Who counts? Challenges and biases in defining ‘households’ in research on poverty

Author: Carlos Oya – SOAS (University of London)

Reader in Political Economy of Development
Department of Development Studies
Thornhaugh St., Russell Square
London WC1H oXG
United Kingdom
Email: co2@soas.ac.uk

ABSTRACT

In an important article published in 2002, H. White made a case for mixed methods in poverty analysis and addressed a number of puzzles arising from conventional oppositions between quantitative and qualitative approaches. One of the examples was the relationship between household size and poverty and the related notion of household economies of scale. This paper revisits this debate and updates it with new contributions on the use and misuse of the „household” in surveys and censuses, particularly in quantitative research designs. Indeed, there is much scope for improving the way quantitative survey designs treat the „household” and for learning from qualitative approaches, especially in poor
agrarian contexts. Some alternative approaches to the household and to sampling decisions in survey design for poverty and labour studies are proposed and their advantages and disadvantages briefly discussed.

**Keywords**: mixed methods; household; poverty; survey design
Introduction

This paper focuses on key challenges in the production of data, by contributing to a reflection on burning methodological questions in research on poverty and labour. The article focuses on a key dimension of evidence on poverty and wellbeing: the unit of analysis. Despite the massive growth of research on poverty measurement, its causes and dynamics, methodological challenges persist. The apparent precision of poverty statistics is at odds with the existence of a chain of imperfections in the generation of data on consumption and assets, from the interview setting to the final global poverty datasets. An important source of potential imprecision is the choice of unit of analysis in poverty studies: the household. Per capita consumption estimates undoubtedly hinge on the precision of consumption calculations at household level and the accuracy of estimates of household size, as will be argued below.

The choice of unit of analysis and implications for the generation and interpretation of evidence on poverty is an important problem that relates to broader debates about the „quantitative-qualitative“ divide and the rise of mixed methods (MM) as an increasingly popular approach in development studies. What scholars like Kanbur and Shaffer (2003 and 2007) call „Q-squared“ (as a form of systematic integration of qualitative and quantitative traditions) has gained in prominence particularly in the context of poverty analysis. White, in his well-known 2002 article on combining quantitative and qualitative approaches to poverty analysis (White 2002), an early contribution to the subsequently popular Q-squared literature, made a compelling case for methodological pluralism,
suggesting that „using the approaches together yields more than the sum of the two approaches used independently“ (White 2002, 513). Earlier debates about the value of combining methods stressed the need for a distinction between types of data (quantitative and qualitative) and types of data collection methods (survey-based or „contextual“), since the notions of „quantitative“ and „qualitative“ can take different, contested, meanings (Hentschel 1997, Booth et al. 1998, Kanbur 2003). A very rich body of literature has since been accumulated on the advantages of MM to overcome barriers to higher quality evidence (Shaffer 2013a and 2013b), and will not be extensively reviewed here. However, and in order to place the particular topic of this paper within a broader epistemological and methodological debate, some reflections and outcomes of these discussions will be mentioned below.

First, despite the growing awareness of the need to build bridges between qualitative and quantitative traditions, the MM/Q-squared literature has highlighted tensions and challenges. In fact, quantitative and qualitative approaches to understand key development challenges can work both in opposition and in dialogue (Kanbur and Shaffer 2007; White 2002; Bryman 2008). Epistemological barriers and defensive „walls“ between disciplines sometimes prevent more productive dialogues and synergies, despite potential value added in cross-disciplinarit (Harriss 2002, Booth et al. 1998). The persistent attachment of disciplines to their own particular epistemologies and preferred methods contributes to slow progress towards inter-disciplinarity and reflects continuous tensions between „context“ in qualitative research and
“standardisation” in large-N quantitative data collection (Harriss 2002, Kanbur 2003).²

Second, the case for methodological pluralism remains nonetheless compelling, as both research traditions have gained and can still benefit from suitable forms of integration and rethinking of one’s own epistemological and methodological straightjacketing, as cogently argued by Shaffer (2013b). As Shaffer (2013b) explains, (a) knowledge is partial; (b) conflicting results between different epistemological and methodological traditions are common (that is, “empirical adjudication is imperfect”); (c) the phenomena under examination are complex and entail a variety of meanings, and (d) there is no single concept of causation in social sciences.³ These factors together imply that a single method is unlikely to offer satisfactory explanations of complex phenomena in diverse contexts. This also means that there is substantial potential for improvements within each methodological tradition.

One area for improvement directly concerns the focus of this paper. In fact, many assumptions underpinning research on poverty and particularly data collection directly relate to the basics of conceptualization of levels and units of analysis, whether in contextual research of large-N survey research. This article tackles this aspect, highlighting the importance of rigorous MM research to better understand the nature, boundaries and contested meanings of the “household”, which is the critical unit of analysis and observation in poverty research.
After this introduction, Section 2 follows with a critical discussion of the use and abuse of the concept of „household” in poverty research, through various illustrations from well-known anthropological debates and recent contributions to the study of poverty and households in African contexts. Section 3 will propose an alternative approach to household rosters in survey design through MM contributions. The paper will end with concluding remarks on possible challenges on methodological innovations.

Understanding the unit of analysis: challenges on the concept of ‘household’

The „household” remains a fundamental unit of analysis and observation in development research and practice, despite an abundant critical literature on its uses and misuses. A basic problem with the rise of uniform concepts of the „household” , particularly in the context of postcolonial Africa, is the way in which this responded to a quest for simplification which had political aims, in contexts where the process of nation state building faced the obstacles of differentiated and weakly integrated societies (Guyer and Peters 1987).

The emergence of „household models” also resulted from the application of basic analytical frameworks from orthodox neoclassical economics to development issues, especially within agricultural economics and poverty analysis, as well as from the emergence of development practice in the context of development projects that needed suitable units for action and targeting.
Universal definitions of the „nuclear household unit” based on basic and rigid residence patterns (i.e. anyone residing for more than six months in a household unit), and the superficial operationalization of such definitions still tend to predominate in national surveys and censuses, micro-level research and development projects. Yet, these definitions eschew the centrality of cultural and social differences for an understanding of household relations and how „households” are organized in different socio-cultural formations, socio-economic contexts, political realities and development trajectories.

An illustration of the clash between context in ethnographic research and standardisation in large-N quantitative data collection in relation to poverty and the household is provided by White (2002). He discusses the contradictory results in poverty research arising from different methodological traditions, is the relationship between poverty and household size. While conventional poverty profiles based on national household surveys tend to find that larger households are poorer, much ethnographic research in rural Africa tends to show the opposite, that smaller households are poorer. White (2002) tackled this discrepancy by referring to household economies of scale. Two leading World Bank researchers also questioned the conventional stylized fact of a positive relationship between household size and poverty reminded us that „even poor households face economies of size” (Lanjouw and Ravallion 1995, 1430). It is obvious that this assumption alone matters for poverty analysis, especially for calculations of poverty headcounts and for the characterisation of the „poor”. In another paper White and Masset (2003, 120, figure 3) illustrate the implications
of different assumptions on household economies of scale in consumption for measurement of headcount indices. Based on the same dataset on consumption expenditure and poverty outcomes, as the parameter of scale economies is changed from zero to positive values, the relationship between poverty status and household size breaks down and becomes negative, i.e. smaller households are relatively poorer. These puzzles add ammunition to the view that there is „false precision“ in widely cited global poverty figures, which are highly dependent on household size as the denominator in basic consumption calculations (Reddy and Pogge 2005). The example also shows that quantitative research could well learn from discrepancies with other disciplines and alternative assumptions in data analysis, as White and Masset do.

Knowing the extent of resulting biases requires some experimental work. In this regard, there has been recent research on the impact of alternative survey design options in relation to operational definitions of the „household“ on the outcomes of variables of interest (household size, consumption per capita, assets, poverty status). For example, Beaman and Dillon (2012) suggest that the „household“ remains something of a „black box“ for development economists. In their experiments, slight changes in definitions of the „household“ led to significant variations in both household size and household composition, with substantial implications for the measurement of basic outcome variables, such as per capita consumption expenditure, asset statistics and per adult equivalent agricultural output measures. Beaman and Dillon (2012) also raise a conundrum:

(A) a consistent household definition is required for comparisons over time and
across populations, but (B) over time and for a given population, „the definition must also identify the correct economic or decision making unit, which may in fact differ according to the research question‟. This issue was already discussed at length in earlier debates on competing concepts of the „household‟ (Guyer and Peters 1987).

Akresh and Edmonds (2010), in another experiment, focus on issues of residence and mobility and effects on the household roster. The main finding is that households are extremely fluid with 10 percent of individuals spending some time away over a three year period, averaging 16 of the 36 months away. This particularly applies to youth, so the pattern has a systematic element that is likely to cause bias. In panel surveys a significant degree of attrition (household members who are no longer there only one-two years after the first round, or entire households that have moved away) is also commonly observed by Akresh and Edmonds (2010), leading to other potential biases. The implication is that a more detailed and thorough measurement of household composition in multi-purpose household surveys is essential to avoid systematic biases in the estimation of key outcome variables.

Randall and Coast (2014) summarise findings of a wider project on the impact of definitions of the household used in household survey research on our ability to understand intergenerational relations for diverse and fluid contexts (Randall and Coast 2014). For example, in Tanzania and Burkina Faso „households‟ may be of different types and organised in particular ways.
Interdependent groups among Masai are frequently split up in surveys but consider themselves to be one economic unit of production and consumption. There is evidence of coexistence of "open" and "closed" households, the latter usually being better captured by national household surveys than the former for obvious reasons. Randall and Coast (2014, 5) also provide examples of "flexible urban households", that is "large heterogeneous households in compounds, with flows of individuals in and out, often circulating from rural areas, bringing rural produce, working in the city and then returning". As the authors emphasise, the absence of a widely accepted local word for "household" reflects the obstacles in operationalising such concept for a survey, whether of a small or large N, and underscores the power of interviewers in translating these concepts into something meaningful (Randall et al. 2013).

Apart from the impact that imperfect definitions of the "household" may have on poverty statistics, there are many other aspects of development and social change that may be affected. For example, there is often evidence of very high mobility of children and young people between households in both rural and urban areas. Leone et al. (2010) and Randall and Coast (2014, 6), reporting on various case studies, show that "children move between households on a weekly or longer-term basis" rendering multiple households "open". The decision of where to include a child is not straightforward if the criterion is one of residence and/or consumption. In a World Bank evaluation of education programmes in Ghana, White and Masset also find that the incidence of child fostering is important, thus the issue of which household is considered matters for statistical
analysis when household characteristics vary between fostering households and the „sending” household. If the decisions are located at the latter, collection of data on household characteristics at the former can introduce biases, „so that the relevant household characteristics are missing for children who are fostered in” (White and Masset 2004, 134). However, the relevant household will depend on the variable of interest. For schooling decisions the „sending” household makes more sense but for test scores characteristics of the fostering household may be more appropriate. This could imply having to consider two different households for the same observation (child).

These experiments and findings therefore raise serious questions about who is included or not in a household roster, whether the concept and related terms are consistently understood by interviewers and respondents in diverse contexts, whether some consumption decisions should be at the centre, or not, of household definitions, and whether conventional residential criteria lead to systematic biases.

Central to conceptualisations of the „household” and their operationalization in the context of quantitative surveys are the understanding of intra-household and especially gender relations and how assumptions about the nature and dynamics of conjugal ties affect the effects of alternative definitions of the „household”. This is particularly challenging in contexts where polygyny prevails, which is often the case in some parts of rural Africa. Important conceptual and operational questions can be raised about whether nuclear
joint households should be considered; whether „mother and her children” as the smallest unit can be then extended to the matrilineal lineage or linked to the husband and/or other breadwinners; whether socio-economic status and class matter for relevant units of consumption and production; or how to grapple with variation in strength and dynamics of conjugal relations, in any given context.

In relation to these and other questions, O’Laughlin (1995) discusses the „myth of the African family”, which refers to the idea that the conjugal bond is traditionally weak „reflecting the predominance of women’s labour in farming systems and importance of lineage ties” (p. 69). So, earlier critiques of conventional household models emerging from descent theory and feminist research (e.g. Sen 1990, Folbre 1986) resulted in new assumptions that abstracted from class relations „that cross-cut town and country”. While there is broader agreement that household relations are characterised by power, what is less clear is how power operates and why its nature varies substantially both between and within cultural contexts (Kabeer 1998; Agarwal 1997). Beyond culture and social norms, conjugal ties are indeed shaped by processes of social and economic change associated with capitalist development, which affect households quite differently depending on the class and geographical locations, making the operational utility of conventional and „universal” definitions of the household limited. These effects transcend conjugal ties and concern relations and conflict between generations within the household. Sometimes violent conflict and related grievances derive from unequal and exploitative household relations as in the case of civil wars in Sierra Leone (Peters 2011). In sum, the
“relational”, in its various manifestations, is central to household composition, contestation, boundaries and dynamics and should be carefully considered in any process of data collection.

Understanding household composition and its dynamics also requires particular care with hard-to-reach populations. In fact, as Carr-Hill shows (2014) argues, basic sources of statistical information such as censuses and national household surveys often fail to account for fluid households in the context of blurred rural-urban boundaries. Pincus and Sender (20078) show how strict residential registration systems in contexts of labour migration – hukou systems in Vietnam (also found in China) - put large numbers of migrants in the fringes of “invisibility” in towns/cities and bias population estimates towards over-counting of rural people, thereby also affecting poverty data. These sampling frames then affect the coverage of national household surveys that inform statistically representative poverty analysis. As a result, groups that are particularly under-represented in poverty analysis, for different reasons, include seasonal migrants in agro-export areas, nomadic or pastoralist populations, fragile, disjointed or split-households, people living in non-household residential units (homeless, those in institutions, hostels, workers’ dormitories), slum populations, and people in areas posing security risks. Operational questions on household membership, notably rules about residential definitions, as well as restrictions in terms of which residential units are visited in surveys and censuses, lead to these important blindspots.
Finally, defining who is a household member also affects our understanding of labour outcomes and its dynamics, which should central to poverty analysis (Wuyts 2011). Indeed, access to labour and the relations of economic dependence probably matter for poverty more than sitting around a common meal. As O" Laughlin (1995, 87) argues, „whether rural families stay poor or develop basis for accumulation depends on access to labour“ , and household size may be both exogenous and endogenous to access to labour, depending on socio-economic status and structural context. Thus differential access to labour is associated with different household compositions and differences in the extent to which particular lineage systems or conjugal ties predominate or not. In the 2002 article, White also pointed out that household definitions matter for understanding labour market outcomes, and particularly the phenomenon of labour exchange and labour hiring in rural Africa. This is precisely one of the aspects in which my own collaborative research has focused in recent years, as the design of rural labour surveys hinges on alternative and „contextual” operational approaches to household rosters and sampling (Cramer et al. 2014; AUTHOR 2013). The following section will briefly summarise some of the main lessons from this field research experience.

Towards alternatives in household survey design: residential units and economic definitions of the household
In light of the challenges described in the previous section a search for alternatives remains imperative. However, the quest is likely to be a complicated one since the issues have been raised for decades, but operational imperatives in household survey design have prevented substantial innovations. Some marginal adjustments have been made over time in relation to definitions along residential lines, but more often the solution has been to collect more information at individual level to compensate for possible problems stemming from a focus on the household ("Laughlin 1999). In relation to gender dynamics, "Laughlin (1999) questions this departure and recommends not ignoring the concept and reality of the household, but instead to think carefully about its operationalization and contextualisation in terms of how class, politics and ideology define social difference and identity, and the nature and place of the "household".

Working with different analytically relevant typologies of households may also be an option. The distinction between "open" and "closed" households, as articulated in the various case studies reported by Randall and Coast (2014) is a useful one in this respect, insofar as it could inform sample stratification in large-N data collection, but may not solve all challenges and could be understood and operationalised in a variety of ways depending on context and survey objectives.

There may be other practical solutions which may work in a wider range of different contexts. Based on several years of experience designing labour and poverty micro-surveys (sometimes of relatively large N size) in rural Africa
(AUTHOR 2013, FTEPR 2014, Cramer et al. 2014), operationalizing suitable definitions of the household and its membership requires prior understanding of a number of issues, namely:

- the need to capture the mobile and „invisible” populations, especially migrant workers whose labour may be crucial for poor households residing at long-distances;

- variation in population density and residential patterns, which results in different types of residential units and forms of accommodation (where people actually sleep, regardless of whether the unit constitutes a household in the conventional sense);

- the nature of predominant socio-economic change and the potential extent of attrition in longitudinal surveys, that is, the risk of not finding the same households sometime after baseline surveys;

- Different forms of poverty, status, and livelihoods, which matter for household definitions since the nature of poverty also affects patterns in household composition and organisation.

All these are research questions in themselves, which can be tackled ex-ante, prior to a large-scale quantitative survey, in order to make sure that survey definitions and tools are designed to capture what matters for the objectives of the research. Going back to the value of mixed methods in poverty analysis, a judicious use of rigorous qualitative scoping research in the early stages of the research project with the aim of addressing some of the questions suggested
above may prove essential to avoid the common biases discussed in this paper. This is what Shaffer (2013, 111) calls „development”, that is, „the use of methods from one approach to assist in the methodological development of another”. However, this is only one of many ways of improving „established” methods in survey research, to borrow a term coined by Chambers (2003) in relation to the „Q-squared” debate.

Another option is to work with different definitions of the household, with implications in terms of who is counted or not in a household roster. Participatory methods sitting at the other extreme of „established” methods (Chambers 2003), whether qualitative or quantitative, could mean working with locally-defined concepts of the household, which could then be used to make large-N sampling methods more flexible. However, if locally-defined concepts depend on multiple criteria, depending on context, comparability across units of observation would be compromised.

Instead, considering a criterion different from physical residence, to be applied consistently, might be compatible with standardisation in large-N studies. For example, the use of „economic definitions” of the household, that is, the focus on relations of economic dependence (monetary or in-kind contributions, short or long-term) rather than on residential norms based on time spent in the „household”, is likely to be more effective for poverty analysis and labour research (Cramer et al. 2014). The operationalization of „economic linkages” between notional household members requires an easy and practical
definition of the „residential unit” to identify key individuals (respondents) who can then act as „anchors” for the identification of the empirically relevant household roster, including members who may never or rarely reside with the respondent in question. This demands a step-by-step identification of *economically linked individuals*, which is intensive in terms of probing needs and therefore requires well-trained and closely supervised interviewers. A key advantage of this step-by-step approach is the independence from existing official lists (government, village authorities, employers, etc.) and therefore greater reliability and likelihood of capturing „hidden” populations. A more accurate and context-specific sampling frame can thus be built to make survey samples more representatives of a wider range of potential respondents.

As it happens with most MM approaches to research on poverty, labour and other key development issues, these „innovations” are not always straightforward or cost-neutral. This is especially important if scaling-up is encouraged, i.e. if these methods could be streamlined in existing national data collection efforts. Resistance to change from established institutions in national data collection, especially in poor countries, is reasonably underpinned by concerns over cost and limited capacities. Indeed, substantive qualitative scoping research requires additional time, energy and (human and financial) resources. Likewise, the training and supervision needs of such an approach may be higher than a quick conventional survey, which takes established routines for granted and focuses on the mechanical aspects of the survey administration. A less ideal but perhaps more realistic option is to iteratively improve survey design through focused
micro-surveys in few pilot areas, which would help designers construct a more context-relevant household definition to be applied to a national data collection process, which would unavoidably involve larger numbers of less trained interviewers. The tension between context and standardisation might not be fully resolved but biases could be reduced.

Independently from the chosen approach and innovation however, improvements in training remain a priority for higher quality evidence. This does not mean a systematic need for specialist ethnographers to undertake and facilitate training, since many of the key issues can be conveyed to an audience that has sufficient knowledge of context and of relevant issues rather than higher education degrees or extensive survey experience.

**Concluding remarks**

This article has built on accumulated evidence on the advantages of MM approaches in development studies, and especially in poverty analysis, of which Howard White has been a long-standing advocate. The article has focused on the uses and misuses of the „household” as the key unit of analysis, a central aspect of research on poverty and labour. There are numerous challenges in the identification of suitable units of analysis and observation, and the debate on the definitions of the „household” has not been settled yet. Inertia and international conventions may also constitute obstacles to innovation in survey design, like the
ones proposed in section 3, which could help us avoid the traps described in the much of the critical literature on concepts of household in poverty analysis. However, the potential gains from revising and changing such conventions outweigh the potential costs and challenges as argued by much of the MM research in development studies. As O” Laughlin (1999, 34) argues, „neither the design nor the interpretation of household surveys is possible without reference to the qualitative and historical information that makes their categories and questions relevant”.

References

AUTHOR. 2013.


*Acknowledgements. I would like to thank the participants in the Colloquium „Improving Lives through Better Evidence“ for a productive and stimulating debate. The
This was preceded by another report, which laid the grounds for a case for mixed-methods approaches in the field of poverty analysis (Carvalho and White 1997). This report and the 2002 World Development article are in the top six of most cited publications by Howard White (Google scholar).

2 See Hentschel (1997) on the need to distinguish between data and methods.

3 The final point is crucial and also cogently argued by Cartwright (2004).

4 See special issue introduced by Guyer and Peters (1987), Folbre (1986) and O’Laughlin (1995) as seminal works questioning established assumptions about and uses of the concept of the household, especially in neoclassical economics.

5 See also project website at http://www.householdsurvey.info/