

2009 Annual Retail Trade Survey

Summary of Changes

The following changes were made with the release of the 2009 ARTS estimates:

- ARTS estimates were revised to reflect benchmarking using final results of the 2007 Economic Census. Previously, estimates were benchmarked using preliminary results of the 2007 Economic Census. For more information, refer to the Benchmarking section of the ARTS Methodology link located at: http://www.census.gov/retail/arts/how_surveys_are_collected.html
- Detailed operating expense estimates for 2007, collected as part of the Business Expenses Supplement (BES) to the 2007 ARTS, and corresponding measures of sampling variability were revised to reflect benchmarking using final results of the 2007 Economic Census.
- Sales for NAICS 7223 (Special Food Services) and aggregate industry levels to which NAICS 7223 contributes were revised back to 1998. The list of aggregate industry levels are:
 - Retail and food services, total
 - Total (excl. motor vehicle and parts dealers)
 - NAICS 722 (Food Services and Drinking Places)

Sales estimates for these industry levels were revised further back than sales estimates for other industries because a revision was made to the 2002 Economic Census sales total for NAICS 722330 (Mobile Food Services). This revision was not included in the preliminary results of the Economic Census and thus was not part of last year's benchmarking.

- The span of years displayed in the pdf versions of the tables for sales, inventories, purchases, gross margin, and gross margin as percent of sales have been updated to include years 1998 through 2009. Excel versions of the tables include all years of our NAICS time series.
- The span of years for the per capita sales table, Excel and pdf versions, have been updated to include years 2000 through 2009. Population estimates from 2000 through 2008 were revised. Per capita estimates prior to 2000 are not available, because the population estimates used in the calculations were not available on a comparable basis. For additional information, see <http://www.census.gov/popest/topics/methodology/2010-relnotes.pdf>. For data from 1992 through 2008, see http://www.census.gov/retail/arts/historic_releases.html
- The publication tables for sales, purchases, inventories, total operating expenses, gross margin and gross margin as percent of sales were expanded to include additional industries. Estimates for these additional levels are presented back to the beginning of the NAICS time series for each data item. Annual NAICS estimates prior to 1999 are derived from data that were collected on the Standard Industrial Classification (SIC) system. For a description of how these estimates were derived, see <http://www2.census.gov/retail/releases/benchmark/annpub00.pdf>. The list of additional NAICS codes and description are presented below by data item.

Additional sales levels:

4412 Other Motor Vehicle Dealers
44229 Other Home Furnishings Stores
4431 Electronics and Appliance Stores
4452 Specialty Food Stores
4461 Health and Personal Care Stores
4483 Jewelry, Luggage, and Leather Goods Stores
4511 Sporting Goods, Hobby, and Musical Instrument Stores
4512 Book, Periodical, and Music Stores
45121 Book Stores and News Dealers
45211 Department Stores
4531 Florists
4539 Other Miscellaneous Store Retailers
45399 All Other Miscellaneous Store Retailers
4542 Vending Machine Operators
4543 Direct Selling Establishments
722211 Limited Service Restaurant
7223 Special Food Services
72231 Food Service Contractors

Additional purchases levels:

4461 Health and Personal Care Stores
4511 Sporting Goods, Hobby, and Musical Instrument Stores
45211 Department Stores

Additional inventories levels:

4411 Automobile Dealers
4413 Automotive Parts, Accessories, and Tire Stores
4441 Building Material and Supplies Dealers
4451 Grocery Stores
4453 Beer, Wine, and Liquor Stores
4461 Health and Personal Care Stores
44611 Pharmacies and Drug Stores
4481 Clothing Stores
44811 Men's Clothing Stores
44812 Women's Clothing Stores
44814 Family Clothing Stores
4482 Shoe Stores
4511 Sporting Goods, Hobby, and Musical Instrument Stores
4512 Book, Periodical, and Music Stores
4529 Other General Merchandise Stores
4541 Electronic Shopping and Mail-Order Houses

Additional total operating expenses levels:

4412 Other Motor Vehicle Dealers
4421 Furniture Stores
4461 Health and Personal Care Stores
4483 Jewelry, Luggage, and Leather Goods Stores
4511 Sporting Goods, Hobby, and Musical Instrument Stores
45211 Department Stores
4532 Office Supplies, Stationery, and Gift Stores
4539 Other Miscellaneous Store Retailers
4543 Direct Selling Establishments
7221 Full-Service Restaurants
7222 Limited-Service Eating Places

Additional gross margin levels:

4461 Health and Personal Care Stores
4511 Sporting Goods, Hobby, and Musical Instrument Stores
45211 Department Stores

Additional gross margin as percent of sales levels:

4461 Health and Personal Care Stores
4511 Sporting Goods, Hobby, and Musical Instrument Stores
45211 Department Stores

Annual Retail Trade Survey Methodology

Overview

The Census Bureau introduced a new sample with the 2005 Annual Retail Trade Survey (ARTS). This new sample is designed to produce estimates based on the 2002 North American Industry Classification System (NAICS).

The U.S. Census Bureau conducts the Annual Retail Trade Survey (ARTS) to produce national estimates of total annual sales, e-commerce sales, end-of-year inventories, inventory-to-sales ratios, purchases, total operating expenses, inventories held outside the United States, gross margins, and end-of-year accounts receivable for retail businesses and annual sales and e-commerce sales for accommodation and food service firms located in the United States. Estimates are summarized by industry classification based on the North American Classification System (NAICS) and are released approximately 15 months after the reference year. Firms without paid employees, or nonemployers, are included in the estimates through imputation or administrative data provided by other federal agencies.

Sampling Frame

The sampling frame used for the ARTS has two types of sampling units: Employer Identification Numbers (EINs) and large, multiple-establishment firms. Both sampling units represent clusters of one or more establishments owned or controlled by the same firm. The information used to create these sampling units was extracted from data collected as part of the 2002 Economic Census and from establishment records contained on the Census Bureau's Business Register as updated to December 2004. The next few paragraphs give details about the Business Register; the distinction between firms, EINs, and establishments; and the construction of the sampling units. Though important, they are not essential to understanding the basic sample design and readers may continue to the Stratification, Sampling Rates, and Allocation section.

The Business Register is a multi-relational database that contains a record for each known establishment that is located in the United States or one of its territories and has paid employees. An establishment is a single physical location where business transactions take place and for which payroll and employment records are kept. Groups of one or more establishments under common ownership or control are firms. A single-unit firm owns or operates only one establishment. A multiunit firm owns or operates two or more establishments. The treatment of establishments on the Business Register differs according to whether the establishment is part of a single-unit or multiunit firm. In particular, the structure of an establishment's primary identifier on the Business Register differs depending on whether it is owned by a single-unit firm or by a multiunit firm.

A single-unit firm's primary identifier is its EIN. The Internal Revenue Service (IRS) issues the EIN, and the firm uses it as an identifier to report social security payments for its employees under the Federal Insurance Contributions Act (FICA). The same act requires all employer firms to use EINs. Each employer firm is associated with at least one EIN and only one firm can use a given EIN. Because a single-unit firm has only one establishment, there is a one-to-one relationship between the firm and the EIN. Thus the firm, the EIN, and the establishment all reference the same physical location and all three terms can be used interchangeably and unambiguously when referring to a single-unit firm.

For multiunit firms however, a different structure connects the firm with its establishments via the EIN. Essentially a multiunit firm is associated with a cluster of one or more EINs and EINs are associated with one or more establishments. A multiunit firm consists of at least two establishments. Each firm is

associated with at least one EIN and only one firm can use a given EIN. However, one multiunit firm may have several EINs. Similarly, there is a one-to-many relationship between EINs and establishments. Each EIN can be associated with many establishments but each establishment is associated with only one EIN. Because of the possibility of one-to-many relationships, we must distinguish between the firm, its EINs, and its establishments. The multiunit firm that owns or controls a particular establishment is identified on the Business Register by way of the establishment's primary identifier.

The primary identifier of an establishment owned by a multiunit firm consists of a unique combination of an alpha number and a plant number. The alpha number identifies the multiunit firm, and the plant number identifies a particular establishment within that firm. All establishments owned or controlled by the same multiunit firm have the same alpha number. Different multiunit firms have different alpha numbers, and different establishments within the same multiunit firm have different plant numbers. The Census Bureau assigns both the alpha number to the multiunit firm and plant numbers to the corresponding establishments based on the results of the quinquennial economic census and the annual Company Organization Survey.

To create the sampling frame, we extract the records for all employer establishments located in the United States and classified in the Retail Trade sector as defined by the 2002 NAICS. For these establishments we extract sales, payroll, employment, name and address information, as well as primary identifiers and, for establishments owned by multiunit firms, associated EINs. To create the sampling units for multiunit firms, we aggregate the economic data of the establishments owned by these firms to an EIN level by tabulating the establishment data for all retail establishments associated with the same EIN. Similarly, we aggregate the data to a multiunit firm level by tabulating the establishment data for all retail establishments associated with the same alpha number. No aggregation is necessary to put single-unit establishment information on an EIN basis or a firm basis. Thus, the sampling units created for single-unit firms simultaneously represent establishment, EIN, and firm information. In summary, the sampling frame is a complex amalgam of establishments, EINs, and firms.

Stratification, Sampling Rates, and Allocation

The primary stratification of the sampling frame is by industry group based on the detail required for publication. We further stratify the sampling units within industry group by a measure of size (substratify) related to their annual sales. Sampling units expected to have a large effect on the precision of the estimates are selected "with certainty." This means they are sure to be selected and will represent only themselves (i.e., have a selection probability of 1 and a sampling weight of 1). Within each industry stratum, we determine a substratum boundary (or cutoff) that divides the certainty units from the noncertainty units. We base these cutoffs on a statistical analysis of data from the 2002 Economic Census. Accordingly, these values are on a 2002 sales basis. We also used this analysis to determine the number of size substrata for each industry stratum and to set preliminary sampling rates needed to achieve specified sampling variability constraints on sales estimates for different industry groups. The size substrata and sampling rates are later updated through analysis of the sampling frame.

Sample Selection

The first step in the sample selection identified firms selected with certainty. If a firm's annual sales or end-of-year inventories were greater than the corresponding certainty cutoff, that firm was selected into the ARTS sample with certainty.

All firms not selected with certainty were subjected to sampling on an EIN basis. If a firm had more than one EIN, we treated each of its EINs as a separate sampling unit. To be eligible for the initial sampling,

an EIN had to have nonzero payroll in 2003. The EINs were stratified according to their major industry and their estimated sales (on a 2002 basis). Within each noncertainty stratum, a simple random sample of EINs was selected without replacement. The selected noncertainty EINs were divided into two approximately equal groups. One group is canvassed for both the monthly and the annual survey, the other group is canvassed for only the annual survey.

Sample Maintenance

Periodically, we update the samples to represent new EINs appearing on the Business Register. These new EINs, called births, are EINs recently assigned by the IRS on the latest available IRS mailing list for FICA taxpayers and assigned an industry classification (if possible) by the Social Security Administration (SSA).

EIN births are sampled on a quarterly basis using a two-phase selection procedure. To be eligible for selection, a birth must either have no industry classification or be classified in an industry within the scope of the ARTS, the Annual Wholesale Survey (AWTS), or the Service Annual Survey (SAS), and it must meet certain criteria regarding its number of paid employees or quarterly payroll. In the first phase, births are stratified by broad industry groups and a measure of size based on quarterly payroll. A relatively large sample is selected using equal probability systematic sampling. The selected births are canvassed to obtain a more reliable measure of size, consisting of sales in two recent months, company affiliation information, and a new or more detailed industry classification code. Births that have not returned their questionnaire after 30 days are contacted by telephone.

Using this more reliable information, the selected births from the first phase are subjected to probability proportional-to-size sampling with overall probabilities equivalent to those used in drawing the initial ARTS sample from the December 2004 Business Register. Because of the time it takes for a new employer firm to acquire an EIN from the IRS, and because of the time needed to accomplish the two-phase birth-selection procedure, births are added to the samples approximately nine months after they begin operation.

If a firm was selected with certainty and had more than one establishment at the time of sampling, any new establishments that the firm acquires, even if under new or different EINs, are included in the sample with certainty. However, if a single-unit firm was selected with certainty, only future establishments associated with that firm's originally-selected EIN are included in the sample with certainty; any new EINs that might later be associated with that firm are subjected to sampling through the quarterly birth-selection procedure.

Singleunit EINs selected into the sample with certainty are not dropped from canvass and tabulation if they are no longer on the IRS mailing list. Rather, the firm that used the EIN is contacted, and if a successor EIN is found, it is added to the survey. For both inactive and reactivated EINs, data are tabulated for only the portion of the reference year that these EINs reported payroll to the IRS.

Births that are selected in the quarterly birth-selection procedure in November of the annual survey reference year are included in the initial mailing of the annual survey questionnaires in January of the following year. To better represent all EIN births in the reference year, and specifically to account for the lag between the time a business starts operation and the time it takes to acquire and EIN and identify and select the EIN into one of our surveys, we add births to the annual survey sample that are selected in February, May, and August of the year following the annual survey reference year. We mail annual survey forms to these births in June and August to supplement the initial annual survey mailings.

Estimation and Sampling Variance

Total estimates are computed using the Horvitz-Thompson estimator (i.e., as the sum of weighted data (reported or imputed) for all selected sampling units that meet the sample canvass and tabulation criteria). The weight for a given sampling unit is the reciprocal of its probability of selection into the ARTS sample. These estimates are input to a benchmarking procedure, as described below. Variances are estimated using the method of random groups and are used to determine if measured changes are statistically significant.

Historical Estimates

Annual estimates prior to 1999 are derived from data that were collected and published based on the Standard Industrial Classification (SIC) system. For a description of how these estimates were derived, see: <http://www2.census.gov/retail/releases/benchmark/annpub00.pdf>

Because of the method used to derive annual estimates prior to 1999, these estimates should be used with caution. It is expected that for estimates for NAICS codes, that, by definition, are the same or nearly the same as a given SIC code, the quality of the estimates will be similar to that of the estimates released on an SIC basis. Estimates may be of less quality for NAICS codes that consist of more than one SIC component. Additionally, for reference years further from 1997, estimates are likely of less quality than for those years close to 1997. Note, however, that estimated year-to-year changes for 1992 through 1998 are dependent on the underlying SIC-based year-to-year changes. Year-to-year changes for 1999 and subsequent years are derived from data collected on a NAICS basis.

Special caution should be exercised when using the end-of-year retail inventory estimates prior to 1999. Retail inventory data has historically been analyzed at much broader industry levels than for sales. Determining clear relationships between NAICS and SIC codes was much more difficult at broader levels.

Benchmarking

Final results of the 2007 Economic Census are now available and are used to benchmark the ARTS sales estimates. During the processing of the 2007 Economic Census, a revision was made to the 2002 Economic Census sales total for NAICS 722330 (Mobile Food Services), and is reflected in this benchmarking operation. This revision was not included in the preliminary results of the Economic Census and thus was not part of last year's benchmarking. Prior to benchmarking to final 2007 Economic Census results, two operations are performed:

- Historical corrections are made to current sample data back to 2004. Corrections are made to replace previously reported data with more accurate data received at a later date or to replace imputed data with reported data obtained from the company.
- Sales estimates from the current sample are linked to the published census-adjusted estimates from the prior sample. For a given detailed industry based on the 2002 North American Industry Classification System (NAICS), the linking is performed by multiplying the sample-based sales estimate by a ratio. The numerator and denominator of the ratio are as follows:
 - The numerator is the 2004 published census-adjusted sales estimate for the industry from the prior sample.

- The denominator is the 2004 sales estimate for the industry from the current sample.

The resulting sales estimates (call these "modified" sales estimates) for 2002 through 2009 are input to the benchmarking program. Using this program, the modified sales estimates for 2002 through 2009 are revised in a manner that:

- Uses the 2002 and 2007 Economic Census sales totals as constraints.
- Minimizes the sum of squared differences between the year-to-year changes of the input and revised estimates for 2003 through 2009.

Because of the revision to the 2002 Economic Census sales total for NAICS 722330, the 1997 Economic Census sales total is also used as a constraint in the benchmarking process described above. For this industry, the modified sales estimates from 1997 through 2009 are revised in the same manner as the other industries.

Refer to the estimates output from the benchmarking operation as "benchmarked."

A method similar to the one for benchmarking sales is used to benchmark end-of-year inventories and purchases estimates. First, the sample-based inventories and purchases estimates for each detailed industry for 2004 and subsequent years are multiplied by the sales ratio described above, resulting in adjusted estimates for these years. Then, the published adjusted estimates for 1998 through 2004 from the prior sample are input to the benchmarking program. Using this program, the estimates for 1999 through 2004 are revised in a manner that:

- Uses the published adjusted estimate for 1998 from the prior sample as a constraint, resulting in no revision to the published 1998 estimate.
- Uses the adjusted estimate for 2004 from the current sample as a constraint.
- Minimizes the sum of squared differences between the year-to-year changes of the input and revised estimates for 1999 through 2004.

The resulting modified inventories and purchases estimates for 1997 through 2009 are input to the benchmarking program. Using this program, the modified estimates for 1997 through 2009 are revised in a manner that:

- Uses 1997, 2002, and 2007 constraints for inventories and purchases, where the constraints are calculated by multiplying the modified inventories and purchases estimates just calculated by the ratio of the benchmarked-to-modified sales.
- Minimizes the sum of squared differences between the year-to-year changes of the input and revised estimates for 1998 through 2009.

Benchmarked total expenses estimates for 2006 and subsequent years for a detailed industry are calculated by multiplying the Horvitz-Thompson estimates of total expenses by the ratio of the benchmarked-to-modified sales estimates for the corresponding year. Benchmark estimates of detailed expense items for 2007 are calculated by multiplying the previously published detailed-to-total expense percentages by the benchmarked total expense estimate for 2007. Benchmarked foreign inventories estimates for 2005 and subsequent years are calculated by multiplying the Horvitz-Thompson estimates of foreign inventories by the ratio of the benchmarked-to-modified total inventories estimates for the corresponding year.

Benchmarked e-commerce estimates are produced in a manner similar to that used for inventories and purchases, except 1998 is used as a constraint because e-commerce was not collected prior to 1998.

For the Electronic Shopping and Mail Order Houses industry group (NAICS 4541), modified merchandise lines sales estimates for 1999 and subsequent years are obtained using methodology similar to that used for inventories and purchases. Benchmarking merchandise lines sales are calculated by multiplying the modified estimates of merchandise lines sales by the ratio of the benchmarked-to-modified sales estimates for the corresponding year. Benchmarking merchandise lines e-commerce estimates are created using the same method, but the ratio of the benchmarked-to-modified estimates is based on e-commerce instead of sales.

Benchmarking estimates at aggregate industry levels are computed by summing the benchmarking estimates for the appropriate detailed industries comprising the aggregate.

Reliability of the Estimates

The published estimates may differ from the actual, but unknown, population values. For a particular estimate, statisticians define this difference as the total error of the estimate. When describing the accuracy of survey results, it is convenient to discuss total error as the sum of sampling error and nonsampling error. Sampling error is the error arising from the use of a sample, rather than a census, to estimate population values. Nonsampling error encompasses all other factors that contribute to the total error of a sample survey estimate. The sampling error of an estimate can usually be estimated from the sample; whereas, the nonsampling error of an estimate is difficult to measure and can rarely be estimated. Consequently, the actual error in an estimate exceeds the error that can be estimated. Further descriptions of sampling error and nonsampling error are provided in the following sections. Data users should take into account the estimates of sampling error and the potential effects of nonsampling error when using the published estimates.

Sampling Error

Because the estimates are based on a sample, exact agreement with results that would be obtained from a complete enumeration of firms on the sampling frame using the same enumeration procedures is not expected. However, because each firm on the sampling frame has a known probability of being selected into the sample, it is possible to estimate the sampling variability of the survey estimates.

The particular sample used in this survey is one of a large number of samples of the same size that could have been selected using the same design. If all possible samples had been surveyed under the same conditions, an estimate of a population parameter of interest could have been obtained from each sample. For the parameter of interest, estimates derived from the different samples would, in general, differ from each other. Common measures of the variability among these estimates are the sampling variance, the standard error, and the coefficient of variation (CV). The sampling variance is defined as the squared difference, averaged over all possible samples of the same size and design, between the estimator and its average value. The standard error is the square root of the sampling variance. The CV expresses the standard error as a percentage of the estimate to which it refers. For example, an estimate of 200 units that has an estimated standard error of 10 units has an estimated CV of 5 percent. The sampling variance, standard error, and CV of an estimate can be estimated from the selected sample because the sample was selected using probability sampling. Note that measures of sampling variability, such as the standard error and CV, are estimated from the sample and are also subject to sampling variability. (Technically, we should refer to the estimated standard error or the estimated CV of an estimator. However, for the sake of

brevity we have omitted this detail.) It is important to note that the standard error and CV only measure sampling variability. They do not measure any systematic biases in the estimates.

The Census Bureau recommends that individuals using published estimates incorporate this information into their analyses, as sampling error could affect the conclusions drawn from these estimates.

The estimate from a particular sample and its associated standard error can be used to construct a confidence interval. A confidence interval is a range about a given estimator that has a specified probability of containing the average of the estimates for the parameter derived from all possible samples of the same size and design. Associated with each interval is a percentage of confidence, which is interpreted as follows. If, for each possible sample, an estimate of a population parameter and its approximate standard error were obtained and using a t-statistic with 15 (16 random groups - 1) degrees of freedom, then:

1. For approximately 90 percent of the possible samples, the interval from 1.75 standard errors below to 1.75 standard errors above the estimate would include the average of the estimates derived from all possible samples of the same size and design.
2. For approximately 95 percent of the possible samples, the interval from 2.13 standard errors below to 2.13 standard errors above the estimate would include the average of the estimates derived from all possible samples of the same size and design.

To illustrate the computation of a confidence interval for an estimate of total sales, assume that an estimate of total sales is \$10,750 million and the CV for this estimate is 1.8 percent, or 0.018. First obtain the standard error of the estimate by multiplying the total sales estimate by its CV. For this example, multiply \$10,750 million by 0.018. This yields a standard error of \$193.5 million. The upper and lower bounds of the 90-percent confidence interval are computed as \$10,750 million plus or minus 1.75 times \$193.5 million. Consequently, the 90-percent confidence interval is \$10,411 million to \$11,089 million. If corresponding confidence intervals were constructed for all possible samples of the same size and design, approximately 9 out of 10 (90 percent) of these intervals would contain the average of the estimates derived from all possible samples.

Nonsampling Error

Nonsampling error encompasses all other factors, other than sampling error, that contribute to the total error of a sample survey estimate and may also occur in censuses. It is often helpful to think of nonsampling error as arising from deficiencies or mistakes in the survey process. Nonsampling errors are difficult to measure and can be attributed to many sources: the inclusion of erroneous units in the survey (overcoverage), the exclusion of eligible units from the survey (undercoverage), nonresponse, misreporting, mistakes in recording and coding responses, misinterpretation of questions, and other errors of collection, response, coverage, or processing. Although nonsampling error is not measured directly, the Census Bureau employs quality control procedures throughout the process to minimize this type of error.

A potential source of bias in the estimates is nonresponse. Nonresponse is defined as the inability to obtain all the intended measurements or responses about all selected units. Two types of nonresponse are often distinguished. Unit nonresponse is used to describe the inability to obtain any of the substantive measurements about a sampled unit. In most cases of unit nonresponse, the questionnaire was never returned to the Census Bureau after several attempts to elicit a response. Item nonresponse occurs either when a question is unanswered or the response to the question fails computer or analyst edits.

The Total Quantity Response Rate (TQRR) for the 2009 annual sales of U.S. retail, accommodation, and food services firms was 92.4 percent, end-of-year inventories of U.S. retail firms was 92.6 percent, annual purchases of U.S. retail firms was 79.4 percent, and annual operating expenses of U.S. retail, accommodation, and food services firms was 85.0 percent. The TQRR is defined as data that has been reported by the respondent or has been obtained from sources determined to be equivalent in quality to reported data. Equivalent quality data is data obtained from another source than the respondent, which have quality equivalent to data reported by the respondent. Equivalent quality data have three possible sources: 1) data directly substituted from another census or survey (for the same reporting unit, question wording, and time period); 2) data from administrative records; or 3) data obtained from some other equivalent source that has been validated by a study approved by the program manager in collaboration with the appropriate Research and Methodology area (e.g., company annual reports, Securities and Exchange Commission (SEC) filings, and trade association statistics).

Estimates with a coefficient of variation greater than 30 percent or with a total quantity response rate less than 50 percent have been suppressed from publication. These estimates have been replaced with an "S" in the published table. For a description of Census Bureau publication standards and the total quantity response rate, see the Census Bureau's statistical quality standard <http://www.census.gov/quality/standards/standardf1.html>