

2007 Survey of Business Owners (SBO)

Public Use Microdata Sample (PUMS)

Data Users Guide

Technical Documentation

Issued August 2012

1. Introduction

This document describes the Public Use Microdata Sample (PUMS) file created for the 2007 Survey of Business Owners and Self-Employed Persons (SBO). The document provides a brief description of the SBO and the differences between the SBO data and the records available in the PUMS file. The document also explains how to determine company-level characteristics from the PUMS and how to create estimates and measures of sampling variability from the PUMS data. For a complete description of the SBO and access to the published tables, please visit the SBO website (<http://www.census.gov/econ/sbo/>).

2. SBO Background

The SBO collects information on the gender, ethnicity, race and veteran status of business owners. The SBO additionally collects characteristics of businesses and business owners such as the sources of capital used to start the business and the age of the business owner. The SBO produces estimates for a variety of industry classifications, geographic levels, and cross-tabulations (for example, state by sector and race by gender). The PUMS file allows data users greater access to these data than standard tabulations and provides data users the ability to create a wider variety of tables tailored specifically to their data

needs.

For special tabulations of the SBO data that the PUMS file will not satisfy, please contact the SBO staff by e-mail (csd.sbo@census.gov) or phone (888-225-4022 or 301-763-3316) to discuss cost estimates and exact specifications for the type and format of the data requested.

Family-owned business and Number of owners

The complete layout for the PUMS file is provided in Appendix A.

Source: U.S. Census Bureau, 2007 Survey of Business Owners

4. Differences Between Published SBO Estimates and the PUMS Data

In the SBO, data are collected at the business level and at the owner level for up to four individual owners. Businesses with more than one domestic establishment are counted in each industry and geographic area in which they operate, but only once in the total for all sectors and the totals at the national and state levels.

Measures were taken in constructing the PUMS file to protect the confidentiality of the SBO data. In the PUMS file, each record corresponds to a business. For businesses operating in multiple states and/or industry sectors, one record exists for each statesector combination in which the firm conducts business. Identifiers to link the component records of a business are not included. Additionally, businesses classified in the SBO as publicly owned or not classifiable by gender, ethnicity, race, or veteran status are not included in the PUMS file because many publicly owned firms are easily

identifiable.

In the PUMS file, states with fewer than 100,000 weighted businesses in the 2007 SBO are combined with similar states. Alaska is combined with Wyoming; Delaware with the District of Columbia; North Dakota with South Dakota; and Rhode Island with Vermont. To further protect the confidentiality of individuals and businesses, the PUMS file provides rounded, noise-infused estimates of receipts, payroll, and employment. Details relating to the disclosure avoidance may be found on the SBO website (<http://www.census.gov/econ/sbo/methodology.html>) along with the sampling and estimation methodologies.

Detailed race and ethnicity groups have been collapsed in the PUMS file (e.g., Hispanic ownership is provided at the aggregate level for persons of Mexican, Puerto Rican, Cuban, and other Hispanic, Latino, or Spanish origin rather than by detailed Hispanic subgroup). Additionally, SBO allows business owners to belong to as many as four different race groups, whereas the PUMS file limits each owner to a maximum of two race groups to maintain to protect the confidentiality of individuals who reported three or more race groups. Furthermore, businesses in states with a small population of a particular group have been placed in a different, less identifiable group in the PUMS file.

Tables of U.S.-level estimates by employer status comparing published SBO data to the PUMS data are available. See Appendix B.

Source: U.S. Census Bureau, 2007 Survey of Business Owners

5. Determining Company-Level Characteristics

The SBO collected information on the gender, ethnicity, race and veteran status of business owners. The SBO defines these characteristics for a business when a particular group owns 51 percent or more of the equity, interest, or stock of the business (e.g., a women-owned business has women owners owning more than half of the business). The PUMS file contains percentage ownership and the characteristics of each owner (up to four owners per business), but does not contain company-level gender, ethnicity, race or veteran status indicators. This allows data users the flexibility to define ownership in accordance with their data needs. To create products most similar to published SBO data, the following are recommended:

Add the owner percentages for owners of a particular group to determine the percent of the business owned by that group. For example, if there are three owners, Owner 1 is female and owned 34 percent, Owner 2 is male and owned 33 percent, and Owner 3 is female and owned 33 percent, then 67 percent of the company is women-owned and 33 percent is men-owned.

When the total percentage for a group is greater than 50, classify the business as belonging to that group. In the example above, that business would be classified as women-owned. Note that a Hispanic owner may be of any race. Moreover, each owner had the option of selecting more than one race and therefore should be included in each selected race group. Therefore, the race percentages for the business may exceed 100 percent.

Continuing with the previous example, suppose Owner 1 responded as white and black, Owner 2

responded as white and Asian, and Owner 3 responded as black and Asian. This business would be classified as white-owned, black-owned, and Asian-owned since 67 percent of the business is white-owned, 67 percent is black-owned, and 66 percent is Asian-owned. Classify the business as equally owned when there is 50-percent ownership within a category. For example, Owner 1 owns 50 percent and is Hispanic, and Owner Hispanic/non-Hispanic-owned.

Determine minority status based on whether an owner has an ethnicity of Hispanic or a race of black, American Indian and Alaska Native (AIAN), Asian, Native Hawaiian and Other Pacific Islander (NHOPI), or Some Other Race (SOR).

Note that separate classifications need to be determined for each major characteristic (gender, ethnicity, race, minority status, or veteran status) for each business.

6. Determining PUMS Estimates

To determine the estimate for the total number of businesses, sum the tabulation weights assigned to each record belonging to the specified category. This formula estimates the number of firms: y is the estimated total number of firms or establishments.

$TABWGT_j$ is the tabulation weight for the j^{th} record in the category.

j is the reference for the record in the category ($j=1,2,\dots,n$).

n is the total number of records in the category.

7. Determining Standard Errors

Standard errors are calculated using the method of random groups. In the PUMS file, all certainty cases are placed in random group 0, and the noncertainty records are divided into 10 noncertainty random groups (1,2,...,10). Start by calculating the finite population correction (FPC) for each record. The FPC is the square root of the difference of 1 and the inverse of the tabulation weight.

Source: U.S. Census Bureau, 2007 Survey of Business Owners

See http://www2.census.gov/econ/sbo/07/sbo_export_report.pdf for the full report.