

Use of Freight Production Functions to Identify Outliers in the Commodity Flow Survey Data

Jose Holguin-Veras

Rensselaer Polytechnic Institute

Shama Campbell

Rensselaer Polytechnic Institute

Lokesh Kalahasthi

Rensselaer Polytechnic Institute

CES Technical Notes

CES-TN-2017-08

Abstract

This research aims at analyzing the Commodity Flow Survey (CFS) data across various years (1993, 1997, and 2007), develop methodologies to identify outliers, and to improve freight planning and modeling to help in policy development to make the freight system more efficient and sustainable. To this end, the team has developed freight production models that estimate the amount (in pounds per year) of cargo produced by the firms as a function of number of employees, industry sector, and geography. Using the most significant model the firms that behave eccentric with respect to freight compared to other firms in the same industry (at 2 digit North American Industry Classification System, NAICS) were identified. The results are presented in the form of summary tables showing the total number of firms, number of outliers per industry sector and percentage of outliers per industry sector. The results show that there is a need for further exploration into current sample and survey design to ensure realistic coverage of the variations in the predominant freight indicators such as firm size, sector, and geography.

Relevant Datasets:

CFS, LBD

Required Dataset Permissions:

CFS, LBD 1993, 1997 & 2007

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 represent those of the U.S. Census Bureau. Researchers (CES staff, SSS 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.