LCS Workshop comments

Stephen P. Jenkins

Email: s.jenkins@lse.ac.uk

“Outline your work on consumption”

- Ahem … I have no research at all that is directly about consumption!
- I am an interested onlooker and teacher, with research to date mainly on ‘income’
- Also, US NAS panel member: *Creating an Integrated System of Data and Statistics on Household Income, Consumption, and Wealth: Time to Build*, Smeeding et al. (eds), 2024, 360pp.
  - Reviews of measurement principles (concepts and data); US-focused but with cross-national awareness – read it! Cite it!
- Research on related topics:
  - Distributional measurement and analysis concepts and tools generally, including …
  - Equivalence scales and implications of different choices
  - Reference periods and related issues such as dynamics and volatility (within- and between-years)
  - Measurement error (unit non-response, item non-response, reporting error)
  - Data and data quality more generally; longitudinal and cross-section data; admin versus survey data; linked admin-survey data

LSE

Plaudits, but …

I like:

- The logical variable structure including alignment with COICOP's 13 divisions: aggregates plus sub-aggregates where feasible
- File organisation with focus on central set of variables/definitions, but also including supplementary and country-specific in separate files – allowing users to potentially do more
- More generally, LCS integration into the LIS DataCenter data environment and using its software tools
- 'Walk before running' strategy: get a database established, then modify and extend it ('expansion scenarios')

But …

- Issues mostly set out in terms of principles to aspire to
  - Which are indeed are relevant and important
- For each of the issues as literally framed, I largely concur with the authors' choices

- But "in principle" $\neq$ "in practice"
- The Devil is in the Detail
- And choices are much constrained by what is available in existing data sources

- Also, some prior questions should be addressed when thinking about the LCS development roadmap ahead …

Answers to the Questions are also contingent on what the research questions (RQs) of most interest are

- Supplement the 3 existing questions by considering what the RQs are that potential users will likely focus on
- Different focuses imply different directions of travel and emphasis in Expansion Scenarios (cf. LCS October note, Table 2)
  - Current emphasis of discussion of LCS development is input orientated (data), not output orientated (research)
- What are the potential research focuses?
- And then, given these, what are the data we'd like to be able to use?

LSE

Potential research focus(es)?

- Focuses affect Expansion Scenarios and the resources are devoted to what and when
- Topics? E.g., …
  - Levels and shares of C, inequality, poverty, Engel curves and fitting demand systems
  - Longitudinal aspects (e.g., lifecycle, ‘insurance’ against shocks)
  - E.g., what degree of detail is needed below aggregate level or division headings?
- Countries?
  - Choice of countries to include: to focus on HIC, LMIC, or both?
    - Cf. Expansion Scenarios with emphasis on broadening (more countries) versus deepening (better estimates or more years for a narrower range of countries)
    - Very different institutional and market contexts; differences line up with many of the ‘difficult questions’ raised in the WP and data sources
    - Enunciate strategy more explicitly and justify it
- How important are *cross-national* perspectives? Spell out your answers …
- From ‘consumption’ researchers’ points of view, why would they want to use LCS rather other potential data sources? What is the LCS ‘USP’?

LSE

Data sources on consumption/expenditure

- Household budget surveys = LCS sources
  - Many HICs and LMICs have household budget surveys, but potential quality problems (see later)
  - Most in the public domain
  - Mostly cross-sectional (few panel surveys, cf. PSID), relatively small samples, "annual", but rich detail on household/individual characteristics
- Private sector admin data from financial aggregator apps or bank account transactions
  - E.g. [bank transactions] US JPMCI or UK SmartData Foundry (NatWest group) or ES BBVA data
  - Account inflows (incomes) as well as outflows (expenditures), large samples; high frequency
  - Constraints on access; potentially non-representative (people, items); not about gold standard "consumption"; problems with "unit" (cf. single versus joint a/cs, versus a/cs elsewhere) and person and unit characteristics
- Government admin data
  - Registers with consumption/expenditure data almost non-existent (Nordic exceptions?); better for 'income' than for 'consumption' (though good for some specific items, including STIK-related)
  - Longitudinal; very large samples; but research access constrained
- [Reprise] What is the USP that LCS intends to offer to researchers? To address what RQs? Including: what can *cross-national perspectives* offer?

Lessons: make connections with ‘income’, income datasets

- What lessons to be learnt from histories of LIS (and LWS) data developments and their research emphases?
  - Surprising that no mention in LIS WP912
  - And what about lessons from WB’s LSMS, and are those data available for LCS inclusion?
  - Cross-sectional versus Longitudinal: cf. LIS versus Cross-National Equivalent File (or SHARE)
- There are many analogies between development of a consumption database (LCS) and development of an ‘income’ database (e.g., LIS!) – indeed, some issues are the same (e.g., treatment of imputed rent from housing, STIK)! – so could exploit this more
  - E.g.: whether to focus on **narrower** concepts about which more researchers agree and closer to what official statistics likely to use and more policy-persuasive, versus use of **broader** concepts involving more controversy and mostly one-off academic research studies
    - My ‘what is the purpose?’ point again, now also referring to research impact as well as the research per se
  - Example: discussion of income measures including or excluding measures of imputed rent from housing by e.g. Sauli & Törmålehto in Atkinson & Marlier (eds), Eurostat NetSILC volume, 2010
  - Relatedly, Tony Atkinson’s remarks (EcoFin 2010) “non-economist neighbours would probably recognise that they do indeed benefit from public services and from not having to pay rent, and that in the future they may benefit from the pension funds. However, these are not spendable income; they are not cash in the bank. It is the latter that they … have more in mind. ”
  - Justify further “Consumption” versus “Expenditure”

LSE

Things to add

- **Harmonised measures of 'I'** (as in 'ICW') to the datasets in a potential Expansion Scenario
  - Income measures are available in most household budget surveys, at least for HIC
  - Can draw on LIS experience and expertise
  - Comparisons of I and C for the same households facilitate …
    - Derivation of measures of savings (I–C) flows
    - Assessments of reporting errors through examination of outlier differences between I and C, especially at the bottom – cf. Meyer & Sullivan (USA); Brewer et al. (UK)
    - Extend research possibilities, e.g., linking welfare state-provided incomes with consumption outcomes
    - As well as discussion of the 'concentration of (dis)advantage', as in LIS WP912 heatmaps
- **Survey design variables** such as cluster and strata IDs, not only weights
  - Essential for derivation of correct SEs, also important given cross-country heterogeneity in survey designs
- **Documentation/metadata** to discuss these features, plus response rates and questionnaire design (including, e.g., balance between direct questions to respondents and diaries and for what items, recall and reference periods for each, etc.), and survey collection mode(s) and within-year timing
  - Essential for assessments of 'total error' and thence apparent cross-national or cross-year differences
  - NB cannot assume that HBS quality necessarily better in HICs than LMICs: see next slide …
    - Moreover we cannot assume that HBS are necessarily the best data for all RQs (reprise of earlier point)

  - Same remarks for LIS and income surveys, btw! Though potential for admin data use (alone or linked to survey) greater
  - I.e., reprise of remarks asking for more re relative strengths of surveys relative to other data sources for what topics?

LSE

Household budget surveys: falling unit response rates

USA various, including CEX, (Abraham, RIW 2022)

UK LCFS (UK Statistics Authority 2025, from ONS)

Review of Income and Wealth, Series 68, Number 4, December 2022

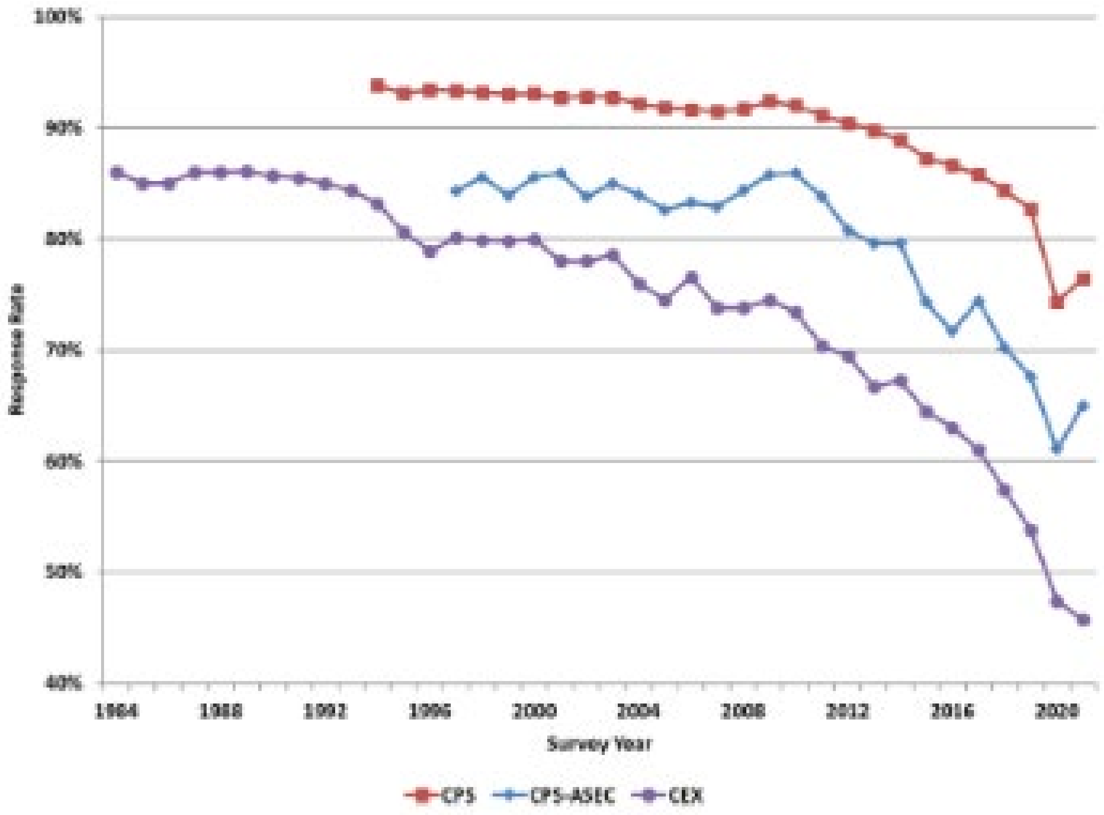
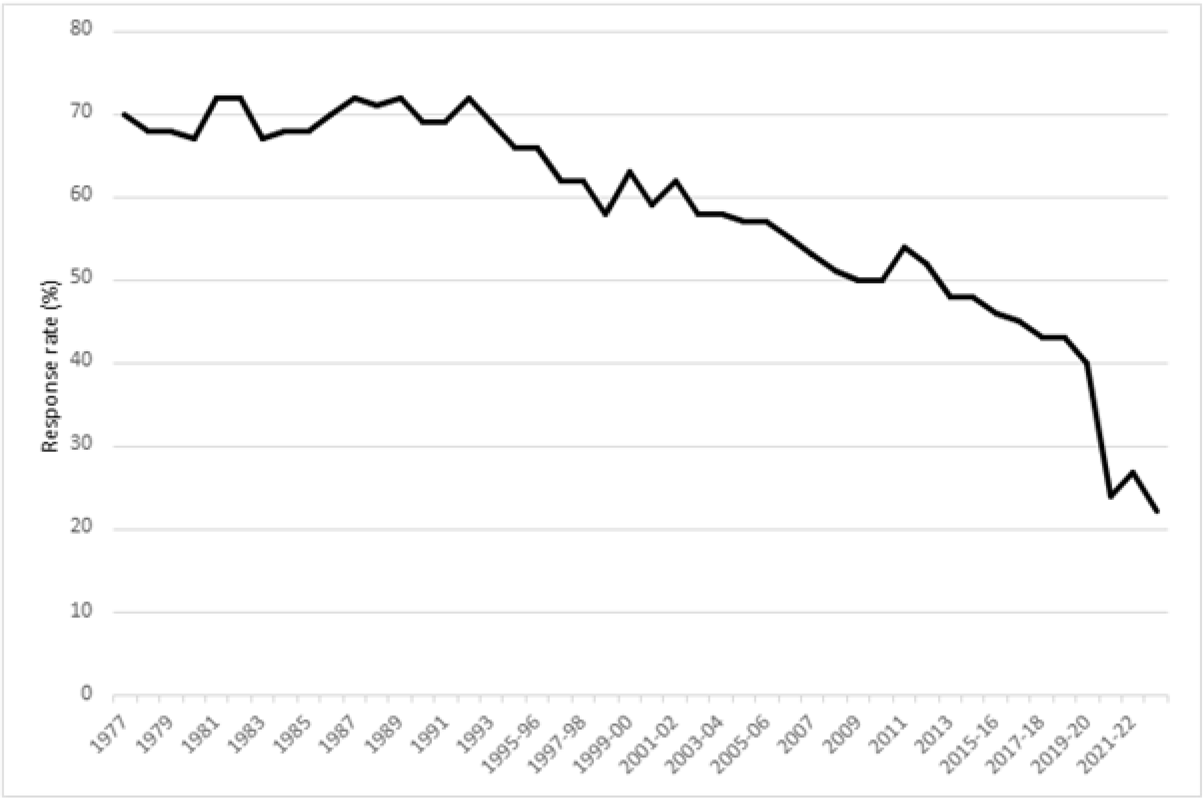


Figure 1. Response Rates for Selected Household Surveys Conducted by the U.S. Census Bureau, 1984–2021.

Source: CPS response rates downloaded using BLS Series Report tool (series ID LNU09300000). CPS-ASEC and CEX rates through 2013 from Meyer *et al.* (2015); later years' response rates for CPS-ASEC from annual supplement documentation and for CEX from BLS website at https://www.bls.gov/osmr/response-rates/. CPS = Current Population Survey. CPS-ASEC=CPS Annual Social and Economic Supplement. CEX = Consumer Expenditure Interview Survey. Annual or annual average response rates reported. [Colour figure can be viewed at wileyonlinelibrary.com].

Chart 1. Response rate for Living Costs and Food Survey.



Source: Office for National Statistics.