

Developing a Residence Candidate File for Use With Employer-Employee Matched Data

Matthew Graham

U.S. Census Bureau

Mark Kutzbach

U.S. Census Bureau

Danielle H. Sandler

U.S. Census Bureau

CES Technical Notes

CES-TN-2017-01

This paper describes the Longitudinal Employer-Household Dynamics (LEHD) program's ongoing efforts to use administrative records in a predictive model that describes residence locations for workers. This project was motivated by the discontinuation of a residence file produced elsewhere at the U.S. Census Bureau. The goal of the Residence Candidate File (RCF) process is to provide the LEHD Infrastructure Files with residence information that maintains currency with the changing state of administrative sources and represents uncertainty in location as a probability distribution. The discontinued file provided only a single residence per person/year, even when contributing administrative data may have contained multiple residences. This paper describes the motivation for the project, our methodology, the administrative data sources, the model estimation and validation results, and the file specifications. We find that the best prediction of the person-place model provides similar, but superior, accuracy compared with previous methods and performs well for workers in the LEHD jobs frame. We outline possibilities for further improvement in sources and modeling as well as recommendations on how to use the preference weights in downstream processing.

Relevant Datasets: LEHD RCF

Required Dataset Permissions:

CES Technical Notes may contain confidential data and, thereby, disclosure is prohibited. The opinions, conclusions and suggestions expressed in these notes are those of the authors and do not necessarily represent those of the CES or the U.S. Census Bureau. Researchers (CES staff, SSEs and other Census

employees) who wish to access the full text of a technical note or submit their own technical note should contact CES.Technical.Notes.List@census.gov.