Part B – **Collections of Information Employing Statistical Methods**

<u>**Question 1. Universe and Respondent Selection**</u>

Target Population

The target population of the Annual Wholesale Trade Survey (AWTS) consists of all United States firms with paid employees that are primarily engaged in wholesale trade, as defined by the 2012 North American Industry Classification System (NAICS).

Sampling Frame

The AWTS sampling frame has two types of sampling units represented: single-establishment firms and multiple-establishment firms. The information used to create these sampling units was extracted from data collected via the 2012 Economic Census and from establishment records contained on the Census Bureau's Business Register (as of October 2015). The Business Register (BR), which is the Census Bureau's master business list, contains basic economic information for more than 7.4 million employer businesses and over 22.5 million nonemployer businesses. It is updated through direct data collections and administrative records from other federal agencies.

To create the sampling frame, the Census Bureau extracts the records for all employer establishments located in the United States and classified in the wholesale trade sector, as defined by the 2012 NAICS. For these establishments, the agency extracts sales, end-of-year inventories, payroll, employment, name and address information, wholesale type of operation code (TOC), primary identifiers, and associated employer identification numbers (EINs).

For the AWTS, the Census Bureau uses the TOC to distinguish between three different types of wholesale establishments:

- merchant wholesale establishments, excluding manufacturers' sales branches and offices;

- manufacturers' sales branches and offices; and

- agents, brokers, and business electronic markets.

To create the sampling units, the agency sums the establishment data for all wholesale establishments associated with the same firm identifier. In some cases, a multi-unit firm has establishments active in more than one wholesale TOC. In these situations, the Census Bureau creates firm-level sampling units for each type of operation. No aggregation is necessary to put

single-unit establishment information on a firm basis. Thus, the sampling units created for single-unit firms simultaneously represent establishment and firm information.

Stratification

The sample for the AWTS consists of three separate samples (one for each wholesale TOC). This survey uses a stratified, one-stage design with primary strata defined by industry. There are 83 primary strata: 56 from the merchant wholesale establishments, excluding manufacturers' sales branches and offices, sample; 19 from the manufacturers' sales branches and offices sample; and 8 from the agents, brokers, and business electronic markets sample. For the AWTS, the Census Bureau further stratifies the sampling units within each industry group by a measure of size related to their annual sales (i.e., substratify). Sampling units expected to have a large effect on the precision of the sales and end-of-year inventories 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, the agency determines a substratum boundary (or cutoff) that divides the certainty units from the noncertainty units. These cutoffs are based on a statistical analysis of data from the 2012 Economic Census. Accordingly, these values are on a 2012 sales and end-of-year inventories basis. For the AWTS, the Census Bureau also uses this analysis to determine the number of size substrata for each industry stratum. Furthermore, the analysis is used to set preliminary sampling rates needed to achieve specified sampling variability constraints on sales and end-of-year inventories 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 process is to identify firms that will be selected with certainty. If a firm's annual sales or end-of-year inventories are greater than the corresponding certainty cutoff, the agency includes that firm in the AWTS sample with certainty.

All firms not selected with certainty are subjected to sampling. To be eligible for the initial sampling, a firm had to have nonzero payroll in 2014 (based on the BR as of October 2015). The Census Bureau stratifies the firms according to their major industry and their estimated sales (on a 2012 basis). Within each noncertainty stratum, the agency selects a simple random sample of firms without replacement. This process is conducted separately for each wholesale TOC. Noncertainty firms from the previous sample are placed in a nonselection category for the new sample if their exclusion would not bias the sample. This is done to reduce burden on these companies. Overall, the effective sampling rate for the small- and medium-sized firms is approximately 2% of the universe.

For the AWTS, a new sample is selected approximately every five to seven years. The AWTS introduced a new sample for survey year 2016. The survey will continue to use this sample throughout this three-year clearance period. Sample sizes are computed to meet multiple coefficient of variation constraints on estimated annual sales totals and end-of-year inventory totals. Constraints are specified at detailed industry levels and at broad industry levels (e.g., durable goods, nondurable goods) up to the total wholesale level. The initial sample consisted of

roughly 8,400 firms (6,700 merchant wholesalers, excluding manufacturers' sales branches and offices; 1,000 manufacturers' sales branches and offices; and 700 agents, brokers, and business electronic markets). This count changes over the course of a sample due to sample maintenance activities, which are discussed in the Sample Maintenance section below.

Sample Maintenance

The sample is updated on a quarterly basis to represent EINs issued since the initial sample selection. These new EINs, called births, are EINs recently assigned by the Internal Revenue Service (IRS) that have an active payroll filing requirement on the IRS Business Master File (BMF). An active payroll filing requirement indicates that the EIN is required to file payroll for the next quarterly period. The Social Security Administration (SSA) attempts to assign an industry classification to each new EIN. EINs with an active payroll filing requirement on the IRS BMF are said to be "BMF active," and EINs with an inactive payroll filing requirement are said to be "BMF inactive." The Census Bureau samples EIN births 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 Service Annual Survey (SAS), the Annual Retail Trade Survey (ARTS), or the AWTS. Additionally, it must meet certain criteria regarding its quarterly payroll. In the first phase, the agency stratifies births by broad industry groups and a measure of size based on quarterly payroll. A relatively large sample is drawn and canvassed to obtain a more reliable measure of size (consisting of revenue in two recent months) and a new or more detailed industry classification code. Births are contacted by telephone if they have not returned their electronic questionnaire within 30 days. Using this more reliable information, in the second phase, the Census Bureau subjects the selected births from the first phase to probability proportional-to-size sampling with overall probabilities equivalent to those used in drawing the initial AWTS sample from the October 2015 BR. Because of the time it takes for a new employer firm to acquire an EIN from the IRS and the time needed to accomplish the two-phase birth-selection procedure, the Census Bureau adds births to the sample approximately nine months after they begin operation.

Similarly, each quarter the agency checks the BR to determine if any EINs on the survey have become BMF inactive. Typically, the Census Bureau does not canvass BMF inactive EINs during the reference year. Likewise, if any EIN on the survey was BMF inactive in a previous reference year or was part of an inactive sampling unit in the survey and is now currently BMF active on the BR, it is included in the canvass again. In both cases, the AWTS only tabulates data for the portion of the reference year that these EINs reported payroll to the IRS.

The sample is also updated to reflect mergers, acquisitions, divestitures, splits, and other changes to the business universe throughout the year. (The Census Bureau learns about these changes through a variety of means, including directly from respondents, and updates the sample accordingly as soon as the information is obtained. For example, the AWTS asks about organizational changes on each of its questionnaires. When the Census Bureau analyst in charge of a company becomes aware of a structural change, he or she will immediately take the appropriate sample maintenance actions to account for this change.) In general, any new establishments that a firm acquires, even if under new or different EINs, are included in the sample with the same sampling status as the original firm (i.e., with the same initial sampling

weight). For noncertainty firms, additional evaluation may be done in some instances to determine the feasibility of adding the new establishments (by evaluating the effect of the new establishments on the industry estimates).

Response Rates

The unit response rate is defined as the percentage of active reporting units in the statistical period, based on unweighted counts, that were eligible for data collection or of unknown eligibility that responded to the survey. It is an indicator of the performance of data collection for obtaining usable responses. The AWTS calculates and monitors the unit response rate during each processing cycle. To be classified as a response, the respondent for the reporting unit must have provided sufficient data, and the data must satisfy all the edits. Specifically, to be considered a respondent for the AWTS:

- merchant wholesalers, excluding manufacturers' sales branches and offices, must provide sales, end-of-year inventories, purchases, or total operating expenses.

- manufacturers' sales branches and offices must provide sales, end-of-year inventories, or total operating expenses.

- agents, brokers, and business electronic markets must provide commissions, sales, sales on own account, gross selling value, or total operating expenses.

Since the current sample was launched in the 2016 survey year, there have been three AWTS releases (i.e., the 2016 AWTS, 2017 AWTS, and 2018 AWTS). Presently, the 2019 AWTS data are still being collected and processed. The reporting instrument for the 2019 AWTS will be deactivated on October 14, 2020. The table below lists the unit response rate for NAICS code 42 (i.e., the entire wholesale trade sector) each year. (For the 2016 AWTS, 2017 AWTS, and 2018 AWTS, the value represents the unit response rate when the data were released to the public for that survey year. For the 2019 AWTS, it is the value as of July 31, 2020.)

2019[1]	2018	2017	2016
66.7%	71.3%	71.1%	74.9%

[1]As of 7/31/20 (collection ends on 10/14/20)

For the AWTS, the Census Bureau also computes and monitors total quantity response rates. This type of response rate is defined as the percentage of the estimated (weighted) total of a given data item reported by the active tabulation units in the statistical period or from sources determined to be equivalent-quality-to-reported data. The total quantity response rate is an item-level indicator of the "quality" of each estimate. In contrast to the unit response rate, these weighted response rates are computed for individual data items. Thus, there are several total quantity response rates per statistical period and release. The total quantity response rate is a weighted measure that takes the size of the tabulation unit into account as well as the associated sampling parameters. To compute the total quantity response rate for a particular estimate, it is necessary to determine the source of the final tabulated value of the associated data item for each

tabulation unit. This value could be directly obtained from respondent data, indirectly obtained from other reported-equivalent data sources, or imputed.

Question 2. Procedures for Collecting Information

In addition to the following information, please refer to the Sampling Frame, Stratification, Sample Selection, and Sample Maintenance segments of the Universe and Respondent Selection section above.

Data Collection

The initial letter to respondents identifies the survey name, due date, instructions for accessing the survey, authority for collection, and burden estimate. It also provides a telephone number for those needing assistance. The initial and follow-up mailings instruct respondents to access the Census Bureau's Respondent Portal (https://portal.census.gov). The Census Bureau will also periodically send an e-mail reminder, or eBlast, to respondents. These e-mail reminders also provide the same instructions. A telephone follow-up is scheduled for nonrespondents, too. Once the user arrives at https://portal.census.gov, the individual will create a new account (i.e., register) or sign in to an existing account. The Respondent Portal allows users to add and view all current Census Bureau surveys assigned to them. Using the Respondent Portal, respondents can access the online reporting system called Centurion (by clicking "Report Now" for the appropriate survey). For the AWTS, electronic reporting is the only advertised reporting option. Nevertheless, the Census Bureau will accept data submitted through other methods. For example, if a company does not have access to the Internet, the Census Bureau can arrange for the business to provide its data to an analyst via telephone.

The quarterly birth/death process, removal of out-of-scope cases, and restructuring of reporting units to account for organizational changes (e.g., mergers, acquisitions, etc.) will inevitably cause the number of respondents to fluctuate over time. Historically, the number of respondents typically decreases each year within a given sample. This trend has held true for this sample so far (survey years 2016, 2017, 2018, and 2019). The table below shows the number of respondents at the time of the initial mailing for the current sample by year.

2019	2018	2017	2016
7,629	7,960	8,417	8,898

Analysis

Census Bureau analysts review the company-level and industry-level data for accuracy. The analysis performed includes (but is not limited to) comparing a company's data for the current year against its historical data, investigating differences between the industry-level estimates generated from the Monthly Wholesale Trade Survey (MWTS) and the AWTS, and reviewing records that are displaying unusual behavior (e.g., operating expenses that are greater than sales). When analysts are unable to validate certain information through their own research, they sometimes have to contact respondents to obtain an explanation.

Nonresponse

The Census Bureau imputes data using survey data and administrative data for unit nonresponse, item nonresponse, and for responses that fail computer or analyst edits.

Estimation Procedures

Total estimates are computed using the Horvitz-Thompson estimator (i.e., the sum of weighted reported or imputed data) 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 AWTS sample. These estimates are then adjusted through a benchmarking process. The Census Bureau estimates variances for published statistics using the method of random groups.

Linking Samples

As previously mentioned, the current sample was introduced with the 2016 AWTS. This sample is designed to produce estimates based on the 2012 NAICS. All published estimates from the 2015 AWTS were restated from 2007 NAICS definitions to 2012 NAICS definitions. (Definitions changed for NAICS 4236 and 4237.)

In order to maintain the time series for each industry, an operation is performed to link estimates from the prior and new samples. For the linking operation to occur, two years of data were collected (2015 and 2016) from units in the new sample.

Sales estimates from the new sample for reference year 2015 and subsequent years are linked to the restated prior sample estimates by multiplying the Horvitz-Thompson estimates from the new sample by a ratio. The ratio is calculated as follows:

- The numerator is the 2015 published, census-adjusted (based on the 2012 Economic Census) sales estimate for the industry restated on a 2012 NAICS basis from the prior sample.

- The denominator is the 2015 Horvitz-Thompson sales estimate for the industry on a 2012 NAICS basis from the new sample.

The resulting sales estimates (called "modified" sales estimates) are implicitly benchmarked to 2012 Economic Census results via this linking procedure.

The following method is used to produce modified estimates for end-of-year inventories, purchases, gross selling value, sales on own account, and e-commerce. First, the sales ratio described above is multiplied by the Horvitz-Thompson estimate for the given item for

2015 and subsequent years. Then, the published estimates for 2007 through 2015 from the prior sample are input into the benchmarking program. Using this program, the estimates for 2008 through 2015 for each detailed industry are revised in a manner that:

- uses the benchmarked estimate for 2007 from the prior sample as a constraint, resulting in no revision to the 2007 estimate.

- uses the modified estimate for 2015 from the new sample as a constraint.

- minimizes the sum of squared differences between the year-to-year changes of the input and revised estimates for 2008 through 2015.

- for agents and brokers, rakes the gross selling value and sales on own account to the sales estimate total for 2008 through 2015.

A similar method is used for total operating expenses, only using the benchmarked estimate for 2012 from the prior sample as a constraint instead of 2007.

For agents, brokers, and business electronic markets, to ensure consistency with total sales, the benchmarked gross selling value and the sales on own account are raked to the benchmarked total sales estimate. This is done by calculating the proportion of gross selling value and sales on own account to total sales and then applying each of these proportions to the benchmarked total sales to get the corresponding benchmarked estimates for gross selling value and sales on own account.

Modified estimates at aggregate industry levels are computed by summing the modified estimates for the appropriate detailed industries comprising the aggregates.

The AWTS estimates are also benchmarked to the 2017 Economic Census results.

Benchmarking

Results of the 2017 Economic Census are used to benchmark the AWTS estimates. Due to the uncertainties of a new sample, some details described herein could change.

Sales estimates are input to the benchmarking program and are revised in a manner that:

- uses the 2017 Economic Census sales total as a constraint, along with the existing 2012 modified sales estimate, which is already linked to the 2012 Economic Census.
- minimizes the sum of squared differences between the year-to-year changes of the input and revised estimates for 2013 through 2019.

The process is applied separately to merchant wholesalers except manufacturers' sales branches and offices and to manufacturers' sales branches and offices. The same process is applied to agents, brokers, and business electronic markets (NAICS 425) using sales estimates defined as

gross selling value plus sales on own account. The estimates output from this operation are referred to as "benchmarked."

A similar method to the one for adjusting sales is used to adjust estimates for inventories, purchases, operating expenses, and e-commerce. Each of these items are revised in the following manner:

- 2012 and 2017 modified estimates are multiplied by the ratio of benchmarked sales divided by modified sales for the same year.
- Modified estimates for each item are input into the benchmarking program using the two constraints calculated above.
- The benchmarking program minimizes the sum of squared differences between the year-to-year changes of the input and revised estimates for 2013 through 2019.

For agents, brokers, and business electronic markets, to ensure consistency with total sales, the benchmarked gross selling value and the sales on own account are raked to the benchmarked total sales estimate. This is done by calculating the proportion of gross selling value and sales on own account to total sales and then applying each of these proportions to the benchmarked total sales to get the corresponding benchmarked estimates for gross selling value and sales on own account.

For commissions, the 2013 through 2019 modified commissions estimates are multiplied by the ratio of benchmarked gross selling value divided by modified gross selling value for the same year.

Benchmarked estimates at aggregate industry levels are computed by summing the benchmarked estimates for the appropriate detailed industries comprising the aggregates, and benchmarked estimates for merchant wholesalers are computed by summing the benchmarked estimates for manufacturers' sales branches and offices and merchant wholesalers except manufacturers' sales branches and offices.

Quality Suppressions

Estimates can be suppressed from publication for quality reasons. In particular, an estimate with a coefficient of variation greater than 30 percent, with a total quantity response rate less than 50 percent, or with other concerns about data quality is suppressed from publication (unless the estimate has consistently been published for prior years and the coefficient of variation and total quantity response rate are acceptably close to the thresholds). A suppressed estimate and its corresponding measure of sampling variability are replaced with an "S" in the published tables when this occurs. For a description of the Census Bureau's Standards for Releasing Information Products, see https://www.census.gov/about/policies/quality/standards/standardf1.html.

Disclosure Avoidance

Disclosure is the release of data that reveals information or permits deduction of information about a particular survey unit through the release of either tables or microdata. Disclosure

avoidance is the process used to protect each survey unit's identity and data from disclosure. The AWTS uses cell suppression for disclosure avoidance.

Cell suppression is a disclosure avoidance technique that protects the confidentiality of individual survey units by withholding cell values from release and replacing the cell values with a symbol ("D"). If the suppressed cell values were known, one would be able to estimate an individual survey unit's data too closely.

The cells that must be protected are called primary suppressions. To make sure the cell values of the primary suppressions cannot be closely estimated by using other published cell values, additional cells may also be suppressed. These additional suppressed cells are called complementary suppressions.

The process of suppression does not usually change the higher-level totals. Values for cells that are not suppressed remain unchanged. Before the Census Bureau releases data, computer programs and analysts ensure primary and complementary suppressions have been correctly applied.

Question 3. Methods to Maximize Response

Respondents are notified of their obligation to report via a mailed letter. This letter provides a telephone number that companies can call if they have any questions or concerns. They receive at least 30 days to complete the AWTS. If needed, businesses can request time extensions through the Respondent Portal. They can send secure messages to the Census Bureau using the Respondent Portal, too.

Approximately four weeks after the initial mailing, the Census Bureau provides nonrespondents with a due date reminder letter. Roughly one week after this, a due date reminder e-mail is transmitted to companies that have not reported still. Once the due date has passed, delinquent companies receive their first follow-up letter. This occurs about three weeks after the due date reminder e-mail and about two weeks after the due date. About two weeks after the follow-up letter is distributed, the first follow-up e-mail is sent to nonresponse cases. Roughly three weeks later, outstanding companies receive their second follow-up letter. This letter is provided via certified or priority mail. The Census Bureau sends the second follow-up e-mail to delinquent companies roughly four weeks after this. The agency then attempts to reach the nonrespondents via telephone about one week after the second follow-up e-mail. The telephone follow-ups are conducted over a period of approximately four weeks. About one week after the telephone follow-ups conclude, the third e-mail follow-up is sent. Then (roughly two weeks later), a third follow-up letter is delivered to respondents. Delinquent companies receive their fourth follow-up e-mail approximately three weeks later. The fifth follow-up e-mail is transmitted roughly four weeks after the fourth follow-up e-mail.

Throughout the year, the Census Bureau monitors the AWTS response rates (previously discussed in the Universe and Respondent Selection section) to ensure publication standards are satisfied. If the response rates fail to meet publication requirements, additional follow-ups, such as analyst phone calls or follow-up e-mails, may be performed.

As discussed in Part A (Project Schedule section), the Odyssey companies, which are a subset of the AWTS respondents, do not share the same follow-up schedule as the cases outlined above. Nevertheless, the general follow-up approach will be the same for these companies (i.e., letters, e-mails, and telephone calls used throughout the survey year to acquire responses).

Samples of some of the correspondence mentioned in this section can be viewed in Attachment 2. (Please note that the exact language on some of these samples is subject to change.)

Question 4. Tests of Procedures or Methods

The Census Bureau checks the procedures employed in every phase of the AWTS – from mailout and data capture to data editing and publication. Employees test the systems used to ensure they are programmed as expected (i.e., satisfy the provided specifications). The staff also regularly reviews the guiding logic behind various aspects of the AWTS to confirm that logic is still relevant in the current economy.

The AWTS will continue to use the same sample and collect virtually the same information as the previous three-year clearance window. The information received from respondents has not revealed issues with the questionnaires. Regardless, the staff still checks the worksheets for accuracy, grammar, and layout each year. During this three-year window, Centurion will still be the online reporting tool for respondents. This system is tested internally each survey cycle to make sure the correct content is displayed, branching occurs properly, etc. Additionally, the staff ensures the data are transferring to the internal processing system correctly. The data edits, imputation, and estimation are also tested annually to confirm they are functioning properly. The staff reviews the output from the AWTS publication tables program to determine if the tables are displaying the correct values and suppressions/disclosures. The general methodology employed for this three-year period will also be the same as the methodology used in the previous three-year clearance window.

Question 5. Contacts for Statistical Aspects and Data Collection

For questions regarding the planning and implementation of this survey, please contact John Dougherty, Chief of the Wholesale Trade Branch (Economy-Wide Statistics Division, U.S. Census Bureau), at 301-763-8936 or John.Dougherty@census.gov.

For inquiries about the methodology for this survey, please contact Deanna Weidenhamer, Chief of the Retail and Wholesale Surveys Statistical Methods Branch (Economic Statistical Methods Division, U.S. Census Bureau), at 301-763-7186 or Deanna.L.Weidenhamer@census.gov.

Attachments to the Supporting Statement

1. Bureau of Economic Analysis Support Letter

2. Initial Letters and Follow-up Correspondence Examples

3. Centurion (Reporting Instrument) Screenshots

4. AWTS Worksheets

5. Legal Authority for the AWTS - Title 13, United States Code