3,4, 1898:10,11 and 1919:12. The data on the tape are corrected to be consistent with the primary source.

E. Cattle Receipts at Chicago

CATREC
1000s of Head Received During the Month
Available 1884:1 - 1940:12

1. Basic Facts about the Series

   This series shows the head of cattle received in Chicago in a given month. Beginning in 1907 there exist actual data on cattle slaughtered. The receipt series appears to be highly correlated with slaughter, suggesting that cattle receipts are a good proxy for the production of dressed beef. The receipt series probably becomes a less good proxy for total slaughter over time because Chicago declines in importance.

2. Original Source

The base data are from the <u>Chicago Board of Trade Annual Report</u>. According to the source, the data show receipts at stock yards in Chicago and are derived from reports of the U.S. Department of Agriculture, terminal elevators, and transportation agencies. We obtained Xeroxes of the <u>Annual Reports</u> that were missing from Baker Library directly from the Chicago Board of Trade Records Center.

3. Transformations

This series is available from the NBER tape (variable 312 of dataset 01A). We have checked the data on the tape against the original source. Minor discrepancies were found in 1930:1, 1935:9, and 1940:9 and these were resolved in favor of the original source.

F. Hog Receipts at Chicago

HOGREC
1000s of Head Received During the Month
Available 1884:1 - 1940:12

1. Basic Facts about the Series

This series shows the number of hogs received in Chicago in a given month. Beginning in 1907 there exist actual data on hogs slaughtered. The receipt series appears to be highly correlated with slaughter, suggesting that receipts are a good proxy for slaughter. The receipt series probably becomes a less good proxy for total slaughter over time because Chicago declines in importance.

2. Original Source

The base data are from the <u>Chicago Board of Trade Annual Report</u>. According to the source, the data show receipts at stock yards in Chicago and are derived from reports of the U.S. Department of Agriculture, terminal elevators, and transportation agencies. We obtained Xeroxes of the <u>Annual Reports</u> that were missing from Baker Library directly from the Chicago Board of Trade Records Center.

3. Transformations

This series is available from the NBER tape (variable 306 of dataset 01A). We have checked the data on the tape against the original source. Discrepancies were found in many years: 1885:7, 1889:2,6, 1893:2, 1894:1,2,7, 1895:3,11, 1898:1, 1899:1, 1900:1,5, 1901:4, 1902:10 1903:5, 1906:9, 1908:6, 1909:3, 1912:1, 1913:12, 1915:5,12, 1918:6, 1921:3, 1922:10, 1923:9, 1926:4, 1929:5, 1932:9, 1933:6, 1934:6. These appear to be due to rounding errors and some very large typographical errors in the NBER series. We resolve all discrepancies in favor of the original source.

G. Connellsville Coke Shipments

COKESHIP
Net Tons (2000 Pounds) Shipped during the Month
Available 1894:1 - 1935:12

1. **Basic Facts about the Series**

This series shows the tons of coke shipped from the Connellsville region of Pennsylvania. This series should provide a good indicator of total coke shipments in the U.S. because the Connellsville region accounted for at least 50 percent of total production in the pre-1940 period. Furthermore, because beehive coke was ordinarily not stored at the ovens in any great quantity, shipments should provide a reasonable proxy for actual production.

2. **Original Sources**

The base data are from Mineral Resources and Minerals Yearbook. For the period before 1928 the data in these sources are taken directly from the Connellsville Courier; for the period 1929-1935 the data are compiled by the Bureau of Mines from the weekly reports of the Courier. No data on Connellsville shipments are available before 1894 or after 1935.

3. **Transformations**

This series is not available from the NBER tape and so was entered by hand directly from the original source. No adjustments needed to be made to the base data.

H. **Minneapolis Flour Shipments**

FLRSHIP
1000S of Barrels Shipped from Minneapolis During the Month
Available 1884:1 - 1940:12

1. **Basic Facts about the Series**

This series shows the shipments of flour out of Minneapolis. Though many of the major mills were based in Minneapolis, it appears that on an annual basis, shipments out of Minneapolis accounted for only 10-20% of total wheat flour production. However, the annual total production data on which this calculation is based are quite imprecise. Furthermore, since flour milling is to some extent affected by agricultural production, one could reasonably expect Minneapolis shipments to be similar to total shipments. Indeed, after World War I monthly data on total production become available and the correlation between this series and Minneapolis shipments is quite high.

2. **Original Source**

The base data are from the Annual Report of the Chamber of Commerce and Board of Trade of the City of Minneapolis. We obtained Xerox copies of the Annual Reports from the Minneapolis Chamber of Commerce.

3. **Transformations**

This series is available from the NBER tapes (variable 379 of dataset 01A). We have checked the data on the tape against the original data. Several discrepancies were found. Those for 1885:3, 1886:9-11, 1887:9, 1889:7-12, 1902:5, 1905:12, 1907:9, 1920:9 and 1937:12 were resolved in favor of the original source. Most of these errors appear to be due to the fact that the NBER took the data from a summary volume in 1899 that had many errors. We have used the yearly reports instead. For 1920:6 we

use the NBER number because it is in accord with the weekly and annual data in the Annual Report, whereas the original monthly number is not. Parts of 1886 and 1905 could not be checked because the original sources were missing or failed to report the necessary numbers.

I. Wool Receipts at Boston

WOOLR
Millions of Pounds Received in Boston (Domestic and Foreign) per Month
Available 1888:1 - 1940:12

1. Basic Facts about the Series

This series shows the amount of raw wool coming into Boston from both domestic and foreign producers. As late as the 1930s receipts at Boston accounted for around one half of the total domestic wool clip and between 20 and 40 percent of foreign imports. Provided that inventories of raw wool are not large and do not have strong cyclical properties, this series should provide a reasonably good proxy for the output of woolen cloth.

2. Original Sources

The base data for 1888-1918 are from the Boston Chamber of Commerce Annual Report. For the period after 1918 the data are from the Survey of Current Business. The data in the Survey are described as being from the U.S. Department of Agriculture, but the U.S.D.A. compiled its data from the Boston Commercial Bulletin and the records of the Boston Grain and Flour Exchange. Thus, the series from the two sources appear to be based on similar underlying reports.

3. Transformations

Data on wool receipts are available from the NBER tape. Variables 163 and 165 of dataset 01A show the receipts of domestic wool in Boston per month. The two series differ in that the earlier series is measured in bales and the later series is measured in pounds. Variables 167 and 169 of dataset 01A show the receipts of foreign wool in Boston per month. Again the two series differ in the units in which wool is measured.

We have checked the data on the tape against the original sources. There were no major discrepancies, but there were many apparent rounding errors. We have corrected all of these discrepancies so that the various series are consistent with the original sources.

The fact that the units of measurement for both the domestic and foreign receipts series changed from bales to pounds in 1900 should not present a problem because bales were of a nearly standard weight and because there is some overlap in the series. Thus, we combine the two series for each type of receipts by ratio splicing the series in bales to the series in pounds in 1900:1. This amounts to multiplying the early domestic series (variable 163) for 1888:1-1899:12 by 7.26/29.03 and the early foreign series (variable 167) for 1888:1-1899:12 by 5.23/10.47.

We aggregate the foreign and domestic series into a series on total wool receipts. Since the domestic and foreign series are in the same units, a simple sum is appropriate.

J. Coffee Imports

COFFEE
Millions of Pounds Per Month
Available 1884:1 - 1940:12

1. Basic Facts about the Series

This series shows the amount of raw coffee imported into the U.S. Because the U.S. does not grow coffee domestically, coffee imports should provide a comprehensive measure of the inputs to the coffee roasting and grinding industry.

2. Original Sources

The data on imports are from the Monthly Summary of the Foreign Commerce of the U.S. that is published by the Bureau of the Census. This source had different titles and different publishing agencies in the period before 1914. Among the alternative titles are Monthly Reports on the Commerce and Navigation of the U.S., Monthly Summary of Commerce and Finance of the U.S., and Monthly Summary of the Imports and Exports of the U.S. Libraries typically bind these various publications into one volume for the entire period that is labeled the Monthly Summary of the Foreign Commerce of the U.S. and there is no apparent change in coverage or the procedures used over time. For simplicity we refer to the original source in the following discussion as simply the Monthly Summary of the Foreign Commerce.

The only problem we have noted with the base data is that because of tariff changes, the observations for 1913:9 and 10 and 1922:9 and 10 may be erroneous. Specifically, imports for the first three days of 1913:10 are included in the September 1913 figure and imports for the last 8 days of 1922:9 are included in the October 1922 figure.

We do not to try to correct these observations because simple prorating of the data seems likely to yield more erroneous observations. The Underwood Tariff of 1913 lowered average rates on October 4th. Hence, one would not expect much importation in the first 3 days of October when tariffs were still high. Thus, it is of no consequence that these observations are included in the September figure. Similarly, in 1919 the Fordney-McCumber Tariff raised average rates on September 22nd. In this case, one would not expect much importation in the last days of September. Therefore, it does not matter that those days are included in October.

3. Transformations

This series is available from the NBER tape (variable 93 of dataset 07A). We have checked the NBER data against the original source. (Note: We have not checked the data for 1885:7-1886:6, 1887:7-1888:6, and 1939:7 because we were unable to obtain the source documents.) There were minor discrepancies between the NBER series and the Monthly Summary in 1903:2, 1906:9, 1909:4, 1915:8, 1932:2, and 1936:3. These observations were changed to be consistent with the original source. No further corrections are needed for the series.

K. Tin Imports

TIN
Long Tons (2240 Pounds) Per Month
Available 1884:1 - 1940:12

1. Basic Facts about the Series

This series shows the amount of bars, blocks, pigs, grain, and granulated tin imported into the U.S. as well as the tin content of imported ore. It is used to proxy for the output of refined tin and tin products produced in the U.S.

2. Original Sources

The data on imports are from the Monthly Summary of the Foreign Commerce of the U.S. See the original sources section for the coffee import series for a more thorough description of this source.

One problem noted by the NBER was that the first 3 days of October 1913 were included with the September number. We do not do anything about this because average tariffs were dropped on October 4, 1913. Hence, one would not expect many imports during the first 3 days of October.

3. Transformations

This series is on the NBER tape (variable 107 of dataset 07A). We have checked the data on the tape against the original source. (Note: We have not checked the data for 1885:7-1886:6, 1887:7-1888:6, and 1939:7 because we were unable to obtain the source documents.) To check the data on the tape, the data from the Monthly Summary of the Foreign Commerce had to be treated in the following way. The data in pounds was first rounded to the nearest 1000 and then divided by 2.24 to yield long tons. This number was then rounded to the nearest integer. This way of dealing with the data yielded a series that matched the NBER in most years. There were minor discrepancies in 1901:6, 1901:11, 1912:4, 1921:11, 1923:6, 1926:4, 1928:5, 1929:10, 1930:2, 1931:1,4,8, 1933:1,8, 1934:8,11, 1938:6, 1939:4,8,and 9. These were corrected to be consistent with the original source. The NBER codes the observation for 1893:7 as missing. However, the data from the Monthly Summary indicate that imports simply rounded to zero in this month. Consequently, we have replaced the missing value code with a zero for this observation.

L. Crude Rubber Imports

RUBBER
Millions of Pounds
Available 1884:1 - 1940:12

1. Basic Facts about the Series

This series shows the amount of crude rubber imported. It is used to proxy for the output of rubber goods such as tires and rubber shoes produced in the U.S.

2. Original Sources

The data on imports are from the Monthly Summary of the Foreign Commerce of the U.S. See the original sources section for the coffee import series for a more thorough description of this source.

The only problem with this series is that the Monthly Summary tracks somewhat different products in different periods. After 1910:7 the Monthly Summary numbers cover the imports of India rubber. For the period 1890:10-1912:12, the Monthly Summary data include both India rubber and guayule gum. For the period 1884:1-1892:12, the Monthly Summary data include India rubber, guayule gum, and gutta percha. A comparison of the various series in the short periods of overlap suggests that the behavior of the different series is very similar. Thus, it seems reasonable to join the series together to form a single rubber imports series. However, some correction for changes in levels is necessary because imports of guayule gum are nontrivial.

3. Transformations

The rubber imports data are available from the NBER tape. The tape includes three series: variable 101 of dataset 07A covers the period through 1892:12 and includes data on India rubber, gutta percha, and guayule gum; variable 103 of dataset 07A covers the period 1890:10-1912:12 and includes data on India rubber and guayule gum; and variable 105 of dataset 07A covers the period 1911:1-1940:12 and includes data only on India rubber.

The Monthly Summary of the Foreign Commerce contains the data necessary to carry the series covering only India rubber (variable 105) back to 1910:7. We update this series with these additional observations.

We have checked the data against the original source. (Note: We have not checked the data for 1885:7-1886:6, 1887:7-1888:6, and 1939:7 because we were unable to obtain the source documents.) Minor discrepancies were found in 1896:4, 1900:7, 1926:5, 1932:8,9,11, 1935:3,4,8,11, 1936:8, 1937:8, and 1940,8,10. These observations were updated to be consistent with the original source.

The three different variants of the rubber import series are spliced together so that there are no jumps in the series. We choose to leave the longest series (variable 105) unspliced and to splice the other series to it. Specifically, we use the corrected and expanded variable 105 without further adjustment for 1910:7-1940:12. For 1890:10-1910:6 we splice variable 103 to variable 105 by multiplying variable 103 by the ratio of variable 105(1910:7) to variable 103(1910:7). For 1884:1-1890:9 we splice variable 101 to the already spliced series. That is, we multiply variable 101 by the ratio of the already spliced series (1890:10) to variable 101(1890:10).

M. Raw Silk Imports

SILK
1000S of Pounds per Month
Available 1884:1 - 1940:12

1. Basic Facts about the Series

This series shows the amount of raw silk imported into the U.S. Since domestic silk production is minimal, this series should provide a comprehensive measure of the raw materials input to the

production of silk cloth.

2. Original Sources

The data on imports are from the Monthly Summary of the Foreign Commerce of the U.S. See the original sources section for the coffee import series for a more thorough description of this source.

The only issue surrounding this series involves a slight change in coverage in the early 1920s. For the period up to 1924:12 the data available on raw silk imports include re-exports. Beginning in 1919:1 data excluding re-exports are available. This change is fairly minor because re-exports are very small, and because the two series behave very similarly in the five-year period when both exist. Nevertheless, it is sensible to ratio splice the two series together to prevent a slight discontinuity in the final data.

3. Transformations

The silk imports data are available from the NBER tape. Variable 131 of dataset 07A covers the period 1884:1-1924:12 and includes re-exports. Variable 133 of dataset 07A covers the period 1919:1-1940:2 and excludes re-exports.

We have used the Monthly Summary to extend variable 133 through 1940:12. In doing this we have followed the NBER procedures: we first round all the components to 1000s, then add raw silk imports and waste (the numbers on cocoons are not included) and then subtract off re-exports.

We have checked the data on the tape against the original source. Many minor discrepancies were found and the data from the tape were corrected to be consistent with the original source.

To deal with the minor change in coverage we use a ratio splice in 1924:12. We choose to leave the longer series (variable 131) unspliced and to splice variable 131 to it. Specifically, variable 131 for 1884:1-1924:12 was left unadjusted and variable 133 for 1925:1-1940:12 was multiplied by the ratio of variable 131(1924:12) to variable 133(1924:12).

II. Procedures Used to Form the Index of Industrial Production

The aggregation of the thirteen series we have collected into an index of industrial production involves several additional steps. These steps include adjustment for production days, conversion of each series to an index, and combination of the individual indexes using value-added weights.

A. Adjustment for Production Days

With the exception of the series on petroleum production (APPAPETR) and pig iron capacity (PICAPM), all of our series show total output over the month. As a result, the series show predictable upticks in long months and downticks in short months. To remove this effect it is desirable to generate a series that is on a per production day basis. This is consistent with what is done today in the derivation of the Federal Reserve Board's seasonally-unadjusted Index of Industrial Production.

To make this adjustment we do the following. First, no correction is needed for the pig iron

series. This is true because PICAPM shows the daily capacity of the average number of furnaces in blast during the month. Thus, it is already on a per day basis. Furthermore, because blast furnaces were not turned off for weekends, no adjustment is needed for the difference between calendar days and production days.

Second, the petroleum series in the original source is calculated as the monthly production of petroleum divided by the number of days in the month. To put it on a production day basis we first multiply the base series by the number of calendar days in the month and then divide by the number of production days.

All other series are put on a production day basis by dividing by the number of production days. The number of production days is calculated as the number of calendar days minus the number of Sundays. We do not exclude Saturdays because it appears that most industries worked six days a week until at least the mid-1930s.

B. Forming Individual Indexes

As described below, we use 1909 as the base year for our index. That is, 1909 is the year that is set equal to 100 for each individual index and for the aggregate index. We form the individual indexes by dividing the monthly observations for each series (adjusted for production days) by the average monthly value of that series in 1909.

C. Value-Added Weights

To combine the individual indexes into an aggregate index of industrial production we use value-added weights. Value-added weights are used because we want the index to represent the output of manufacturing industries. Thus, it should be irrelevant whether the inputs were expensive or cheap.

The original data on value added are from the Census of Manufactures or the Census of Mines and Quarries. We use the version of these data given in Solomon Fabricant, The Output of Manufacturing Industries, 1899-1937 (1940, Appendix C) and Historical Statistics (1975). For most categories Fabricant is identical to the Census.

1909 is chosen as the base year because it is the Census year nearest the middle of our sample period and because it is an ordinary year. Unlike 1904 and 1914 which were recession years, 1909 represents a point of mid-expansion in the cycle. Using a base year near the center of the sample period is desirable because it limits the importance of the compositional changes over time. Using a year of mid-expansion ensures that the differential effects of depressions will not affect the weights for the entire sample period.

The allocation of value added to various series is fairly straight-forward. For most of our series there is an obvious analog in the Census. For example, for our series on flour shipments from Minneapolis, we use the value added in flour milling from the Census. Following this procedure, we end up weighting our series by the value added in approximately the corresponding two- or three-digit SIC industry. The following list shows the category from Fabricant used for each of our series. It also discusses any special adjustments that needed to be made to the Fabricant data.

1. Pig Iron Capacity. We assign this series the value added in blast furnace products and steel mill products. We include the value added of steel mills because pig iron capacity is a good proxy for steel production and because the value added in pig iron production alone does not adequately reflect the importance of this series. It is useful to note that the steel mill products category only includes very crude steel goods; it does not include highly fabricated goods such as machines.

2. Anthracite Coal Shipments. Historical Statistics gives the value added in coal mining. We divide this between anthracite and bituminous coal using the data on the gross value of anthracite and bituminous coal output also given in Historical Statistics. This procedure indicates that anthracite should be given 26.9% of the total value added in coal production.

3. Crude Petroleum Production, Appalachian Region. We assign this series the value added in petroleum refining, lubricants not elsewhere mentioned, and oils not elsewhere classified.

4. Sugar Meltings in Four Ports. We use the value added in cane sugar refining and cane sugar, not elsewhere mentioned.

5. Cattle Receipts at Chicago. Fabricant lists the value added in meat slaughtering. Even the original Census documents do not divide this value added among the various types of animals. To derive separate weights for cattle and hogs, we allocate the total value added in meat packing according to the gross value of the beef and pork (fresh and cured) produced in 1909. These data are available from the Census of Manufactures for 1914. Based on the gross value of the output, the total value added in meat packing is divided by giving 41.1% to cattle and 58.9% to hogs.

6. Hog Receipts at Chicago. See the description of cattle receipts.

7. Connellsville Coke Shipments. We assign this series the value added of coke oven products.

8. Minneapolis Flour Shipments. We assign this series the value added in the flour industry.

9. Wool Receipts at Boston. We assign this series the value added in woolen goods, worsted goods, wool pulling, and wool scouring. We chose these series because they seemed related to very basic wool cloth, not the more finished goods like clothing and carpet also listed in Fabricant.

10. Coffee imports. We use the value added in coffee roasting and spices. This series is listed on p. 637 of Fabricant (1940).

11. Tin Imports. We assign this series the value added of tin and other foils, tin cans, tinware not elsewhere classified, and tin and terne plate.

12. Crude Rubber Imports. We assign this series the value added of all rubber products. This seems appropriate because the only subcategories listed are rubber shoes and tires and tubes. There is no category such as basic rubber.

13. Raw Silk Imports. Fabricant and the Census only give the value added of silk and rayon goods combined. To isolate the part appropriate to silk, we use Fabricant's data on the gross value of silk and rayon products (from Appendix B). We divide the products according to fabric content; those products for which fabric content were unclear were simply excluded from the calculations. This procedure may understate the importance of silk in the earlier period because a larger than average amount of velvet and

upholstery material may have been made of silk. This procedure suggests that 74.9% of the value added in silk and rayon is safely attributable to silk.

The value-added figures assigned to each series are used to derive weights that sum to one. The base data and the weights are given in Table A1. These weights are then used to combine the individual indexes into an aggregate index.

D. Adjustment for Missing Values

While many of our series exist from 1884-1940, there is a certain amount of attrition at both ends of the index. There is also some attrition in the middle because of missing data on coal shipments in the 1920s. The pattern of available series is the following:

1884:1-1884:5	Sugar, Wool, Coke and Pig Iron are missing.
1884:6-1887:12	Sugar, Wool, and Coke are missing.
1888:1-1889:12	Sugar and Coke are missing.
1890:1-1893:12	Coke is missing.
1894:1-1922:3	All series are available.
1924:1-1926:12	Coal is missing.
1927:1-1935:12	All series are available.
1936:1-1937:12	Coke is missing.
1938:1-1940:12	Coke and Sugar are missing.

To deal with missing series we do the following. First, we recalculate the percentage weights, setting to zero the weight for the missing series. Second, to prevent discontinuous jumps in the series, we ratio splice the series whenever there is a change in coverage. Specifically, we take the index for 1894:1-1923:12, for which all the series exist, as the base series. Then, on either end, we ratio splice the less complete series to it. For example, for 1890:1-1893:12 no data on coke shipments exist. Therefore, we derive new weights and construct this less complete series through 1894:1. Then we multiply this less complete series by the ratio of the complete series in 1894:1 to the less complete series in 1894:1. This procedure is done repeatedly, working out chronologically from the complete series.

The one complication to this procedure arises for coal for which there is a gap in the data in 1924:1-1926:12. Since we have data for all series for several years after 1926, it seems reasonable to use this complete series unspliced. Therefore, what we do is ratio splice the less complete series to the complete series in 1923:12, but then use no splice in 1926:12. This should be fine, unless the trend of coal production over the missing years is very different from that of other series.

E. Adjustment for Tariff Effects on Wool Receipts

For most purposes, it is desirable to use our index in its least adulterated form. However, our index has a dramatic spike in mid-1897 that is due to changes in the tariff on the wool receipts that we use to proxy for wool textile production. This spike is sufficiently peculiar that it seems reasonable to present an alternative index for part of this year. To do this, we do not use the wool receipts series for the six months around the definite spike in wool receipts (1897:2-1897:7). In making this adjustment, we ratio splice the series excluding wool receipts to the complete series in 1897:8 but then use the complete series before 1897:2 without a splice. This ensures that the alternative only changes the observations for 1897:2-7, not all earlier observations.

F. Final Indexes

INDXP
Miron-Romer Index of Industrial Production, 1909=100
Available 1884:1 - 1940:12

INDXPW
Miron-Romer Index with Adjustment for Wool Receipts, 1909=100
Available 1884:1 - 1940:12

TABLE A1

VALUE-ADDED WEIGHTS

Series	Value Added in 1909 Millions of Current $	Weights
Pig Iron Capacity	399.00	.3191
Anthracite Coal Shipments	124.05	.0992
Crude Petroleum Production	48.10	.0385
Sugar Meltings	31.60	.0253
Cattle Receipts	67.73	.0542
Live Hog Receipts	97.07	.0776
Coke Shipments	31.70	.0254
Flour Shipments	116.00	.0928
Wool Receipts	148.60	.1189
Coffee Imports	27.30	.0218
Tin Imports	26.70	.0213
Rubber Imports	74.60	.0597
Silk Imports	57.76	.0462

Sources: See text for a description of the sources of data on value added in 1909.