EMPLOYMENT OUTLOOK: 2010-2020

Layout and Description

For 195-Order Input-Output Tables:

1993 through 2010 Historical and Projected 2020

Prepared in the
BUREAU of LABOR STATISTICS
OFFICE of OCCUPATIONAL STATISTICS and
EMPLOYMENT PROJECTIONS

February 2012

Richard W. Graham graham.richard@bls.gov **(202) 691-5692**

CONTENTS

I. I-O Layout and Description

Introduction

This document describes input-output, final demand, and value added data developed by the Bureau of Labor Statistics in the Office of Occupational Statistics and Employment Projections. For a complete discussion of the projections presented here, please refer to Employment Outlook: 2010-2020, a series of five articles that appeared in the January 2012 **Monthly Labor Review (MLR)**.

For the 2020 projections, input-output, final demand, and value added data were developed for the years 1993 through 2010 and projected year 2020. Historical tables are provided in both nominal (current) dollars and in 2005 chain weighted real dollars. The projected tables are provided in real dollars only.

In addition to the data tables, sectoring plan files "sect_industry_stubs_x195.xls" and "sect_final_demand_x190.xls" are needed because the data files have no labels or codes to signify columns and rows. The "sect_industry" file provide our industry sectoring plan with industry codes and titles. The "sect_final_demand" file provide our final demand column codes and titles.

In order to determine the coordinates for each data cell, one must rely on the software of a spreadsheet to determine row and column sector. For example, Excel has an option that allows one to change column letters to column numbers.

The matrix tables described here are in tab-delimited ASCII (text) files, which can be readily converted to spreadsheet use. Matrices include only data. As stated above, **there are no row or column labels**. Each row contains all the columns for that row, including zeroes, separated by tab characters. An ASCII carriage-return character denotes the end of a given row of data.

Dollar value matrices are expressed in millions of dollars rounded to three decimal places. Therefore, they may not add exactly to their totals due to rounding error. In the real tables, the data do not add up to published totals like gross domestic product because of chain weighting.

These data are based on the 2007 North American Industrial Classification System (NAICS) and the U.S. Department of Commerce's Bureau of Economic Analysis (BEA) 1997 and 2002 benchmark input-output tables and the BEA Annual input-output tables for 1998-2009 as published as part of the July 31, 2009 Comprehensive revisions. The BLS industry sectoring plan is shown in the industry sectoring plan files discussed above.

Finally, please be careful when using the intermediate portion of the input-output tables with detailed final demand (FD) columns from the "FD" matrices, discussed in Section B. Intermediate data are in producers' value. However, most detailed final demand columns are in purchasers' values; the imports of goods column is in foreign port values. As explained in Section B, reallocation and adjustment columns allow final demand rows to add to producers' and domestic port values. Aggregate columns in the "FDAGG" matrices are in producers' values, consistent with data in other input-output tables.

A. Naming conventions for input-output tables

All of the file names begin with either “NOMINAL_” for nominal (current) dollars, “REAL_” for chain weighted real dollars, or “PROJECT_” for projections 2020. It should be noted that all input-output projections files only contain real dollar data. No current dollar input-output tables were projected.

Following “NOMINAL_”, “REAL_”, or “PROJECT_” are either “USE”, “MAKE”, “FD”, “FDAGG”, or “OUTPUT”.

The "USE" matrix contains the sales of commodities sold to intermediate consumers and final demand. In addition, it contains the intermediate inputs and value added factors of production to industries for the production of their product.

Each column sums to its respective industry output. Each row sums to its respective commodity output.

The "MAKE" matrix details the production of commodities by industries. Each row sums to industry output and each column sums to commodity output.

The "FD" matrix is a detailed set of 190 final demand types. Each of the 190 columns is distributed across the 195 input-output commodity rows identified in “sect_industry_x195.xls” as mentioned in page 1. The "FDAGG" is the "FD" matrix collapsed from 190 columns to 12 columns. The column sectors are shown in “sect_final_demand_x190.xls” as mentioned in page 1.

The “OUTPUT_COM” and “OUTPUT_IND” matrices detail the respective historical commodity and industry outputs by each of the 195 sectors identified in the input-output system for the years 1993 through 2010. The time series starts with 1993 outputs in the first column, the next year’s outputs in the second column, and continues through 2010 outputs in the last column.

For the projection year, the “OUTPUT_2020” matrix contains two columns. The first column details commodity output for each sector. The second column details the industry output for each sector.

All of the file names are followed by “_”, the data year, and the extension “.DAT” which means that the file is a data file.

B. Standard matrix types

"USE" Matrix (196x196)

This matrix contains intermediate inter-industry inputs plus value added (row 196) for the year specified in the matrix name. This matrix also contains intermediate inter-industry sales plus final demand (column 196) for the year specified in the matrix name. Each column sums to its respective industry output. Each row sums to its respective commodity output.

"MAKE" Matrix (195x195)

This matrix contains the production of commodities by industries for the year specified in the matrix name. Each row sums to industry output and each column sums to commodity output.

"OUTPUT_IND" Matrix (195x18)

This matrix contains the industry output time series for the 195 industry sectors for the years starting from 1993 (column 1) and continuing through 2010 (column 18).

"OUTPUT_COM" Matrix (195x18)

This matrix contains the commodity output time series for the 195 commodity sectors for the years starting from 1993 (column 1) continuing through 2010 (column 18).

"PROJECT" Matrix (195x2)

This table contains the total outputs for the projected year 2020. The first column is the total commodity output and the second is the industry output for the sector designated by the row number.

"FD" Matrix (195x190)

This is the final demand matrix for the year specified in the title. Each "FD" matrix has 195 rows and 190 columns. Rows represent commodity output. Columns represent final demand by detailed category, margin reallocations, or import valuation adjustments. The column types are described below.

Each "FD" matrix has 116 columns containing final demand data by detailed category. Their column sums are obtained from BEA's National Income Product Accounts (NIPA) and are sub-aggregates of GDP[1]. Using distributional data from BEA's benchmark Input-Output (IO) tables, the data in each column are allocated to the 195 commodity rows. Summing these data by row yields commodity output on a purchasers' value basis.

The composition of purchasers' value data differs among aggregate categories of final demand. There are 12 such categories (see "FDAGG" below). In 10 of them, the purchasers' value data may include margin costs, which represent the trade and transportation costs incurred in bringing commodities to market. In the import goods category, purchasers' value data may include import valuation adjustments, which reconcile different accounting methods used by BEA to

[1] Chain weighted aggregation methods must be used for tables containing values in chain weighted dollars.

classify goods imports[2]. Data in the remaining category, import services, include neither margin costs nor import valuation adjustments.

Commodity output must be transformed from purchasers' value to producers' value. Producers' value excludes margin costs and import valuation adjustments, which are removed in the margin reallocation and import valuation adjustment columns, respectively. These columns and their functions are described below. The transformation is necessary, because data in all other input-output tables are on a producers' value basis. The "FD" matrix row sums must yield data on a consistent basis to carry out the projections.

There are 70 margin reallocation columns in each "FD" matrix. These are associated with 10 aggregate final demand categories, and grouped seven (7) columns per category. Within each category, the seven (7) columns are used to reallocate margin costs for two (2) trade and five (5) transportation commodities, one column per commodity. The function of the columns is to remove the margin costs from purchasers' value data, and reallocate them to these trade and transportation commodities.

To do so, margin rates[3] are calculated for each cell within the columns. Each rate is multiplied by the purchasers' value within its row and category. The resulting products appear as negative entries. Within each column, the sum of these entries is exactly offset by a positive value inserted at the intersection with the corresponding commodity row, such that the column sum is zero. This offsetting positive value represents that commodity's margin value for that final demand category.

There are four (4) import valuation adjustment columns, all in the goods imports category. The columns are used to calculate the import valuation adjustment for the following commodities: air transportation, couriers and messengers, water transportation, wholesale trade, and insurance carriers. The first two commodities are adjusted in a single column; the remaining are each adjusted in their own columns.

To carry out the adjustments, import valuation adjustment rates[4] are calculated for each cell within the four (4) columns. Each rate is multiplied by the purchasers' value within its row in the goods import category. The resulting products appear as negative entries. Within each column, the sum of these entries is exactly offset by one or two positive values[5] inserted at the intersection with the corresponding commodity row(s). Each offsetting positive value represents that commodity's import goods valuation adjustment.

"FDAGG" Matrix (195x12)

---

[2] Import valuation adjustments reconcile differences arising from the inconsistent valuation of purchasers' value for import goods: NIPA accounting uses foreign port value, whereas input-output accounting uses domestic port value.

[3] Margin rates are based on data from BEA's benchmark IO tables.

[4] Import valuation adjustment rates are calculated from detailed trade data.

[5] The air transportation column offset is split between the air transportation and the couriers and messengers commodities.

This matrix is the collapsed version of the "FD" matrix above for the year specified in the matrix name. The 190 categories are aggregated to 12 general categories as shown in either of the files "sect_final_demand_x190.xls".

1 Personal consumption expenditures
2 Private investment in equipment and software
3 Private investment in nonresidential structures
4 Private investment in residential structures
5 Change in private inventories
6 Exports of goods
7 Exports of services
8 Imports of goods
9 Imports of services
10 Federal Government defense consumption and investment
11 Federal Government non-defense consumption and investment
12 State and local government consumption and investment

II. How to Access Files

Data files:

All files are compressed using WINZIP version 8.1. They are not self-extracting files.

Hence, one must save the ZIP file first. Then, use an unzip program to unzip it. The ZIP files named below are not self-extracting. Clicking the name will not open any of these ZIP files.

The file "ionom.zip" contains nominal (current) dollar historical I-O/FD data. The file "ioreal.zip" contains chain weighted real 2005 dollar historical I-O/FD data. The file "io2020.zip" contains chain weighted real 2005 dollar I-O/FD projections. No current dollar I-O table projections were made.

The data files contained in these ZIP files are described in Part I. The data are in tab-delimited ASCII (straight text) files, which are readily converted to spreadsheet use. These matrices contain only data.

The data files contain no row or column labels. Each row contains all the columns for that row, including zeroes, separated by tab characters. An ASCII carriage-return character denotes the end of a given row of data.

To determine the coordinates for each cell, rely on the software of a spreadsheet to determine row and column labels. For example, Excel has an option that allows you to change column letters to column numbers. One way to do this in Excel is to do the following:

1. Click the "Tool" option on the main menu on the top
2. Select "Option".
3. Click the "General" packet or stub
4. On the top left hand corner, click a check in the little box for "R1C1 reference style" under "settings".

The sectoring plan files are needed to understand the data files because the data files have no labels or codes to signify columns and rows.

Sectoring plan files:

The sectoring plan files are included outside of the ZIP files. The sectoring plan files are not included inside the ZIP files.

The BLS industry sectoring plan is provided by the MS EXCEL file called **"sect_industry_x195.xls"**.

In addition**, "sect_final_demand_x190.xls"** provide the final demand sectoring plan.

III. Conversion of Tables to Inverse and Other Coefficient Matrices

December 10, 2002

Mathematical Derivation of the Total Requirements Tables for Input-Output Analysis[i]

From the make and use tables, the following are defined:

q: Total commodity output. A column vector in which each entry shows the total amount of commodity output. It is a commodity-by-one vector.

g: Total industry output. A column vector in which each entry shows the total amount of each industry's output, including its production of scrap. It is an industry-by-one vector.

U: Intermediate portion of the use matrix in which the column shows for a given industry the amount of each commodity it uses, including noncomparable imports and scrap, used and secondhand goods. This is a commodity-by-industry matrix.

V: Make matrix, in which the column shows for a given commodity the amount produced in each industry. It is an industry-by-commodity matrix. V has columns showing only zero entries for noncomparable imports and for scrap.

^: A symbol that, when placed over a vector, indicates a square matrix in which the elements of the vector appear on the main diagonal and zeros elsewhere.

B: Direct input coefficients matrix (also known as the direct requirements matrix) in which entries in each column show the amount of a commodity used by an industry per dollar of output of that industry. It is a commodity-by-industry matrix.

$$B = U\hat{g}^{-1} \qquad (1)$$

D: A matrix in which entries in each column show, for a given commodity (excluding scrap), the proportion of the total output of that commodity produced in each industry. It is assumed that each commodity (other than scrap) is produced by the various industries in fixed proportions (*industry technology assumption*). D is an industry-by-commodity matrix. D also is referred to as the market share matrix or transformation matrix.

$$D = V\hat{q}^{-1} \qquad (2)$$

i: Unit (summation) vector containing only l's.

I: Identity matrix, where I = $\hat{i}$. Each of the diagonal elements of the matrix contain the value 1, and zeros elsewhere.

e: A column vector in which each entry shows the total final demand purchases for each commodity from the use table.

h: A column vector in which each entry shows the total amount of each industry's production of scrap. Scrap is separated to prevent its use as an input from generating output in the industries in which it originates.

p: A column vector in which each entry shows the ratio of the value of scrap produced in each industry to the industry's total output.

W: An industry-by-commodity matrix in which the entries in each column show, for a given commodity, the proportion of the total output of that commodity produced in each industry adjusted for scrap produced by the industry.

From the above definitions, the following identities are derived:

$$q = Ui + e \qquad (3)$$

$$g = Vi + h \qquad (4)$$

Scrap output in each industry is proportional to total output of the industry, then:

$$h = \hat{p}g \qquad (5)$$

The model expressed in equations (1) through (5) thus involves three constants (*B, D, p*) and six variables (*U, V, h, e, q, g*). The model solution is derived as follows:

From (1) and (3), we derive:

$$q = Bg + e \qquad (6)$$

From (2) and (4), we derive:

$$g - h = Dq \qquad (7)$$

Substituting (5) into (7) and solving for g:

$$g - \hat{p}g = Dq$$
$$(I - \hat{p})g = Dq$$

$$g = (I - \hat{p})^{-1} Dq \qquad (8)$$

Let $(I - \hat{p})^{-1} D = W$, then

$$g = Wq \qquad (9)$$

Substituting (9) into (6) and solving for *q*:

$$q = BWq + e$$

$$(I - BW)q = e$$

$$q = (I - BW)^{-1} e \qquad (10)$$

Substituting (10) into (9) gives:

$$g = W(I - BW)^{-1} e \qquad (11)$$

Therefore, three total requirements coefficients matrices are derived[ii]:

Commodity-by-commodity total requirements matrix:

$$(I - BW)^{-1} \qquad (12)$$

which shows commodity output required per dollar of each commodity delivered to final users.

Industry-by-commodity total requirements matrix:

$$W(I - BW)^{-1} \quad (13)$$

which shows the industry output required per dollar of each commodity delivered to final users.

And the industry-by-industry total requirements matrix:

$$(I - WB)^{-1} \qquad (14)$$

which shows the industry output required per dollar of each industry product delivered to final users.

[i] The notation and derivation of the tables presented follow the System of National Accounts recommended by the United Nations. See: A System of National Accounts Studies in Methods, Series F No. 2 Rev. 3, United Nations, New York, 1968; also, Stone, R., Bacharach, M. & Bates, J., "Input-Output Relationships, 1951-1966," Programme for Growth, Volume 3, London, Chapman and Hall, 1963.

[ii] Tables are prepared at the detailed, summary and sector levels of aggregation.