LIS
Technical Working Paper

No. 14

The Luxembourg Consumption Study (LCS): Feasibility and First Steps

G. Mancini, G. Vecchi

August 2024

Luxembourg Income Study (LIS), asbl

The Luxembourg Consumption Study (LCS):

Feasibility and First Steps

Giulia Mancini
University of Sassari

Giovanni Vecchi
Tor Vergata University of Rome

This draft:

December 16, 2023

Acknowledgements: We would like to thank Thesia Garner for useful comments, as well as Gintare Mazeikaite and Jörg Neugschwender for their contribution to data analysis.

1 Introduction

The objective of this report is to evaluate the feasibility of a Luxembourg Consumption Study (LCS) database, a consumption-focused counterpart to the Luxembourg Income Study (LIS) and Luxembourg Wealth Study (LWS) databases.

The prospective LCS would be, like its predecessors, a repository of harmonized microdata spanning multiple countries and years, including one or more harmonized welfare indicators that are comparable across surveys. The focus on consumption would add an important dimension to the analysis of living standards in LIS core countries, and it would create common ground between areas of the world that subscribe to different monetary welfare indicators: by and large, Europe, the Western Offshoots and Latin America tend to favor income as a welfare measure, while Africa and Asia tend to privilege consumption (Mancini and Vecchi 2022, p.14). A number of recent exercises delving into the joint analysis of consumption, income and wealth also attest the need for a repository of household consumption data that is truly harmonized across countries (Balestra and Oeheler 2023, Bebczuk et al. 2015, Fisher et al. 2021).

For this feasibility study, we consider a set of 25 countries, listed in Table 1. The choice of surveys has been guided by practical considerations, i.e., by the material that is currently already available as part of the LIS repository. However, it is not uncommon for countries to implement separate surveys for the purposes of collecting income and expenditure data, and the surveys listed in Table 1 naturally skew on the income side.[1] Expanding the assessment to consumption- and expenditure-focused surveys in the target countries is one of the next steps highlighted in the conclusions of this document.[2]

[1] For example, for the case of Italy, the Survey of Household Income and Wealth (SHIW) that we consider here contains only a few questions on consumption, but the main instrument used to collect consumption information in the country is the Italian SILC Survey, which is not included in this assessment. India is in a similar situation.

[2] Expenditure and consumption are both mentioned here to highlight the fact that countries tend to focus on one or the other when collecting data for surveys.

Table 1. List of surveys included in the feasibility assessment

	Country	Year	Survey
1	Australia	2016	Household Expenditure Survey and Survey of Income and Housing (HES and SIH)
2	China	2013	Urban and Rural Residents Income and Living Conditions Survey
3	Dominican Republic	2007	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)
4	Estonia	2000	Household Budget Survey
5	France	2010	Enquête Budget de Famille
6	Georgia	2019	Household Incomes and Expenditures Survey
7	Guatemala	2014	Encuesta Nacional de Condiciones de Vida (ENCOVI)
8	India	2011	Human Development Survey
9	Israel	2020	Household Expenditure Survey
10	Italy	2016	Survey of Household Income and Wealth (SHIW)
11	Ivory Coast	2015	Enquête sur le Niveau de Vie des Menages (ENV)
12	Japan	2008	Japan Household Panel Survey (JHPS)
13	Mali	2020	Enquête Modulaire et Permanente aupres des Menages (EMOP)
14	Mexico	2018	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)
15	Palestine	2017	Household Expenditure and Consumption Survey
16	Peru	2019	Encuesta Nacional de Hogares (ENAH)
17	Poland	2020	Household Budget Survey
18	Russia	2010	Russia Longitudinal Monitoring Survey (RMLS)
19	Serbia	2016	Household Budget Survey
20	Slovenia	2015	Household Budget Survey
21	South Africa	2017	National Income Dynamics Study (NIDS)
22	South Korea	2014	Household Income and Expenditure Survey
23	Switzerland	2004	Enquete sur les Revenus et la Consommation
24	Taiwan	2016	Family Income and Expenditure Survey
25	Vietnam	2013	Household Living Standards Survey

The process of building an LCS database starts with the definition of a concept of total household consumption expenditure – that is, the definition of a consumption aggregate. For this, we follow Deaton and Zaidi (2002) and Mancini and Vecchi (2022); these references underlie the global and regional poverty rates officially reported in the World Bank's Poverty and Inequality Platform (PIP) (World Bank 2023), and hinge on a common theoretical framework (Deaton and Muellbauer, 1980; Ravallion 2016) that is well-established among scholars and practitioners focused on the analysis of welfare through household survey data (Deaton 1997; Arndt and Tarp 2017). Two additional and related references are OECD (2013) and UNECE (2017).

The next step to determine the feasibility of an LCS database requires addressing several questions. The first one concerns *data availability*: do surveys provide the information needed to construct a consumption aggregate as per the references above? A second question concerns *data quality*, meant here in a broad sense: it refers to the accuracy, completeness, consistency, reliability, and timeliness of survey data, in turn influenced by the quality of the methods and instruments for data collection and processing. So, the second judgment to be made is: do enough surveys provide good quality data on consumption?

Finally, a few thoughts on *comparability*. The ability to make cross-country and intertemporal comparisons is a crucial feature of a prospective LCS database: on top of clearing minimum comprehensiveness and quality criteria, LCS consumption aggregates should be harmonized. Comparability extends to questionnaire design, as well as to the methods used to estimate components of expenditure that are not directly reported by respondents (imputed rent expenditure is a typical example). Realistically, ex-post harmonization is unlikely to lead to full comparability in a context such as the one at hand. Therefore, comparability is intended here in an *operational* sense, meaning that, for instance, the methodologies used to estimate a certain sub-component need not be identical, as long as they achieve the same conceptual goal. In other words, two consumption aggregates are comparable when they capture the same "achievement" associated to a particular welfare level. As an example, consider two countries where imputed rent for homeowners is included as a self-reported value in one case, and predicted via hedonic model in another: while the use of different methods is undoubtedly responsible for some of the observed difference between the housing expenditure components between the two countries, such a difference will be ignored, and the resulting consumption aggregates will be taken as "operationally comparable" (they both capture the functioning of housing).[3] Still, differences between measures that are "operationally comparable" will be acknowledged and carefully documented.[4]

The strategy we adopt in the rest of this document is as follows. In section 2, we outline a conceptual framework for the construction of the consumption aggregate: we identify its main components and define an aggregation plan. In sections 2.1-2.5, for each of the main components of the consumption aggregate, we define "checklists" that account for both data availability and data quality: the idea is to set out the criteria that a survey should meet for entering the LCS, and to document any relevant differences in survey design and expenditure estimation methods. Section 3 explains how the decision on whether a survey meets the criteria or not is made: we use a scoring system with some built-in flexibility. Section 4 presents the results, that is, the partial

[3] Section 2.5 provides more details on the housing component of the consumption aggregate.

[4] In the same spirit, the present document does not aim to establish a set of detailed rules for consumption aggregation that are meant to be applied identically across countries. Rather, it identifies some minimum criteria that an LCS-compliant consumption aggregate should meet, with the awareness that, even within the bounds of "compliance", the use of different methods for computing the aggregates may affect comparability. To explore these issues, Section 5 presents the results of a preliminary sensitivity study focused on alternative implementation approaches.

and overall LCS suitability scores for each survey. Section 5 presents a pilot study that computes comparable consumption aggregates for three countries, and assesses the effects of a few alternative definitions of the aggregate on poverty profiles. Section 6 concludes and outlines the next steps.

2 Conceptual framework

The assumption underlying the traditional approach to welfare measurement is that an individual's welfare depends on the *consumption* of a bundle of goods and services. Accordingly, the LCS focuses on consumption expenditure. In what follows we shall refer to it as "the consumption aggregate".

Ideally, analysts aim at constructing a *real* consumption aggregate, i.e., a monetary measure adjusted for both (1) purchasing power differences across households (within-survey inflation and geographical cost-of-living differences), and (2) differences in household size and composition. With some abuse of notation, we describe this as:

$$real \text{ consumption aggregate} = \frac{nominal \text{ consumption aggregate}}{CPI \times SPI \times ES} \qquad (1)$$

The numerator in equation 1 is the nominal consumption aggregate (CA), which is defined in detail in the rest of this section; we use simplified notation at the denominator, to denote a temporal price deflator (CPI) needed to compare households interviewed in different months during the survey period, a spatial price deflator (SPI), and an equivalence scale (ES) to account for economies of scale and different needs across households.[5]

In this document we do *not* discuss price adjustments. When a reliable national CPI is available, it should be used, and we expect it to pose no problems for most countries. Regarding SPI, deflation is more controversial from a theoretical standpoint, and empirically challenging, for most countries. This suggests that we leave the discussion of spatial deflation to a future step of the project (section 6 touches on this).

Regarding the equivalence scale in equation 1, in the rest of this document we use the

[5] A correct formulation of equation (1) would not assume that adjustments for cost-of-living differences and household characteristics be performed through simple multiplication: this functional form is adopted here to simplify exposition. In fact, the details of how these adjustments are performed often matter a great deal (see for instance Amendola et al. 2023).

consumption aggregate in per capita terms, whenever required. This is a convenient starting point (Deaton and Zaidi 2002). The definition of household size is not trivial, though, and it is important for a project involving countries with high heterogeneity in their demographic structure to consider it carefully. We briefly return to this issue in the conclusions.

In the rest of this section, we focus on the nominal CA, the numerator of equation 1. The nominal CA is the value of all goods and services consumed by members of the household during a certain reference period (typically one year). Consequently, the nominal CA should be 1) comprehensive (it should cover *all* consumption), 2) relevant (it should cover only consumption, that is, it should exclude outlays that represent investment, for instance), 3) typical (it should exclude contingent expenditures that are not representative of the household's consumption during a typical year), and 4) it should evaluate consumption at market prices[6] (Mancini and Vecchi 2022, p. 22). Schematically, the nominal CA can be defined as follows:

$$CA = x^F + x^{NFND} + \hat{x}^{HOUS} + \hat{x}^{DUR} \qquad (2)$$

where:

1. x^F denotes the food consumption aggregate
2. x^{NFND} denotes non-food non-durable expenditures
3. $\hat{x}^{HOUS}$ denotes the consumption flow from owner-occupied housing
4. $\hat{x}^{DUR}$ denotes the consumption flow from durable goods.

The consumption flow from durable goods and housing are denoted with a hat, because they require estimation. Each of the main sub-components of the nominal CA in equation 2 is, in turn, made up by finer expenditure groups. Table 2 details the "aggregation plan" for the nominal CA, down to the level of detail that is required to ascertain whether the coverage of the aggregate is satisfactory.

[6] This is true for food and non-food goods that are purchased, but also for own-produced food and in-kind transfers, which must be priced appropriately before their value can enter the consumption aggregate.

Table 2. The aggregation plan: nominal consumption aggregate

X			**Nominal consumption aggregate**
	XF		**Food consumption expenditure** (including alcoholic beverages)
		XFP	Food purchases
		XF0	Food own-produced
		XFI	Food in-kind
		XFF	Food away from home
	XN		**Non-food non-durable expenditure**
		XN2	Tobacco and narcotics
		XN3	Clothing and footwear
		XN4	Water, electricity, gas and other fuels
		XN5	Furnishings, household equipment and routine household maintenance
		XN6	Health
		XN7	Transport
		XN8	Information and communication
		XN9	Recreation, sport and culture
		XN10	Education services
		XN11	Restaurants and accommodation services
		XN12	Insurance and financial services
		XN13	Personal care, social protection and miscellaneous goods and services
		XNO	Lumpy expenditures (-)
	XD		**Consumption flow from durable goods**
	XH		**Shelter**
		XHR	Actual rent
		XHI	Imputed rent
X = XF + XN + XD + XH			

Notes: 1) The minus sign associated to XNO denotes that the entry should be subtracted from the consumption aggregate; operationally, "lumpy" expenditures can either be subtracted from each individual sub-aggregate, or grouped together and subtracted from total non-food non-durable expenditures. To improve readability, the table depicts this second option. 2) XN2 is not COICOP 02, as it excludes alcohol. 3) XN4 is not COICOP 04, as it excludes shelter (which is included separately as XH).

Knowing whether a survey contains the information necessary to construct a CA as in Table 2 is helpful, but it is not enough. The following sections discuss each component in detail and add some more dimensions to assess comparability between surveys.

2.1 Survey quality

For all countries included in the LCS, we set *minimum quality criteria* related to the

implementation of the survey. For example, all surveys should be nationally representative; the sampling frame should be up to date; the non-response rate should be "acceptable"; and so on.[7] Survey implementation and data processing are also to be accounted for: we do not want to include datasets that suffer from any major issues regarding field operations (sometimes areas in a country become inaccessible due to natural disasters, security, etc.), nor do we want to include datasets with deficient documentation regarding data editing and imputation. The latter are determined from our own subjective assessments and do not necessarily reflect the position of the national statistical offices that collected and released the data.

The proposal is to document these key implementation features via the grid in Table 3. More time and documentation, and possibly some interaction with the Statistical Offices of participating countries, would be needed to fill it in a satisfactory way. The sketch in Table 3 is intended to convey the plan.

[7] The threshold for deeming a low response rate 'unacceptable' is a matter of discussion; there are no hard-and-fast rules to set such a threshold. This aspect should be evaluated on a case-by-case basis in more advanced steps of the LCS project.

Table 3. Key survey features and data quality

Country	Year	Data collection period (months)	Data collection mode	Total survey sample size (hh)	Non-response rate (%)	Refusal rate (%)	Year of sampling frame	National coverage
Australia	2016	12		10,046	33.7	24.0		yes
China	2013							yes
Dominican Republic	2007	12		8,363	5.0		2002	yes
Estonia	2000							yes
France	2010							yes
Georgia	2019							yes
Guatemala	2014		PAPI					yes
India	2011		PAPI		8.0			yes
Israel	2018							yes
Italy	2016		CAPI					yes
Ivory Coast	2015	12						yes
Japan	2008							yes
Mali	2020	12		7,398	9.4	n/a		yes
Mexico	2018							yes
Palestine	2017			4,692	28.6	n/a	2007	yes
Peru	2019							yes
Poland	2020		CAPI					yes
Russia	2010							yes
Serbia	2016			6,457	27.1	n/a		yes
Slovenia	2015							yes
South Africa	2017		CAPI	13,719				yes
South Korea	2014							yes
Switzerland	2004							yes
Taiwan	2016		n/a					yes
United States	2021	12	CAPI	~11,000	54.4		2010	yes
Vietnam	2013		n/a					yes

Note: The table has been filled on the basis of publicly available documentation. Information on the year of the sampling frame year must be interpreted with caution: e.g. for the US, the frame is augmented on a regular basis to account for new construction and deconstruction.

2.2 Food expenditure

The food consumption aggregate (x^F in equation 2) "should include the (annualized) value of food consumed during the reference period, coming from all possible sources: 1) purchased in the marketplace (including meals purchased away from home, for consumption at or away from home); 2) produced by the household itself (food own-production is common among rural households); 3) received in-kind (as a transfer from other households, charities, or the government, or as payment in exchange for services rendered)" (Deaton and Zaidi, 2002, p. 27), and 4) food consumed away from home. Operationally, this can be expressed as follows:

$$x^F = x^P + x^O + x^{IK} + x^{FAFH} \quad (3)$$

where:

1. x^P denotes purchased food
2. x^O denotes own-produced food
3. x^{IK} denotes food received in kind
4. x^{FAFH} denotes food prepared away from home.

Note that while own-produced food may make up a significant share of household consumption in developing countries, this component tends to be negligible in advanced economies, so that household income and expenditure surveys often collect virtually no data on it. Therefore, while the need for a comprehensive x^F remains a general principle, it should be acknowledged that absence of information on x^O is not problematic in many contexts.

The proposed checklist to assess whether a survey is LCS-compliant for what concerns the food component is summarized in Table 4. Questions 1-4 establish the minimum data availability requirements: a survey *must* provide the information needed to estimate x^F, inclusive of all four components (net of allowances considering the prevailing consumption patterns in low- and high-income countries).

On top of this, questions 5-9 target a few important characteristics of the questionnaire, that have been shown to impact the estimated distribution of food expenditure (FAO and World Bank 2018). Questions 10-13 have to do with the need to potentially estimate the food component, or part of it: the use of non-standard measurement units, for instance, often calls for a rather complex conversion process in order to obtain the food aggregate; some surveys collect quantities rather than expenditures, so that the food component needs to be estimated via unit values or market prices, which comes with its own set of quality issues; it is useful to know whether it is possible to compute calorie intake, because that is often a valuable indicator of data quality; and so on. We also check whether the country at hand has a public distribution or *ration system* in place, because of the implications this would have on the estimation of a food expenditure component that ranks both recipients and non-recipients correctly.

Answers to questions 5-13 do not determine exclusion from the LCS: that is because the literature does not indicate that, for instance, a diary module is necessarily and, in all cases, better than a recall module, or vice-versa. However, it is important to keep track of these features when evaluating comparability, and to figure out the best strategy in the presence of dramatic differences.[8]

[8] For instance, surveys recording expenditures as broad aggregates, through global questions (such as "How much did

Table 4. Food expenditures: key questions to assess comparability

	Data availability
1	Can food consumption from purchases be included?
2	Can food consumption from own-production be included?
3	Can food consumption from in-kind receipts be included?
4	Can consumption of food away from home be included?
	Questionnaire design and other background information
5	Is food expenditure recorded via recall, diary, or a mixture of the two?
6	What is the recall period (in days), if data are collected via recall?
7	What is the reference period (in days) for the diary?
8	How many food items does the questionnaire list, approximately?
9	Does the questionnaire record food consumed, food acquired, or a mixture of the two?
10	Are respondents allowed to report quantities in non-standard measurement units (e.g. bunches)?
11	Is it possible to compute household calorie intake?
12	Does consumption expenditure need to be estimated from quantities (as opposed to being reported by respondents)?
13	Is there a public distribution/food rations program in the country?

In the instance of advanced economies with minimal food own-production, requirement 2 might not be considered stringent. Further discussion is warranted on this matter.

2.3 Nonfood nondurable goods

The second component of the nominal consumption aggregate in equation 2 refers to non-food non-durable goods (NFND):

$$x^{NFND} = \sum_{j \in J} x_j - x^{LUMPY} \tag{4}$$

where x^{LUMPY} is a catch-all variable denoting expenditures that do not represent consumption during the reference period (such as the full purchase price of consumer durables), and/or that are

you spend on clothing last month?"), versus item-by-item (for example, listing each item of clothing purchased last month); or surveys using a long diary with extensive supervision versus one with just a few visits (Beegle et al. 2012).

not part of the "habitual" household consumption pattern and therefore not representative of long-run living standards (such as expenditures related to a one-time special event).

Our proposed checklist for non-food non-durable expenditures is summarized in Table 5. As for the food component, on top of the minimum data availability requirements (questions 1-13), we consider a few additional requirements questionnaire design features: the use of the COICOP classification is an indication of how difficult it might be to reclassify non-food expenditures into standard subgroups, and the use of a diary or recall module is informative in terms of how comparable the data might be across surveys (questions 14-15).

Table 5. Non-food non-durable expenditures: key questions to assess comparability

	Data availability
1	Can alcohol and tobacco expenditure (COICOP group 02) be included?
2	Can clothing expenditure (COICOP group 03) be included?
3	Can expenditures for housing and utilities (COICOP group 04) be included?
4	Can expenditure for furnishings (COICOP group 05) be included?
5	Can health expenditure (COICOP group 06) be included?
6	Can transportation expenditure (COICOP group 07) be included?
7	Can information and communication expenditure (COICOP group 08) be included?
8	Can recreation and culture expenditure (COICOP group 09) be included?
9	Can education expenditure (COICOP group 10) be included?
10	Can hotel and accommodation expenditure (COICOP group 11) be included?
11	Can insurance expenditure (COICOP group 12) be included?
12	Can expenditure for personal care (COICOP group 13) be included?
13	Can "lumpy" expenditures be excluded?
	Questionnaire design and other background information
14	Does the questionnaire use the COICOP classification for non-food items?
15	Are NFND expenditures recorded via recall, diary, or a mixture of the two?

2.4 Housing

The fourth component in equation 2 is:

$$\hat{x}^{HOUS} = \begin{cases} x_{\text{rent}} & \text{if renter} \\ x_{\text{imputed rent}} & \text{if non} - \text{renter} \end{cases} \quad (6)$$

where

- x_{rent} is the rent reported by renters
- $x_{\text{imputed rent}}$ can be either self-reported or estimated by the analyst, for instance based on a hedonic regression model.

Imputed rents are needed to represent the flow of services for rental housing that is valued at less than market rents (e.g., public and subsidized rental units, living rent free) and for owner-occupants. The choice between self-reported rent and estimations based on a hedonic model (or other approaches) crucially depends on data availability: sometimes self-reported assessments are simply not recorded in the questionnaire.[9] However, the choice also depends on the context of the rental market in the country: for example, self-assessments may not be reliable in countries where owners and non-market occupants have little information on the going market prices due to a thin rental market (Balcazar et al. 2017, Ceriani et al. 2019, Garner and Short 2009,). In some cases, when the information is seriously deficient, other methods may be considered (e.g. capitalization based on reported purchase prices.

Table 6 summarizes the checklist for assessing the availability and comparability of information on dwelling expenditures. Once again, the proposed minimum requirement for LCS compliance is that a survey must allow for the estimation of rent for both renters and non-renters, regardless of the specific method (questions 1-2). Questions 3-9 assess which estimation methods are available to the analyst, based on the way the questionnaire is designed.

Table 6. Dwelling expenditures: key questions to assess comparability

	Data availability
1	Can actual rent for renters be included?
2	Can rent for non-renters be imputed (any method)?
	Questionnaire design and other background information
3	Does the questionnaire record actual rent for renters?
4	Does the questionnaire record self-reported imputed rent?
5	Type of self-reporting (willingness to accept/willingness to pay)
6	Does the questionnaire record the year of purchase of the dwelling?
7	Does the questionnaire record the original purchase amount?
8	Does the questionnaire record the current market value?
9	Number of dwelling characteristics available

[9] The specification of the hedonic model itself is context- and survey-dependent.

2.5 Durable goods

The third component in equation 2 is

$$\hat{x}^{DUR} = \sum_{k \in K} x_k^{CF} \tag{5}$$

where x_k^{CF} is the estimated consumption flow for durable good k.

The choice of the estimation model depends to a large extent on the information available in the questionnaire: for this reason, the checklist in Table 7 is designed to document which variables the analyst may count on. The minimum data availability requirement, in this case, is the possibility of including *any* reliable estimate of $\hat{x}^{DUR}$ in the consumption aggregate (question 1). Methods will almost certainly have to differ across countries, which means that the resulting aggregates may not be comparable in the strictest sense, but rather in the operational sense discussed in Section 1. Normally, x_k^{CF} is estimated using one of the approaches discussed in Amendola and Vecchi (2022), that is the 'geometric model', or the 'economic life model', or others – with the first two approaches being preferable, in general (Mancini and Vecchi 2022).

Table 7. Consumption flow from durable goods: key questions to assess comparability

	Data availability
1	Can we estimate a consumption flow from durable goods (any method)?
2	Is there sufficient information to apply the geometric depreciation model?
3	Is there sufficient information to apply the economic life depreciation model?
4	Is there sufficient information to apply another model?
	Questionnaire design and other background information
5	How many different durable goods are respondents asked about?
6	Are respondents asked about durable ownership (and/or access)?
7	Does the questionnaire record how many goods are owned by the household, by type? (e.g. 2 cars, 1 bike, etc.)
8	Does the questionnaire record the year of purchase?
9	Does the questionnaire record the year of production ("vintage")?
10	Does the questionnaire record whether durable is new or second-hand?
11	Does the questionnaire record the original purchase amount?
12	Does the questionnaire record the current market value?

3 How LCS feasibility is assessed

Each of the surveys included in this pilot is assessed by going through the checklists described in section 2. Each of the four components of the consumption aggregate receives a score that describes the extent to which that component can be constructed within the bounds of the LCS: in some cases, construction might be a matter of simply combining information that is collected by the survey; in other cases, a certain component might need to be estimated, based on available data; it might be that data gaps require workarounds or imputations; or that the information is simply not there. To convey the range of possible outcomes, we adopt the scoring system summarized in Table 8.

Table 8. Scoring system

Score	Interpretation	Color
1	All information available, no major data quality issues	
2	Minor information gap/data quality or comparability not optimal, but acceptable	
3	Significant information gap/data quality or comparability is an issue	
4	No information available/data quality or comparability seriously defective	

Both the definitions of the scores and the way they are assigned leave some room for flexibility and subjectivity. For example, considerations on data quality are necessarily partial and tentative at this stage. Given the heterogeneity of data collection approaches in the surveys considered by this study, the approach illustrated in Table 8 was considered preferable to a more rigid one, at least in the feasibility assessment phase.

As a final step, an overall LCS feasibility score (again, on a scale of 1 to 4) is assigned to the survey, taking the four CA components into consideration. The score indicates the extent to which the survey is suitable for inclusion in the LCS.

The following sections provide more information on how the scores are attributed, overall and for each component.

4 Results

Table 9 provides an overview of the results of the feasibility assessment. Results for each individual component are reported in Tables 10 to 13. Results are based on the data currently available in the LIS data files. This means that although expenditures and consumption related data might have been collected by a country's national statistical office in the household income and expenditure survey, only the income data were made available to LIS. As noted earlier additional data would need to be requested for further stages of LCS development.

Let us start from Table 9, which introduces two different consumption aggregates:

- "CA" (column 3) is the comprehensive aggregate described in equation 2, which includes expenditures for food, non-food non-durables, housing, and an estimate of the consumption flow from durable goods.
- "CA no durables" (column 4) excludes the consumption flow from durable goods.

The results in Table 9 indicate that the single most problematic component in terms of data availability is, indeed, the consumption flow from durables: only 5 out of 25 surveys provide fully satisfactory information on that front (where "fully satisfactory" means that the $\hat{x}^{DUR}$ component in equation 2 may be estimated using the geometric or economic life model), and another 8 provide some information ($\hat{x}^{DUR}$ may be estimated in some other way; more details are provided in Table 13). If the inclusion of the durables' component were to be considered a requirement for LCS suitability, then the constraints would be prohibitive. For this reason, we also consider an alternative "CA no durables".

Overall, about 20% of assessed surveys are perfectly suitable for the construction of a full coverage CA (including durables), and about 40% are not; the remaining 40% are somewhere in between (it would be possible to estimate the CA, but at some cost in terms of quality and/or comparability). If we reduce requirements, and target a CA that does not include durables, then about 60% of assessed surveys are perfectly suitable, 15% are not, and the remaining 25% fall in between.

Table 9. Component scores and overall feasibility score

		Overall		Components			
Country	**Year**	**CA**	**CA no durables**	**Food**	**Non-food non-durables**	**Housing**	**Durables**
				(table 10)	(table 11)	(table 12)	(table 13)
Australia	2016	4	2	1	1	3	4
China	2013	4	4	4	4	4	4
Dominican Republic	2007	1	1	1	1	1	1
Estonia	2000						
France	2010	2	1	1	1	1	3
Georgia	2019	2	1	1	1	1	3
Guatemala	2014	1	1	1	1	1	1
India	2011	4	2	2	1	3	4
Israel	2018						
Italy	2016	4	2	2	2	1	4
Ivory Coast	2015	3	3	3	2	2	3
Japan	2008	4	4	3	4	2	4
Mali	2020	2	1	1	1	2	3
Mexico	2018	2	1	1	1	1	3
Palestine	2017	1	1	1	1	1	1
Peru	2019	1	1	1	1	1	1
Poland	2020			1	1		
Russia	2010	2	1	1	1	2	3
Serbia	2016	2	1	1	1	1	3
Slovenia	2015	4	1	2	1	2	4
South Africa	2017	4	1	1	1	1	4
South Korea	2014						
Switzerland	2004	4	4	1	1	4	4
Taiwan	2016	3	3	3	2	2	3
United States	2021	4	2	2	1	1	3
Vietnam	2013	1	1	1	1	2	1

Note: The empty rows in the table (Estonia, Israel and South Korea) are surveys for which the available documentation did not allow to reach a definitive assessment.

The following paragraphs provide some details on the scoring of individual components, following the checklists described in Section 2.

Regarding food (Table 10), 14 out of 25 surveys score at 1, meaning that they provide information on all four components of the aggregate. A score of 2 or 3 is given to surveys that do not ask respondents explicitly about food received in kind or own-produced, or consumed away from home. A score of 4 was given only to China, where the survey seems to contain no information on consumption at all. Regarding the quality of the information on food consumption, questionnaire designs are disparate: about as many surveys use food diaries as they do recall modules, reference periods vary from 7 to 30 days, and respondents are prompted about food consumed or food acquired. Perhaps the dimension where we observe the widest differences in design is the level of aggregation of expenditures: some surveys, like Japan, Italy, Taiwan, India, or Mali, prompt respondents to report food expenditure as an aggregate or for broad categories (through the use of global questions), while most other surveys record food consumption at the item level, with varying levels of detail. This has been shown to make a difference to final results (the literature suggests that aggregate questions tend to underestimate consumption, see Jolliffe 2001, Pradhan 2009, Beegle et al. 2012, De Weerdt et al. 2016). Finally, for a majority of surveys examined in Table 10, the questionnaire suggests that it is possible to estimate household calorie intake, which can be used as an indicator of data quality.

Regarding non-food non-durables (Table 11), the vast majority of surveys (17 out of 25) allow for the inclusion of expenditures as discussed in section 2, and for the exclusion of "lumpy" ones. The only surveys with less-than-perfect scores are, once again, those that collect expenditures in very aggregated categories (these coincide with surveys that do not use the COICOP system), making it difficult to ensure that the non-food component has the correct composition. Overall, this component does not appear to be problematic.

Regarding housing (Table 12), 10 surveys received a score of 1, meaning that they provide information on actual rent for renters, self-reported imputed rent for owners, and the housing module is detailed enough to allow for the estimation of a hedonic rent model, if need be. 7 more surveys received a score of 2, meaning that homeowners are not asked to self-report imputed rent, but that the questionnaire allows for the estimation of a hedonic model. This restricts the choice set of the analyst when it comes to estimating housing expenditure, but still complies with the data availability requirements. A score of 3 is given to surveys where the estimation of imputed rent would be problematic and would need to rely on sub-optimal approaches (this is the case for Australia and India). For 2 countries, the information is simply not there, and the housing

component cannot be estimated, unless extra-survey information is available.

Finally, the durables component is the most problematic in terms of data availability (Table 13). Only 5 surveys provide the information needed to estimate a consumption flow based on the geometric or economic life depreciation model (these score 1). Another 8 surveys record ownership of durable goods, and sometimes the purchase amount for purchases made in the last 12 months, but not much else. These qualify as thorny cases, where the estimation of a consumption flow would require strong assumptions on the part of the analyst (these surveys score 3). Another 8 surveys do not even record ownership, which makes the estimation of a consumption flow all but impossible (these score 4).

Table 10. Food

		Questionnaire design								Aggregation plan				
Country	**Year**	**Recall/ diary**	**Recall period (days)**	**Diary period (days)**	**No. food items**	**NSU?**	**Consumption/acquisition**	**Possible to estimate kcal?**	**Need to estimate expenditure?**	x^P	x^O	x^{IK}	x^{FAFH}	**Score**
Australia	2016	diary	n/a	7	open list	no	acquisition	yes	no	yes	yes	yes	yes	1
China	2013	none	n/a	n/a	n/a	n/a	n/a	no	n/a	no	no	no	no	4
Dominican Republic	2007	diary	n/a	7	open list	yes	acquisition	yes	no	yes	yes	yes	yes	1
Estonia	2000													
France	2010	diary		7			acquisition			yes	yes	yes	yes	1
Georgia	2019	diary	n/a	7	open list	yes	acquisition	yes	yes (off-market)	yes	yes	yes	yes	1
Guatemala	2014	recall	30	n/a	116	yes	acquisition	yes	no	yes	yes	yes	yes	1
India	2011	recall	30	n/a	19	no	consumption	no	no	yes	yes	no	yes	2
Israel	2018													
Italy	2016	recall	30	n/a	0	no	acquisition	no	no	yes	no	no	yes	2
Ivory Coast	2015	recall	n/a	7 & 30	120		both	yes	no	yes	no	no	yes	3
Japan	2008	recall	30	n/a	recorded as a total	no	acquisition	no	no	yes	no	no	yes	3
Mali	2020	recall	7	n/a	30	no	consumption	no	no	yes	yes	yes	yes	1
Mexico	2018	diary	n/a	7	open list	yes	acquisition	yes	no	yes	yes	yes	yes	1
Palestine	2017	diary	n/a	30	357	no	both	yes	no	yes	yes	yes	yes	1
Peru	2019	recall	15	n/a	207	no	both	yes	yes (off-market)	yes	yes	yes	yes	1
Poland	2020	diary	n/a	30	110	no	both	yes		yes	yes	yes	yes	1
Russia	2010	recall	7	n/a	57	no	acquisition	yes	no	yes	yes	yes	yes	1
Serbia	2016	diary	n/a	30	open list	yes	both	yes	yes (off-market)	yes	yes	yes	yes	1
Slovenia	2015	diary	n/a	14	open list	n/a				yes	yes	yes	not clear	2
South Africa	2017	recall	30	n/a	32	n/a	both	no	no	yes	yes	yes	yes	1
South Korea	2014													
Switzerland	2004	diary	n/a	30	open list	no	acquisition			yes	yes	yes	yes	1
Taiwan	2016	diary	n/a	n/a	4	no	acquisition			yes	no	no	yes	3
Vietnam	2013	recall	30	n/a	80	no	consumption	yes	no	yes	yes	yes	yes	1

Note: NSU stands for non-standard measurement units. The empty rows in the table (Estonia, Israel and South Korea) are surveys for which the available documentation did not allow to reach a definitive assessment.

Table 11. Non-food non-durable goods

		Questionnaire design		Aggregation plan												
country	**year**	**COICOP**	**Recall / diary**	**Alcohol**	**Cloth.**	**Utilities**	**Furn.**	**Health**	**Transp**	**Comm.**	**Recr.**	**Educ.**	**Accom.**	**Insur.**	**Personal**	**Lumpy**	**Score**
Australia	2016		both	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
China	2013	no	none	no	no	no	no	no	no	no	no	no	no	no	no	no	4
Dominican Republic	2007	no	both	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Estonia	2000																
France	2010	no	recall		yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Georgia	2019	no	both	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Guatemala	2014	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
India	2011	no	recall	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Israel	2018																
Italy	2016	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	2
Ivory Coast	2015	no	recall	not clear	yes	yes	not clear	yes	yes	yes	yes	yes	yes	yes	yes	yes	2
Japan	2008	no	recall	not clear	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	4
Mali	2020	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Mexico	2018		both	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Palestine	2017	yes	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Peru	2019		recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Poland	2020		recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Russia	2010	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Serbia	2016	yes	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Slovenia	2015	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
South Africa	2017	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
South Korea	2014																
Switzerland	2004	no	both	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1
Taiwan	2016	no	n/a	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	2
Vietnam	2013	no	recall	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	1

Note: The empty rows in the table (Estonia, Israel and South Korea) are surveys for which the available documentation did not allow to reach a definitive assessment.

Table 12. Housing

		Questionnaire design						Aggregation plan		
country	**year**	**Actual rent recoded**	**Self-reported rent recorded**	**Willingness to accept/pay**	**Year dwelling was purchased**	**Purchase amount**	**Current value**	**Number of dwelling chars.**	x_{rent}	x_{irent}	**Score**
Australia	2016	yes	no	n/a	yes	yes	no	4	yes	yes	3
China	2013	no	no	n/a	no	no	no	0	no	no	4
Dominican Republic	2007	yes	yes	wta	yes	no	no	15	yes	yes	1
Estonia	2000										
France	2010	yes	yes	wtp	yes	yes	no	19	yes	yes	1
Georgia	2019	yes	yes	wtp	yes	no	no	11	yes	yes	1
Guatemala	2014	yes	yes	wtp	no	no	no	12	yes	yes	1
India	2011	yes	no	n/a	no	no	no	5	yes	not clear	3
Israel	2018										
Italy	2016	yes	yes	both	yes	yes	no	4	yes	yes	1
Ivory Coast	2015	yes	no	n/a	no	no	no	11	yes	yes	2
Japan	2008	yes	no	n/a	yes	yes	yes	7	yes	yes	2
Mali	2020	yes	no	n/a	no	no	no	15	yes	not clear	2
Mexico	2018	yes	yes	wtp	no	no	no	24	yes	yes	1
Palestine	2017	yes	yes	wtp	no	no	no	7	yes	yes	1
Peru	2019	yes	yes	wtp	no	no	no	16	yes	yes	1
Poland	2020	yes							yes		
Russia	2010	yes	no	n/a	no	no	yes	10	yes	yes	2
Serbia	2016	yes	yes	wtp	yes	no	no	7	yes	yes	1
Slovenia	2015	yes	no	n/a	no	no	no	20	yes	yes	2
South Africa	2017	yes	yes	wtp	no	no	yes	15	yes	yes	1
South Korea	2014										
Switzerland	2004	yes	no	no	no	no	no	0	yes	no	4
Taiwan	2016	yes	not clear	n/a	no	no	no	5	yes	not clear	2
Vietnam	2013	yes	no	n/a	yes	no	yes	7	yes	yes	2

Note: The empty rows in the table (Estonia, Israel and South Korea) are surveys for which the available documentation did not allow to reach a definitive assessment.

Table 13. Durable goods

		Questionnaire design							Aggregation plan					
country	**year**	**Ownership recorded**	**List of durables**	**Number items owned**	**Purchase year**	**Vintage**	**New/used status**	**Purchase amount**	**Current value**	x_k^{CF}	**Geometric feasible**	**Economic life feasible**	**Other models feasible**	**Score**
Australia	2016	yes	1*	no	no	no	no	no	yes	no	no	no	no	4
China	2013	yes	1*	no	no	no	no	no	yes	no	no	no	no	4
Dominican Republic	2007	yes	23	yes	yes	no	no	no	yes	yes	yes	yes	yes	1
Estonia	2000													
France	2010	yes	66	no	no	no	if bought last 12 months	if bought last 12 months	no	yes	no	no	yes	3
Georgia	2019	yes	29	yes	no	no	no	if bought last 12 months	no	yes	no	no	yes	3
Guatemala	2014	yes	50	yes	no	yes	yes	no	yes	yes	no	yes		1
India	2011	yes	24	no	no	no	no	no	no	no	no	no		4
Israel	2018													
Italy	2016	no	n/a	no	no	yes*	no	no	yes	no	no	no		4
Ivory Coast	2015	yes	24	yes	no	no	no	if bought last 12 months	no	yes	no	no	yes	3
Japan	2008	no	n/a	no	no	no	no	no	no	no	no	no		4
Mali	2020	yes	26	no	no	no	no	if bought last 12 months	no	yes	no	no	yes	3
Mexico	2018	yes	29	yes	yes	no	no	no	no	yes	no	no	yes	3
Palestine	2017	yes	35	yes	yes	yes	no	yes	yes	yes	yes	yes	yes	1
Peru	2019	yes	21	yes	yes	no	no	yes	yes	yes	yes	yes	yes	1
Poland	2020													
Russia	2010	yes	27	no	no	yes	no	no	no	not clear	no	no	yes	3
Serbia	2016	yes	15	yes	no	no	no	no	no	not clear	no	no	yes	3
Slovenia	2015	no	no	no	no	no	no	no	no	no	no	no		4
South Africa	2017	yes	27	no	no	no	no	no	no	no	no	no	no	4
South Korea	2014													
Switzerland	2004	no	no	no	no	no	no	yes	no	no	no	no		4
Taiwan	2016	yes	30	no	no	no	no	not clear	not clear	not clear	no	no		3
Vietnam	2013	yes	37	yes	yes	yes	no	yes	yes	yes	yes	yes	yes	1

*only family car

Note: The empty rows in the table (Estonia, Israel and South Korea) are surveys for which the available documentation did not allow to reach a definitive assessment.

5 Sensitivity analysis

This section presents the results of a sensitivity exercise conducted using recent household surveys from four countries: Bhutan, Italy, Malawi, and Peru.[10] The objective of the analysis is to assess the impact of changes in the definition of the consumption aggregate on key poverty estimates, particularly on poverty profiles (that is, poverty estimates computed by relevant population subgroups, such as urban and rural households).

Several considerations motivated the choice of the four surveys used for the exercise: first, they represent four very different contexts, national income levels, and survey instruments; second, they all contain sufficient information to allow for the computation of different versions of the consumption aggregate; and third, practical considerations played a role (the data were relatively accessible).

The rationale for the analysis is as follows: "compliance" with the criteria set out in sections 2 and 3, so that a survey's overall score could be as high as a 1 or a 2, allows for different methodological choices in the computation of the consumption aggregate, especially when it comes to the estimation of consumption flows for owner-occupied dwellings and durable goods (components $\hat{x}^{HOUS}$ and $\hat{x}^{DUR}$ in equation 2). The feasibility study has highlighted how the inclusion of an allowance for durable goods, in particular, is problematic. Therefore, it is important to understand the extent to which variations in the way these components are included in the consumption aggregate threaten the comparability of poverty profiles extracted from the harmonized data. This would help answer two broad questions: first, it would give an indication as to whether the ex-post harmonization required by the LCS should lean towards strictness, or rather should embrace a broader concept of comparability, in the "operational" sense, as discussed in section 1; second, it would reassure that even if the LCS went for a lower-coverage welfare indicator – one that does not include a consumption flow from durable goods – key results would remain relatively stable.

For each survey, three alternative consumption aggregates were constructed: the food and non-food, non-durable expenditure components were computed in the same way for each version (following the principles outlined in section 2), while the housing and durables components were

[10] The surveys are Bhutan's Living Standards Survey 2022, Italy's Indagine sulle Spese delle Famiglie 2021, Malawi's Fifth Integrated Household Survey 2019, and Peru's Encuesta Nacional de Hogares 2019.

obtained using different methods. Table 14 summarizes the definitions of the aggregates.

Table 14. Three alternative consumption aggregates for sensitivity analysis

Equation	Description
$PCE_1 = x^F + x^{NFND} + \hat{x}_{self}^{HOUS}$	PCE_1 includes the housing component as self-reported rent for owner-occupiers (and actual rent for renters). No allowance for durable goods is made.
$PCE_2 = x^F + x^{NFND} + \hat{x}_{hed}^{HOUS}$	For PCE_2 the housing component is computed using a hedonic model for owner-occupiers (renters are still associated with actual rent). No allowance for durables is made.
$PCE_3 = x^F + x^{NFND} + \hat{x}_{self}^{HOUS} + \hat{x}^{DUR}$	PCE_3 is the same as PCE_1, except an estimate of the consumption flow from durable goods is added.

Note: Hedonic models used to estimate $\hat{x}_{hed}^{HOUS}$ differ slightly across surveys, depending on the information on dwelling characteristics collected in each case. $\hat{x}^{DUR}$ is estimated using the "straight line" depreciation model (see section 2.5).

PCE_1 and PCE_2 compare the two main ways to include an allowance for the consumption flow of a household's dwelling, when no market rental price is paid. Self-reported values are typically collected by the survey, but if they are not available or if they are deemed unreliable, analysts can resort to hedonic estimation. Comparing results obtained for the same survey on the basis of PCE_1 vs. PCE_2, helps assess whether, all other things being equal, a lack of (or imperfect) harmonization of the housing component may lead to faulty comparisons. The same reasoning can be applied to the comparison between PCE_1 and PCE_3, which aims at exploring the impact of including or excluding the consumption flow from durable goods, all other things being equal.

A few technical clarifications are in order. First, for obtaining poverty estimates, we use each country's official national absolute poverty line, where available. In cases where a national poverty line is not available (such as Italy, which uses multiple household-specific poverty lines) a national value is recovered by inverting the quantile function for PCE_1 and selecting the value yielding the official national poverty headcount. Moreover, all comparisons are obtained by keeping the poverty line fixed. In principle, when the definition of the welfare indicator changes (from PCE_1 to PCE_2 to PCE_3) the poverty line should change accordingly: as a result, findings are likely to indicate an upper bound of the sensitivity of the poverty profile to the definition of

the consumption aggregate. Finally, the results presented in the rest of the section are *not* in any way comparable to official national estimates, given that the consumption aggregates were constructed from scratch for the purpose of this analysis, disregarding each country's own documented or undocumented procedures for computing official statistics.

Table 15 reports poverty headcount ratios computed by the households' area of residence (urban or rural), based on the three alternative definitions of PCE, for each survey. Although, as expected, the magnitude of poverty estimates varies with the adoption of different welfare indicators, the urban-rural rankings remain unchanged (with the exception of Italy, where the urban-rural difference is minimal in the first place). This suggests that the poverty profile is robust.

Table 15. Robustness of the urban-rural poverty profile (poverty headcount ratios, %)

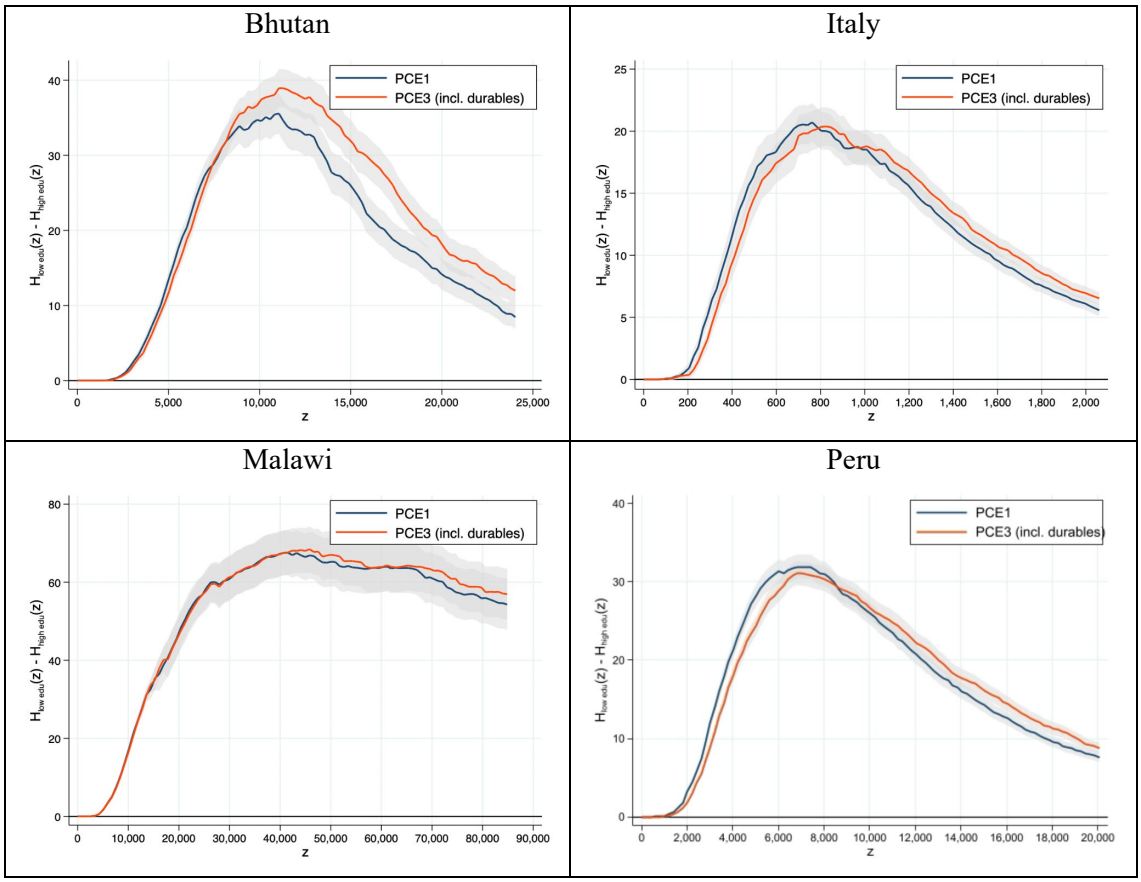
Country	Household subgroup	PCE_1	PCE_2	PCE_3
		self-reported rent, durables excluded	hedonic rent, durables excluded	self-reported rent, durables included
Bhutan	urban	5.0	4.9	3.7
	rural	29.6	28.4	25.6
Italy	urban	11.3	12.4	8.4
	rural	11.6	12.1	8.0
Malawi	urban	18.3	17.5	17.8
	rural	56.7	57.6	55.9
Peru	urban	11.8	6.9	6.3
	rural	34.6	30.1	28.9

Other household characteristics used to draw profiles are the gender, education, and employment status of the head or the household. Appendix A reports the full results, including additional poverty measures (poverty gap index, poverty gap squared index). In each case, estimates indicate an overall robustness of poverty profiles to variations in the definition of the consumption aggregate.

A fuller picture of the robustness of poverty comparisons can be gleaned by using *headcount difference curves* (see Mancini and Vecchi 2022) which plot, for any possible value of the poverty line, the difference between any two poverty headcount ratios H, for example $H_{urban} - H_{rural}$. In turn, headcount difference curves obtained from different consumption aggregates can be

compared to each other, to assess the sensitivity of the poverty profile. Figure 1 reports headcount difference curves obtained from PCE_1 and PCE_3 (durable goods excluded or included); the curves plot the difference between poverty headcount ratios for households whose head has low vs. high education (where "low" means less than secondary). For instance, the top-left curve indicates that, if the poverty line were 20,000 NU and the welfare indicator were PCE_1, then the poverty headcount ratio for households headed by someone with low education would be about 15 percentage points higher than the headcount for households with high-educated heads; the same difference would be about 20 percentage points if we used PCE_3 as a welfare indicator.

Figure 1. Headcount difference curves, household head low vs. high education, PCE_1 vs. PCE_3



The difference between headcounts for low- and high-educated headed households is always positive (households headed by someone who has lower-than-secondary education are always poorer); in no case does the choice of the welfare indicator reverse the ranking of the two population subgroups, regardless of the choice of the poverty line. The magnitude of the

difference does change according to the welfare indicator, but overall, the impact is rather small. Including or excluding durables seems to matter most in Bhutan, where the inclusion of durables magnifies the difference between poverty for high- and low-education households, though this is only true for relatively high values of the poverty line. In Malawi, the inclusion or exclusion of durables has the smallest impact on the profile. For both Peru and Italy, results are ambiguous: including durables decreases the high-low education difference in poverty rates when the poverty line is relatively low, while it increases it at high levels of the poverty line. Differences, however, are barely significant. Appendix A reports the complete set of poverty difference curves, which roughly paint the same picture as figure 1.

Overall, results seem to suggest that basic poverty profiles are, by and large, robust to some of the key methodological choices made when constructing consumption aggregates. The exercise is necessarily preliminary, but these early indications are reassuring.

6 Discussion and conclusions

Outlining the conceptual framework for constructing a consumption-based monetary welfare indicator is a relatively easy task. Several recent publications help identify the key elements that enter the definition of a consumption aggregate, and a core set of international institutions share a common understanding of how such an indicator should be constructed (Amendola and Vecchi, 2022; Balcazar et al 2017; Ceriani, Olivieri and Ranzani 2022; Belotti, Mancini and Vecchi, 2022; Deaton and Zaidi 2002; Mancini and Vecchi 2022, OECD 2013, UNECE 2011). Section 2 proposes an "aggregation plan" for the LCS consumption aggregate, summarized in Table 2.

The key challenge is the operational implementation of the framework presented in Table 2. The feasibility analysis described in sections 3 and 4 has considered three dimensions: i) data *availability* (do enough surveys contain the information needed to construct a consumption aggregate as in Table 2?), ii) data *quality* (do these surveys suffer from major flaws?), and iii) *comparability* (would the resulting welfare measure be comparable – even in a broad sense – across surveys?).

None of these questions, much less their combination, lead to Boolean conclusions: the heterogeneity of household questionnaires and survey designs is better handled with a multi-tier scoring system, a tool that allows us to introduce some nuances in the assessment. The results are clear enough, though, and have been summarized using the scoring presented in Table 8. Out of the 21 surveys that could be assessed based on the available documentation, 5 allow for the

construction of a comprehensive consumption aggregate, and 13 are suitable for the construction of a "reduced coverage" aggregate (one that does not include an allowance for durable goods). These numbers could be increased to 11 and 16, respectively, by including surveys that do not allow for "first best" estimation methods for the durables and housing components, but that are still viable. Overall, the evidence suggests that it is possible to go forward with building the LCS: within the limits of an ex-post harmonization process, most of the examined surveys do contain the information required to produce a consumption aggregate that fits the current international standards.

Looking ahead, several issues that emerged from this assessment will require further reflection. The rest of this section offers a summary of these open issues, and some suggestions for next steps.

The first set of considerations concerns data quality, and the need to compare results from surveys that are not harmonized ex-ante. The assessment of data quality summarized in Table 3 is to be considered highly preliminary and incomplete; it needs additional work in strict coordination with LIS headquarters, particularly in light of the fact that any conclusions drawn from it would apply to the LIS and LWS databases as well. The 'easy' part of this task is to ascertain that any survey entering LCS does not suffer from major flaws (lack of coverage of the entire national territory, exceptionally high nonresponse rates, and so on). Difficulties arise when it comes to comparing surveys characterized by major differences in their designs, for instance, PAPI-based surveys (e.g., India) with CAPI-based ones (e.g., Poland). While the literature does not usually offer definitive rankings between designs (FAO 2018, and World Bank 2018), experience suggests that comparisons are likely to be significantly affected by these differences. A systematic discussion of how to deal with major differences in survey questionnaires and survey implementation methods is left to future stages of the project.

Another issue concerns ex-post harmonization, that is, the procedure used to construct a welfare indicator that is comparable in terms of composition and estimation methods. This report has leaned on the concept of "operational comparability", that is, welfare indicators would be considered comparable as long as they include the main consumption components listed in the "aggregation plan" of Table 2, regardless of any methodological differences in the estimation of said components. This, however, does not mean that there shouldn't be a protocol for creating comparable consumption aggregates. For example, this report has mentioned that some household expenditures must be excluded from the consumption aggregate, but has refrained from specifying which ones, exactly. Some exclusions are relatively uncontroversial (if an allowance

is made for the consumption flow of durable goods, then the purchase of those same goods should not enter the aggregate), while others are debated: for instance, some expenditures for health and education goods and services may be considered investment, rather than consumption; health expenditures may not be considered indicative of well-being; the exclusion of infrequent and irregular expenditures is a point of contention; and so on. Similarly, alternative methods may be used to obtain a given component (think of imputed rent, for instance, or the consumption flow from durables), and again, these methods should be ranked from first best to last resort, in order to guide the ex-post LCS harmonization process.

A related problem has to do with the potential need to detect and treat outliers, or deal with missing values. This report is based on the analysis of questionnaires and survey documentation only: datasets (and their flaws) have not yet entered the picture. As soon as the project moves to the implementation phase, when consumption aggregates are constructed, there will be a need to design a consistent strategy to tackle data issues, if they were to emerge.

Other issues that need additional reflection are listed below, in no special order:

1. Links between LIS, LWS and LCS. The surveys included in this assessment were selected both for convenience (they are all part of the LIS repository), and because it was assumed that it would be interesting to see how measures of income, consumption and wealth interact with each other *within* a country (as well as across countries). Linking the three indicators is, however, potentially difficult: some surveys that collect both income and expenditure data advise against the joint analysis of the two (Cifaldi and Neri 2013). Some reflection on this point is needed.
2. Household definition. The definition of household membership may change across surveys (e.g., are boarders, servants, or partakers considered part of the household?). Should the harmonization of the definition of household, if it is possible, be part of the LCS (and potentially LIS) agenda?
3. Equivalence scales. A recent publication by the United Nations shows that Germany's average household size 2.05 compares to Senegal's 9.4. The discussion of the adjustment for household size and needs is left to future developments of the project, but identifying a strategy seems critical.
4. Rations and subsidies. The global coverage of LCS implies that countries with a variety of institutional settings may enter the database. In particular, special attention should be given to countries where expenditures reflect controlled (i.e., subsidized) prices. The large-scale food subsidies and public distribution systems that are in place in Bangladesh, India, Iraq,

Mexico, Indonesia, Sri Lanka, among others, are an example. These arrangements and other non-market valuing systems must be accounted for when consumption aggregates are compared internationally.

5. Consumption of health and education services. The inclusion of health and education expenditures in the consumption aggregate requires careful consideration of how to value services that are provided through public health and education systems. Most surveys collect data on out-of-pocket expenditures only, and given the cross-country differences in the public provision of these services, this point likely to affect comparability, and deserves some thought.
6. COICOP. The use of the (2018?) COICOP system as a way of organizing and eventually disseminating the breakdown of household expenditures in major groups (section 2.3) poses a problem for older non-COICOP based surveys. This needs additional thought.
7. Breakdown into sub-aggregates. While the consumption aggregate excludes some household expenditures (for instance, the purchase value of cars and other durable goods) some users may feel that a COICOP-group based breakdown of expenditures and consumption should include everything (expenditure for new cars included in transportation, etc.). On the other hand, it may not be desirable to have a wedge between the aggregate welfare measure and its sub-components. The strategy for dissemination of harmonized data needs additional thought.
8. Temporal deflation. Adjustment for within-survey inflation is desirable, particularly for countries experiencing double-digit yearly inflation, such as Argentina, Sri Lanka, Turkey, Iran, Pakistan, Ethiopia (leaving aside countries in hyperinflation, Zimbabwe, Venezuela, Sudan, etc.). In practice, the adjustment assumes the availability of reliable CPIs, which is not the case in a number of low-income countries. How to proceed? Any decisions would require consistency with LIS and LWS.
9. Spatial deflation. The adjustment for geographical cost-of-living differences is important for country-level welfare comparisons and might be significant for international comparisons as well. While further research is needed, this seems to be a low-priority task in the near future of the LCS.

Appendix A. Sensitivity analysis: additional results

Note: H indicates the poverty headcount ratio, PG stands for poverty gap index, PG2 stands for poverty gap squared index. All measures are reported as percentages (%). The three consumption aggregates (PCE1, PCE2 and PCE3) are defined in section 5.

Table A1. Selected poverty estimates, Bhutan 2022

	PCE1			PCE2			PCE3		
	H	PG	PG2	H	PG	PG2	H	PG	PG2
National	20.1	5.0	1.8	19.3	4.7	1.7	17.2	4.2	1.5
Urban	5.0	0.9	0.3	4.9	0.9	0.3	3.7	0.7	0.2
Rural	29.6	7.6	2.8	28.4	7.1	2.6	25.6	6.5	2.4
Male head	20.4	5.1	1.9	19.7	4.8	1.7	17.5	4.3	1.6
Female head	19.3	4.7	1.7	18.5	4.5	1.6	16.5	4.0	1.4
Head with secondary education and above	4.1	0.7	0.2	4.0	0.7	0.2	2.5	0.4	0.1
Head below secondary education	25.8	6.6	2.4	24.9	6.2	2.2	22.4	5.6	2.0
Head with no education	28.4	7.3	2.7	27.4	6.9	2.5	24.9	6.2	2.3
Head with primary education	11.9	2.3	0.7	11.0	2.2	0.6	9.9	1.8	0.5
Head with post-primary education	11.1	2.5	0.9	10.8	2.4	0.8	8.3	2.0	0.7
Head with secondary education	6.8	1.3	0.4	6.2	1.3	0.4	4.8	1.0	0.3
Head with tertiary education	2.9	0.4	0.1	2.9	0.4	0.1	1.5	0.2	0.0

Note: Bhutan's 2022 BLSS does not include any labor market information.

Table A2. Selected poverty estimates, Italy 2021

	PCE1			PCE2			PCE3		
	H	PG	PG2	H	PG	PG2	H	PG	PG2
National	11.4	2.6	0.9	12.2	2.8	1.0	8.2	1.7	0.6
Urban	11.3	2.7	0.9	12.4	3.0	1.1	8.4	1.9	0.6
Rural	11.6	2.5	0.9	12.1	2.7	0.9	8.0	1.6	0.5
Male head	12.2	2.8	1.0	13.1	3.0	1.1	8.7	1.8	0.6
Female head	9.6	2.1	0.7	10.2	2.3	0.8	7.1	1.4	0.5
Head employed	6.3	1.3	0.4	7.0	1.5	0.5	4.0	0.8	0.3
Head not employed	16.8	4.0	1.4	17.7	4.2	1.5	12.6	2.7	0.9
Head with secondary education and above	6.3	1.3	0.4	7.0	1.5	0.5	4.7	0.9	0.3
Head below secondary education	16.8	4.0	1.4	17.7	4.2	1.5	13.6	2.8	0.9
Head with no education	22.1	5.9	2.4	21.8	5.5	2.2	18.1	4.7	1.9
Head with primary education	15.8	3.7	1.4	16.6	3.9	1.4	11.3	2.6	0.9
Head with post-primary education	16.9	4.0	1.3	18.0	4.3	1.5	12.8	2.6	0.8
Head with secondary education	7.7	1.5	0.5	8.3	1.8	0.6	4.8	0.9	0.3
Head with tertiary education	3.2	0.6	0.2	3.9	0.7	0.3	2.0	0.4	0.1

Table A3. Selected poverty estimates, Malawi 2019

	PCE1			PCE2			PCE3		
	H	**PG**	**PG2**	**H**	**PG**	**PG2**	**H**	**PG**	**PG2**
National	50.7	18.5	8.9	51.3	18.5	8.7	50.0	18.2	8.7
Urban	18.3	4.8	1.8	17.5	4.3	1.6	17.8	4.6	1.8
Rural	56.7	21.1	10.2	57.6	21.1	10.1	55.9	20.8	10.0
Male head	48.5	17.4	8.2	49.2	17.4	8.1	47.7	17.0	8.0
Female head	56.6	21.5	10.6	56.9	21.3	10.4	56.0	21.4	10.5
Head employed	50.2	18.3	8.7	50.6	18.2	8.6	49.4	18.0	8.6
Head not employed	54.8	20.6	10.0	56.7	20.9	10.0	54.0	20.4	9.9
Head with secondary education and above	7.4	1.2	0.3	7.5	1.0	0.3	6.8	1.0	0.3
Head below secondary education	48.4	17.2	8.1	48.9	17.2	8.0	47.5	16.9	7.9
Head with no education	56.9	21.0	10.1	57.6	21.0	10.0	56.2	20.7	10.0
Head with primary education	42.3	13.7	5.9	42.8	13.7	5.8	41.3	13.3	5.7
Head with post-primary education	29.7	8.7	3.5	29.4	8.7	3.6	28.2	8.4	3.4
Head with secondary education	15.6	4.1	1.5	15.8	4.1	1.5	14.8	3.9	1.5
Head with tertiary education	16.4	2.2	0.4	16.4	1.7	0.3	15.3	1.7	0.3

Table A4. Selected poverty estimates, Peru 2019

	PCE1			PCE2			PCE3		
	H	**PG**	**PG2**	**H**	**PG**	**PG2**	**H**	**PG**	**PG2**
National	16.1	3.3	1.0	11.3	2.2	0.6	10.6	2.0	0.6
Urban	11.8	2.1	0.6	6.9	1.1	0.3	6.3	1.0	0.3
Rural	34.6	8.2	2.9	30.1	6.5	2.1	28.9	6.2	2.0
Male head	17.4	3.6	1.1	12.5	2.4	0.7	11.7	2.3	0.7
Female head	12.8	2.5	0.8	8.0	1.5	0.4	7.6	1.4	0.4
Head employed	17.1	3.5	1.1	12.2	2.3	0.7	11.4	2.2	0.7
Head not employed	11.0	2.1	0.7	6.3	1.2	0.3	6.1	1.1	0.3
Head with secondary education and above	9.2	1.7	0.5	6.2	1.1	0.3	5.6	1.0	0.3
Head below secondary education	24.0	5.1	1.7	17.1	3.4	1.1	16.2	3.2	1.0

Figure A1. Headcount difference curves, rural vs. urban households

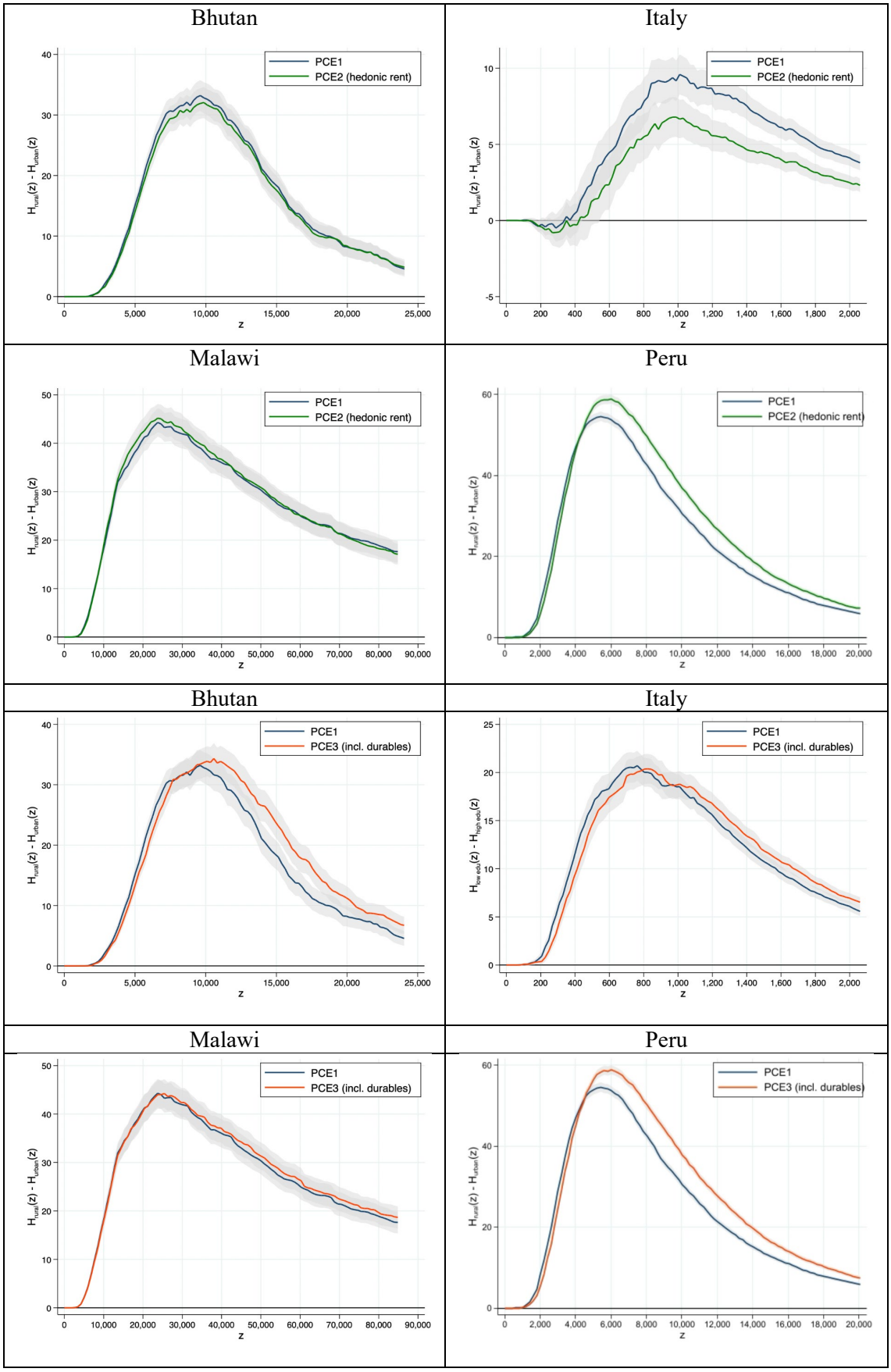


Figure A2. Headcount difference curves, household head low vs. high education

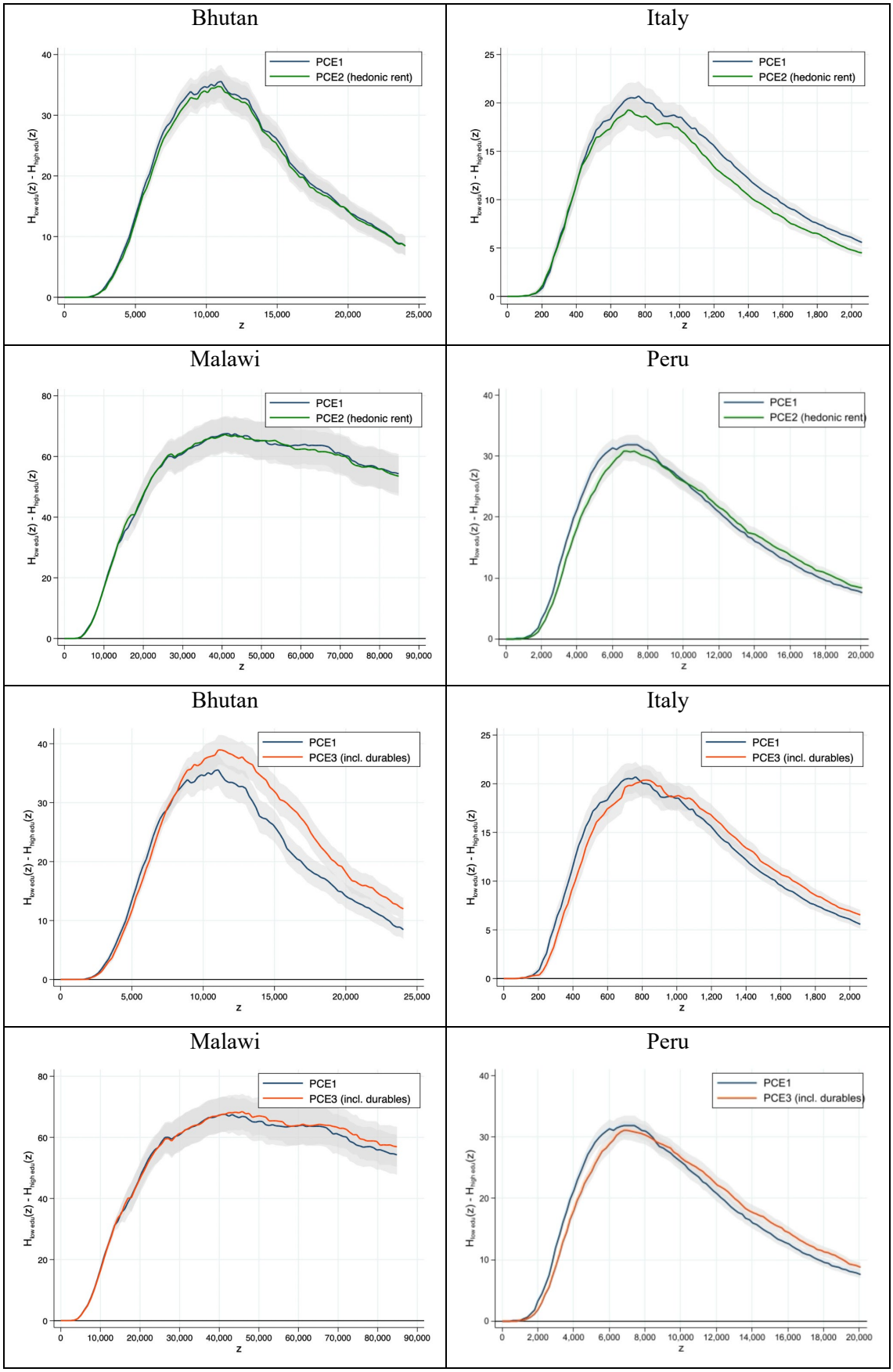
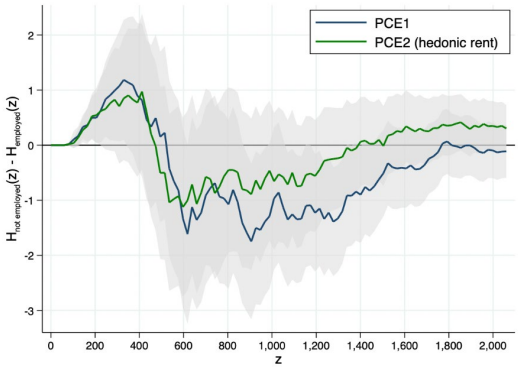


Figure A3. Headcount difference curves, household head not employed vs. employed

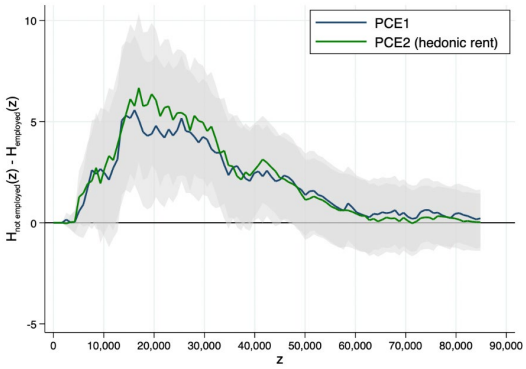
Bhutan

Bhutan's 2022 BLSS does not include labor market information.

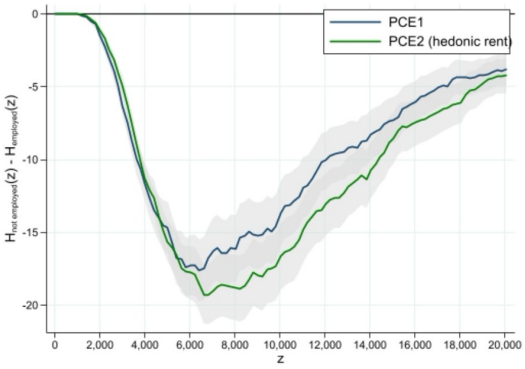
Italy



Malawi



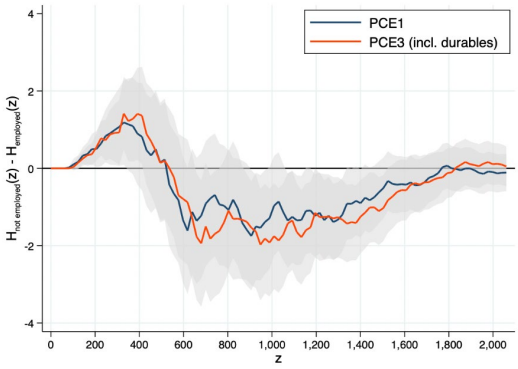
Peru



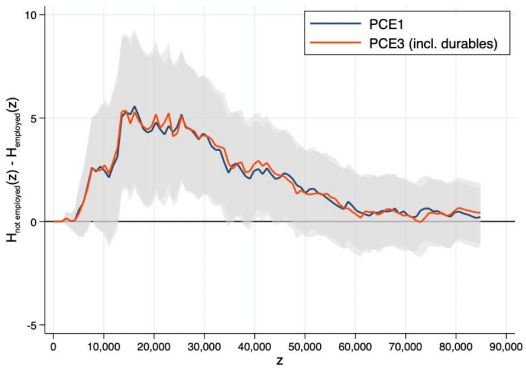
Bhutan

Bhutan's 2022 BLSS does not include labor market information.

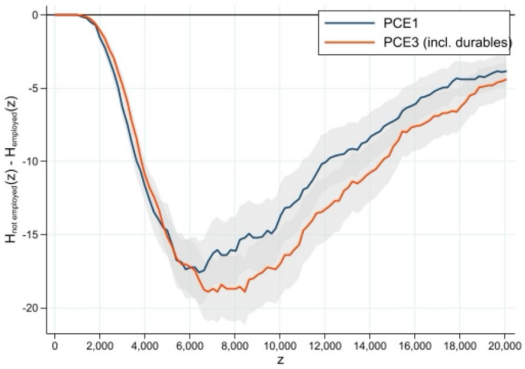
Italy



Malawi



Peru



References

Amendola, N. and G. Vecchi (2022), "Durable Goods and Welfare Measurement." *Journal of Economic Surveys* 36, no. 4: 1179-1211.

Amendola, N., Mancini, G., and Vecchi, G. (2023), "Deflation by expenditure components: a harmless adjustment?" *Review of Income and Wealth*, forthcoming.

Arndt, C. and F. Tarp (2017), (Eds.). *Measuring Poverty and Wellbeing in Developing Countries*, WIDER Studies in Development Economics. Oxford, New York: Oxford University Press.

Balcazar, C. F., L. Ceriani, S. Olivieri, and M. Ranzani (2017), "Rent-Imputation for Welfare Measurement: A Review of Methodologies and Empirical Findings." *Review of Income and Wealth* 63: 881–898.

Balestra, C. and F. Oeheler (2023) "Measuring the joint distribution of household income, consumption and wealth at the micro-level – Methodological issues and experimental results – 2023", OECD Papers on Well-Being and Inequalities no. 11.

Bebczuk, R., L. Gasparini, J. Amendolaggine, and N. Garbero (2015), "Understanding the Determinants of Household Saving: Micro Evidence for Latin America", IDB Technical note n. 843.

Beegle, K., J. De Weerdt, J. Friedman, and J. Gibson (2012), "Methods of Household Consumption Measurement Through Surveys: Experimental Results from Tanzania." *Journal of Development Economics* 98: 3–18.

Belotti, F., G. Mancini, and G. Vecchi (2022), "Outlier Detection for Welfare Analysis." *World Bank Policy Research Working Paper* no. 10231.

Ceriani, L., S. Olivieri, and M. Ranzani (2019b), "Housing, Imputed Rent, and Households' Welfare." *World Bank Poverty and Equity Global Practice Working Paper*, no. 213.

Cifaldi, G. and A. Neri. 2013. "Asking income and consumption questions in the same survey: what are the risks?", Bank of Italy Working Paper no. 908.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington, DC: The World Bank.

Deaton, A. and S. Zaidi (2002), "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." LSMS Working Paper no. 135.

Deaton, A. and J. Muellbauer (1980), *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

De Weerdt, J., K. Beegle, J. Friedman, and J. Gibson (2016), "The Challenge of Measuring Hunger through Survey." *Economic Development and Cultural Change* 64, no. 4: 727–758.

FAO and The World Bank (2018), Food Data Collection in Household Consumption and Expenditure Surveys. Guidelines for low- and middle-income countries. Rome: FAO.

Fisher, J.D., Johnson, D.S., Smeeding, T.M. and Thompson, J.P. (2022), "Inequality in 3-D: Income, Consumption, and Wealth." *Review of Income and Wealth*, 68: 16-42.

Garner, T. I. and K. Short (2009), "Accounting for owner-occupied dwelling services: Aggregates and distributions," *Journal of Housing Economics* 18, no. 3, 233-248.

Jolliffe, D. (2001), "Measuring Absolute and Relative Poverty: The Sensitivity of Estimated Household Consumption to Survey Design." *Journal of Economic and Social Measurement* 27, no. 1 and 2: 1–23.

Lanjouw, P. (2009), "Constructing a Consumption Aggregate for the Purpose of Welfare Analysis: Principles, Issues and Recommendations Arising from the Case of Brazil." Unpublished manuscript.

Lanjouw, J. O. and P. Lanjouw (2001), "How To Compare Apples and Oranges: Poverty Measurement Based on Different Definitions of Consumption." *The Review of Income and Wealth* 47, no. 1: 25–42.

Mancini, G. and G. Vecchi (2022), *On the Construction of the Consumption Aggregate for Inequality and Poverty Analysis.* Washington, DC: World Bank.

Pradhan, M. P (2009), "Welfare Analysis with a Proxy Consumption Measure: Evidence from a Repeated Experiment in Indonesia." *Fiscal Studies* 30, no. 3/4: 391–417.

OECD (2013), *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth.* Paris: OECD Publishing.

Ravallion, M. 2016. *The Economics of Poverty. History, Measurement and Policy*. New York and Oxford: Oxford University Press.

UNECE. 2017. *Guide on Poverty Measurement.* New York: United Nations.

UNECE (2011). *Canberra Group Handbook on Household Income Statistics*, 2nd edition (Chapter 2).

World Bank (2023), *Poverty and Inequality Platform Methodology Handbook.* Edition 2023-09. Available at https://datanalytics.worldbank.org/PIP-Methodology/.