

## Application for SynLBD Analysis Validation

The Census Bureau has placed the SynLBD Beta Data Product on a restricted remote access site requiring a password for data access and not permitting data downloads.

The Census Bureau will reproduce analyses conducted using the SynLBD on the LBD Gold Standard File. This activity will be undertaken by Census Bureau employees authorized to use these confidential data. Once an analysis has been repeated on the LBD Gold Standard File, the Gold Standard result summaries will be reviewed by the Census Disclosure Review Board (DRB). Data products and output approved by the DRB will be released to the users and SOI/IRS. The Census Bureau will validate results obtained from the SynLBD on the internal, confidential version of these data.

*IMPORTANT: Request for disclosure of tabular output will not be accommodated unless the requests involves simple descriptive means/deviations of the kind you would find in a research paper (min/max, pctiles are not allowed unless masked). Tabular datasets will not be released under any circumstance. Requests for regression output will be considered as long as they are reasonable (meaning a limited number of models/subsets such as those you would find in a research paper). Reseachers are required to provide observation counts for each of the statistics they want released (number of establishments/firms). These counts are used in the disclosure review. Counts should be provided in a separate file and in a format that is easily relatable to the desired statistics. Counts for regression coefficients are not necessary unless they are dummy variables. All output files submitted for disclosure review have to be submitted in ascii text file format.*

Users who wish to obtain validated results should follow the protocol outlined here.

1. Make sure your programs have successfully run without error on the synthetic data.
2. Make sure your programs generate counts of establishments/firms for every statistics you would like disclosed.
3. Complete and submit your Application for SynLBD Analysis Validation.

Please submit this application to [ces.synthetic.data.use@census.gov](mailto:ces.synthetic.data.use@census.gov).

Name

VRDC Account

Affiliation

Address 1

Address 2

City  State  Postal Code

Email

Phone

**Source Code Statement:** Provide the file name and the location of the analysis code (or sequence of code) and describe the software used for the analysis.

**Verify the code runs without error on the synthetic data before submitting your application. Verify counts for each statistic you would like released are provided and that they can be clearly cross referenced to the actual statistics of interest.**

**Output Statement:** Describe each of the output files generated by the code. For simple statistics describe parameters over which they are produced (industry detail, years, etc...); do same for regression analysis. For regression output please provide the regression model. Describe any potential disclosure risks. Please make sure to identify output files containing counts of establishments/firms for each of the statistics and how they relate to output files containing your statistics of interest. **Requests for validation of tabular output (beyond summary statistics) will be denied.**

Census Response: Approval/Denial

Approved

Denied

VRDC Account

Project Number

Approval/Denial Date

Comments: