Linking the 1940 U.S. Census to Five Modern Aging Surveys

John Robert Warren, Dafeng Xu, and Jonas Helgertz (University of Minnesota)
Fabian T. Pfeffer (University of Michigan)

Background and Objective

• Modern surveys of older people often lack data on early life conditions and exposures.
• Respondents to several modern aging surveys were children or young adults at the time of the 1940 U.S. Census.
• Goal: Link the 1940 Census to records from:
  1. Health and Retirement Survey (HRS)
  2. Wisconsin Longitudinal Study (WLS)
  3. Panel Study of Income Dynamics (PSID)
  5. National Health and Aging Trends Study (NHATS)
• Product: individual-, family-, and spatial measures of social, economic, familial, and environmental circumstances as of 1940.

Linking Strategy

• Using information about sample members’ names, years and places of birth, and family members' information, we use both automated and hand linking techniques.
• Automated linking uses machine learning algorithms. Hand linking confirms and supplements automated linking.
• Biased methods minimize false positives.

How Often Are We Able to Link Sample Members to the 1940 U.S. Census?

<table>
<thead>
<tr>
<th>Survey Measure</th>
<th>WLS</th>
<th>HRS</th>
<th>PSID</th>
<th>NHATS &amp; NSHAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940 U.S. Census Measure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 to 8 Years</td>
<td>Unlinked</td>
<td>Hand Linked</td>
<td>Machine Linked</td>
<td></td>
</tr>
<tr>
<td>9 to 12 Years</td>
<td>Unlinked</td>
<td>Hand Linked</td>
<td>Machine Linked</td>
<td></td>
</tr>
<tr>
<td>&gt;12 Years</td>
<td>Unlinked</td>
<td>Hand Linked</td>
<td>Machine Linked</td>
<td></td>
</tr>
</tbody>
</table>

Examples of Available 1940 U.S. Census Measures

Measures based on the 100%-count 1940 U.S. Census file are available at IPUMS.org. For example:

• Individual (for panel & family members)
  - Education
  - Occupation
  - Income

• Family/Household
  - Nativity
  - Marital Status
  - Income

• Spatial/Neighborhood
  - Family size, composition
  - Household size
  - Home value

Illustrative Analyses

<table>
<thead>
<tr>
<th>WLS: Mother’s Education</th>
<th>WLS: 1940 Household Income &amp; Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940 U.S. Census Measure</td>
<td></td>
</tr>
<tr>
<td>0 to 8 Years</td>
<td>0 to 8 Years</td>
</tr>
<tr>
<td>9 to 12 Years</td>
<td>9 to 12 Years</td>
</tr>
<tr>
<td>&gt;12 Years</td>
<td>&gt;12 Years</td>
</tr>
</tbody>
</table>

Potential Uses of the Linked Data

• Analyses of the effects of early-life social and economic circumstances on later-life health, cognition, labor market behaviors, economic outcomes, and mortality.
• Analyses of the effects of environmental and spatial exposures on such outcomes.
• Analyses of social and economic mobility across multiple generations.

How to Access the Linked Data

Data linked to the 1940 Census will be available via each project’s restricted-use data system:

• WLS
• HRS
• PSID
• NHATS
• NSHAP

Acknowledgements

• The project is supported by the National Institute on Aging (1R01AG050300-01A1).
• Support was also provided by the Minnesota Population Center, which receives core funding (P2CHD041023) from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD).