Measuring Household Consumption Expenditure

Thomas Crossley (Essex and IFS)

Drawing on joint work with

Joachim Winter and Martin Browning

HFCN, November 2014
Motivation: Why Measure Household Consumption Expenditure?

• How has the well-being of the poor evolved over time?
• How well insured are households against job loss? Disability? Major changes in the economy?
• How do consumption and saving respond to interest rates?
• Do tax-favored savings accounts generate net new savings?
• Do house price movements have a causal effect on consumer spending?
• How effective are tax rebates and other payments for fiscal stimulus?
• How can we have confidence in the models we use for macro policy analysis?
Those with the lowest cash incomes do not have the lowest cash outlays.

Notes: LCFS 2009; Great Britain only
Source: see Brewer, Etheridge and O’Dea, “Why are households that report the lowest incomes so well-off?”
http://ideas.repec.org/esx/essedp/736.html
Aren’t Budget Surveys Sufficient?

- Limited information in other domains (WEALTH, health, employment, time use…..)
- Not longitudinal (often interested in changes)
- Some concerns with sustainability of budget surveys
  - Will budget surveys be able to continue to meet many needs?
Response Rates and Coverage of Household Expenditure in National Accounts, UK and US budget surveys

Source: Barrett et al. A comparison of micro and macro expenditure measures across countries using different survey methods, NBER Working Paper 19544
Preliminary Remarks:

- Will focus on Household Consumption Expenditure
- Will focus on survey-based methods
  - Lots of interesting possibilities with administrative data
  - But we will continue to need survey data
- Will focus on multiple domain surveys (not budget surveys)
Household Consumption Expenditure

- Prices
- Start of period durables stocks, depreciation, relationship between stocks and services flows
- Returns to scale in consumption, household composition

Household Consumption

Individual Consumption
Preliminary Remarks:

- How do we know what works?
  - The problem of benchmarks
  - National Accounts (but reconciliation important)
  - Other survey and administrative data
  - Beegle et al. (JDE, 2012), Tanzania: Each Adult a diary; each dependent assigned to an adult; Daily visits.
  - Higher totals probably better (evidence of underreporting)
  - Lower variance might be better
  - Theory and guesses about the nature of measurement error
Possible Approaches:

1. Capture total expenditure with a single (‘one shot’) question or small, but complete set of categories.

2. Use the inter-temporal budget constraint: \( \text{income minus saving} \)

3. Ask a subset of expenditure categories and impute total expenditure at the household level.

4. Ask a subset of expenditure categories and estimate objects of interest directly.
One-Shot Questions

- The Italian SHIW has asked the following:
  - What was your family’s average monthly expenditure in 1995 for all consumption items?
  - Consider all expenses, including food, but excluding those for: housing maintenance; mortgage installments; purchases of valuables, automobiles, home durables and furniture; housing rent; insurance premiums.

- Also: COEP, Spanish Survey of Household Finances (also French), AHEAD pilots, Centre Panel (Netherlands)
One-Shot Questions

• Theoretical underpinning: two-stage budgeting.
• High response rates (often better than household income).
  – Except AHEAD pilots.
  – Respondents view questions about broad categories of expenditure as being less sensitive than comparable income questions (see focus groups, below).

• One-shot questions generate useful data.
  – Engel Curves look good (Browning et al. 2003, Bottazzi et al. 2008).
  – Data from one-shot questions have been successfully employed in a number of research papers (e.g., Browning & Crossley 2001, 2008).
What to Worry About?

• One-shot questions always give significantly lower estimates of total consumption expenditure than more disaggregated data collection.

• Focus groups and cognitive interviews have also documented problems with one-shot questions (Gray et al. 2008, d’Ardenne & Blake 2012).

• Recall of total expenditure is challenging for many respondents.
  – But they appear to use a variety of methods for estimation (see below).

• Complex households a particular problem.
Short but Complete Sets of Expenditure Categories

- Some trials of ‘short breakdown’ approach in web mode (US, Netherlands)
  - Useful data
  - Evidence that a reconciliation screen or expenditure check improves the data
New Evidence On Quick Expenditure Questions

- Blake, Crossley, D’Ardenne, Oldfield, Winter, “Testing Quick Expenditure Questions” (Preliminary)*

* With funding from the Nuffield Foundation
Some Lessons from the Focus Groups

• Importance of wording ("household" example)
• Problems with complex households
  – Higher aggregation easier?
• Keep questions short; avoid excessive examples
• Heterogeneity in household financial management
  – Heterogeneity in response strategies
• Total expenditure less intrusive than income or expenditure categories
• Income minus saving intrusive for some
Some Lessons from Cognitive Testing

- Essential/nonessential disaggregation unhelpful
- Income minus savings did not work well
- One-shot feasible
  - Strategy showcard unhelpful
  - example showcard rather than long question
- Break-down feasible
  - Reconciliation helpful
- ‘Benefit unit’ works well in UK (though: implementation)
US-IP6 Experiment

• Understanding Society is a major longitudinal survey of UK households.
  – 40,000 households, annual, mixed mode.
  – Follows on from (and incorporates) the British Household Panel Survey from 2008.

• Separate *Innovation Panel* of 1500 households for testing and development.

• Our experiment in Wave 6 of the Innovation Panel (IP6).
  – Field work Feb-July 2013.
  – 1137 benefit units.
US-IP6 Experimental Design

<table>
<thead>
<tr>
<th>Mode</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-shot/ web</td>
<td>Breakdown/ web</td>
</tr>
<tr>
<td>One-shot/ f2f</td>
<td>Breakdown/ f2f</td>
</tr>
</tbody>
</table>
US-IP6 Experimental Design

• The one-shot question, in F2F mode was:

  “About how much did you [and [NAME OF PARTNER/SPOUSE]] spend on EVERYTHING in the LAST MONTH? Please exclude work expenses for which you are reimbursed, money put into savings and repayment of bank loans. Examples of what to include and exclude are shown on this card.”

• The web mode was similar, with the exclusions and examples show below the answer box.

• The breakdown approach asked for spending in 12 specific categories plus an “other” category.
  – Categories developed from focus groups
One-shot

Examples and exclusions shown underneath the answer box

Strategy: How did you work out your answer to the spending question?

Usual spending follow up: Would you say your spending last month was: higher than usual, lower than usual, typical of a usual month’s spending?

If higher/lower: how much do you [and X] spend on everything in a usual month?

Breakdown

Reconciliation: So in total in the last month you [and X] spent [total] pounds. Does that sound right?

If no: How much did you [and X] spend in the last month?
# US-IP6 Experiment – Main Results

<table>
<thead>
<tr>
<th>Month:</th>
<th>Web  `One-shot’</th>
<th>Web  `One-shot’</th>
<th>Web  `Break-down’</th>
<th>F2F  `One-shot’</th>
<th>F2F  `One-shot’</th>
<th>F2F  `Break-down’</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>84%</td>
<td>85%</td>
<td>90%</td>
<td>93%</td>
<td>93%</td>
<td>99%</td>
</tr>
<tr>
<td>mean</td>
<td>2260</td>
<td>1646</td>
<td>1807</td>
<td>1312</td>
<td>1219</td>
<td>1801</td>
</tr>
<tr>
<td>median</td>
<td>1600</td>
<td>1500</td>
<td>1570</td>
<td>1000</td>
<td>980</td>
<td>1373</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>4239</td>
<td>896</td>
<td>1265</td>
<td>1580</td>
<td>1544</td>
<td>1654</td>
</tr>
</tbody>
</table>

LCFS mean *household* monthly spend £2040
### US-IP6 Experiment – Response Strategies

<table>
<thead>
<tr>
<th>Strategies for answering the spending question (not mutually exclusive)</th>
<th>Web</th>
<th>F2F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checked statements</td>
<td>28%</td>
<td>10%</td>
</tr>
<tr>
<td>Added up categories</td>
<td>35%</td>
<td>67%</td>
</tr>
<tr>
<td>Income minus saving</td>
<td>24%</td>
<td>8%</td>
</tr>
<tr>
<td>Recall without checking</td>
<td>21%</td>
<td>15%</td>
</tr>
<tr>
<td>Other</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>
US-IP6 Experiment – Additional Results

- Reconciliation question in the breakdown approach improved data in two ways
  - Reduced variance/outliers
  - In web mode, overcame item nonresponse (raised fraction with usable total spend response from 69% to 90%).
Preliminary Conclusions From the Experiment

• Focus groups and cognitive testing also identified key improvements to the ‘one-shot’ question
  – Showing examples
  – Choice of response unit
  – Avoiding problem language (ex ‘household’ spending)
• Breakdown still gives better data (but takes longer)
• Web mode seems attractive
  – One-shot: less under-reporting
  – Breakdown: reconciliation helps
Use the Inter-temporal Budget Constraint

- Some surveys collect information on wealth and income
- An identity: \( x_{t,h} \equiv y_{t,h} - s_{t,h} \),

- The inter-temporal budget constraint:

\[
 w_{t+1,h} = (w_{t,h} + y_{t,h} - x_{t,h})(1 + r_{t,h}),
\]

- Inverted:

\[
 x_{t,h} = y_{t,h} - [(1 + r_{t,h})^{-1} w_{t+1,h} - w_{t,h}] .
\]

- Might be approximated by:

\[
 x_{t,h} \approx y_{t,h} - [w_{t+1,h} - w_{t,h}] .
\]
Use the Inter-temporal Budget Constraint

• Zilliak (1998): PSID.

• Administrative (tax) wealth data (Browning and Leth-Peterson, 2003; Kriener et al 2014; Koijen et al 2014; Browning et al., 2013)
  – Browning and Leth-Peterson (2003) use budget survey data to validate the administrative data; Kriener et al, (2014) and Koijen et al (2014) use the administrative data to validate budget survey data….

• Bozio, Laroque, O’Dea (2014): ELSA
What to Worry About?

- Resulting measure very noisy (Zilliak, 1998; Browning et al., 2013)
- Koijen et al (2014) show that ignoring capital gains and losses induces substantial errors.
- Estimating effects of wealth and income shocks?

\[
\Delta x_{t,h} = \alpha + \beta \Delta y_{t,h} + \gamma \Delta w_{t,h} + u_{t,h}, \\
\Delta[y_{t,h} - (w_{t+1,h} - w_{t,h})] = \alpha + \beta \Delta y_{t,h} + \gamma \Delta w_{t,h} + u_{t,h}, \\
\Delta y_{t,h} - \Delta w_{t+1,h} + \Delta w_{t,h} = \alpha + \beta \Delta y_{t,h} + \gamma \Delta w_{t,h} + u_{t,h}.
\]
Imputing Total Expenditure from a Subset of Categories

• **Skinner (1987):**

\[
CE : \quad x_{t,h} = \beta_0 + \sum_j x_{t,h}^j \beta_j + \nu_{t,h},
\]

\[
PSID : \quad \hat{x}_{t,h} = \hat{\beta}_0 + \sum_j x_{t,h}^j \hat{\beta}_j.
\]

• much employed

• Zilliak (1998) compares this measure to Y-ΔW in the PSID
Imputing Total Expenditure from a Subset of Categories

- Blundell, Preston, Pistaferri (BPP; 2004, 2008)
  - Engel Curve (theory and experience with demand modelling)
  - Consistency of parameter estimates (IV)

\[
\tau(x_{t,h}^f) = Z'\alpha + \gamma\phi(x_{t,h}) + e_{t,h},
\]

\[
\hat{\phi} = \frac{1}{\hat{\gamma}}(\tau(x_{t,h}^f) - Z'\hat{\alpha}),
\]

\[
\hat{x}_{t,h} = \phi^{-1}\left[\frac{1}{\hat{\gamma}}(\tau(x_{t,h}^f) - Z'\hat{\alpha})\right].
\]
Imputing Total Expenditure from a Subset of Categories

- BPP (2004): when are $M[\hat{x}]$ and $V[\hat{x}]$ consistent estimates of $M[x]$ and $V[x]$?

- IF parameters of Engel curve estimated consistently then $V[\hat{x}]$ converges to $V[x]$ plus an additive term (allowing for comparisons over time.)
  - IV important

- Attanasio & Pistaferri (2014) use a version of this procedure with post 1997 PSID and earlier PSID.
  - Comparisons of (movements in) $V[x]$ and $V[\hat{x}]$ post 1997 are encouraging.
What to worry about?

• Requires external data

• BPP conditions are strong
  – Correct specification of Engel Curve
  – ME magnitudes and relationships stable
  – All sources of endogeneity dealt with (Campos and Reggio, 2014)

• Easy to show that prediction error variance depends on level of food expenditure
  – ME in household expenditure heteroscedastic

• Engel curve mis-specified (approximation error)
  – Expected ME depends on true value
What to worry about?
Estimating Other Moments Directly

- Might not be necessary (desirable) to estimate household level directly
- Basic idea: 2 measures with uncorrelated, classical measurement error

\[ x_{t,h}^i = x_{t,h}^i + \varepsilon_{t,h}^i \]

\[ V[x^1] = V[x] + 2C[x\varepsilon^1] + V[\varepsilon^1] = V[x] + V[\varepsilon^1] \]

\[ C[x^1x^2] = V[x] + C[x\varepsilon^1] + C[x\varepsilon^2] + C[\varepsilon^1\varepsilon^2] = V[x] \]

- Consistency doesn't depend on reliability of measures (though precision does)
Estimating Other Moments Directly

- Browning and Crossley (2009):  
  - $x_{t,h}$ is log total nondurable expenditure;  
  - $x_{t,h}^i$ are expenditure categories (in logs)  
  - Target is variance of log nondurable consumption

\[
x_{t,h}^i = \alpha_i x_{t,h} + \eta_i (x_{t,h}) + e_{t,h}^i.
\]

\[
x_{t,h}^i \equiv x_{t,h} + \varepsilon_{t,h}^i.
\]

\[
\varepsilon_{t,h}^i \equiv x_{t,h}^i - x_{t,h} = (\alpha_i - 1)x_{t,h} + \eta_i (x_{t,h}) + e_{t,h}^i
\]
Estimating Other Moments Directly

- BC (2009):

\[ \varepsilon^i_{t,h} \equiv x^i_{t,h} - x_{t,h} = (a_i - 1)x_{t,h} + \eta_i(x_{t,h}) + e^i_{t,h} \]

\[ C[x^1 x^2] = V[x] + C[x \varepsilon^1] + C[x \varepsilon^2] + C[\varepsilon^1 \varepsilon^2]. \]

- \( C[x \varepsilon^i] \) small if income elasticity close to 1, approximation error small; positive (negative) for luxury (necessity)

- \( C[\varepsilon^1 \varepsilon^2] \) tends to positive (negative) for strong compliments (substitutes)

- Choose goods accordingly
Estimating Other Moments Directly

- BC 2009 Simulation Study
- Estimating $V[\log C]$
- Canadian Budget Survey
- Food, recreation (services and non/semi-durable goods)

![Graph showing the distribution of % bias for different estimates]

- direct estimate
- 2 good estimate
- direct estimate with error (R=0.6)
- BPP Estimate
Estimating Other Moments Directly

- Related procedures in Attanasio, Hurst, Pistaferri (2014)
- Could be done with Skinner/BPP measure and Y-ΔW measure.
- Aguiar and Bils (2013) propose a somewhat different approach to exploiting multiple goods.
- As another example, the EIS for nondurable consumption can be recovered from the EIS for a good and knowledge of the Engel curve (See Browning and Crossley, 2000)
What to worry about?

• Requires good quality budget data to choose goods
  – Are these relationships stable?
• Poor choices lead to poor results
• How general?
  – General de-convolution methods to get whole distribution, but
  – Need one measure with mean zero ME to get location
  – Specific solutions
What to worry about?

Poorly chosen goods

- direct estimate
- food in and out
- direct estimate with error (R=0.6)
- food and tel.
Conclusions and Recommendations

• Data from these methods can be very useful in research
  – General release a harder question
  – Provide tools/guidance?

• If asking multiple (but not comprehensive) categories, consider how they might be used jointly.
  – May be tension with other uses (e.g., food security)
Conclusions and Recommendations

• Designing questions: wording is very important
  – Lots of cognitive testing
  – Language specific

• Longer run: consider web mode?