Paul Fisher

Does repeated measurement improve data quality?

IARIW 34th General Conference

Discussion by Brian Bucks

Consumer Financial Protection Bureau

August 22, 2016

The views expressed in this presentation are mine alone and not necessarily those of the Consumer Financial Protection Bureau or the United States.
Overview

- Prior studies find income is under-reported in surveys (especially government transfers)

- Paper makes use of UKHLS design to gauge:
  1. Changes in income mis-reporting across waves
  2. Explanations for under-reporting & changes in it

- Key findings: Under-reporting . . .
  1. is driven by “false negatives” for unearned income
  2. is greatest in the earliest waves
  3. improves mainly due to panel conditioning, particularly increased respondent trust over time
Under-reporting of income sources

Refusal + don’t know rates by income source

Also, diffs in income quantiles vs benchmark generally largest below median & in 1st wave
UKHLS design offers a quasi-experiment

- Waves are fielded for 24 months
- Households were randomly assigned a survey month and are interviewed annually

<table>
<thead>
<tr>
<th>Year</th>
<th>HH in mos 1-12</th>
<th>HH in mos 13-24</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Wave 1</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>Wave 2</td>
<td>Wave 1</td>
</tr>
<tr>
<td>2011</td>
<td>...</td>
<td>Wave 2</td>
</tr>
</tbody>
</table>
Identifying the effect of an add’l interview

- **Idea:** Compare 2010 income for waves 1 & 2
- **Control for diffs in demographic charac.**
- **Assume attrition (cond’l on observables & model) the same across svy years**

<table>
<thead>
<tr>
<th>Year</th>
<th>HH in mos 1–12</th>
<th>HH in mos 13–24</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Wave 1</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>Wave 2</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>. . .</td>
<td>Wave 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wave 2</td>
</tr>
</tbody>
</table>
Wave 2 HH’s report ≈ 8% more income

- **Total Income** [£1588.15]
- **Benefits & Unearned Income** [£507.22]
  - **Social security benefits** [£203.98]
  - **Pensions** [£288.15]
  - **Other unearned** [£45.10]
- **Earnings** [£1078.14]
- **Investment Income** [£2.79]

Wave 2 effect: 124.12, 104.86, 23.37, 70.91, 10.58, 19.44, -0.19

A possible reason for better (more) data: Dependent interviewing

▶ Saying "No" to income source reported in prior wave prompts follow-up: “Can I just check...”

▶ UKHLS flags instances where DI was triggered; Setting these to 0, shows effect of DI

▶ DI accounts for only $\approx \frac{1}{3}$ of measured effect
  $\Rightarrow \approx \frac{2}{3}$ attributed to panel conditioning
Digging deeper

1. Greater wave 2 income is driven almost entirely by reported receipt, not larger amounts
   ▶ Notable exception: Employer pensions

2. Same analytical approach for, e.g., waves 3 and 4 shows no significant differences after wave 2

3. Some evidence of similar patterns in BHPS based on refreshment samples
What’s behind the panel conditioning?

- It’s not iwers or Rs getting better at the survey
  - Findings unchanged if control for iwer traits
  - Iwer ratings of Rs’ understanding no different at Wave 2

- Instead, Rs seem more willing to answer due to greater trust:
  - Confidentiality concerns less common in wave 2
  - Iwers rated wave 2 Rs as less suspicious of svy
  - Confidentiality queries predict nonresponse to income questions
A twist on what I “know”? 

- My initial sense: most income misreporting stems from stigma or ambiguity
- Largest effects here are for pensions

- Might split benefits based on degree of stigma
- Larger question: Distinguishing between
  1. stigma (some benefits; drug use)
  2. complexity or variability (self-emp income)
  3. over-precision (day or month started job)
  4. sensitivity (high incomes)
To my mind, attrition is the toughest knot

- I trust the sample design (and size) gets comparable HH in wave 1 in each year

- We can’t similarly ensure attrition is ignorable

- Model attrition in year 1 and 2 separately, compare out-of-sample predictions or reweight

- Note: Emp statuses among the few signif diffs in W1 traits across years in analysis sample
  - Is this true if don’t drop wave-2 dropouts?
Filling in the story with individual-level data

Two types of tables I’d be curious to see

<table>
<thead>
<tr>
<th>Wave 1</th>
<th>Wave 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reported</td>
</tr>
<tr>
<td>Reported</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Change: Wave 2-Wave 1 (real £)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
</tr>
<tr>
<td>Soc sec benefits</td>
<td>...</td>
</tr>
<tr>
<td>Pensions</td>
<td>...</td>
</tr>
<tr>
<td>Wage/salary</td>
<td>...</td>
</tr>
<tr>
<td>Self-employment</td>
<td>...</td>
</tr>
</tbody>
</table>