EDITING STRATEGIES USED BY THE U.S. BUREAU OF LABOR STATISTICS IN DATA COLLECTION OVER THE INTERNET

By

Stephen H. Cohen
United States
Bureau of Labor Statistics
Internet Data Collection Philosophy

• Decrease in perception of burden
• Improved response rates
• Control of costs
• Improved data quality and
• For surveys with multiple closings--smaller revisions between preliminary and final estimates
Internet Edits

- Start with a limited number of edits and build
- Limit type of edits
- Borrow edits from post-data collection processing
Current Employment Statistics Program

- All employees
- Women employees
- Production or nonsupervisory workers
- Production/Nonsupervisory payroll
- Production/Nonsupervisory hours
- Commissions collected for the Service Providing industry
- Overtime hours collected in Manufacturing
1997 CES Internet
Edit Rules

• Logic errors--e.g. All Employment greater than Production Workers
• Range checks--Average hourly earnings between $1.00 - $150.00
• Validity checks--numeric entry
• Data entry errors
• Mandatory field (All employees) must be completed
Current Internet Edit Rules

- Expanded basic and longitudinal edits
- Critical values vary by major industry division
- Over-the-month changes compared for each data element and for several “calculated” averages
CES Internet
Edit Experiences

- Approximately 40% of current web sample has failed at least one edit check over a 12 month span
- Approximately 3% of web reports fail one edit check each month
- In 88% of all edit failures, respondent corrected and submitted data during same session
# Monthly Unit Cost of Data Collection

<table>
<thead>
<tr>
<th>Messaging Function</th>
<th>CATI</th>
<th>Mail</th>
<th>EDI</th>
<th>TDE/VR/FAX</th>
<th>Web</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outbound messages (advance notices)</td>
<td>$0.23</td>
<td>$0.34</td>
<td>$0.01</td>
<td>$0.07</td>
<td>$0.00</td>
</tr>
<tr>
<td>Inbound messages (data)</td>
<td>$1.00</td>
<td>$0.34</td>
<td>$0.28</td>
<td>$0.14</td>
<td>$0.00</td>
</tr>
<tr>
<td>Non-response prompting</td>
<td>$0.00</td>
<td>$0.10</td>
<td>$0.01</td>
<td>$0.04</td>
<td>$0.00</td>
</tr>
<tr>
<td><strong>TOTAL TRANSMISSION COSTS</strong></td>
<td><strong>$1.23</strong></td>
<td><strong>$0.78</strong></td>
<td><strong>$0.30</strong></td>
<td><strong>$0.25</strong></td>
<td><strong>$0.00</strong></td>
</tr>
<tr>
<td>Hardware</td>
<td>$0.03</td>
<td>$0.01</td>
<td>$0.03</td>
<td>$0.03</td>
<td>$0.01</td>
</tr>
<tr>
<td>High speed communication lines</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td><strong>$0.12</strong></td>
</tr>
<tr>
<td>Digital Certificates</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td><strong>$0.10</strong></td>
</tr>
<tr>
<td>Labor</td>
<td>$5.00</td>
<td>$0.50</td>
<td>$0.43</td>
<td>$0.28</td>
<td>$0.10</td>
</tr>
<tr>
<td><strong>TOTAL MONTHLY UNIT COSTS</strong></td>
<td><strong>$6.18</strong></td>
<td><strong>$1.29</strong></td>
<td><strong>$0.76</strong></td>
<td><strong>$0.56</strong></td>
<td><strong>$0.33</strong></td>
</tr>
</tbody>
</table>
Research Issues Associated with Editing Schedules on the Web

- Respondent behavior
- Data quality
- Overall strategy