Improving the Job Openings and Labor Turnover Survey's Sampling Procedure October 2009

Sarah Goodale, Darrell Greene
Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington, DC 20212
goodale.sarah@bls.gov, greene.darrell@bls.gov

Abstract
The Job Openings and Labor Turnover Survey (JOLTS), conducted by the U.S. Bureau of Labor Statistics, is a panel sample of 16,000 establishments, in which participants remain in the sample for two consecutive years. Every year a sample of approximately 8,000 establishments is drawn from the 1st quarter data from the Longitudinal Database. Each month one new panel is rolled into the sample while the oldest panel in the sample is rolled out. Previously, each annual sample was weighted to the current sample frame while previous samples remained unchanged. There was also no consideration made for younger units, which are thought to be more dynamic and should be included in the sample as early as possible. With the new sampling procedures, all establishments will be updated to reflect the current frame, and quarterly birth samples will be taken to update JOLTS with younger establishments.

Key Words: panel sample, post stratification, re-weight, birth sample, birth establishment

1. JOLTS Sampling Procedure Prior to April 2009

1.1 Description of Sampling Procedures

The Job Openings and Labor Turnover Survey (JOLTS) is an establishment survey conducted by the Bureau of Labor Statistics (BLS) since December 2000. Each month JOLTS produces estimates on the nation's job openings, hires, and separations (quits, layoff and discharges, and other separations) rates.

JOLTS is a stratified random sample that is selected without replacement. The JOLTS strata are defined by 4 census regions, 20 industry classifications, and 6 employment size classes. Each year a sampling frame is developed from the Quarterly Census of Employment and Wages (QCEW) using 1st quarter establishment data, located on the BLS's Longitudinal Database (LDB). A sample of 16,000 non-farm establishments is allocated using the Neyman allocation, based on the QCEW employment, with a condition that the number of establishments selected per stratum be a multiple of 24. The establishments selected for JOLTS are weighted to the sampling frame.

Equation 1: JOLTS Neyman Allocation

$$(1) \quad n_h = \frac{N_h S_h}{\sum\limits_{i=1}^{H} N_i S_i} \text{ where stratum h = 1, 2,..., H; Note: } n_h \text{ must be a multiple of 24}$$

Equation 2: JOLTS Sample Weight

$$(2) \quad weight_h = \frac{N_h}{n_h}$$

All large establishments are selected with certainty (a probability of 1.00) and are placed in a certainty panel. Also selected with certainty are the establishments in the strata that contain at most 24 establishments. The remaining establishments selected as a part of the JOLTS sample are randomly placed into the 24 non-certainty panels. Each month one non-certainty panel rolls into estimation while another non-certainty panel rolls out. Each non-certainty panel will be used in estimation for 24 months, meaning an establishment will be a part of JOLTS for 2 years. Since only one non-certainty panel is rolled in each month only 12 panels will be rolled in before the next annual sample is drawn. However, all sampling procedures are done on the basis of 24 panels.

All the non-certainty panels of the annual sample are stratified random samples with identical composition. Therefore, one non-certainty panel may represent the entire sample (with mathematical adjustment). By doing this, the number of establishments or the sampling weights should not vary greatly each sampling year. Little changes between the annual samples is desirable because with panel rotation, most of the time there are three different annual samples present in JOLTS estimation.

Each month one new non-certainty panel is used in JOLTS estimation with the older 23 non-certainty panels. During the course of a sampling year only 12 new non-certainty panels are added to JOLTS estimation. The remaining non-certainty panels come from previous annual samples. There can be up to three different annual samples present in JOLTS estimation at a given time. Example 1 below illustrates this point looking at the first month a new annual sample is rolled into JOLTS estimation.

Example 1

A	B	B	B	B	B	B	B	B	B	B	B	B	C	C	C	C	C	C	C	C	C	C	C

A = the current annual sample
B = the annual sample from a year ago
C = the annual sample from two years ago

Each month a new panel from the annual sample will roll into the estimation sample, replacing an older panel from the prior year's sample, or a panel from sample A will roll in and replace a panel from sample C. With the 12th month of the annual sample rolling into the estimation sample the 24 panels will have this distribution:

A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B

Then a new annual sample is selected and the cycle starts again with a new panel from the new annual sample being introduced into the estimation sample each month. A panel stays within the JOLTS estimation sample for 24 months.

1.2 Issues within the Sampling Procedure

This section discusses the major issues associated with the JOLTS sampling procedures. These issues can be divided into three parts:

1) The previously selected annual samples are not updated
2) The age of an establishment is not taken into account
3) The amount of time that lapses for an establishment when it appears on the sampling frame until it is used in JOLTS estimation

1.2.1 Previously selected annual samples are not updated
The annual sample reflects the sampling frame from when it was selected, in terms of its strata definitions and the sampling weights for the establishments. However, since there are as many as three different annual samples at a given time in JOLTS estimation, there are up to three different sampling frames represented. Currently, there are no procedures to update the older sampled establishments. Once an establishment (panel) enters the JOLTS estimation, no changes are made. If the establishment's strata information changes (changes industry classification or employment class size) or the establishment goes out-of-business, the change is not reflected in the samples.

The panels are designed to minimize structural changes from year to year. However, establishments from the previous annual samples may move to different strata due to either employment growth or a change in industry classification. Since there are no updates within the prior samples, the sampling weights remain constant. This means that the establishments in estimation will weight to different sampling frames most of the time.

1.2.2 Age of an establishment
The stratification of establishments takes into consideration, the six employment size class, the 20 2-digit industry classification, and the four census regions. Research has shown that the age of the establishments should also be considered (Crankshaw, 2008).

Younger establishments are more dynamic than the older establishments in their churn rates (hires + separations). Since JOLTS estimates job openings, hires, and separations, JOLTS should distinguish between the younger establishments and older establishments.

1.2.3 Amount of time lapse
From the time an establishment's information is referenced on the sampling frame until it can first be used in estimation, over a year has lapsed. It takes about 9 months for the establishment's information to be usable on the LDB. There is also an additional 3-4 months from when an establishment is selected to be a part of the annual sample until the first panel rolls into JOLTS estimation. Finally, it takes a year for all the panels of the annual sample to be used in JOLTS estimation. The amount of time lapsed from the sampling frame until the establishment is used in JOLTS estimation is anywhere from a year to two years.

2. JOLTS Sample Improvements

The purpose of the JOLTS sample improvements are to update the previous annual samples, take the age of establishments into consideration, and to decrease the time between the sample frame and estimation. To improve the JOLTS sample, all establishments to be used in estimation need to be considered, not only the annual sample.

2.1 Addressing the Issues

2.1.1 Update previous samples
 The current annual sample is selected from all in-scope establishments on the sampling frame for the JOLTS survey. Therefore, to update the previously selected annual samples, the establishments previously selected are compared to the current sampling frame. By doing this, the stratum information will reflect the current sampling frame. Also, all out-of-business and out-of-scope establishments will be identified and removed from JOLTS estimation. Many establishments that the data collection center is unable to contact are coded as non-respondents, but in actuality they are out-of-business. Proper coding and removal of the out-of-business establishments results in a more accurate response rate. Updating the establishments used in estimation should allow for a more efficient and accurate sample, but requires a sampling weight adjustment, which will be discussed later.

2.1.2 Age of establishments and time lapse
Next the age of the establishments is examined. Research has shown that the younger establishments are more dynamic than older establishments in the churning rate (hires + separations) of the labor force (Crankshaw, 2008). Since this information is important for JOLTS to capture, the age of the establishments needs to be taken into consideration. To include the age of the establishment, the establishments used in JOLTS estimation and the annual sample will be post-stratified by an age variable. The age of an establishment will be determined by the date of its first positive employment. JOLTS now has three classifications of the age variable: 0 (less than a year old), 1 (greater than a year old but less than two years old), and 2 (two years old and older). Since the annual sample is drawn from the QCEW data, the date used as the cut-off for the age variable classification will be the last month of the quarter (which is currently March of the sampling year).

Post-stratification of the establishments illustrates the following distribution: only the current annual sample contains establishments with an age of 0, 1, or 2. The annual sample selected the previous year has only establishments with an age of 1 or 2. The annual sample selected two years ago has only establishments with an age of 2. Therefore, there are far more age 2 establishments in the entire sample than there are of younger establishments.

When the annual sample for the full 24 panels is selected, it contains the proper distribution of establishments by age in comparison to the frame. After selection, the younger establishments are distributed randomly and evenly among the panels. Only 12 of the panels will be used during the year for JOLTS estimation. There is at least a year lapse in time from when the establishments are a part of the sampling frame until they are introduced into the JOLTS estimation. Since the younger establishments have a more dynamic churn rate, the earlier the information is obtained from these establishments to be used in JOLTS, the better. However, there is a 6 to 9 months delay from the time the information is reported to QCEW until it is available for use. Then there is another 3 month or more delay until the establishment can be enrolled into JOLTS estimation. Finally, if the younger units are in one of the later panels then almost another 12 months has passed. At that point, almost two years may have passed since the younger establishment was identified and valuable information may be missed. To improve the coverage of the younger establishments in JOLTS estimation, these units are moved into the first non-certainty panel of the new annual sample, so that they are used in JOLTS estimation earlier. The drawback is that by moving all the younger establishments of the annual sample to the front panels, the panels will no longer have identical compositions.

To improve the distribution of the younger establishments in estimation and to ensure that the data is obtained in a timely manner, a quarterly birth sample will be drawn. The quarterly birth sample will contain only establishments that are new in that quarter and fulfill the JOLTS in-scope conditions. The establishments selected as part of the quarterly birth sample will be added to the end of those quarters' panels from the current annual sample. This allows the younger units to be enrolled in the estimation sample earlier. Also, the younger units will be more evenly distributed throughout the estimation sample, since each quarter (excluding quarter in which the annual sample is drawn) will have a quarterly birth sample.

To include the age variable and to update the previously selected establishments in JOLTS, all the eligible establishments for the current year's JOLTS estimation must be re-weighted. The age of the establishments is included by post-stratifying. Since the younger units (age = 0 and age = 1) have fewer establishments than the older establishments, they will be post-stratified by age/industry classification/employment class size. The older establishments (age = 2) will be post-stratified by age/region/industry classification/employment class size.

Since all estimates in JOLTS are produced by using 24 non-certainty panels, the re-weights also need to be based on 24 panels. However, combining the panels used in current JOLTS estimation and the annual sample produces 36 panels. The age distribution can be used to determine which panels to use in the re-weight. In Example 2, as in the Example 1, sample A is the annual sample, while sample B is the annual sample selected a year ago, and sample C is the annual sample selected two years ago.

Example 2

Sample C:

| 2 | 2 | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* |

Sample B:

| *2/1* | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* | *2* |

Sample A: (Note: The Age=0 and Age=1 can be in one or two panels, for the example they are put into two panels)

| *0* | *1* | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

When computing the re-weights, the 24 panels used are the panels that are bolded, since the age of the establishments in the remaining panels are all the same. The weights (re-weights) are then computed by the post-stratification:

Equations 3 – 5: JOLTS Re-weights

$$(3) \quad weight_{age=0,industry,size} = \frac{N_{age=0,industry,size}}{n_{age=0,industry,size}}$$

$$(4) \quad weight_{age=1,industry,size} = \frac{N_{age=1,industry,size}}{n_{age=1,industry,size}}$$

$$(5) \quad weight_{age=2,region,industry,size} = \frac{N_{age=2,region,industry,size}}{n_{age=2,region,industry,size}}$$

2.2 Supplemental Sample Programs

The original JOLTS sampling procedures were not changed, to implement the changes to the sampling procedure. Instead, new supplemental programs were developed.

2.2.1 Annual Sample Update

The first supplemental program is the sample update program in which all the establishments currently used in JOLTS estimation are updated to reflect the current sampling frame. The younger establishments of the annual sample are moved to the first panels to rotate into JOLTS estimation. All the establishments to be used in JOLTS estimation are post-stratified to incorporate the age variable. Finally, all the establishments are re-weighted.

2.2.2 Quarterly Birth Sample

The quarterly birth sample is the supplemental program to add a sample of the quarter's new establishments into JOLTS estimation. In the quarterly birth sample, a quarterly birth frame is built from that quarter's QCEW data. Since most birth establishments will be from the lower employment size classes, the quarterly birth sample is limited to those establishments in employment size classes 1 (0-9 employees), 2 (10-49 employees), and 3 (50-249 employees). To determine the number to be sampled, the re-weight values of the corresponding strata are used:

Equation 6: Birth Sample Size Determination

$$(6) \; n_{birth,age,industry,size} = \frac{N_{birth,age,industry,size}}{weight_{age=0,industry,size}}$$

If the n(birth) includes a decimal amount, a rounding function is employed. Any decimal amount 0.5 or greater will be rounded to the next highest integer. To ensure that if a stratum has birth establishments available for selection a minimum amount to be selected is set. At least 3 birth establishments should be selected from every stratum when possible. The new quarterly birth sample is given a weight:

Equation 7: Birth Sample Weight

$$(7) \; weight_{birth} = \frac{N_{birth,age,industry,size}}{n_{birth,age,industry,size}}$$

Finally the new quarterly births are assigned their JOLTS identification numbers and appended to that quarter's panels.

3. Conclusion

The new JOLTS sampling procedures were implemented during the 2009 annual sample. The first panel of the new annual sample was used in JOLTS estimation starting in April 2009. The impact of the new sampling procedures is unknown as of yet. The full impact of the sampling procedures will be an area of further study and research including: How are the estimates affected? If so, is it the desired outcome?

Further study should also be performed on the younger/birth establishments. Are these establishments being captured early enough to collect the desired information? Also, what would happen if the JOLTS estimation procedure was applied to only the younger/birth establishments? Would there be a difference in the estimates? If so, would the difference be enough to justify the addition of the quarterly birth samples?

The changes to the sampling procedures were implemented so that the original procedures remained intact, but to correct for the issues perceived. Should the new procedures implemented remain or does the entire sampling system need to change? It is possible to update the JOLTS sample on a quarterly basis, so that instead of using 12 panels of the annual sample, use 3 panels from each quarterly sample. This would change the sample weights every quarter. However, it would be difficult to know if there were changes in the JOLTS estimates whether it was from a weight change or an economic change. Through simulation, the use of a quarterly sample could be examined.

4. References

Crankshaw, Mark and Stamas, George (2000), "Sample Design in the Job Openings and Labor Turnover Survey," In *JSM Proceedings,* Survey Research Method Section. Alexandria, VA: American Statistical Association. 371 - 376

Crankshaw, Mark (2008), "Simulating JOLTS Hires and Separations Data Using the LDB," Bureau of Labor Statistics Internal Documentation

JOLTS Program (2009), "Improving JOLTS Methodology," Bureau of Labor Statistics JOLTS website
http://www.bls.gov/jlt/methodologyimprovement.htm

Lohr, Sharon (1999), "Sampling: Design and Analysis," Brooks/Cole Publishing Company