NADDI 2018
Implementing DDI to Document the Consumer Expenditure Surveys

Reginald Noël
Parvati Krishnamurty

5 April 2018
The Consumer Expenditure Surveys

National surveys on U.S. households’ spending

- Survey sponsor & data processing: BLS
- Data Collection: Census Bureau
- Two independent survey instruments: Diary, Interview Survey
- Data Processing: BLS – across 3 subsystems
  - Over 80 data processing jobs/edits for each survey of each year
  - Over 3,400 variables for each survey of each year
- Data Products release schedule: BLS
  - Tabular data (rolling 4 quarters of data every 6 months)
  - Public Use Microdata (annually)
  - Input to CPI (quarterly)
This Presentation

The CE program is attempting to move towards a more efficient and comprehensive documentation of survey metadata. We provide 2 illustrations of current inefficiencies, and how we are attempting to use DDI to address them:

1. Questionnaire development

2. Investigating data inconsistencies
1. Questionnaire Development
CE Questionnaire Development Process

- The Interview Survey is long and contains complex skip patterns designed to reduce respondent burden.
- Changes to questionnaire items such as wording, order, skip patterns, or new questions are often needed.
- CE provides detailed requirements, which Census then enters into their own specification system (SPIDER), and then programs the Blaise instrument. Changes are usually made every 2 years.
- CE Interview Survey instrument is being redesigned.
Interview Survey Questionnaire Development

Challenges

- No question bank
- No integrated mapping of questionnaire metadata
  - Multiple documents in different locations
- The changes document for each revision is in word/pdf/csv
  - Difficult to navigate: current full specifications ~ 4,000 pages
  - Iteratively making design changes is difficult

Solution: Testing the use of Colectica Questionnaires to create requirements for the new CAPI instrument, to be programmed into a Blaise instrument by Census Bureau.
Colectica Questionnaires: Advantages

- Creates DDI based survey specifications
- Surveys can be published in multiple formats including Blaise 5
- Produces machine readable DDI 3.2 format xml files (source code for CAI systems)
- Rapid design iterations: Easy to make changes in the software and then generate new specifications
- Uses a repository of questions so we can retrieve questions and blocks
- Integrates with the rest of the Colectica software suite including Repository, Portal, and Designer
Some advantages: Output formats

- Blaise 5
  - One or more questions are required to generate Blaise
  - Paper Form
  - Specification
  - DDI 3.2
Some advantages: Repositories
Colectica Questionnaire: Limitations

- Trial version cannot handle some of the more complex questionnaire design elements.
- Some examples are: grid questions, reusable lists of response options, globally adding fields for specifications, setting don’t know and refusals globally, ability to store and cumulate display logic, rosters and loops, adding comments to questions.
- We got around the limitations by programming what we could in the software and providing a cover sheet for each section with additional information that could not be entered in the software.
  
e.g. text fills, refusal/don’t know options, instructions for looping, and grid questions.
Cluster screener (grid question) in the CE

42. ? [F1]
Since the first of November, have you had expenses for any of the
following, either for you or for someone outside your household?
* Read each item on list.

1. Fresh flowers or potted plants?
2. Professional photography?
3. Services of lawyers or other legal professionals?
4. Accounting fees?
5. Occupational expenses, such as union dues or
   professional licenses?
6. Gardening or lawn care services?
7. Housekeeping services?
8. Home security system service fees?
9. Other home services or small repair jobs around
   the house, not previously reported?
10. Moving, storage, or freight?
11. Stamp or coin collecting?
12. Lotteries or games of chance?
95. Continue list

<table>
<thead>
<tr>
<th>Item</th>
<th>Misc exp/item</th>
<th>Description</th>
<th>Month</th>
<th>Who</th>
<th>Amount</th>
<th>C</th>
<th>Combined Codes</th>
<th>Store</th>
<th>Online</th>
<th>Location</th>
<th>Any more</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Garden/lawn serv</td>
<td>Garden/lawn serv</td>
<td>1</td>
<td>1</td>
<td>$500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Housekeep serv</td>
<td>Housekeep serv</td>
<td>2</td>
<td>1</td>
<td>$300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Housekeep serv</td>
<td>Housekeep serv</td>
<td>1</td>
<td>1</td>
<td>$200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Home security</td>
<td>Home security</td>
<td>2</td>
<td>1</td>
<td>$50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>
Grid question in Colectica Questionnaires

Household operations

**OPSEC**

Now I will ask you about expenditures for household maintenance, repairs and service contracts. Please remember to include any payments you made online or had automatically deducted. Also, please include any shipping and handling charges with the cost of any item that was shipped.

**OPSCSC**

During [Month1], [Month 2], or [Month 3], did (you/you and any members of your household) have any expenses for--

1. Gardening or lawn care services?
2. Housekeeping services?
3. Termite or pest control services?
4. Water softening services?
5. Home security system service fees?
6. Moving, storage or freight?

Follow up questions
Display if OPSCSC Contains Any 1 Gardening or lawncare services? or 2 Housekeeping services? or 3 Termite or pest control services? or 4 Water softening services? or 5 Home security system service fees? or 6 Moving, storage or freight?

**OPSCSCFL**

12 items
2. Investigating Data Inconsistencies
Fragmented metadata across time and processes

Challenges

- Maintained separately
  - by surveys (Diary or Interview)
  - by years (... 2015, 2016, 2017 ...)
  - by 3 data processing subsystems
  - Initial Edit System (IES), Edit and Estimation System (EES), Microdata, and Tables

- Multiple points of manual input and alteration

- Many different folders, different file formats, separate database

Solution: Testing Colectica Designer, Repository, and Portal to put all metadata across time and subsystem onto one web-based portal
2. Investigating Data Inconsistencies

**Inquiry Request:** Research Universal Classification Code 580311

- What variables are used to map to this UCC?
- How is the associated question asked in the survey?
- Get descriptive statistics, i.e. average, frequencies, etc.
2. Investigating Data Inconsistencies

- Current Production Method
  1. Excel mapping file
  2. Web-based P&C data dictionary
  3. PDF public-use microdata dictionary
  4. BLS website to get data
  5. SAS enterprise to get stats

- Using Colectica and DDI
  Colectica Repository and Portal through web-based prototype
### File 1: Excel mapping file for UCC

<table>
<thead>
<tr>
<th>CONDSPEC</th>
<th>NAME_IN</th>
<th>REC_SEC</th>
<th>SOURCE</th>
<th>TIME_ADJ</th>
<th>UCC_NO</th>
<th>STARTDATE</th>
<th>COMMENT</th>
<th>IDCNT</th>
<th>SEQCNT</th>
<th>SEQENDT</th>
<th>TRANSDE</th>
<th>LASTUPDAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1FURNPUT</td>
<td>FURNCFCU</td>
<td>2FURNCFCU</td>
<td>FRAS0BA</td>
<td>520221</td>
<td>20132</td>
<td>69</td>
<td>1</td>
<td>9123</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1393</td>
<td>NA</td>
<td>QREP3MNX</td>
<td>FRBS0BA</td>
<td>54630</td>
<td>20012</td>
<td>2</td>
<td>1</td>
<td>9563</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1144</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500112</td>
<td>199601</td>
<td>10</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1316</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500114</td>
<td>199601</td>
<td>12</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1318</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500312</td>
<td>199601</td>
<td>14</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1320</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500304</td>
<td>199601</td>
<td>16</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1322</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500315</td>
<td>199601</td>
<td>19</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1323</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500311</td>
<td>199601</td>
<td>19</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1324</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500322</td>
<td>199601</td>
<td>20</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1326</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500311</td>
<td>199601</td>
<td>21</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1327</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500311</td>
<td>199601</td>
<td>23</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1328</td>
<td>HHICODE 2 &amp; HHIBCBS 2 &amp; HHIFEET 2</td>
<td>QHICMX</td>
<td>HHBS148</td>
<td>500311</td>
<td>199601</td>
<td>23</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1333</td>
<td>NA</td>
<td>QCMADMX</td>
<td>HHBS148</td>
<td>500011</td>
<td>199331</td>
<td>6</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1335</td>
<td>NA</td>
<td>QCDEMDX</td>
<td>HHBS148</td>
<td>500011</td>
<td>199331</td>
<td>6</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1336</td>
<td>NA</td>
<td>QDQPMX</td>
<td>HHBS148</td>
<td>500007</td>
<td>200602</td>
<td>1</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1377</td>
<td>NA</td>
<td>QDQPMX</td>
<td>HHBS148</td>
<td>500007</td>
<td>200602</td>
<td>1</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1392</td>
<td>NA</td>
<td>QPRCMX</td>
<td>HHBS148</td>
<td>790060</td>
<td>201302</td>
<td>12</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1353</td>
<td>NA</td>
<td>QPRCMX</td>
<td>HHBS148</td>
<td>790060</td>
<td>201302</td>
<td>12</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
<tr>
<td>1354</td>
<td>NA</td>
<td>QADDOMX</td>
<td>LSDS10</td>
<td>540313</td>
<td>199132</td>
<td>8</td>
<td>1</td>
<td>20131</td>
<td>19123</td>
<td>20131</td>
<td>20131</td>
<td>20131</td>
</tr>
</tbody>
</table>
### File 1: Excel mapping file for UCC

2013 Interview Time Adjustment Expenditure Mapping file – (Excel file located in year-designated folder)

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONDSPEC</td>
<td>NAME_IN</td>
<td>REC_SEC</td>
<td>SOURCE</td>
<td>TIME_ADJ</td>
<td>UCC_NO</td>
<td>STARTDATE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FURNPPUR-&quot;171&quot;</td>
<td>FURNPPURX</td>
<td>FRD508A</td>
<td>508A</td>
<td>FURNMO</td>
<td>320221</td>
<td>20132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>GFRP3MCX</td>
<td>FRD508B</td>
<td>Recodes</td>
<td>&amp;B</td>
<td>340630</td>
<td>200112</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>GFRP3MCX</td>
<td>FRD508B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>340994</td>
<td>200112</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800112</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800114</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800112</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800114</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800112</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800114</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800112</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800114</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800112</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800114</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICOVQ-&quot;00&quot;</td>
<td>GH3MCX</td>
<td>HBS14B</td>
<td>Recodes</td>
<td>&amp;A</td>
<td>5800112</td>
<td>19961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Highlighted Rows:**
- HHICOVQ-"00" row indicates the mapping for UCC tracking.
- GH3MCX row indicates the UCC code.

---

19 — U.S. BUREAU OF LABOR STATISTICS • bls.gov
Dictionary Search Page

Search by Variable Information

Survey: Interview

Dictionary: 2013Q2

RELN:

Field Name: QHI3MCX

Description:

Phase:

Submit
Variable QHI3MCX

RELN: EIHB
Survey Section: 14B Hospitalization and Health Insurance

Variable Description: Health insurance payments, reference period minus current month

Recode:

Code Description:

Flag Name: QHI3MCX_
 Occurs: 1

Source: Recodes

Variable Phase: 2n3

P2 Attribute: NUMB

P3 Attribute: NUMB

P2 Primary Keys: (FAMID,SEQNO)
P3 Primary Keys: (FAMID,SEQNO,ALCNO)

Information queried from IY2013syb_v26_41.mdb
## Variable HHICODE

**RELN:** ETHB  
**Survey Section:** 14B Hospitalization and Health Insurance

**Variable Description:** Insurance type code

**Code Description:**
1. Health Maintenance  
2. Fee for service plan  
3. Commercial Medicare Supplement  
4. Other Special purpose plan

**Flag Name:** HHICODE  
**Occurs:** 1

**Source:** S14B HHICODE  
**Variable Phase:** 2n3

**P2 Attribute:** CODED1_R  
**P3 Attribute:** CODED1_R

**P2 Primary Keys:** (FAMID, SEQNO)  
**P3 Primary Keys:** (FAMID, SEQNO, ALCNO)

Information queried from IY2013syb_v26_41.mdb
Variable HHIBCBS

RELN: EH
Survey Section: 148 Hospitalization and Health Insurance

Variable Description: Is the insurance company Blue Cross/Blue Shield?

Recode:

Code Description:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
</tr>
</tbody>
</table>

Flag Name: HHIBCBS_

Source: S148 HHIBCBS

Variable Phase: 2n3

P2 Attribute: CODED1

P3 Attribute: CODED1

P2 Primary Keys: (FAMID,SEQNO)

P3 Primary Keys: (FAMID,SEQNO,ALCNO)

Information queried from IV2013syb_v26_41.mdb
## IHB — Detailed Expenditures Files (EXPN)

### Hospitalization and Health Insurance

<table>
<thead>
<tr>
<th>VARIABLE NAME</th>
<th>DESCRIPTION</th>
<th>FLAG</th>
<th>FORMAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHIP.MMPD</td>
<td>What period of time is covered by the regular payment? (HHIPRMMPD = 1 or 2 only) CODED</td>
<td>HHIP_MMPD</td>
<td>CHAR(1)</td>
</tr>
<tr>
<td>Coded Options</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Month</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Quarter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPI Section 14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QHBMCMX</td>
<td>Amount paid for health insurance premiums during the reference period BLS derived</td>
<td>QHBMCMX_</td>
<td>NUM(8)</td>
</tr>
<tr>
<td>QHBMCMX</td>
<td>Amount paid for health insurance premiums during the reference period BLS derived</td>
<td>QHBMCMX_</td>
<td>NUM(8)</td>
</tr>
<tr>
<td>HHICBS</td>
<td>Is insurance company Blue Cross/Blue Shield? CODED</td>
<td>HHICBS_</td>
<td>CHAR(1)</td>
</tr>
<tr>
<td>Coded Options</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPI Section 14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHICODE</td>
<td>What type of insurance plan is it? CODED</td>
<td>HHICODE_</td>
<td>CHAR(1)</td>
</tr>
<tr>
<td>Coded Options</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Health maintenance organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Fee for service plan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Commercial Medicare supplement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Other special purpose plan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPI Section 14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHIP05</td>
<td>Under normal circumstances, if you go to a doctor who is not part of your plan without a referral, will your insurance pay for the cost? (HHICODE = 1 only) CODED</td>
<td>HHIP05_</td>
<td>CHAR(1)</td>
</tr>
<tr>
<td>Coded Options</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Consumer Expenditure Survey

PUMD Data Files

From 1996 forward, public-use microdata (PUMD) are available in four formats: SAS, SPSS, STATA and Comma Delimited (ASCII). Each format presents the data for each year in two zipped files. Data collected with the Interview Survey are in the "Interview" file and those with the Diary Survey are in the "Diary" file.

Pre-1996 PUMD are available as Comma Delimited files. The Inter-university Consortium for Political and Social Research (ICPSR) provides the data for free to its members. Here is their member list. BLS also sells the data on USB flash drives (PDF).

If you are new to CE PUMD data, you may want to explore the data availability with the Getting Started Guide. For those using income data in their analyses, we also recommend reading the User's Guide to Income Imputation in the CE. Additionally, we supply training programs and sample data (ZIP).

The NEW State Weights for New Jersey are available with their documentation.

The next dataset will be released on September 11, 2018. To be notified when new datasets become available, please sign up for the CEX e-mail updates. If you have comments, suggestions, and questions about this page and its contents, please contact us.

If you are having difficulty downloading the data, please see the instructions on how to download zip files.

SAS  SPSS  STATA  CSV

<table>
<thead>
<tr>
<th>Year</th>
<th>Interview (zip)</th>
<th>Diary (zip)</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>(51.6 MB)</td>
<td>(10.5 MB)</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>(87.7 MB)</td>
<td>(10.5 MB)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>(99.1 MB)</td>
<td>(12.3 MB)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>(101.1 MB)</td>
<td>(11.5 MB)</td>
<td></td>
</tr>
</tbody>
</table>
File 4. BLS website to download dataset

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Compressed size</th>
<th>Password protection</th>
<th>Size</th>
<th>Ratio</th>
<th>Date modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>apb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>331 KB</td>
<td>No</td>
<td>3,456 KB</td>
<td>91%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>cla13.sas7bdat</td>
<td>SAS Data Set</td>
<td>1,293 KB</td>
<td>No</td>
<td>11,584 KB</td>
<td>89%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>clc13.sas7bdat</td>
<td>SAS Data Set</td>
<td>25 KB</td>
<td>No</td>
<td>256 KB</td>
<td>91%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>cnt13.sas7bdat</td>
<td>SAS Data Set</td>
<td>251 KB</td>
<td>No</td>
<td>1,600 KB</td>
<td>85%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>cra13.sas7bdat</td>
<td>SAS Data Set</td>
<td>18 KB</td>
<td>No</td>
<td>192 KB</td>
<td>91%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>crb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>205 KB</td>
<td>No</td>
<td>2,752 KB</td>
<td>93%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>eda13.sas7bdat</td>
<td>SAS Data Set</td>
<td>135 KB</td>
<td>No</td>
<td>1,280 KB</td>
<td>90%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>eqb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>55 KB</td>
<td>No</td>
<td>512 KB</td>
<td>90%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>fra13.sas7bdat</td>
<td>SAS Data Set</td>
<td>265 KB</td>
<td>No</td>
<td>2,304 KB</td>
<td>89%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>frb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>10 KB</td>
<td>No</td>
<td>128 KB</td>
<td>93%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>hel13.sas7bdat</td>
<td>SAS Data Set</td>
<td>32 KB</td>
<td>No</td>
<td>256 KB</td>
<td>88%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>ihb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>409 KB</td>
<td>No</td>
<td>2,624 KB</td>
<td>85%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>ihc13.sas7bdat</td>
<td>SAS Data Set</td>
<td>166 KB</td>
<td>No</td>
<td>1,280 KB</td>
<td>88%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>ihd13.sas7bdat</td>
<td>SAS Data Set</td>
<td>97 KB</td>
<td>No</td>
<td>704 KB</td>
<td>87%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>inb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>965 KB</td>
<td>No</td>
<td>12,672 KB</td>
<td>93%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>isl13.sas7bdat</td>
<td>SAS Data Set</td>
<td>50 KB</td>
<td>No</td>
<td>384 KB</td>
<td>88%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>mdb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>489 KB</td>
<td>No</td>
<td>4,672 KB</td>
<td>90%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>mdc13.sas7bdat</td>
<td>SAS Data Set</td>
<td>18 KB</td>
<td>No</td>
<td>192 KB</td>
<td>92%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>mis13.sas7bdat</td>
<td>SAS Data Set</td>
<td>477 KB</td>
<td>No</td>
<td>4,480 KB</td>
<td>90%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>mor13.sas7bdat</td>
<td>SAS Data Set</td>
<td>705 KB</td>
<td>No</td>
<td>3,264 KB</td>
<td>79%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>opb13.sas7bdat</td>
<td>SAS Data Set</td>
<td>430 KB</td>
<td>No</td>
<td>3,648 KB</td>
<td>89%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>opd13.sas7bdat</td>
<td>SAS Data Set</td>
<td>10 KB</td>
<td>No</td>
<td>128 KB</td>
<td>93%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>oph13.sas7bdat</td>
<td>SAS Data Set</td>
<td>25 KB</td>
<td>No</td>
<td>192 KB</td>
<td>87%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>opj13.sas7bdat</td>
<td>SAS Data Set</td>
<td>889 KB</td>
<td>No</td>
<td>13,312 KB</td>
<td>94%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
<tr>
<td>ovg13.sas7bdat</td>
<td>SAS Data Set</td>
<td>1,972 KB</td>
<td>No</td>
<td>18,304 KB</td>
<td>90%</td>
<td>9/2/2016 12:06 PM</td>
</tr>
</tbody>
</table>
File 5. SAS to get dataset descriptive statistics
1. Colectica Portal: Prototype Home Page
2. Find and click the UCC in question
3. Obtain desired information

Health maintenance organization (not BCBS)

- **Label**: Health maintenance organization (not BCBS)
- **Is Missing**: False
- **Conditional Variables**
  - EIHB.HHBCBS - Is the insurance company Blue Cross/Blue Shield?
  - EIHB.HTCODE - Insurance type code
- **Value Source Variables**
  - EIHB.QH3MCOX - Health insurance payments, reference period minus current month
4. Get additional information

Amount paid for health insurance premiums during the reference period

Health insurance payments, reference period minus current month

Variable | Details | Lineage | Concordance Variables
--- | --- | --- | ---
Statistics | Code Comparison | Correspondence Tree

- **QHI3MCX**: ihb12
- **QHI3MCX**: ihb13

Values:
- 10,521
- 690
- 305
- 103
- 0
Health insurance payments, reference period minus current month

Dataset | Variable | Valid | Invalid | Min | First Quartile | Median | Third Quartile | Max | Mean | StDev
---|---|---|---|---|---|---|---|---|---|---
Ab12 | QH3MCX | 25991 | 8755 | 0 | 108 | 505 | 660 | 10,521 | 501 | 655
Ab13 | QH3MCX | 22949 | 9311 | 0 | 130 | 381 | 728 | 11,302 | 576 | 709
4B. Get additional info for different subsystems
A User’s Perspective

Colectica can help us by:

- Limiting the different folders and files need to research
- Saving on the amount of time to look up info
- Making all metadata readily accessible in one platform
- Facilitating easy access to descriptive data
Consumer Expenditure Surveys

Contact Us:
krishnamurty.parvati@bls.gov
noel.reginald@bls.gov

CE DDI Team
Daniel Gillman
Evan Hubener
Parvati Krishnamurty
Reginald Noël
Bryan Rigg
Arcenis Rojas
Lucilla Tan
Taylor Wilson
Extra slides
DDI Structure for Consumer Expenditure Quarterly

Iteration I

Series

Interview Survey

Global Lists
- Question Sets
- Category Sets

2012 Interview

2013 Interview

Studies

2012 Package

2013 Package

Metadata Packages

2012 Code Set

2013 Code Set

Variable Structure

Instance Variable (P2)

2012 Instance Variables

Instance Variable (P3)

Instance Variable (P4)

Represented Variable {Characteristic Set 1}

Represented Variable {Characteristic Set 2}

Represented Variable {Characteristic Set 3}

Conceptual Variable
DDI Structure for Consumer Expenditure Quarterly

Iteration I
DDI Structure for Consumer Expenditure Quarterly

Iteration I

Series

Studies

Metadata Packages

Variable Structure

Interview Survey

2012 Interview

Global Lists

Question Sets

Category Sets

2013 Interview

2012 Package

2013 Package

2012 Code Set

2013 Code Set

Instance Variable (P2)

Instance Variable (P3)

Instance Variable (P4)

Instance Variable (P2)

Instance Variable (P3)

Instance Variable (P4)

Represented Variable

Characteristic Set 1

Represented Variable

Characteristic Set 2

Represented Variable

Characteristic Set 3

Conceptual Variable
CE DDI/Colectica Experience

Iteration 1
- Explored use of a few basic DDI constructs to describe the CE Interview Survey metadata. [Colectica Designer]
- Colectica implements DDI

Iteration 2
- Proof-of-concept prototype [Colectica Portal and Repository]
- Eight survey questions across 2 years of the Interview Survey.
  - Establish and view the linkage of a survey question to its end as a variable in the public-use Microdata file
  - Identify changes to a variable: for e.g., a change in the response options to the EDUCA (member education attainment) variable
What’s Next?

Iteration 3

- Will focus solely on Microdata and Tables (Ph4)
  - Inputting metadata from 1996 to 2017
- Continue development work with Colectica
- Scale up prototype, including Diary
- Incorporate the use of Colectica Questionnaire
- Create one portal for the public and another for interoffice use
Potential Hurdles

Iteration 3

- Requires IT support
- Setting up a server
- Colectica programming language (C-sharp, .NET)
- Equipment requests
- Limited funding
Household Operations: Cover Page

A. Section Instructions:
This section is a laundry list section that follows the laundry list standards. Design will be a table/grid format. Repeat OPSCSC through OPSOTH until OPSCSC = 99. After OPSOTH, the FR is expected to continue reading the list of expenditures from the cluster screener (OPSCSC).

B. Section Format:
Table with a maximum of 14 rows.

C. Don’t knows and refusals:
1. Do not allow Don’t Know or Refusal on: OPSCSC, OPSIDE, OPSECE, OPSCOM, OPSSEC.
2. Do not display Don’t Know or Refusal options on screen.
3. Where Don’t Know and Refusal are allowed, use keyboard shortcuts:
   - Don’t know = CTRL + D
   - Refusal = CTRL + R
   - Bookmark = CTRL + B
D. Fills: 1. OPSCOM, OPSOUN and OPSOTH:

<table>
<thead>
<tr>
<th>OPSCSC</th>
<th>Category Name</th>
<th>OPSIEX Less than</th>
<th>OPSIEX Greater than</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gardening or lawn care services?</td>
<td>15</td>
<td>600</td>
</tr>
<tr>
<td>2</td>
<td>Housekeeping services?</td>
<td>20</td>
<td>1260</td>
</tr>
<tr>
<td>3</td>
<td>Termite or pest control services?</td>
<td>16</td>
<td>742</td>
</tr>
</tbody>
</table>

E. Fills: For OPSCOM, OPSOUN and OPSOTH: fill in...

F. Outlets eligibility: Outlets questions will be assigned based on a rotation matrix with POPS categories by PSU by quarter. The following categories are outlets eligible:

G. Item description: For OPSIDE: The text field should be a blank, non-required field to allow entry of the item description by the FR. After the FR enters a list number the instrument should place the cursor in the blank item description field. FRs can choose to hit enter to skip this. In case the item description is needed for a subsequent fill, the item code title will be used.