

One Author's Walk Down Privacy Lane

John M. Abowd
Cornell University

(thanks to Lars Vilhuber for suggestions used here)

Early Disclaimers and SDL

- Abowd and Zellner (*JBES*, 1985): special unpublished tabulations from BLS and Census; SDL procedures undisclosed; no archiving nor curation; data no longer usable.

Article Stable URL: <http://www.jstor.org/stable/1391596>

- Abowd (*AER*, 1989): proprietary data from the Bureau of National Affairs (news agency) acknowledged; no SDL required by the provider; detailed appendix published; data not archived nor curated; data no longer usable.

Article Stable URL: <http://www.jstor.org/stable/1827932>

First Full Disclaimer and Curation

- Abowd (*ILRR*, 1990): used confidential data from an undisclosed source and proprietary stock price and financial data; SDL performed by the author; first use of disclaimer:

“... the compensation data are confidential, but the author’s access is not exclusive.”

Data subsequently archived and curated by the author, still available from the Cornell Restricted-Access Data Center.

Article stable URL: <http://www.jstor.org/stable/2523571>

First Government Agency Curation

- Abowd, Kramarz and Margolis (*Econometrica*, 1999): used confidential data from the French national statistical institute (INSEE); SDL performed by the authors; detailed appendix published; curation acknowledgement:

“The data used in this paper are confidential but the authors' access is not exclusive. Other researchers interested in using these data should contact the Centre de Recherche en Economie et Statistique, INSEE, 15 bd Gabriel Péri, 92245 Malakoff Cedex, France.”

Still true. Input files were archived and curated; analysis files were not.

Article Stable URL: <http://www.jstor.org/stable/2999586>

Confidential Data Curated; Access Usually Denied

- Abowd and Vilhuber (*JBES*, 2005): used confidential Unemployment Insurance wage records with Privacy Act-protected identifiers not removed; SDL supervised by the Census DRB; detailed data appendix published; disclaimer:

“The data used in this article are confidential data from the LEHD program. The U.S. Census Bureau is preparing to support external researchers' use of these data; please contact U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233. Visit lehd.dsd.census.gov for additional data details and www.ces.census.gov for information about accessing these data through the Census RDC network.”

URLs are dead. No RDC support is provided for the files used in this paper. The method is used in the production system for the Census Bureau LEHD data products (Quarterly Workforce Indicators, OnTheMap); input files archived and curated by the Census Bureau; not discoverable externally; not provided to external researchers.

Article Stable URL: <http://www.jstor.org/stable/27638803>

First Curation with Synthetic Public-use Data

- Abowd and Stinson (*REStat*, 2013): used confidential Census, IRS, and SSA data; SDL supervised by the Census DRB; detailed data appendix published online; curation acknowledgement:

“All data used in this paper are confidential. The U.S. Census Bureau supports external researchers' use of some of these data through the Research Data Center network (www.census.gov/ces). For public-use data please visit www.census.gov/sipp and click Access SIPP Synthetic Data.”

The URLs are OK. Input and analysis files archived and curated but not externally discoverable. Some replication possible with public-use synthetic data. Synthetic data are externally discoverable.

[doi:10.1162/REST_a_00352](https://doi.org/10.1162/REST_a_00352)

Something Better, Maybe

A Proposed Solution to the Archiving and Curation of
Confidential Scientific Inputs

eCommons URL <http://hdl.handle.net/1813/30923>