

RESEARCH & REPORTS

# IMPLAN

# Data

# Overview & Sources

# IMPLAN Data Overview

Constructing the annual IMPLAN datasets requires gathering data from numerous sources, estimating non-disclosed data elements, and converting them all to a consistent year and sectoring scheme, all while benchmarking them against other data to maintain accuracy. Construction of the annual IMPLAN datasets requires the use of roughly 100 datasets, each of which may contain multiple files and tables. IMPLAN adds value to publicly available data by:

1. Providing estimates for non-disclosed data
2. Providing estimates for non-census and non-survey years
3. Disaggregation into finer geographic scales
4. Disaggregation into finer sector detail
5. Projecting lagged raw data
6. Making adjustments for under-coverage

Additionally, IMPLAN provides inter-county trade flow data, for which there are no raw data sources. This is done using a calibrated and constrained gravity model. These trade flow data (along with commuting data) allow for multi-regional input-output (MRIO) analysis, which shows the indirect and induced effects of economic activity in one region across any number of linked regions.

## Production Functions

The U.S. production functions, i.e., the list of intermediate goods and services used by an industry to produce its output, are based on the BEA's Benchmark input-output (I-O) tables (released every 5 years). The U.S. gross absorption coefficients are adjusted to sum to new output and value-added totals each year.

1. We start with the latest BEA Benchmark I-O tables.
2. We derive current industry output, Value Added (VA), and final demands. VA and final demands each are equal to GDP, by definition.
3. Using the benchmark byproducts and the current *industry* output, we derive current *commodity* output.
4. Multiplying the current *industry* output by the benchmark absorption vector gives us a first approximation of the current Use Matrix (i.e., current production functions).
5. The columns of the current Use Matrix are forced to sum to *Industry* Output - Value Added (columns).
6. Then the rows of the current Use Matrix are forced to sum to the *Commodity* Output - Final Demand control totals (rows). This will be the first time the proportions of the columns change.
7. This is repeated until no further adjustment is needed.

To obtain sub-national absorption coefficients, the national absorption coefficients are adjusted to match regional Value Added-to-Output ratios. The assumption is that the local data are correct and the national coefficients need to adjust to fit the local situation. Applying the trade flow assumptions, i.e., the rate at which each good and service is procured locally, will then further modify the *regional* absorption coefficients by removing the imports (which vary by region and year).

## Employment and Earnings

In general, BLS's annual totals from the Quarterly Census of Employment and Wages (QCEW) data provide the county-level industry structure for the IMPLAN database. The Census Bureau's County Business Patterns (CBP) data, along with other data, are used to estimate values for non-disclosed data. QCEW and CBP include only wage and salary (W&S) employment and income data. Projected BEA Regional Economic Accounts (REA) data are used to incorporate proprietors, non-covered sectors, and supplements to wages.

## Value Added

Other Property Income (OPI) primarily consists of corporate income and depreciation, and is one of the most difficult items to estimate. The BEA releases state-level 3-digit NAICS GDP data, which include the break-out of GDP into Employee Compensation (EC), Taxes on Production and Imports (TOPI), and Gross Operating Surplus (GOS). GOS includes both Proprietor Income (PI) and OPI, so OPI for each 3-digit NAICS sector is derived by subtracting our estimate of PI from GOS. These control values are distributed to IMPLAN industries based on the BEA benchmark I-O characteristics for GOS.

## Output

In IMPLAN, Total Industry Output (TIO) is the value of annual calendar year production. It can be measured either as the total value of purchases by intermediate and final consumers or as intermediate outlay plus value-added. Output data for most sectors come from the BEA's Annual Industry Accounts and the Annual Survey of Manufacturers. Retail output data come from the U.S. Census Bureau's Annual Census of Retail Trade. Certain other sectors use information from various other surveys and censuses.

## Personal Consumption Expenditures (PCE)

### NIPA PCE DATA

- Annual and current.
- National level.
- Only one spending pattern (i.e., not separated by income class).
- Roughly 100 expenditure categories. The BEA benchmark I-O is used to distribute these expenditure categories among IMPLAN sectors.
- The PCE data are in *purchaser* prices, so margining and re-sectoring are necessary to obtain *producer* prices for use in IMPLAN.

### CENSUS BUREAU CONSUMER EXPENDITURE SURVEY (CES) DATA

- Annual but lagged.
- National level.
- Gives us expenditures by income class (we then control these to the NIPA PCE totals).

Since QCEW data capture only employees covered by public unemployment insurance, the data set cannot capture self-employed persons, railway employment, religious organizations, military, elected officials or any other establishments that have their own social insurance program and/or do not participate in public unemployment insurance. Since most farm employment is self-employment, we have alternative methods for estimating farm employment and labor income.

## Control to NIPA Totals

- All sub-national elements are controlled to national NIPA values

disaggregated or disambiguated by

### Regional Economic Accounts

- From BEA.
- Proprietor employment, W&S employment, Proprietor Income, EC, W&S income.
- Granulated to national, state, and county levels.

### Quarterly Census of Employment and Wages

- From BLS.
- Includes W&S employment, W&S income.
- Granulated to national, state and county levels.

disaggregated or disambiguated by

## County Business Patterns

- From U.S. Census Bureau.
- Includes number of firms by firm-class size (helps disclose QCEW employment data).
- Granulated to national, state, county, and ZIP-code levels.

## Institutions

### HOUSEHOLDS

National PCE estimates are distributed to states and counties based on state- and county-level households-by-income-class data and spending-by-income-class data.

### GOVERNMENT

Federal sales and expenditure data are estimated using NIPA control totals and the benchmark I-O distribution, with the exception of timber sales data, which come directly from the U.S. Forest Service. Data for state and local government sales are obtained from the current Annual Survey of State and Local Governments Finances data series, while state and local government expenditures are estimated using NIPA control totals and the benchmark I-O distribution.

### INVENTORY

For the manufacturing sectors, the Annual Survey of Manufacturers provides inventory data. Inventory for other sectors are derived from benchmark I-O ratios.

### IMPORTS/EXPORTS

For the U.S., data for foreign imports and exports of shippable commodities come from the U.S. Census Bureau. Foreign imports and exports of services are based on the BEA benchmark I-O controlled to current NIPA import and export totals, respectively.

After accounting for foreign trade, the gravity-based trade flow or econometric model is used to determine domestic trade (trade among states and counties).

### CAPITAL

BEA data provide total capital expenditures by commodity, grouped by private investment, federal government investment, and state and local government investment. The national data are distributed to states and counties based on output in the construction industries.

# IMPLAN Data Sources

## IMPLAN Adds Value to Public Data by:

### **PROVIDING ESTIMATES FOR NON-DISCLOSED DATA**

The government is not allowed to disclose data that would allow you to identify a specific business. Thus, the government purposely “hides” many values—so we have to estimate those missing values.

### **DISAGGREGATION INTO FINER GEOGRAPHIC SCALES**

The smaller the geographic entity (e.g., counties, zip-codes), the fewer raw data available for them. Thus, we have procedures to distribute national values among the states, state values among the counties, and county values among the zip-codes, using other data as distributors.

### **DISAGGREGATING INTO FINER SECTOR DETAIL**

The vast majority of our data sources have less sector detail than IMPLAN. We have procedures to distribute to IMPLAN sectors based on other data that do have the necessary sector detail.

### **PROJECTING LAGGED DATA**

Many of our data sources are lagged one year relative to the current IMPLAN data year (that is, there “current” values come out too late for us to use until the following data year).

## Main Raw Data Sources

### **BUREAU OF ECONOMIC ACCOUNTS (BEA)**

#### *National Income and Product Accounts (NIPAs)*

- These are at the U.S. level and serve as our governing controls for the majority of data elements (e.g., total U.S. employment, GDP, capital investment, PCE spending)
- National level only, released annually

#### *Benchmark I-O tables*

- This is where we get our production functions!
- Released roughly every 5 years

#### *Regional Economic Accounts (REA)*

- This is our source of EC and Proprietor Employment and Income
- Released annually but lagged a year
- Only ~100 sectors

#### *GDP-by-State Series*

- Also released annually, also lagged a year
- Output for most service sectors (we get output for farm, manufacturing, and other sectors elsewhere)

#### *Other data*

- Past-year deflators

- State-level tax data
- County-level personal income
- Net commuting rates

## USDA

### *Census of Agriculture*

- County-level farm-sector output
- Released every 5 years

### *National Agricultural Statistics Service (NASS)*

- Annual state-level value of production for farm sectors

### *Economic Research Service (ERS)*

- Annual state-level sales for farm sectors

## BUREAU OF LABOR STATISTICS (BLS)

### *Quarterly Covered Employment and Wages (QCEW) Data*

- Wage and salary employment and income (no proprietors, no EC)
- County-level, 6-digit NAICS level

### *Consumer Expenditure Survey (CES)*

- Allows us to break out the NIPA PCE data among IMPLAN's 9 household income categories
- National level, lagged at least one year, sometimes more

## CENSUS BUREAU

### *County Business Patterns (CBP)*

- Very similar to QCEW data but a) one-year lagged and b) always disclose establishment counts by employment size-classes
- Used to disclose the QCEW data
- One of the few data sources that is available at the zip-code level; this data source allows us to have zip-code-level IMPLAN data

### *Annual Survey of Manufactures (ASM)*

- Output and inventory for manufacturing sectors

### *U.S.-level construction sector output*

### *U.S.-level foreign exports and imports*

### *Census of Government Finances*

- Revenue and spending by state, county, and city governments

## Supplemental Raw Data Sources

### RAILROAD RETIREMENT BOARD

BLS QCEW data do not cover railroad employees since they are not covered by Federal unemployment insurance programs.

### FEDERAL TIMBER SALES DATA FROM THE USFS

- We have a special relationship with the USFS so they give us county-level, current-year data with no non-disclosures for federal timber sales.
- For other commodities sold by Federal Government, we have only U.S.-level data and have to use other distributors (e.g., Federal EC).

### NCES INTEGRATED POSTSECONDARY EDUCATION DATA SYSTEM

#### *Higher education employment*

### NOAA FISHERIES STATISTICS DIVISION

#### *Commercial fishing output*

## Process Overview

Each data source is released on its own schedule, spanning from May to October. These are the general steps that must be carried out for each data set:

- Download raw data
- Upload into SQL Server
- Estimate non-disclosures
- Bridge to IMPLAN sectoring scheme
- Distribute U.S.- and state-level values to counties and zip-codes
- Perform QC checks along each step in the process

We write our own VB.NET programs to automate just about everything other than QC work. There are some manual spreadsheet processes due to the format of the raw data (NASS and ERS data) or to the need for hand-adjustments and conditional formatting (National I-O Controls).

## Periodic Challenges

- Raw data sources change web addresses, table formats, sector coding, scope of coverage, etc.
- Temporary government shut-downs
- Changes to NAICS schemes (every 5 years) and BEA BM sectoring (every 5 years)
- Disappearing and appearing counties and zip-codes; renaming of counties

## Miscellaneous Notes

IMPLAN data must be taken as a whole—zip-codes must sum to counties, counties must sum to states, and states must sum to the U.S. NIPA controls. Thus, we cannot just update one or a few elements without having to re-run everything else. If we updated the IMPLAN data every time one of our data sources is updated, we would spend all year long running and re-running a single year's data without any time to work on data improvements and new data products.

The smaller the geography class the fewer data we have for it, so the less confident we can be in our estimates. Thus, we are most confident in the U.S. values and least confident in the zip-code values.

Although CDs are built from groups of zip-codes, we have somewhat more confidence in the CD data than the zip-code data since it may smooth out some imperfect county-zip matches.



## About IMPLAN

IMPLAN is the leading provider of economic impact data and analytical software. But we're so much more than just that. Having spent decades serving the economic data needs of academics, governments, economic developers, corporations, non-profits, consultants, and more, the trust we've garnered and the relationships we've built throughout domestic and international communities is what we value most.

**Take the first step now! To start your economic impact analysis today, visit [IMPLAN.com](http://IMPLAN.com).**

For more information about how economic impact analysis can benefit your organization or to inquire about IMPLAN's host of exciting products, give us a call at (800) 507-9426.

16905 Northcross Drive, Suite 120

Huntersville, NC 28078

(800) 507-9426

[IMPLAN.com](http://IMPLAN.com)

**IMPLAN**