Stratification

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 2007 Economic Census. Accordingly, these values are on a 2007 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 Annual Wholesale Trade Survey sample with certainty. Any multiunit firm that had activity inscope to Monthly Wholesale Trade Survey (i.e., merchant wholesalers excluding Manufacturers Sales Branches and Offices) and was selected as certainty in Annual Wholesale Trade Survey was also selected into the Monthly Wholesale Trade Survey sample with certainty.

All firms not selected with certainty were subjected to sampling on an Employer Identification Number basis. If a firm had more than one Employer Identification Number, we treated each of its Employer Identification Numbers as a separate sampling unit. To be eligible for the initial sampling, an Employer Identification Number had to have nonzero payroll in 2009. The Employer Identification Numbers were stratified according to their major industry and their estimated sales (on a 2007 basis). Within each noncertainty stratum, a simple random sample of Employer Identification Numbers was selected without replacement. This process was done separately for each wholesale type of operation. The selected noncertainty Employer Identification Numbers were divided into two approximately equal groups. For merchant wholesalers excluding Manufacturers Sales Branches and Offices, one group is canvassed for both the monthly and the annual survey, the other group is canvassed for only the annual survey. The Manufacturers Sales Branch or Office and Agents and Brokers samples are only canvassed in the annual survey. However, EINS selected into the Manufacturers Sales Branch or Office or Agent and Broker sample are included in the Monthly Wholesale Trade Survey sample if that Employer Identification Number had activity inscope to Monthly Wholesale Trade Survey.

Sample Maintenance

Periodically, we update the samples to represent new Employer Identification Numbers appearing on the Business Register. These new Employer Identification Numbers, called births, are Employer Identification Numbers 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).

Employer Identification Number 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 Annual Wholesale Trade Survey, the Annual Retail Trade Survey (ARTS), 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 Annual Wholesale Trade Survey and Monthly Wholesale Trade Survey samples from the Business Register. Because of the time it takes for a new employer firm to acquire an Employer Identification Number 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. The births selected for the Monthly Wholesale Trade Survey sample are a subset of the births selected for the Annual Wholesale Trade Survey sample.

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 Employer Identification Numbers, 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 Employer Identification Number are included in the sample with certainty; any new Employer Identification Numbers that might later be associated with that firm are subjected to sampling through the quarterly birth-selection procedure.

Single-unit Employer Identification Numbers 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 Employer Identification

Number is contacted, and if a successor Employer Identification Number is found, it is added to the survey. For both inactive and reactivated Employer Identification Numbers, data are tabulated for only the portion of the reference year that these Employer Identification Numbers reported payroll to the IRS.

Because births are not represented in the monthly survey until they go through the two-phase selection procedure, an interim procedure is used to account for births during the period between the onset of activity and the time of birth selection. This interim procedure consists of imputing data for all Employer Identification Numbers currently in the monthly survey that go out-of-business, but are still on the IRS mailing list.

Births are added to the monthly survey in February, May, August, and November of each year. At the same time, deaths are removed from the survey. To minimize the effect of births and deaths on the month-to-month change estimates, we phase-in these changes by incrementally increasing the sampling weights of the births and decreasing the sampling weights of the deaths in a similar fashion. In the first month, we tabulate the births at one-third their sampling weight and tabulate the deaths at two-thirds their sampling weight. In the second month, we tabulate the births at two-thirds their sampling weight and tabulate the deaths at one-third their sampling weight. In the third month, we tabulate the births at their full sampling weight and the deaths are dropped (sampling weight equal zero).