EMPLOYMENT OUTLOOK: 2021-2031

Layout and Description

Input-Output Tables:

1997 through 2021 Historical and Projected 2031

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

September 8, 2022

Youyang Li	li.youyang@bls.gov	**(202) 691-5700**
Kathryn Laurence	laurence.kathryn@bls.gov	**(202) 691-5691**

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 the 2031 projections, input-output, final demand, and value added data were developed for the years 1997 through 2021 and projected year 2031. Historical tables are provided in both nominal (current) dollars and in 2012 chain weighted real dollars. The projected tables are provided in 2012 chain weighted real dollars only.

In addition to the data tables, sectoring plan files "SectorPlan.xlsx" and "sect_final_demand_.xlsx" are needed because the data files have no labels to signify columns and rows. The "sector_industry" files provide our industry sectoring plan with industry codes and titles. The "sect_final_demand" files provide our final demand column codes and titles. Note, the SectorPlan.xlsx file contains summary sector codes and titles that DO NOT appear in the I-O tables.

The matrix tables described here are in comma separated values (CSV) files, which can be readily converted to spreadsheet use. Matrices include only data and sector code column and row identifiers. As stated above, **there are no row or column labels**. Each row contains all the columns for that row, including zeroes, separated by commas.

Data 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 2017 North American Industrial Classification System (NAICS); the U.S. Department of Commerce's Bureau of Economic Analysis (BEA) 1997, 2002, 2007, and 2012 benchmark input-output tables; and the BEA Annual input-output tables for 1997-2020. The BLS industry sectoring plan is shown in the industry sectoring plan files discussed above.

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 "PROJECTED_" for projections 2031. 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 "PROJECTED_" 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 153 final demand types. Each of the 153 columns is distributed across the 194 input-output commodity rows identified in "SectorPlan.xlsx" as mentioned in page 1. The "FDAGG" is the "FD" matrix collapsed from 153 columns to 11 columns. The column sectors are shown in "sector_final_demand.xlsx" as mentioned in page 1.

The "OUTPUT_COM" and "OUTPUT_IND" matrices detail the respective historical commodity and industry outputs by each of the 194 sectors identified in the input-output system for the years 1997 through 2021. The time series starts with 1997 outputs in the first column, the next year's outputs in the second column, and continues through 2021 outputs in the last column.

For the projection year, the "OUTPUT_2031" 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 ".CSV" which means that the file is a comma separated values data file.

B. Standard matrix types

"USE" Matrix (195x195)

This matrix contains intermediate inter-industry inputs plus value added (row 195) for the year specified in the matrix name. This matrix also contains intermediate inter-industry sales plus final demand (column 195) 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 (194x194)

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 (194x25)

This matrix contains the industry output time series for the 194 industry sectors for the years starting from 1997 (column 1) and continuing through 2021 (column 25).

"OUTPUT_COM" Matrix (194x25)

This matrix contains the commodity output time series for the 195 commodity sectors for the years starting from 1997 (column 1) continuing through 2021 (column 25).

"PROJECTED_OUTPUT" Matrix (194x2)

This table contains the total outputs for the projected year 2031. The first column is the total commodity output and the second is the total industry output for the sector designated by the row number.

"FD" Matrix (194x153)

This is the final demand matrix for the year specified in the title. Each "FD" matrix has 194 rows and 153 columns. Rows represent commodity final demand. Columns represent final demand by detailed category, margin reallocations, or import valuation adjustments. The column types are described below.

Each "FD" matrix has 121 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 194 commodity rows. Summing these data by row yields commodity final demand on a purchaser value basis.

Commodity output in the input-output system is on producers value rather than purchaser value. Margin costs, i.e. transportation, wholesale, retail and insurance, are removed and reallocated to the sectors providing the margin services. In the intermediate portion of the Use table, for a given industry, the input cells reflect the producer cost of the input commodity. Margin costs for commodity inputs are reallocated to the margin sectors.

In the "FD" matrix the columns containing the detailed category data are purchaser value. For imports of goods, the column data is foreign port value. Producer value excludes margin costs and import valuation adjustments, which are removed in the margin reallocation and import valuation adjustment columns, respectively. There are 30 margin reallocation columns and two import valuation adjustment columns[2] in each "FD" matrix. These columns and their functions are described below.

Three margin columns are associated with each of 10 aggregate final demand categories. The three margin columns are used to reallocate margin costs for retail and wholesale trade and five transportation commodities. Wholesale and retail trade have one reallocation column each and the reallocation of the five transportation modes are combined into one column.

Commodity margin rates[3] are calculated for the margins for each final demand category. Each rate is multiplied by the purchaser value of the commodities within the associated final demand category. The resulting margins appear as negative entries in the margin columns for that final demand category. Within each margin column, the sum of these entries is exactly offset by positive values inserted in the row cells corresponding to the appropriate margin sectors, so that the column sum is zero.

There are two import valuation adjustment columns in the goods and services imports category. The columns are used to reallocate the import valuation

---

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

[2] Import valuation adjustments reconcile differences in the valuation of import goods: NIPA accounting uses foreign port value, whereas input-output accounting uses domestic port value as the producer value. Because imports are negative values in the NIPAs, and because the import valuation adjustments increase the absolute value of imports, the import adjustment values are negative and the reallocation offsets are positive.

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

adjustment on imported goods for transportation and insurance in one column, and wholesale trade in the other.

Import valuation adjustment rates[4] are calculated on a commodity basis for each of the two import valuation adjustment columns. Each rate is multiplied by the foreign port (or purchaser) value of the commodity row for the goods import category. The resulting import valuation (margin) adjustments appear as negative entries in the adjustment columns. Within each adjustment column, the sum of these entries is exactly offset by positive values inserted in the row cells corresponding to the appropriate margin sectors, so that the column sum is zero.

"FDAGG" Matrix (194x11)

This matrix is the colum-wise collapsed version of the "FD" matrix above for the year specified in the matrix name. The 153 categories are aggregated to 11 categories as shown in the file, "sect_final_demand.xlsx". The 11 columns corresponding to the 11 aggregate categories are on a producer value basis.

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

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

II. How to Access Files

Data files:

The folder "IOnom" contains nominal (current) dollar historical I-O/FD data. The folder "IOreal" contains chain weighted real 2012 dollar historical I-O/FD data. The folder "IO2031" contains chain weighted real 2012 dollar I-O/FD projections. No current dollar I-O table projections were made.

The data files contained in these folders are described in Part I. The data are in comma separated values files, which are readily converted to spreadsheet use. These matrices contain only data and row and column sector code identifiers.

The data files contain no row or column labels. Each row contains all the columns for that row, including zeroes, separated by commas.

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

Sectoring plan files:

The BLS industry sectoring plan is provided by the EXCEL file called "SectorPlan.xlsx".

In addition, "sect_final_demand.xlsx" provides the final demand sectoring plan.

III. Conversion of Tables to Inverse and Other Coefficient Matrices

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

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

^: 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.

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.

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.

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

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.

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

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.

From the above definitions, the following identities are derived:

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

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

The model expressed in equations (1) through (4) thus involves two constants (*B, D*) and five variables (*U, V, e, q, g*). The model solution is derived as follows:

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

$$q = Bg + e \qquad (5)$$

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

$$g = Dq \qquad (6)$$

Substituting (6) into (5) and solving for g:

$$q = B(Dq) + e$$

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

$$q = (I\text{-}BD)^{-1}e \qquad (7)$$

The matrix $(I\text{-}BD)^{-1}$ is known as the commodity-by-commodity total requirements matrix and it shows, on a per-dollar basis, the commodity output the economy generates in order to provide commodities to final users.

Substituting (5) into (6) and solving for *g* gives:

$$g = D(Bg + e)$$

$$(I - DB)g = De$$

$$g = (I - DB)^{-1}De \qquad (8)$$

The matrix $(I - DB)^{-1}$ is known as the industry-by-industry total requirements matrix and it shows, on a per-dollar basis, the industry output the economy generates in order to provide an industry's commodities to final users. The vector *De* is a final demand vector where each entry shows the final demand for an industry's output.

Substituting (7) into (6) and solving for q gives:

$$g = D(I - BD)^{-1}e \qquad (9)$$

The matrix $D(I - BD)^{-1}$ is known as the industry-by-commodity total requirements matrix and it shows, on a per-dollar basis, the industry output the economy generates in order to provide commodities 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.