LED Data Quality Process

Presented by:
Stephen Tibbets
Economist, LEHD
LED Data Quality

Data quality review is an integral part of LED procedures. There are a number of factors that affect final output data quality. Quality of state input files are a major contributor to the quality of LEHD data products.
Contributors to LED Data Quality

Program design
- Incomplete coverage

Analytical methods
- Imputation
- Noise infusion, synthetic model variation

Operational steps
- Computational or procedural error

Incoming data
- State data, other data sources
Overview of LED Quality Assurance Review

Production is divided into a sequence of processes.

Each process has a set of automatically run QA reports.

- Basic reports are reviewed by production staff.
- More detailed reports are reviewed by staff economists or QA analysts.
QWI Production Process
## Examples of LED QA Tests

<table>
<thead>
<tr>
<th>Data Quality Check</th>
<th>Issues Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count of firms; firm birth and death</td>
<td>Nonreporting, Identifier issues</td>
</tr>
<tr>
<td>Concordance in firm level reporting between UI Wage and QCEW data</td>
<td>Identifier issues (esp. public sector)</td>
</tr>
<tr>
<td>Longitudinal employment and wage totals: growth, concordance of UI and QCEW</td>
<td>Nonreporting, wage misreporting</td>
</tr>
<tr>
<td>Comparison of smaller cell patterns between BLS and final LEHD data (NAICS sector/county totals)</td>
<td>Firm geography, establishment impute, successor/predecessor, other firm oddities</td>
</tr>
</tbody>
</table>
Correcting Data Quality Issues

When issue derives from state data submission, revised data is requested. Ad-hoc data fixes can be a troublesome to integrate into LEHD systems. Longer term code upgrades can improve the ability of the system to handle idiosyncratic issues.
State Partner Role in QA

Audit reports are a first step on establishing a regular feedback loop.
Many additional eyes on the data are the best QA.
- Partners have been extremely helpful in identifying data issues.
The best data cleanup is done using local knowledge.
- QCEW vs. UI wage data have distinct challenges.
Contact us:
LEHD program
CES.Local.Employment.Dynamics@census.gov

Stephen Tibbets
stephen.r.tibbets@census.gov
(301) 763-2125