Public Disclosure Authorized

Public Disclosure Authorized

Policy Research Working Paper 10698

Are Unit Values Reliable Proxies for Prices?

Implications of Better Price Data for Household Consumption Measurement in a Low-Income Context

Omoniyi Alimi
Wilbert Drazi Vundru
Talip Kilic

WORLD BANK GROUP
Development Economics
Development Data Group
February 2024

Reproducible Research Repository

A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

Policy Research Working Paper 10698

Abstract

Household Consumption and Expenditure Surveys are key to consumption-based monetary poverty measurement. In the absence of market price surveys that are linked to Household Consumption and Expenditure Surveys, unit values are used as proxies for market prices in estimating nominal consumption aggregates, price deflators, poverty lines, and poverty statistics. This practice relies on the Hicksian separability assumption: within-commodity group relative prices are constant across space and the price of a single good is an accurate proxy for the commodity group price. To test, for the first time in a low-income context, whether Hicksian separability holds, this paper uses the price data collected for an extensive list of food items, including several variety/quality-differentiated products for specific items, in a national market survey that was conducted in Malawi in sync with the Household Consumption and Expenditure Survey that is the source of official poverty statistics. The analysis demonstrates that Hicksian separability fails to hold across space and time and that unit values are biased proxies for prices. Integrating the Household Consumption and Expenditure Survey and market survey data based on location and timing of fieldwork permits an assessment of consumption and poverty estimation based on market prices versus unit values. Relative to unit values, using market prices leads to higher food and overall consumption expenditures—both in nominal and real terms—while generating higher poverty lines and higher food and overall poverty rates. Compared to their counterparts based on unit values, spatially-disaggregated poverty estimates based on market prices exhibit a stronger correlation with nightlights —an objective proxy for living standards.

This paper is a product of the Development Data Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at oalimi@worldbank.org and tkilic@worldbank.org. A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Are Unit Values Reliable Proxies for Prices?

Implications of Better Price Data for Household Consumption Measurement in a Low-Income Context

Omoniyi Alimi[†], Wilbert Drazi Vundru[‡], and Talip Kilic[⁑1]

.

[1] The authors are listed alphabetically. † Senior Lecturer, Department of Economics, University of Waikato, New Zealand, niyia@waikato.ac.nz. ‡ Survey Specialist, Living Standards Measurement Study, Development Data Group, World Bank, wvundru@worldbank.org. ⁑ Senior Program Manager, Living Standards Measurement Study, Development Data Group, World Bank, tkilic@worldbank.org. The authors would like to thank (i) Yuri Dikhanov and Giovanni Tonutti for their comments on the earlier draft, (ii) Japan Ministry of Agriculture for providing funding in support of the design and implementation of the Malawi National Market Survey, and (iii) the Malawi National Statistical Office management and survey teams for their dedication and hard work towards the successful implementation of the Malawi Fifth Integrated Household Survey (IHS5) 2019/20 and the Malawi National Market Survey. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the Government of Malawi, the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1 Introduction

Household Consumption and Expenditure Surveys (HCES) allow for the computation of comprehensive food and non-food consumption aggregates that underlie consumption-based monetary poverty and inequality estimates in low- and middle-income countries. These estimates typically rely on a measure of prices to value household consumption; construct poverty lines such as in the Cost of Basic Needs approach (Ravallion, 1992); and to feed into price deflators for temporally and spatially consistent measures of welfare.

In practice, there are two main competing proxies for prices.[2] The first is the unit value proxy, calculated as the ratio of the "consumption line item-specific" self-reported expenditures, or self-valuation of consumption, divided by the associated quantity of consumption – commonly specific to food consumption at home and based on the most knowledgeable respondents' self-reports. The length and specificity of HCES food consumption modules vary significantly across countries, and "line items", henceforth referred to as "commodity groups", that are included in these modules could be a mix of (i) variety/quality-differentiated versions of specific food types, (ii) specific food types listed separately within aggregate food groups, however without variety/quality differentiation, and/or (iii) aggregate food groups that are specified single entries but that combine different food types in their description.

The alternative to the unit value proxy is the market price proxy, derived from a market survey that captures prices for consumption items at market locations where surveyed households are expected to shop. The market survey can provide an opportunity to collect price data for a more disaggregated list of commodity groups vis-à-vis the associated HCES (i.e., across a finer breakdown of food types as well as their variety/quality-differentiated versions within aggregate food categories). Research has shown that market survey data collection across a more disaggregated list of commodity groups vis-à-vis the associated HCES would be valuable, even when the length and specificity of HCES consumption items are unchanged (Deaton, 1990; Gibson, 2013). However, in low- and middle-income countries, the use of unit values is the prevailing practice while estimating nominal consumption aggregates, price deflators and poverty lines (Mancini and Vecchi, 2022), since they can readily be estimated from household survey data, without needing to incur separate visits to markets which would bring about higher data collection costs and complexity in design and implementation of large-scale surveys. The implementation of a market survey that is spatially and temporally linked with an HCES is in fact a rare event (Gibson, 2013).

One assumption that underpins the use of unit values as proxies for prices is the concept of Hicksian separability. This is the assumption that relative prices of goods within a commodity group are constant across space. At the same time, within a commodity group, consumers can purchase several quality/variety-differentiated products, even concerning items such as "maize flour," "rice", or "soft drinks," that often appear in HCES food consumption modules as line items without quality/variety differentiation. If within-group relative prices vary across space, the composition of consumption

[2]There are other price proxies that have been explored in the literature including price opinions of the most knowledgeable household member or key informant about food items (see Gibson and Rozelle, 2005), and other existing price collection efforts for ongoing surveys like the CPI (Gibson, 2007).

within the group will shift towards locally cheaper goods and subsequently unit values cannot represent the same quality mix in all areas and cannot be treated as consistent proxies for prices across space (Gibson and Kim, 2019). For example, if consumers in urban areas consume higher-quality items, unit value differences between these areas and elsewhere will not only represent the price differences but also the quality differences. The violation of the Hicksian separability assumption would have direct implications for poverty and inequality statistics to the extent that unit values inform the underlying nominal consumption aggregates; the deflators used to make spatial and temporal price adjustments to these aggregates; and the poverty lines.[3]

It may be argued that the assumption of Hicksian separability may no longer be required with more detailed data, including scanner data on individual transactions or data from expanded survey modules and consumption item lists to better capture product specificity and quality variation. However, particularly in low-income countries, scanner data remain a long way away from constituting the majority of transactions (Gibson and Kim, 2019) and expanding consumption item lists in surveys to better capture variety and quality-specificity could non-negligibly increase the costs of survey data collection and respondent burden, with potential adverse effects on data quality – particularly in countries where consumption item lists are already considered "relatively detailed" (Beegle et al., 2012). Furthermore, even in the context of survey modules where important consumption items are sufficiently variety and/or quality-differentiated, as arguably in the case of Malawi, unit values are underlined by self-reported consumption expenditures or self-valuation of consumption originating from purchases. As such, they may exhibit measurement errors vis-à-vis local prices for the identical commodities that can theoretically be elicited in a market survey that is implemented in parallel with the household survey in question.

The assumption of Hicksian separability remains central to household consumption-based welfare estimates in developing countries, and despite its importance, there is a dearth of evidence on the validity of this assumption, except for the findings of Gibson and Kim (2015) who previously demonstrated that the assumption does not hold for a range of products in the Vietnamese context and McKelvey's (2011) rejection of the assumption based on data from a district of Central Java, Indonesia for a single food group (rice).[4] To contribute to this scant literature and provide a comprehensive examination of the validity of this assumption and its implications for poverty measurement for the first time in a low-income setting, we use detailed price data for an extensive list of food items, including a substantial number of quality-differentiated products within specific food groups, that were collected through a 12-month national market survey, which was conducted in Malawi by the National Statistical Office (NSO) in sync with the nationally representative household survey that is the source of official poverty and inequality statistics, namely the Fifth Integrated Household Survey (IHS5) 2019/20. The Market Survey relied on dedicated field teams that were trained, deployed, and managed by the NSO together with the IHS5 field teams. The selection of the markets and the timing of market visits were also informed by the timing of the visits to the IHS5 enumeration areas for consumption data collection. Overall, the Market Survey elicited price data for

[3] The Hicksian separability assumption would be invoked even in integrated household-market survey data collection, if the marker survey collects price information for aggregate commodity groups that combine different food items in their description, as opposed to specific commodities within these aggregate groups.
[4] Minten and Kyle (1999) also provide evidence of failure of the Hicksian separability assumption in Former Zaire in the context of marketing margins, but do not examine the implications for welfare estimation.

a larger number of food products and with greater spatial and temporal resolution vis-à-vis the above referenced studies.

The analysis shows that Hicksian separability fails to hold across space and time and that unit values constitute biased proxies for food prices in a low-income setting like Malawi. Mirroring the earlier findings of Gibson and Kim (2015) from Viet Nam, we provide evidence in support of at least one mechanism through which Hicksian separability fails, namely the Alchian-Allen effect – in other words, the fact that relative prices of high-quality and low-quality goods shipped from a common location (such as the nearest port, for imports, or a key geographic location for production and aggregation, for specific domestic products) are not constant over space. Mechanisms such as the Alchian-Allen effect and weak spatial markets integration can be expected to be stronger in low-income countries, thus making Malawi an ideal location to examine this assumption.

In view of these findings and taking advantage of the highly disaggregated market price data that were collected in tandem with the IHS5, we conduct a comparative assessment of consumption and poverty estimation based on unit values versus market prices. Relative to unit values, market price data showcase a different pattern of spatial price variation; lead to higher food and overall consumption expenditures – both in nominal and real terms – and higher food and overall poverty lines, culminating in higher food and overall poverty rates but without any effect on inequality estimates. In an attempt to provide operational guidance on the most cost-effective approach to market price data collection in tandem with an HCES, we subsequently examine the sensitivity of the market price-based results to (a) reduced spatial coverage in markets and (b) reduced item coverage in the market survey questionnaire. We find that estimates of food and total poverty rates are sensitive to reducing the spatial scope as well as the item coverage of the market price data collection effort. As such, for generating statistically consistent poverty estimates that are informed by market prices, it is essential to have a market survey that has a sizeable spatial and temporal overlap with household survey data collection, inclusive of price information on the same consumption items that are also featured in the household survey questionnaire.

The rest of the study proceeds as follows: Section 2 describes the Malawi national market survey – our source of market price data. Section 3 tests for Hicksian separability over space and time and examines the role of Alchian-Allen as the mechanism by which Hicksian separability fails. Section 4 reports the alternative estimates of poverty and inequality using the market price proxy and compares this with the traditional unit value-derived results. Section 5 reports the findings from the sensitivity analyses and Section 6 concludes.

2 The Malawi National Market Survey

The Malawi National Market Survey (henceforth referred to as the Market Survey) was implemented by the Malawi National Statistical Office from April 2019 to April 2020, in parallel with the Integrated Household Panel Survey (IHPS) 2019 and the Fifth Integrated Household Survey 2019/2020 – the latest round of the cross-sectional, nationally representative, multi-topic household survey used, among others, for official poverty and inequality statistics. The Market Survey was designed to achieve two

key objectives: (i) collect, from a national sample of markets, monthly price data for an extensive list of item-measurement unit combinations that are captured as part of the IHS5/IHPS data collection on food consumption, and, separately, agricultural production, and (ii) capture kilogram-equivalent conversion factors for all non-standard measurement units used in quantification of the aforementioned food consumption and agricultural production items. The latter improved the existing set of region-specific conversion factors and were used in the computation of the IHS5/IHPS consumption aggregates. The determination of the item-unit combinations that were featured in the Market Survey questionnaire was based on the analysis of the data from the past rounds of the IHS and IHPS in 2010, 2013 and 2016. The Market Survey collected data on Android tablets and a Computer-Assisted Personal Interviewing (CAPI) application that was designed using *Survey Solutions*.

Moreover, the selection of the markets and the decisions regarding their temporal distribution were finalized prior to the start of the IHS5/IHPS fieldwork. This process sought to (i) maximize the visits to the specific markets that were associated with the sampled IHS5/IHPS enumeration area (EA) locations and (ii) ensure that the sampled markets were visited in sync with the IHS5/IHPS field teams' visits to the associated EAs. This meant that some markets may have been visited once while others could have been visited multiple times during the fieldwork. Furthermore, the Market Survey was piggybacked on the monthly consumer price index (CPI) fieldwork operations and was implemented by the survey teams that were co-supervised by the NSO Prices Department (which is also responsible for the CPI) and the NSO IHS5/IHPS Survey Management Team. In specific months, some markets ended up being visited for both the CPI data collection and the Market Survey – albeit using different questionnaires.[5] Overall, a total of 1,071 visits were made to 377 markets over the 12-month survey period, spread across all 32 districts in Malawi. 47 percent of the selected markets were visited once during the survey period; 40 percent were visited between 2 to 5 times; and 13 percent were visited more than 5 times, mainly because they are the main markets in the different regions or districts. Table 1 provides an overview of the markets visited by month and region.

Table 1
Distribution of Market Visits by Month and Region

Month	Northern	Central	Southern
April 2019	9	14	19
May 2019	19	30	47
June 2019	10	33	37
July 2019	8	26	45
August 2019	17	48	52
September 2019	13	36	37
October 2019	14	30	41
November 2019	20	38	46
December 2019	5	19	29
January 2020	23	41	55
February 2020	13	24	40
March 2020	16	36	46
April 2020	7	19	9
Total	**174**	**394**	**503**

[5] The CPI Survey collects data on food prices and weights (only in standard units of measurement) and non-food prices for a fixed basket of foods specifically to measure inflation.

The Market Survey aimed to collect measurements of unit-specific prices and weights for a total of 163 food consumption and agricultural produce items. On food consumption specifically, the market survey collected information regarding 138 items and their associated measurement units – for a total of 534 item-unit combinations (i.e., approximately 4 measurement units per item). These 138 items represented 96 percent of the items listed in the IHS5 food consumption module and included 49 were elementary goods that were quality or variety-differentiated versions of 18 items that were included in the IHS5 food consumption module. Table A9 in Appendix 2 provides a full list of these elementary goods that were included in the market survey. For instance, IHS5 "Beef" entry was disaggregated, as part of the market survey, into "Beef with bones", "Beef without bones", "Ham beef", "Liver beef", "Minced beef", "Rump steak beef", and "Fillet beef" because these categories are expected to have differences in prices that would otherwise have been lost by only capturing the single item "Beef".

During each market visit, the survey team sought to collect kilogram weight and price information from up to three different vendors for each of the 534 item-unit combinations in the market survey questionnaire. This gives us highly disaggregated price data and ensures high precision in capturing spatial and temporal differences in prices across a wide range of items in Malawi. A total of 435,956 data points were collected across the markets and the item-unit combinations during the 12-month fieldwork. The full description of the market data collection and processing is discussed in Appendix 1.

3 Hicksian Separability: Are Relative Prices of Elementary Goods Constant across Space and Time?

The measurement of household living standards and poverty in many developing countries, including Malawi, relies on unit values to proxy for prices faced by households in different localities and time periods. Unit values are ratios of "item-level" expenditures to associated quantities of consumption - commonly specific to food consumption at home and based on most knowledgeable respondents' self-reports. The length and specificity of HCES food consumption modules exhibit significant cross-country heterogeneity. As such, food consumption items, referred to as "commodity groups", can include a mix of (i) variety/quality-differentiated versions of specific food types (e.g., broken rice, medium grain rice and long grain rice, being listed as separate items), (ii) specific food types broken down within aggregate food groups, however without variety/quality differentiation (e.g., Cassava, Sweet Potato, Yellow Potato, being listed as separate items under the Roots and Tubers food group), and/or (iii) aggregate food groups that may appear as single line items in a module but that may combine different types of food items (e.g. Vegetables and Fruits).

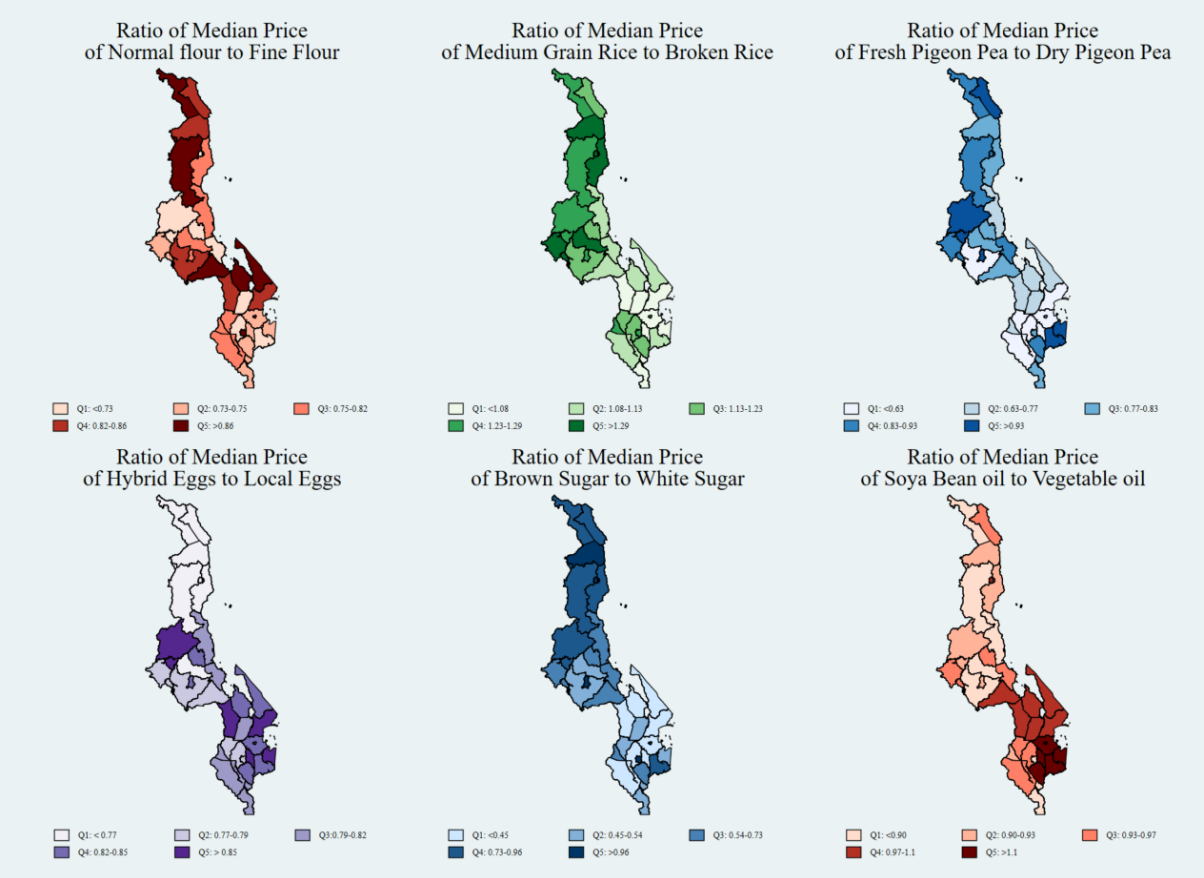
Under the assumption of Hicksian separability, if the price vector for the elementary goods within a group is decomposed into (i) a scalar term that raises or lowers the price level of all goods in the group across locations, and (ii) a reference price vector of the relative price of each good within the group, it is the inter-area scalar variation that dominates intra-group variation in relative prices (Deaton, 1988). Were this not true, when the unit value from one household is compared to the unit value of a household elsewhere (or to the unit values of the same household in a different season), we would not know if the difference in the unit value is because prices are different in the two places (or two seasons), or because the households are consuming a different mix of elementary goods within the group.

Taking maize flour as an example, one driving force for changing the composition of elementary goods within the group is that relative prices of the various types of maize flour may not be constant over space and time. If Hicksian separability does not hold, the poverty analysis risks confusing quality of living (buying higher-quality maize flour) with the cost of living (facing a higher price for all maize flour). There may be several reasons why relative prices within a food group may vary over time and space. If transport, storage, and processing costs are specific and not ad-valorem, then relative prices are unlikely to be constant over time and space. This effect is well documented as the Alchian-Allen effect (Borcherding and Silberberg, 1978; Gibson and Kim, 2015) and it will lead to relative price of quality varying over space and time. The integrated implementation of the Market Survey and the IHS5 gives us an opportunity to examine the evidence for Hicksian separability over space and time.

Before formally testing for Hicksian separability, we provide descriptive evidence on relative prices across space (at the district-level) for selected IHS5 consumption food items.[6] Figure 1 shows the variation across space in the price ratio of elementary versions of the selected food items. The figure suggest that within-item price ratios are not constant over districts. Using rice as example, while there is 1.3 times premium in the district median price of medium grain rice relative to broken rice across the Northern districts, the premium declines and is even non-existent in some of the districts in the South.

Figure 1

Price ratio of selected food items across districts in Malawi



[6] These goods are selected based on their importance in Malawian diets as determined by their budget share.

3.1 Method

To test for Hicksian separability over space, we check if estimated differences in prices between areas or over time for a given commodity group are the same regardless of which elementary good within the consumption item/group is used to represent the local price. The spatial level of all analysis presented is the district-level. This is because of sparse data at levels below the district. We implement the following regression:

$$lnp_{gm} = \alpha_{gm} + \beta D_{gm} + \gamma X_{gm} + \pi X_{gm} D_{gm} + \mu_{gm} \qquad (1)$$

where *g* and *m* denote elementary good and market, respectively; *p* is market-specific elementary good median price computed over time; *D* is the vector of district-fixed effects; *X* is the vector of elementary good identifiers; and α and μ denote constant and error term, respectively. This regression is estimated by consumption item or *food group* (e.g., rice) using the pooled observations of elementary products (e.g., different types of rice) in that food group. The elementary products in the market survey range from two (sugar) to eight (beef) depending on the food group. If we reject the coefficient $\pi = 0$ for all districts, we can conclude that the pattern of district price differences is sensitive to which of the elementary goods is used – in violation of the Hicksian separability assumption. This would suggest that using a single representative good to estimate inter-district price differences may cause bias since a different pattern of area-fixed effects results from using one elementary good rather than the others.

For Hicksian separability across time, we implement the following regression:

$$lnp_{gmt} = \alpha_{gmt} + \beta T_{gmt} + \gamma X_{gmt} + \pi X_{gmt} T_{gmt} + \mu_{gmt} \qquad (2)$$

where *g*, *m,* and *t* denote elementary good, market, and month, respectively; *p* is month-specific median price in each market; *T* is the vector of month fixed effects; *X* is the vector of elementary good identifiers; and α and μ denote constant and error term, respectively. This regression is again estimated by food group (e.g., rice) using the pooled observations of elementary products (e.g., different types of rice) in that food group. If we reject $\pi = 0$ for all months, we can conclude that the pattern of price differences over time is sensitive to which of the elementary goods is used – again in violation of the Hicksian separability assumption. This would suggest that using a single representative good to estimate inter-period price differences may cause bias since a different pattern of time effects results from using one elementary good rather than the others.

3.2 Results

Table 2 reports the results from the estimation of Equation 3 (spatial test) by food group – specifically the F-statistic and the p-value associated with the test of $\pi = 0$ for all districts. We reject this hypothesis for all food groups, which implies the assumption of Hicksian separability across space is not defensible. We do not find that relative prices of goods within a commodity group are constant over space, as can be seen from the sensitivity of the area-fixed effect to the choice of elementary good.

Table 2

Spatial Hicksian Separability Test

Dimension = Spatial	# of Elementary Goods in Food Group	F-stat	P-value	R-sq	N
Maize Flour	5	49.33	0.0000	0.96	1174
Rice	3	8.59	0.0000	0.66	780
Bread	2	13.64	0.0000	0.55	684
Buns	2	9.45	0.0000	0.54	739
Cassava	2	24.75	0.0000	0.68	544
Irish Potato	2	5.41	0.0000	0.43	717
Pigeon Pea	2	25.51	0.0000	0.81	670
Beans	3	9.35	0.0000	0.60	575
Onion	2	10.31	0.0000	0.53	677
Chinese Cabbage	2	18.77	0.0000	0.53	405
Eggs	2	126.29	0.0000	0.91	678
Beef	8	26.62	0.0000	0.81	1359
Pork	2	8.50	0.0000	0.81	586
Chicken	5	18.40	0.0000	0.67	1399
Banana	2	8.09	0.0000	0.55	753
Fresh Milk	2	11.50	0.0000	0.71	273
Sugar	2	133.72	0.0000	0.93	687
Cooking Oil	4	62.99	0.0000	0.86	1471

Note. This table reports the results of whether pattern of district price differences is sensitive to the choice of elementary goods used. F-stat, P-value, and R-sq are for the hypothesis of no difference in patterns of prices across districts.

Table 3 reports the results from the estimation of Equation 4 (temporal test) by food group – specifically the F-statistic and the p-value associated with the test of $\pi = 0$ for all months. We reject this hypothesis for all food groups except for Irish Potato and Beans, and this result implies that the assumption of Hicksian separability across time is also not defensible for most items.

Table 3

Temporal Hicksian Separability Test

Dimension = Temporal	# of Elementary Goods in Food Group	F-stat	P-value	R-sq	N
Maize Flour	5	19.64	0.0000	0.86	1857
Rice	3	6.02	0.0000	0.31	1357
Bread	2	3.24	0.0001	0.21	1221
Buns	2	3.41	0.0001	0.09	1296
Cassava	2	3.01	0.0015	0.11	1017
Irish Potato	2	1.38	0.1674	0.21	1360
Pigeon Pea	2	7.89	0.0000	0.43	869
Beans	3	2.29	0.1307	0.12	971
Onion	2	4.32	0.0000	0.15	1282
Chinese Cabbage	2	3.30	0.0694	0.12	992
Eggs	2	67.76	0.0000	0.63	1236
Beef	8	2.93	0.0000	0.40	1670
Pork	2	4.26	0.0000	0.13	748
Chicken	5	1.83	0.0005	0.09	1963
Banana	2	11.13	0.0000	0.31	1621
Fresh Milk	2	3.24	0.0003	0.26	405
Sugar	2	21.85	0.0000	0.32	1247
Cooking Oil	4	39.43	0.0000	0.54	2780

Note. This table reports the results of whether pattern of district price differences is sensitive to the choice of elementary goods used. F-stat, P-value, and R-sq are for the hypothesis of no different in patterns of prices across districts over time.

Taking advantage of the spatial and temporal dimensions of our data, we combine both space and time dimensions by testing for Hicksian separability across space (i.e., estimating Equation 3) separately with observations elicited during the harvest season (April 2019 to September 2019 and including April 2020) versus the lean season (October 2019 to March 2020). Table 4 reports the results from the estimation of Equation 1 by food group during the harvest season and lean season, respectively. We also reject the assumption of Hicksian separability across space in both harvest and lean seasons for most food groups.

Table 4

Spatial Hicksian Separability Tests in Harvest and Lean seasons

		Harvest period				Lean period			
	# of Elementary Goods in Food Group	F-stat	P-value	R-sq	N	F-stat	P-value	R-sq	N
Maize Flour	5	10.32	0.000	0.94	955	17.15	0.000	0.92	902
Rice	3	3.47	0.000	0.54	648	2.92	0.000	0.46	709
Bread	2	3.48	0.000	0.21	624	12.45	0.000	0.39	597
Buns	2	7.63	0.000	0.42	724	6.63	0.000	0.49	572
Cassava	2	5.73	0.000	0.34	491	7.71	0.000	0.39	526
Irish Potato	2	4.96	0.000	0.33	672	2.7	0.000	0.29	688
Pigeon Pea	2	6.2	0.000	0.58	475	0.07	0.793	0.44	394
Beans	3	1.23	0.254	0.37	495	NA	NA	NA	NA
Onion	2	2.68	0.000	0.3	638	2.28	0.002	0.27	644
Chinese Cabbage	2	9.21	0.003	0.4	497	NA	NA	NA	NA
Eggs	2	15.15	0.000	0.42	620	28.86	0.000	0.69	616
Beef	8	5.83	0.000	0.69	974	9.13	0.000	0.67	696
Pork	2	2.16	0.005	0.6	433	12.17	0.000	0.61	315
Chicken	5	2.89	0.000	0.33	1052	6.69	0.000	0.41	911
Banana	2	4.31	0.000	0.39	805	11.32	0.000	0.55	816
Fresh Milk	2	2.27	0.082	0.39	217	NA	NA	NA	NA
Sugar	2	47.6	0.000	0.78	668	4.68	0.000	0.25	579
Cooking Oil	4	5.00	0.000	0.33	1388	9.36	0.000	0.54	1392

Note. This table reports the results of whether pattern of district price differences is sensitive to the choice of elementary goods used. F-stat, P-value, and R-sq are for the hypothesis of no difference in patterns of prices across districts in Harvest and Lean periods. Beans, Chinese cabbage, and Fresh milk are not included in the results for the lean season due to insufficient number of price observations for the associated elementary goods during that time.

We find strong evidence that within-group prices are not constant over time and space for most items. Inter-district and inter-month differences captured in unit values will thus be capturing both differences in quality mix and differences in prices. Our results are similar to the existing evidence on Hicksian separability found in Viet Nam by Gibson and Kim (2015). Their study found that relative prices within food groups were not constant across space in Viet Nam and suggests the Alchian-Allen effect as the mechanism for why Hicksian separability fails.

The Alchian-Allen effect implies that if the prices of two substitutable goods, such as high- and low-quality versions of the same item, are both increased by a fixed per-unit amount, consumption will shift toward the higher-grade product (Alchian and Allen, 1967; Gould and Segall, 1969; Borcherding and Silberberg, 1978). This means that when the prices of two elementary goods in the same food group are both increased by a fixed per-unit amount such as a transportation or storage cost, consumption will shift toward the higher-grade product. In other words, with fixed charges for transportation and storage, relative prices will be expected to vary across time and space, in violation of what is to be expected with Hicksian separability. For goods with a quality gradient over which it costs similar fixed amounts to transport such goods over space and store them over time, we can

expect to see a clear pattern of decline in the price ratio of quality in areas far away or over time from excess supply points:

$$\frac{P_{Bh}}{P_{Bl}} > \frac{P_{Bh} + f}{P_{Bl} + f} \qquad (3)$$

Where P_{Bh} is price of high-quality gradient elementary items of good B high-quality gradient of good B; P_{Bl} is price of low-quality gradient elementary items of good B; and f is a fixed amount to transport (store) goods over space (time). In what follows, we check for evidence of the Alchian-Allen effect as the mechanism by which the Hicksian separability assumption is violated in Malawi. To do so, we use beef as an example of a product with a quality gradient and for which the cost of transporting the high- and low-grade versions should be similar. Next, we define a set of markets that should be the excess supply points for beef in Malawi. To identify excess supply points, we relied on (i) an NSO-supplied list of main markets and the items considered to be a main product in the market, and (ii) FEWS NET (2018). Subsequently, we estimate the following regression:

$$PR_{im} = \alpha_{im} + \beta_1 Distance_{im} + \mu_{im} \qquad (4)$$

where *i* and *m* denote food group (beef) and market, respectively; PR_{im} is the price ratio of high/low quality observations for a given food group (*beef* in this case) in market m**;** $Distance$ is the Euclidian distance in kilometers between the market that a given price ratio value is associated with and the reference market for beef (based on GPS coordinates for the markets); and α and μ denote the constant and error term, respectively. Table 5 presents the results of our test, based on distance measures with respect to three different excess supply points (i.e., reference markets), namely Nsanje Boma Market, Blantyre City Market, and Lilongwe Area 36 Market).

Table 5
Test of Alchian-Allen effect: Beef

	Beef		
	Nsanje Boma Market	Blantyre City market	Lilongwe Market
Distance (Kms)	-0.0151***	-0.0155***	-0.0144**
	(0.005)	(0.005)	(0.007)
Constant	1.239***	1.216***	1.202***
	(0.025)	(0.019)	(0.021)
Observations	302	302	302
R-squared	0.034	0.032	0.013

Note. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The price ratio of quality (elementary goods) used is the ratio of Beef without Bones to Beef with Bones. Reference markets are the main markets in Blantyre City, Nsanje Boma, and Lilongwe Market-Area 36, as determined by the NSO-supplied list of main markets for specific products and FEWS NET (2018).

The results indicate that the price ratio of high-quality beef to low-quality beef falls by around 1.4 to 1.5 percentage points for every 100 kilometers away from the reference markets of Nsanje Boma, Blantyre City and Lilongwe Market-Area 36. This result is consistent with what we would expect based on Equation 3 and indicates that transport cost represents an important part of transaction costs in Malawian markets (Fafchamps and Gabre-Madhin, 2006).

Given the evidence that Hicksian separability fails, this implies that unit values will constitute biased proxies for prices in Malawi, as unit values will be capturing both differences in item quality and prices. We now turn to the comparative assessment of household consumption and poverty estimation based on unit values versus market prices.

4 Comparison of Unit Value- versus Market Price-Based Welfare Analyses

The prevailing approach to household consumption and poverty estimation based on the Integrated Household Survey in Malawi has relied on unit values as proxies for prices since 1997/98 (Malawi National Statistical Office, 2021). The results from the previous section have shown that relative prices of elementary goods within a group vary, which in turn means unit values will inaccurately represent food group price levels across locations and can bias consumption and poverty estimation. Given the availability of an alternative source of price data at our disposal, we compare unit value-based estimates of household consumption, price deflators and poverty to those that are obtained by an alternative estimation strategy that is based on market prices. We focus on three areas where prices come into welfare estimation, as summarized in Table 6.

Table 6

Framework for the Use of Price Proxy in Welfare Estimation

Options	Description	Estimation of Price deflators	Valuation of Household consumption	Estimation of Poverty line
A	Current Unit Value Approach	Unit value-based price deflators	Unit value-based household consumption valuation	Unit value-based estimation of poverty lines
B	An alternative Market Price based approach	Market price-based deflators	Market Price-based household consumption valuation	Market Price-based estimation of poverty lines

Note. This framework describes the different stages of using prices in the estimation of welfare.

For our analysis, we keep consistent with the methodology for the estimation of poverty and inequality as described by the Malawi National Statistical Office (2021), with the only difference being that we use market prices instead of unit values in our calculations. Further information on the design of the Fifth Integrated Household Survey (HIS5) 2019/2020 can be found in Malawi National Statistical Office (2020). We discuss the comparison of the estimates of price deflators, consumption valuation, and derivation of poverty lines across our unit value and market price approaches below.

4.1 Estimation of Price Deflators

The prevailing choice of welfare measure in Malawi is the total household consumption that comes from aggregating value of food consumption and non-food expenditures. To ensure comparability across households across the country surveyed in different periods and different areas, consumption estimates must be adjusted for differences in price over time and across space. The IHS5 survey was from April 2019 to April 2020 and adjustment must be made for changes in purchasing power over

this time, as well as adjustment for differences in purchasing power over space. In Malawi, a spatial–temporal Paasche price index is employed to account for differences across space and time.

The temporal adjustment is a regional-specific adjustment implemented using a combination of regional monthly food index derived from unit values of food items in IHS5 and official non-food CPI for non-food items. The food index in region r and month m is a Paasche price index calculated as:

$$Food\ index_{r,m} = \sum_{j=1}^{N} w_{j,r} RP_{j,r,m} \quad (5)$$

Where $RP_{j,r,m} = \frac{Price_j^m}{Price_j^{April|May,2019}}$ i.e., the relative price of item j in region r and month m. $w_{j,r}$ is the average budget share for item j in region r. This food index is then combined with official non-food CPI using the food and non-food consumption shares from official data. The regions are (i) Urban, (ii) Rural North, (iii) Rural Center, and (iv) Rural South.

Similar to the temporal price adjustment, a regional spatial price index is also estimated to account for differences in prices across regions, the regional food index is derived from unit values from IHS5, and these are combined with the official regional non-food CPI to generate the regional spatial price index. The spatial food index in region r is a Paasche price index calculated as:

$$Food\ index_r = \sum_{j=1}^{N} w_{j,r} RP_{j,r} \quad (6)$$

Where $RP_{j,r}$ is the relative price of item j in region r and $w_{j,r}$ is the average budget share for item j in region r.

The Spatial–temporal Paasche price indices are then created by combining the temporal indices with the regional spatial indices to help capture the variation in price over both space and time. Further details on the methodology for generating the spatial and temporal price index are described by the National Statistics Office and World Bank Poverty and Equity Global practice (2018) and the Malawi National Statistical Office (2021).

We follow the same approach, but instead of deriving the food index from unit values of food items like typically done in Malawi, we derive these indices using market prices observed from the Market Survey. For items that had multiple elementary products[7] captured in the Market Survey, there will not be a 1:1 concordance between market prices and the IHS5 consumption module used to derive the budget shares $w_{j,r}$. For these items, we take a geometric mean of the regional relative prices[8] of the elementary items in these food groups and then combine this with the budget share of the food group when forming the market price-based price deflators. The averaging of price relativities rather than prices treats each multiple elementary market survey item equally when forming the group-level

[7] For all the consumption items with multiple elementary products in the market survey, 16 items had more disaggregation in the market survey than the consumption module of IHS5. Maize flour (Ufa) and Maize (Not as Ufa) had the same level of disaggregation in the consumption module of the IHS5 as the market survey and thus had a 1:1 concordance when combining the market price with the budget share weights. See Appendix 2 for details on the consumption items and their elementary products (disaggregation) in the Market Survey.

[8] This elementary aggregation issue is common in CPI estimation (Graf, 2020) and it is suggested that a Jevons index (geometric mean) should be used in the aggregation of multiple items in the elementary aggregate when there are no weights.

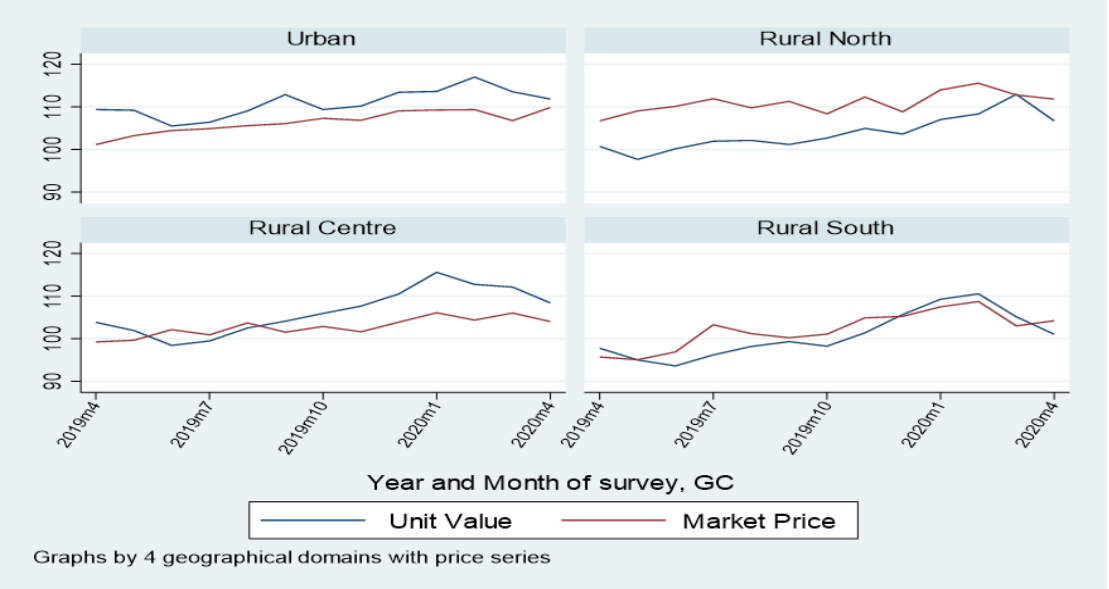
relative price. Figures 2 and 3 present a comparison of the spatial–temporal price deflators from the unit value approach (current method) with our alternative market price-based deflators.

Figure 2

Comparison of Spatial and Temporal Deflators Based on Unit Values and Market Prices

Figure 3

Spatial and Temporal Deflators by Areas: Comparison of Unit Value Proxy and Market Price



Comparing the unit value and market price-derived spatial–temporal price deflators, some interesting patterns emerge. In Urban Malawi, the unit value-derived deflator values are higher than their

comparators based on market prices. In Rural North, the opposite is true. In Rural South and Rural Center, the results are mixed. In Rural Center, the unit value-derived deflator values are higher than their comparators based on market prices during the last three quarters of the fieldwork. In Rural South, the market price-based deflator values are higher than their comparators based on unit values for the first three quarters of the fieldwork, after which there appears to be more of a convergence.

Our results imply a different ranking of regional price differences when comparing unit value or market price-derived price deflators. Indicative of the fact that unit values could be capturing both differences in prices and quality, Urban Malawi shows the highest cost of living when unit values are used, but for market prices it is the Rural North that is the most expensive region. This pattern of price deflators is suggestive of the fact that urban residents are likely to consume higher-quality products, and the inter-area price differences captured by unit values are a mix of both quality and price differences.

Finally, as part of the comparison of the unit value and market price-based price indices, we run a regression of the market price-based index on the unit value-based index, after limiting the computation of the indices to the *common food items at the same level of disaggregation in the market survey and the household survey* (i.e., excluding items with multiple elementary specification in the market survey and items not included in the market survey – primary those that are listed in "other specify categories" as part of the household survey).

If unit value and market prices were perfect proxies, we could expect a coefficient of 1 on the unit value index in a regression of the market price index on the unit value price index. Our results in Table 7 show that for each region, the coefficient on the unit value is significantly different from 1 which indicates differences in derived price indices for seemingly identical items across household and market surveys. These results show that differences in price index derived from the market price and unit values are not driven by the disaggregation of some items in the market survey and generally the process of consumers' estimation of unit value through the valuation of consumption of items leads to different conclusions about price indices than what is obtained from observing market prices.

Table 7

Regression of Market Price-Based Spatial-Temporal Deflators on Unit Value-Based Deflators for Items with One-to-One Mapping in Market Survey and Household Survey

Dependent Variable: Market Price-Based Index	Urban	Rural North	Rural Center	Rural South
Unit Value-Based Index	0.5251***	0.4491***	0.3673***	0.6828***
	[0.1485]	[0.1420]	[0.0711]	[0.1126]
Constant	36.1523***	53.8786***	53.1815***	29.4820***
	[11.2004]	[10.7284]	[6.2153]	[9.1587]
Observations	13	13	13	13
R-squared	0.5321	0.4762	0.7080	0.7697
F-Test (Unit value Coeff. =1)	10.23	15.04	79.15	7.94

Note: Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1. The price indices are derived using items that can be mapped one-to one between the Market Survey and IHS5 consumption module. It excludes that have multiple elementary specifications in the market survey, namely Rice, Bread, Buns, scones, Cassava tubers, Irish Potato, Bean, Pigeon pea (Nandolo), Onion, Chinese cabbage, Eggs, Beef, Pork, Chicken, Banana, Fresh milk, Sugar, Cooking oil, Spices. It also excludes unspecified items in IHS5 survey that are not captured in the Market Survey e.g., "other specify categories".

4.2 Valuation of Household Consumption

Next, we examine the implication of our choice of price proxy for the valuation of consumption. The current approach in Malawi derives unit values by dividing (a) the self-reported monetary valuations of purchased food items that were consumed at home with (b) the associated quantity of consumption that is calculated in kilogram-equivalent terms. These unit values are then used to impute the monetary value for the total quantity of food items that were consumed at home during the last 7 days. This takes the following cascading medians approach where median unit values are computed at the district level and if missing are imputed with median values from the immediate upper level provided there are sufficient observations to obtain a reliable estimate (in this case more than 10 observations). From purchases of items, the unit value used in the valuation is calculated as the median unit value for that item in that district during the same month provided there were sufficient observations to obtain a reliable unit value. If median unit value is missing at this level, the missing median unit value is replaced by the median unit value at the region-month level provided there are sufficient observations in that region in that period in which case the unit value is set to missing. If missing at the region-month level, the missing median unit value is replaced by the median unit value at the region level. If missing at this level, the missing unit value is imputed by the median unit value of the item at the national level. This imputation process is illustrated below:

$$uv_i^h = \begin{cases} uv_i^{d,t} = Median\left[uv_i^h | district, time\right] & \\ uv_i^{ur,t} = Median\left[uv_i^h | region, time\right] & \text{If } uv_j^{d,t} \text{ is missing or less than 10 observations} \\ uv_i^{ur} = Median\left[uv_i^h | region\right] & \text{If } uv_i^{ur,t} \text{ is missing or less than 10 observations} \\ uv_i = Median\left[uv_i^h |, item\right] & \text{If } uv_i^{ur} \text{ is missing or less than 10 observations} \end{cases}$$

Where uv_i^h is the unit value of item i household h, $uv_i^{d,t}$ is the median unit value of item i in district d at month t. $uv_i^{ur,t}$ is the median unit value of item i in region ur at month t. uv_i^{ur} is the median unit value of item i in region ur and uv_i is the median unit value of item i nationally.

For the alternative market price measure, we follow a similar cascading medians approach as the unit value valuation. For each item at the household level, the item median market price at the district level in the same month of visit is used to value household consumption in that district. If the median market price at the district-month level is missing, it is imputed with the item median market price at the region in the same month. If missing at the region month level, it is imputed with the item median market price at the region level and if missing at the region level, it is imputed with the median item price at the national level. This imputation process is as follows:

$$MP_i^h = \begin{cases} MP_i^{d,t} = Median\left[uv_i^h | district, time\right] & \\ MP_i^{ur,t} = Median\left[uv_i^h | region, time\right] & \text{If } MP_j^{d,t} \text{ is missing} \\ MP_i^{ur} = Median\left[uv_i^h | region\right] & \text{If } MP_i^{ur,t} \text{ is missing} \\ MP_i = Median\left[uv_i^h |, item\right] & \text{If } MP_i^{ur} \text{ is missing} \end{cases}$$

Where MP_i^h is the market price of item i household h, $MP_i^{d,t}$ is the median market price of item i in district d at month t. $MP_i^{ur,t}$ is the median market price of item i in region ur at month t. uv_i^{ur} is the median market price of item i in region ur and uv_i is the median market price of item i nationally.

We compare the consumption aggregate obtained using the current approach with the alternative aggregate that is based on prices from the market survey. Like in the formation of price deflators, for the 16 consumption items where there is no 1:1 concordance between the elementary good in the market survey and the food group in IHS5, we use a geometric mean of the elementary prices of the food group. Using geometric means allows us to be consistent with the way elementary prices have combined in the derivation of price indices, and items that are more expensive will have a lesser influence on the geometric mean than on the arithmetic mean.

Table 8 presents the comparison of mean nominal household consumption by COICOP classification category. For nominal consumption, by design, there are no differences between unit value and market price analysis for alcohol and tobacco and non-food items since they were not covered in the Market Survey. For items not covered in the market survey, we use unit values to compute nominal consumption for these items in the market price-based analysis. Only the Food and Beverage and the Hotels and Restaurants COICOP categories differ between both approaches (due to vendor food items, which fall under the Hotels and Restaurants COICOP category being covered in the Market Survey).

Tests of differences in means for the Food and Beverage and Hotels and Restaurants categories show statistically significant difference between the unit value-derived valuation and the market price-derived valuation. Mean household consumption for Food and Beverage in the market price approach is 32 percent higher than the unit value approach. Mean household consumption for the Hotels and Restaurants COICOP category is around 5 percent higher than in the unit value approach, although this also reflects that vendor foods only form a small proportion of this COICOP category.

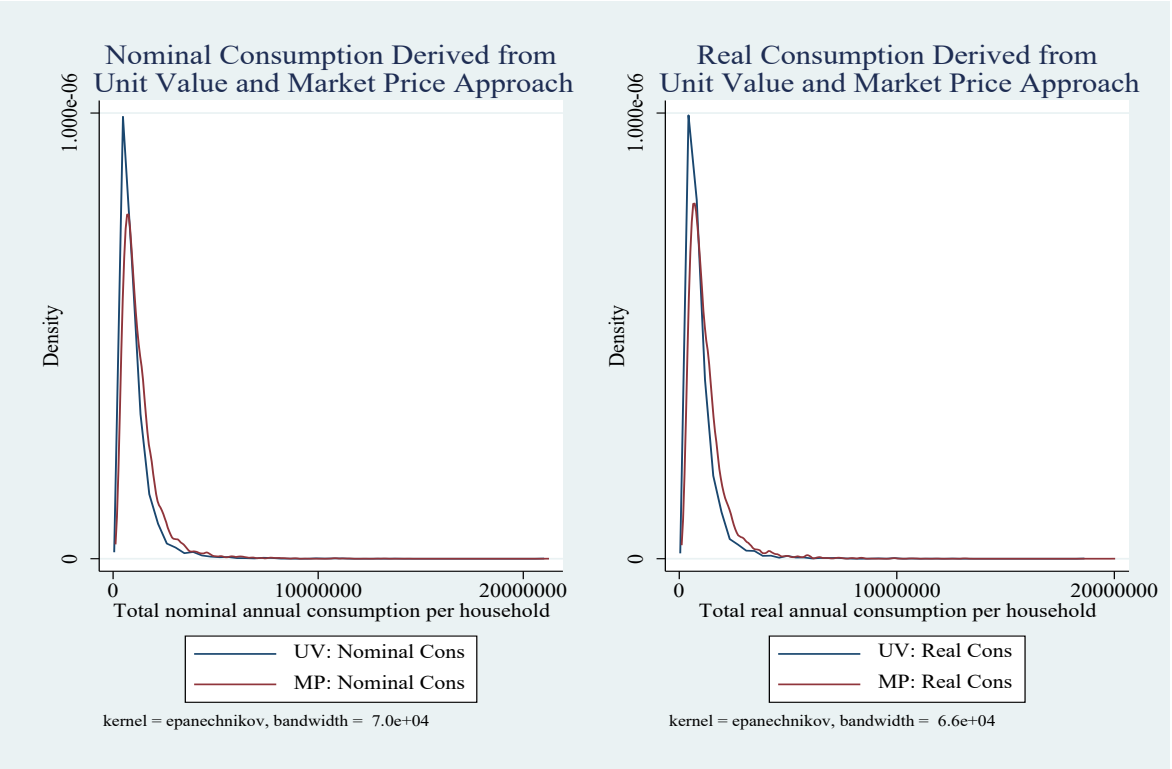
Table 9 presents the comparison of real mean household consumption by COICOP classification category. To calculate real consumption, household consumption is adjusted by the spatial–temporal price deflators. For real household consumption, there are differences in both food and non-food consumption estimates. This is because the non-food sub-aggregates are now being deflated by different price deflators derived from unit values or market price (of food). We find statistically significant differences in real means in all COICOP categories at conventional significance levels.[9] The difference in the non-food categories is indicative of the impact that the differently derived price deflators have on the estimation of welfare in Malawi.

Figure 4 presents the density distribution of nominal and real total household consumption derived using the unit value and market price approaches. The total consumption derived from the market price approach lies to the right and has a longer right tail than that derived from unit value approach. This is consistent with the results in Tables 8 and 9 which shows higher mean household consumption for households from the market price approach. The results from the Kolmogorov-Smirnov (KS) tests of the equality of distributions, though not reported here, also reject the hypothesis (at the 1 percent level) of the equality of distributions based on the unit value versus market price approach.

[9] Tests of differences in means are statistically significant for all COICOP categories at the 99 percent level except for Health (90 percent) and Clothing and Footwear (95 percent).

Figure 4

Density distribution of Nominal and Real Household Consumption from Unit Value and Market Price Approach



Finally, to understand the impact of prices on valuation of consumption, in Tables 10 to 13, we group individuals into deciles using the real household aggregate consumption derived from each price proxy and compare the decile ranking of individuals derived from each method. This shows whether the choice of price proxy has implications for the ranking of individuals in the welfare distribution. Our results show consistent ranking between the two approaches with more than 90 percent of observations from unit value rankings within one decile group ranking in the market price approach.[10]

[10]Results for quintile classifications are reported in Appendix Tables A10 to A13.

Table 8

Comparison of Nominal Consumption Estimates: Unit Value vs Market Price Approaches

	Unit Value			Market Price			Test of difference		
COICOP Category	Mean	Std. Err	Share	Mean	Std. Err	Share	Std. Err	T-Stat	P-value
Food and Beverage	606,135	4,852	55%	803,421	6,247	62%	2,268.8	87.0	0.000
Alcohol and Tobacco	8,548	540	1%	8,548	540	1%	NA	NA	NA
Clothing and Footwear	24,254	497	2%	24,254	497	2%	NA	NA	NA
Housing and Utilities	218,583	2,327	20%	218,583	2,327	17%	NA	NA	NA
Furnishings	42,455	1,185	4%	42,455	1,185	3%	NA	NA	NA
Health	17,211	449	2%	17,211	449	1%	NA	NA	NA
Transport	67,564	2,786	6%	67,564	2,786	5%	NA	NA	NA
Communication	28,957	743	3%	28,957	743	2%	NA	NA	NA
Recreation	6,984	364	1%	6,984	364	1%	NA	NA	NA
Education	41,713	1,859	4%	41,713	1,859	3%	NA	NA	NA
Hotels and Restaurants	10,123	227	1%	10,657	251	1%	71.3	7.5	0.000
Misc. Goods & Services	28,206	308	3%	28,206	308	2%	NA	NA	NA
Total Consumption	1,100,735	11,122		1,298,555	12,151		2,272.8	87.0	0.000

Note. Mean nominal household consumption by COICOP categories based on prices from market price survey and unit values and test of differences. Only Food and Beverage, and Hotels and Restaurants (due to vendor foods) categories differ in both approaches because items in other categories were not priced in the market survey.

Table 9:

Comparison of Real Consumption Estimates: Unit Value vs Market Price Approaches

	Unit values			Market Price			Test of difference		
COICOP Category	Mean	Std. Err	Share	Mean	Std. Err	Share	Std. Err	T-Stat	P-value
Food and Beverage	574,660	4,480	55%	767,693	5,887	62%	2,090.5	92.3	0.000
Alcohol and Tobacco	8,045	495	1%	8,116	506	1%	23.0	3.1	0.002
Clothing and Footwear	23,125	470	2%	23,176	472	2%	25.7	2.0	0.049
Housing and Utilities	206,632	2,104	20%	208,959	2,187	17%	148.5	15.7	0.000
Furnishings	40,068	1,076	4%	40,502	1,120	3%	65.2	6.7	0.000
Health	16,576	434	2%	16,616	437	1%	21.6	1.9	0.063
Transport	63,471	2,549	6%	64,283	2,618	5%	116.6	7.0	0.000
Communication	27,336	698	3%	27,548	706	2%	38.9	5.4	0.000
Recreation	6,448	330	1%	6,593	342	1%	16.9	8.6	0.000
Education	38,857	1,707	4%	39,498	1,755	3%	82.6	7.8	0.000
Hotels and Restaurants	9,617	213	1%	10,206	241	1%	70.1	8.4	0.000
Misc. Goods & Services	26,731	286	3%	26,916	290	2%	21.1	8.8	0.000
Total Consumption	1,041,565	10,142		1,240,107	11,431		2,134.3	93.0	0.000

Note. Mean real household consumption by COICOP category based on prices from market price survey and unit value proxy and test of differences.

Table 10

Number of Individuals in Aggregate per Capita Expenditure Decile Group: Unit Value vs Market Price Approach (Unit Value approach classification is used as the base)

	Unit Value Deciles										
Market Price Deciles	1	2	3	4	5	6	7	8	9	10	Total
1	1,473,951	306,263	22,906	6,378	4,244	-	-	-	-	-	1,813,743
2	285,757	1,030,676	462,556	30,668	3,111	-	1,528	-	-	-	1,814,297
3	51,076	357,569	882,694	473,637	33,046	10,774	4,157	-	-	-	1,812,953
4	4,019	83,379	352,899	845,130	495,554	28,862	3,253	-	-	-	1,813,096
5	-	27,191	71,570	330,501	862,758	483,631	36,910	963	-	-	1,813,524
6	-	7,157	14,816	105,806	354,474	888,996	435,285	9,345	-	-	1,815,879
7	-	-	2,543	15,140	58,673	342,903	1,022,440	366,229	3,260	-	1,811,188
8	-	-	3,215	6,433	5,018	51,968	290,071	1,180,225	276,045	-	1,812,976
9	-	-	-	-	-	4,322	18,918	252,423	1,400,106	137,383	1,813,153
10	-	-	-	-	-	-	-	4,013	133,480	1,675,804	1,813,296
Total	1,814,803	1,812,235	1,813,200	1,813,693	1,816,878	1,811,455	1,812,562	1,813,197	1,812,892	1,813,187	18,134,103

Table 11

Number of Individuals in Aggregate per Capita Expenditure Decile Group: Unit Value vs Market Price Approach (Market Price approach classification is used as the base)

	Market Price Deciles										
Unit Value Deciles	1	2	3	4	5	6	7	8	9	10	Total
1	1,473,951	285,757	51,076	4,019	-	-	-	-	-	-	1,814,803
2	306,263	1,030,676	357,569	83,379	27,191	7,157	-	-	-	-	1,812,235
3	22,906	462,556	882,694	352,899	71,570	14,816	2,543	3,215	-	-	1,813,200
4	6,378	30,668	473,637	845,130	330,501	105,806	15,140	6,433	-	-	1,813,693
5	4,244	3,111	33,046	495,554	862,758	354,474	58,673	5,018	-	-	1,816,878
6	-	-	10,774	28,862	483,631	888,996	342,903	51,968	4,322	-	1,811,455
7	-	1,528	4,157	3,253	36,910	435,285	1,022,440	290,071	18,918	-	1,812,562
8	-	-	-	-	963	9,345	366,229	1,180,225	252,423	4,013	1,813,197
9	-	-	-	-	-	-	3,260	276,045	1,400,106	133,480	1,812,892
10	-	-	-	-	-	-	-	-	137,383	1,675,804	1,813,187
Total	1,813,743	1,814,297	1,812,953	1,813,096	1,813,524	1,815,879	1,811,188	1,812,976	1,813,153	1,813,296	18,134,103

Table 12

Proportion of Individuals in Aggregate per Capita Real Expenditure Decile Group: Unit Value vs Market Price (Unit Value approach classification is used as the base)

	Unit Value Deciles										
Market Price Deciles	1	2	3	4	5	6	7	8	9	10	Total
1	81.2%	16.9%	1.3%	0.4%	0.2%	-	-	-	-	-	1,813,743
2	15.7%	56.9%	25.5%	1.7%	0.2%	-	0.1%	-	-	-	1,814,297
3	2.8%	19.7%	48.7%	26.1%	1.8%	0.6%	0.2%	-	-	-	1,812,953
4	0.2%	4.6%	19.5%	46.6%	27.3%	1.6%	0.2%	-	-	-	1,813,096
5	-	1.5%	3.9%	18.2%	47.5%	26.7%	2.0%	0.1%	-	-	1,813,524
6	-	0.4%	0.8%	5.8%	19.5%	49.1%	24.0%	0.5%	-	-	1,815,879
7	-	-	0.1%	0.8%	3.2%	18.9%	56.4%	20.2%	0.2%	-	1,811,188
8	-	-	0.2%	0.4%	0.3%	2.9%	16.0%	65.1%	15.2%	-	1,812,976
9	-	-	-	-	-	0.2%	1.0%	13.9%	77.2%	7.6%	1,813,153
10	-	-	-	-	-	-	-	0.2%	7.4%	92.4%	1,813,296
Total	1,814,803	1,812,235	1,813,200	1,813,693	1,816,878	1,811,455	1,812,562	1,813,197	1,812,892	1,813,187	18,134,103
	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	
% Ranked Within One Decile	97.0%	93.5%	93.7%	90.9%	94.3%	94.7%	96.4%	99.2%	99.8%	100.0%	

Table 13

Proportion of Individuals in Aggregate per Capita Real Expenditure Decile Group: Unit Value vs Market Price (Market Price approach classification is used as the base)

	Market Price Deciles										
Unit Value Deciles	1	2	3	4	5	6	7	8	9	10	Total
1	81.3%	15.8%	2.8%	0.2%	-	-	-	-	-	-	1,814,803
2	16.9%	56.8%	19.7%	4.6%	1.5%	0.4%	-	-	-	-	1,812,235
3	1.3%	25.5%	48.7%	19.5%	3.9%	0.8%	0.1%	0.2%	-	-	1,813,200
4	0.4%	1.7%	26.1%	46.6%	18.2%	5.8%	0.8%	0.4%	-	-	1,813,693
5	0.2%	0.2%	1.8%	27.3%	47.6%	19.5%	3.2%	0.3%	-	-	1,816,878
6	-	-	0.6%	1.6%	26.7%	49.0%	18.9%	2.9%	0.2%	-	1,811,455
7	-	0.1%	0.2%	0.2%	2.0%	24.0%	56.5%	16.0%	1.0%	-	1,812,562
8	-	-	0.0%	-	0.1%	0.5%	20.2%	65.1%	13.9%	0.2%	1,813,197
9	-	-	-	-	-	-	0.2%	15.2%	77.2%	7.4%	1,812,892
10	-	-	-	-	-	-	-	-	7.6%	92.4%	1,813,187
Total	1,813,743	1,814,297	1,812,953	1,813,096	1,813,524	1,815,879	1,811,188	1,812,976	1,813,153	1,813,296	18,134,103
	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	
% Ranked Within One Decile	98.2%	98.1%	94.5%	93.4%	92.5%	92.4%	95.6%	96.3%	98.7%	99.8%	

4.3 Derivation of Poverty Lines

Our choice of price proxy also plays an important role in the estimation of the poverty line. The poverty line in Malawi is a national absolute poverty line derived using a Cost of Basic Needs (CBN) approach (Haughton and Khandker, 2009). Under this CBN approach, a food poverty line is determined by estimating the cost of meeting the food energy requirements from the foods that are eaten by the poor. The total poverty line is then derived by adding the food poverty line to the average value of non-food spending by households whose food consumption is close to the food poverty line. A household with consumption below the food poverty line is considered food poor.

The choice of price proxy plays an important role in the estimation of the poverty line in two respects. Firstly, it directly informs the costing of the daily food basket of the poor. Secondly, it informs the spatial and temporal price deflators which are used to adjust household consumption prior to ranking of households to determine the reference group of the poor. The evidence laid out in Section 4.1 has shown different patterns of spatial and temporal price deflators emerge depending on whether our price deflators are derived from unit values or market prices. Thus, we estimate the poverty line using the market price proxy and compare it to the unit value approach.

The current unit value-based poverty line in Malawi assumes a daily energy requirement of 2,215 calories and uses a reference population between the 5th and 6th decile rankings of the real aggregate household consumption distribution (National Statistical Office, 2021). We keep both assumptions for our market price analysis but in Appendix Table A14 and Table A15 (Appendix 4) examine the sensitivity of our market price derived results to these choices.

Table 14 presents the daily unit value- and market price-based poverty lines from both approaches. We find that market price-based food and total poverty lines are around 49 percent higher than the unit value-based lines. However, aggregate consumption is also higher with the market price approach (Tables 8 and 9) and market prices lead to a different pattern of spatial and temporal price deflators (Figure 2). As such, the ultimate effects of the market price approach on the poverty estimates are not clear ex-ante. Next section uncovers these effects.

Table 14

Comparison of Daily Poverty Lines from Market Price- and Unit Value-Based Price Proxies

	Daily Food Poverty Line		**Daily Total Poverty Line**		**Non-Food Shares**		**Reference group Deciles**
Calorie Req.	Unit Value	Market Price	Unit Value	Market Price	Unit Value	Market Price	
2215	MK 277.4	MK 413.0	MK 453.4	MK 679.2	0.63	0.64	Deciles 5 and 6

Note. Daily food and total poverty line estimates (in Malawian Kwacha) derived from market price and unit value price proxies.

4.4 Poverty and Inequality Estimates

In the previous sections, we have shown that the choice of price proxy does matter for price deflators, valuation of consumption, and poverty lines. In this section, we examine the implications for measures of poverty and inequality across space and time in Malawi. We compare the estimates of both food and total poverty from our unit value- and market price-based approaches. Our measure of poverty is the poverty headcount ratio, which can be estimated as part of the family of poverty indices proposed by Foster et al. (1984):

$$P_\alpha = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Z - y_i}{Z}\right)^\alpha * I(y_i < Z) \quad (9)$$

where y_i is the consumption of individual i; N is the total population; Z is the poverty line derived from either using the market price or unit value approach; $I(y_i < Z)$ is an indicator function which is equal to 1 when individual consumption is below the poverty line and 0 when the consumption is above the poverty line. α is a non-negative parameter. The poverty headcount index (α = 0) is the percentage of population whose consumption is below the total poverty line. We also report the food poverty line (food poverty) which is the percentage of population whose consumption is below the food poverty line. Our measure of inequality is the well-known Gini index, which measures the extent to which the distribution of consumption expenditure among individuals deviates from a perfectly equal distribution. A Gini index of 0 represents perfect equality, while an index of 1 implies perfect inequality. Additional inequality statistics are presented in Appendix 8.

Table 15 presents the poverty rates by region and across all of Malawi.[11] For both food and total poverty, the current unit value approach shows that poverty is highest in the Rural Center, followed by the Rural South, and is lowest in urban areas. These spatial patterns are also confirmed by the market price analysis, but the market price results show a statistically significant higher level of poverty rates (in both food and total poverty) than the unit value approach. Nationally, the unit value approach estimates a poverty head-count rate of 51 percent, which is less than the market price approach of 64 percent. Total poverty rates derived from the market price approach are about 9 to 15 percentage points higher, with the biggest gap in the Rural Center region where the unit value approach estimates around 63 percent poverty head count rate compared to 76 percent for the market price approach.

Taking a temporal view (Table 16), poverty is lowest in the first quarter and increases in subsequent quarters, peaking in the 4th quarter. These patterns are confirmed in the market price approach. Like in the spatial estimates, we find a statistically significant higher level of poverty in the market price approach. The choice of price proxy, however, does not seem to matter for inequality as we find similar spatial and temporal patterns in the levels of inequality from both approaches.

[11] We provide a sensitivity analysis of food and total poverty estimates to changes in the daily calorie requirements from between 2,215 kcal and 2,815 kcal in Appendix Tables A10 and A11.

Table 15

National and Regional Poverty and Inequality Estimates: Comparison of Unit Value and Market Price Approach

	Food Poverty					Total Poverty					Inequality			
	Estimates		Test for difference			Estimates		Test for difference			Gini-Unit Value		Gini-Market Price	
	Unit Value	Market Price	T-Stat	Std. error	P-value	Unit Value	Market Price	T-Stat	Std. error	P-value	Estimate	Std. Error	Estimate	Std. error
Urban	3.3%	7.7%	7.13	0.006	0.0000	19.2%	28.6%	10.46	0.009	0.0000	0.3901	0.0093	0.3762	0.0090
Rural-North	9.8%	16.0%	5.59	0.011	0.0000	35.8%	51.3%	10.39	0.015	0.0000	0.3196	0.0079	0.2972	0.0073
Rural-Center	29.5%	44.6%	18.91	0.008	0.0000	62.6%	75.9%	18.76	0.007	0.0000	0.3265	0.0052	0.3281	0.0050
Rural-South	22.0%	33.6%	17.10	0.007	0.0000	56.6%	70.2%	21.45	0.006	0.0000	0.3254	0.0049	0.3222	0.0046
National	20.5%	31.6%	26.77	0.004	0.0000	50.6%	63.7%	31.82	0.004	0.0000	0.3789	0.0039	0.3691	0.0037

Note. National and regional food and total poverty headcount rates derived from market price and unit value price proxy and test of differences between them.

Table 16

Quarterly Poverty and Inequality Estimates: Comparison of Unit Value and Market Price Approaches

	Food Poverty					Total Poverty					Inequality			
	Estimates		Test for difference			Estimates		Test for difference			Gini-Unit Value		Gini-Market Price	
	Unit Value	Market Price	T-Stat	Std. error	P-value	Unit Value	Market Price	T-Stat	Std. error	P-value	Estimate	Std. error	Estimate	Std. error
1st Quarter	17.1%	25.9%	10.11	0.009	0.0000	46.2%	58.9%	13.65	0.009	0.0000	0.3692	0.008	0.3542	0.007
2nd Quarter	19.5%	30.3%	13.93	0.008	0.0000	49.9%	63.1%	17.12	0.008	0.0000	0.3824	0.009	0.3702	0.009
3rd Quarter	22.0%	31.4%	11.05	0.009	0.0000	50.6%	63.3%	15.28	0.008	0.0000	0.3836	0.007	0.3710	0.007
4th Quarter	22.1%	36.0%	17.41	0.008	0.0000	53.8%	67.2%	17.59	0.008	0.0000	0.3772	0.006	0.3732	0.006

Note. Quarterly food and total poverty headcount rates derived from market price and unit value price proxy and test of differences between them.

Furthermore, the shares of individuals living in poor households, as estimated though the two approaches, show a high degree of consistency. For food (Table 17), we find that 97 percent of individuals identified as living in food poor households according to the unit value approach are equally identified as living in food poor households under the market price approach, and that 85 percent of individuals identified as living in non-food-poor households according to the unit value approach also identified as living in non-food-poor households under the market price approach. About 15 percent of individuals identified as living in non-food poor households according to the unit value approach are classified as living in food poor households under the market price approach.

Table 17

Comparison of Food-Poor Individuals Based on Unit Value and Market Price Approaches

	Unit Value		
Market Price	Not Food Poor	Food Poor	Total
Not Food Poor	12,297,040 (85%)	108,556 (3%)	12,405,596
Food Poor	2,128,341 (15%)	3,600,166 (97%)	5,728,507
Total	14,425,381	3,708,722	18,134,103

Similar results emerge for total poverty (Table 18). We find that 99 percent of individuals identified as living in poor households according to the unit value approach are also identified as living in poor households under the market price approach. On the other hand, only 73 percent of individuals identified as living in non-poor households according to the unit value approach are also classified as living in non-poor households under the market price approach. About 27 percent of individuals identified as living in non-poor households according to the unit value approach are in fact classified as living in poor households under the market price approach.

Table 18

Cross-tabulation of CBN Poverty Numbers from Unit Value and Market Price Approaches

	Unit value		
Market Price	Non-Poor (CBN)	Poor (CBN)	Total
Non-poor (CBN)	6,516,895 (73%)	63,003 (1%)	6,579,898
Poor (CBN)	2,433,996 (27%)	9,120,209 (99%)	11,554,204
Total	8,950,891	9,183,212	18,134,103

4.4.1 Validation of poverty and consumption estimates with nightlights

We attempt to validate our estimated measures of poverty and consumption from both unit value and market price approaches by comparing them with an exogenous measure of economic activity – nightlights. Nightlights are increasingly gaining prominent in the remote sensing and economic literature as a good proxy for economic performance/activity especially some of the data available from the Visible Infrared Imaging Radiometer Suite (VIIRS) (Gibson, 2021). They have been used to estimate economic activity in data-poor settings including estimation of sub-national GDP level and have found to be good proxy for economic activity (Chen and Nordhaus, 2019; Sutton et al.,2007; Gibson, 2021; Gibson and Boe-Gibson, 2021). We examine the correlation between the competing poverty estimates and night lights at the district level, i.e., the sub-regional, lowest geographic area

for which the IHS5 consumption and poverty estimates are deemed to be representative. To do so, we estimate the following regression:

$$Poverty\ rates_{i,District} = \alpha + \beta_1 Lights_{m,District}$$

We take a log-log and levels approach where:

- $Poverty\ rates_{i,District}$ are the district poverty rates derived from method i (i.e., either using unit values or market prices as price proxies)
- $Lights_{m,District}$ are average VIIRS measured radiance for the month which data was collected from households in the district.

The results reported in Table 19 are in line with the previous findings in the literature that indicate a strong association between night lights and measures of economic performance such as GDP per capita and poverty rates. As indicated by the R-squared estimates, in comparison to their unit value-based counterparts, the market price informed estimates of consumption and poverty consistently show a higher level of association with our exogenous measure of economic activity. Given that night lights will be picking up much more than just household sector activity, our main interest here is to show that both our market price derived estimates show similar level of association with an exogenous measure of economic activity as the unit value approach at sub-national levels. These results are also consistent at lower spatial scales at the territorial area (TA) and enumeration area (EA) levels which are reported in Appendix 6 (Appendix Tables A17 and A18).

Table 19

Regression of Poverty Rates (CBN and Food) at district level on Night Lights: Comparison of Market Price vs Unit Value Estimates

	Log District CBN Poverty Rates (Unit Value)	Log District CBN Poverty Rates (Market Price)	Log District Food Poverty Rates (Unit Value)	Log District Food Poverty Rates (Market Price)	District CBN Poverty Rates (Unit Value)	District CBN Poverty Rates (Market Price)	District Food Poverty Rates (Unit Value)	District Food Poverty Rates (Market Price)
Log District Average Night Lights	-0.44***	-0.37***	-0.79***	-0.64***				
	(0.07)	(0.05)	(0.20)	(0.13)				
District Average Night Lights					-7.82***	-8.98***	-3.79***	-5.58***
					(1.56)	(1.49)	(1.05)	(1.47)
Constant	3.45***	3.80***	1.95***	2.68***	53.99***	68.89***	21.18***	33.52***
	(0.07)	(0.06)	(0.22)	(0.14)	(2.82)	(2.66)	(1.87)	(2.61)
Obs.	32	32	32	32	32	32	32	32
R-squared	0.60	0.64	0.34	0.44	0.46	0.55	0.30	0.33
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1								

5 Sensitivity Checks

Comprehensive market price questionnaires are an exception in household surveys for a majority of low- and middle-income countries as household surveys in these countries prioritize collecting expenditure or income data over price data (Gibson 2013; Gibson and Rozelle 2011). The additional technical and financial requirements of fielding these surveys have often been an important concern preventing the implementation of detailed market price surveys alongside household surveys.

This section examines the sensitivity of our market price derived poverty and inequality estimates to a reduction in the spatial coverage of the market survey (visiting less markets across districts) and reduction in item coverage (collecting prices of less items). We examine whether our market price results are different if we assume scenarios of "more modest" data collection both in terms of market coverage and item coverage. For the reduced market coverage, we consider scenarios where we visit the top 50 to 90 percent of markets in each district based on the number of available price observations from each market. Appendix Table A19 presents the distribution of the number of markets in the reduced market coverage scenarios. For the reduced item coverage, we consider scenarios where market prices are only collected for the most important items i.e., items that make up the top 50 to 90 percent of the household budget share.

5.1 Sensitivity of Poverty and Inequality Estimates to Item Coverage

We first examine the sensitivity of our poverty and inequality estimates to reducing the item coverage in the market survey, i.e., instead of pricing almost all the items from the household survey in the market price survey, we examine how our estimates change if the market survey focused on pricing only a subset of main items, as determined by their budget share, ranging from 50 percent, 60 percent, 70 percent, 80 percent, and 90 percent of household consumption. In this context, a 50 percent item coverage means market prices are used for items that make up the top 50 percent of budget share at the national level with unit values used as the price proxy for the rest. This is meant to simulate a scenario in which the market price collection focused on the collection of prices of key items that are important in household consumption instead of all items. The derived poverty lines from each item coverage scenario are reported in Appendix Table A16.

Table 22 presents the national and regional food and total poverty rates derived from the market price item coverage assumptions. As with the market coverage scenarios, we find very limited impact of reduced item coverage on inequality as measured by the Gini coefficient. However, for poverty rates, while the results show similar magnitudes under different scenarios, tests of statistical significance show significant differences between the poverty rates derived from the complete coverage and the reduced item coverage scenarios at the 95 percent level – except for food poverty rates in Urban areas, and Rural Center at 70-90 percent item coverage. For total poverty, we find also statistically significant difference between the complete survey poverty rates and the reduced item coverage scenarios at the 95 percent level – except for Rural North at 50 and 70 percent item coverage. Taken together, the sensitivity analyses pertaining to market survey indicate that there are no viable shortcuts in our existing market survey database that can provide us with poverty rates that are statistically indistinguishable from the estimates based on the complete dataset.

Table 20

Sensitivity of Poverty Rates to Item Coverage of Market Survey

	Food Poverty					CBN Poverty					Inequality				
			Test of difference between food poverty at complete item coverage and other coverage levels					Test of difference between CBN poverty at complete item coverage and other coverage levels							Item Coverage Level
	Full Item Coverage	Reduced item coverage	T-stat	Std. Err	P-value	Full Item Coverage	Reduced item coverage	T-stat	Std. Err	P-value	Full Item Coverage Gini	Std. Err	Reduced Item Coverage Gini	Std. Err	
National	31.6	32.3	-5.07	0.00	0.000	63.7	61.0	12.52	0.00	0.000	0.369	0.004	0.366	0.004	90
National	31.6	32.8	-7.50	0.00	0.000	63.7	58.6	17.80	0.00	0.000	0.369	0.004	0.367	0.004	80
National	31.6	33.1	-7.54	0.00	0.000	63.7	62.3	7.06	0.00	0.000	0.369	0.004	0.368	0.004	70
National	31.6	33.1	-6.41	0.00	0.000	63.7	59.2	17.02	0.00	0.000	0.369	0.004	0.365	0.004	60
National	31.6	33.0	-5.63	0.00	0.000	63.7	60.9	12.59	0.00	0.000	0.369	0.004	0.366	0.004	50
Rural-Center	44.6	45.0	-1.86	0.00	0.062	75.9	72.7	8.24	0.00	0.000	0.328	0.005	0.326	0.005	90
Rural-Center	44.6	45.0	-1.73	0.00	0.084	75.9	70.4	11.76	0.00	0.000	0.328	0.005	0.326	0.005	80
Rural-Center	44.6	44.5	0.22	0.00	0.827	75.9	73.5	6.73	0.00	0.000	0.328	0.005	0.325	0.005	70
Rural-Center	44.6	43.6	2.51	0.00	0.012	75.9	70.0	12.40	0.00	0.000	0.328	0.005	0.321	0.005	60
Rural-Center	44.6	42.9	4.21	0.00	0.000	75.9	71.2	11.13	0.00	0.000	0.328	0.005	0.321	0.005	50
Rural-North	16.0	16.7	-1.98	0.00	0.048	51.3	48.0	3.62	0.01	0.000	0.297	0.007	0.293	0.007	90
Rural-North	16.0	17.8	-3.48	0.01	0.001	51.3	44.1	5.53	0.01	0.000	0.297	0.007	0.293	0.007	80
Rural-North	16.0	19.4	-5.37	0.01	0.000	51.3	52.1	-1.29	0.01	0.196	0.297	0.007	0.296	0.008	70
Rural-North	16.0	20.0	-5.11	0.01	0.000	51.3	47.6	3.92	0.01	0.000	0.297	0.007	0.295	0.007	60
Rural-North	16.0	19.4	-5.03	0.01	0.000	51.3	50.6	1.14	0.01	0.255	0.297	0.007	0.295	0.007	50
Rural-South	33.6	34.7	-4.47	0.00	0.000	70.2	67.7	7.95	0.00	0.000	0.322	0.005	0.319	0.005	90
Rural-South	33.6	35.8	-6.82	0.00	0.000	70.2	65.6	11.36	0.00	0.000	0.322	0.005	0.319	0.005	80
Rural-South	33.6	36.5	-7.37	0.00	0.000	70.2	69.1	3.46	0.00	0.001	0.322	0.005	0.321	0.005	70
Rural-South	33.6	37.1	-8.31	0.00	0.000	70.2	66.6	9.20	0.00	0.000	0.322	0.005	0.321	0.005	60
Rural-South	33.6	37.7	-8.92	0.00	0.000	70.2	68.4	5.13	0.00	0.000	0.322	0.005	0.323	0.005	50
Urban	7.7	8.0	-1.18	0.00	0.237	28.6	26.9	4.47	0.00	0.000	0.376	0.009	0.375	0.009	90
Urban	7.7	8.2	-1.92	0.00	0.056	28.6	25.1	6.20	0.01	0.000	0.376	0.009	0.376	0.009	80
Urban	7.7	8.2	-1.76	0.00	0.079	28.6	27.4	3.62	0.00	0.000	0.376	0.009	0.377	0.009	70
Urban	7.7	8.4	-1.65	0.00	0.100	28.6	24.6	6.81	0.01	0.000	0.376	0.009	0.375	0.009	60
Urban	7.7	8.2	-1.04	0.00	0.296	28.6	26.1	5.54	0.00	0.000	0.376	0.009	0.376	0.009	50

Note: Sensitivity of item coverage reported above with complete market coverage

5.2 Sensitivity of Poverty and Inequality Estimates to Market Coverage

Using market prices of items from each of our sensitivity analysis scenarios, we compute poverty rates and compare these rates with the complete market price survey results. Table 20 presents the daily food and total poverty lines from each of the scenarios. Food price lines under the sensitivity scenarios were between 1 and 14 percent higher compared to the complete market price line. Total poverty lines were between 12 percent less and 6 percent higher than the complete market price survey total poverty line. The estimated poverty rates based on the market prices from the reduced market survey locations are compared with the complete market survey estimates in Table 21.

Table 21

Sensitivity of Poverty Lines to Spatial Coverage of Market Survey

Price Proxy Approach	Calorie Req.	Reference Group Deciles	Daily Food Poverty Line	Daily Total Poverty Line	Non-food Share
Complete Price Survey	2215	Deciles 5 and 6	MK413	MK679	0.64
Reduced Price Survey-90% coverage	2215	Deciles 5 and 6	MK414	MK569	0.38
Reduced Price Survey-80% coverage	2215	Deciles 5 and 6	MK416	MK627	0.51
Reduced Price Survey-70% coverage	2215	Deciles 5 and 6	MK417	MK599	0.44
Reduced Price Survey-60% coverage	2215	Deciles 5 and 6	MK417	MK621	0.49
Reduced Price Survey-50% coverage	2215	Deciles 5 and 6	MK417	MK599	0.44

Note. Daily food and total poverty line estimates (in Malawian Kwacha) derived from unit value price proxy and market price surveys (complete and reduced samples).

The results for food poverty rates derived under each reduced survey scenarios are slightly higher in magnitude than the estimates from the complete market coverage. Depending on which region, the figures indicate around 1 to 2 percent more people in food poverty in the reduced coverage scenarios. The food poverty results are closest in magnitude in the Rural Center region where the reduced scenarios food poverty estimates are between 0.1 and 0.6 percent of the complete coverage results and furthest in the Rural South where it is between 1.0 and 1.7 percent of the complete price survey results. However, although close in magnitude, tests of statistical significance between the reduced market coverage scenarios and the complete price survey reveal that these estimates are statistically significantly different from the complete market coverage results. Except for Rural Center, we reject the hypothesis that the food poverty rates are not statistically significantly different from the food poverty rates derived under the complete market price survey approach. For total poverty rates, the results under the reduced survey scenarios are lower in magnitude than the complete market survey derived results. Tests of statistical significance also show that the derived total poverty estimates are statistically significantly different from the complete scenario results in all coverage scenarios and at both national and regional levels. For inequality, we find very limited impact of reduced market coverage on inequality as measured by the Gini coefficient.

Taken together, the findings indicate that for generating statistically consistent poverty estimates that are informed by market prices, it is essential to have a market survey that has a sizeable spatial and temporal overlap with household survey data collection, inclusive of price information on the same consumption items that are also featured in the household survey questionnaire. Indeed, given the evidence on the failure of Hicksian separability, the main benefit of a market survey is to provide proxies of local prices that people face for a well-defined set of items across a wide geographic area.

Table 22

Sensitivity of Poverty Rates to Spatial Coverage of Market Survey

	Food Poverty					CBN Poverty						Inequality			
			Test of Difference					Test of Difference							
	Full Market Survey	Reduced Market Survey	T-stat	SE	P-value	Full Market Survey	Reduced Market Survey	T-stat	SE	P-value	Full Market Survey Gini	Std.Err	Reduced Market Survey Gini	Std.Err	Market coverage
National	31.6%	32.2%	4.35	0.001	0.000	63.7%	52.8%	-28.7	0.004	0.000	0.369	0.004	0.369	0.004	90%
National	31.6%	32.3%	4.91	0.001	0.000	63.7%	59.1%	-18.1	0.003	0.000	0.369	0.004	0.368	0.004	80%
National	31.6%	32.3%	4.66	0.001	0.000	63.7%	55.7%	-24.1	0.003	0.000	0.369	0.004	0.368	0.004	70%
National	31.6%	32.2%	3.71	0.002	0.000	63.7%	58.0%	-19.9	0.003	0.000	0.369	0.004	0.367	0.004	60%
National	31.6%	32.7%	5.77	0.002	0.000	63.7%	56.0%	-23.5	0.003	0.000	0.369	0.004	0.367	0.004	50%
Rural-Center	44.6%	44.8%	1.13	0.002	0.259	75.9%	65.4%	-16.8	0.006	0.000	0.328	0.005	0.327	0.005	90%
Rural-Center	44.6%	45.2%	2.53	0.002	0.012	75.9%	71.5%	-10.7	0.004	0.000	0.328	0.005	0.327	0.005	80%
Rural-Center	44.6%	44.8%	1.18	0.002	0.238	75.9%	67.8%	-14.6	0.006	0.000	0.328	0.005	0.327	0.005	70%
Rural-Center	44.6%	44.6%	0.26	0.002	0.798	75.9%	70.1%	-11.9	0.005	0.000	0.328	0.005	0.326	0.005	60%
Rural-Center	44.6%	44.8%	0.90	0.003	0.366	75.9%	68.0%	-14.3	0.006	0.000	0.328	0.005	0.326	0.005	50%
Rural-North	16.0%	16.8%	2.57	0.003	0.010	51.3%	36.4%	-9.6	0.015	0.000	0.297	0.007	0.297	0.007	90%
Rural-North	16.0%	16.8%	3.10	0.003	0.002	51.3%	44.4%	-6.1	0.011	0.000	0.297	0.007	0.297	0.007	80%
Rural-North	16.0%	16.8%	3.00	0.003	0.003	51.3%	40.5%	-7.7	0.014	0.000	0.297	0.007	0.297	0.007	70%
Rural-North	16.0%	16.8%	2.67	0.003	0.008	51.3%	42.8%	-6.6	0.013	0.000	0.297	0.007	0.296	0.007	60%
Rural-North	16.0%	17.9%	3.01	0.006	0.003	51.3%	40.7%	-7.6	0.014	0.000	0.297	0.007	0.297	0.007	50%
Rural-South	33.6%	34.7%	3.58	0.003	0.000	70.2%	58.6%	-20.1	0.006	0.000	0.322	0.005	0.322	0.005	90%
Rural-South	33.6%	34.6%	3.25	0.003	0.001	70.2%	65.5%	-12.4	0.004	0.000	0.322	0.005	0.321	0.005	80%
Rural-South	33.6%	34.8%	3.66	0.003	0.000	70.2%	62.2%	-16.4	0.005	0.000	0.322	0.005	0.321	0.005	70%
Rural-South	33.6%	34.8%	3.41	0.003	0.001	70.2%	64.6%	-13.7	0.004	0.000	0.322	0.005	0.321	0.005	60%
Rural-South	33.6%	35.4%	5.01	0.003	0.000	70.2%	62.6%	-16.2	0.005	0.000	0.322	0.005	0.321	0.005	50%
Urban	7.7%	8.0%	1.72	0.002	0.085	28.6%	21.3%	-8.8	0.008	0.000	0.376	0.009	0.376	0.009	90%
Urban	7.7%	8.0%	2.05	0.002	0.040	28.6%	25.6%	-6.3	0.005	0.000	0.376	0.009	0.376	0.009	80%
Urban	7.7%	8.1%	2.42	0.002	0.016	28.6%	23.0%	-8.0	0.007	0.000	0.376	0.009	0.376	0.009	70%
Urban	7.7%	8.0%	1.69	0.002	0.092	28.6%	24.9%	-6.8	0.005	0.000	0.376	0.009	0.376	0.009	60%
Urban	7.7%	8.5%	2.25	0.003	0.024	28.6%	23.5%	-7.4	0.007	0.000	0.376	0.009	0.376	0.009	50%

6. Conclusion

Household Consumption and Expenditure Surveys (HCESs) are key to consumption-based monetary poverty measurement in low- and middle-income countries. In the absence of market price surveys that are linked to HCESs, unit values (commonly limited to food items) are used as proxies for market prices while estimating nominal consumption aggregates, price deflators, poverty lines and poverty statistics. This practice relies on the Hicksian separability assumption: within-commodity group relative prices are constant across space and the price of a single good is an accurate proxy for the commodity group price. Our analysis tests, for the first time in a low-income context, whether Hicksian separability holds. To do so, we use novel price data that were collected for an extensive list of food items, including several variety/quality-differentiated products for specific items, as part of a national market survey that was conducted in Malawi by the National Statistical Office (NSO) in sync with the Fifth Integrated Household Survey (IHS5) 2019/20 – i.e., the HCES that is the source of official poverty statistics. The key features of the market survey were that (a) the selection of the markets and the timing of market visits were informed by the timing of the visits to the IHS5 enumeration areas for consumption data collection and that (b) the IHS5 and the market survey were managed in an integrated fashion by the NSO, albeit relying on separate but coordinated field teams.

We show that Hicksian separability fails to hold across space and time and that unit values constitute biased proxies for prices. The analysis also suggests the Alchian-Allen effect as a mechanism through which Hicksian separability fails. The weak spatial market integration in settings like Malawi (Chitete et al., 2021) can imply that the relative price of quality may not be constant over space and time as required by Hicksian separability. In view of the evidence regarding the untenability of the Hicksian separability assumption, we conduct a comparative assessment of consumption and poverty estimation based on market prices versus unit values – the business as usual. Relative to unit values, market price data showcase a different pattern of spatial price variation and lead to higher food and overall consumption expenditures – both in nominal and real terms, while also generating higher poverty lines and culminating in higher food and overall poverty rates. Our results show that 98 percent of individuals identified as living in poor households according to the unit value approach are also classified as living in poor households under the market price approach, although an additional 18 percent of people classified as living in non-poor households according to the unit value approach are now classified as living in poor households under the market price approach. Compared to their counterparts based on unit values, district-level poverty estimates based on market prices also exhibit a greater degree of correlation with night lights – a geospatial proxy for living standards. Finally, the analysis reveals that the poverty estimates based on market prices are sensitive to any degree of simulated reduction in (a) the spatial coverage of markets and (b) the list of items for which prices are collected.

Overall, the results are indicative of the benefits of market price data collection that is spatially and temporally in sync with household consumption data collection and that covers the full scale of consumption items in the household survey, with a focus on capturing quality- and variety-differentiated versions of key food items. The analysis shows that the resulting data can inform not only the computation of deflators to account for within-survey price differences across space and time but also the creation of nominal consumption aggregates and poverty lines.

In view of (a) the critical role that unit values play in the computation of consumption-based welfare estimates in low- and middle-income countries and (b) the lack of empirical support, at least in Malawi, for the Hicksian separability assumption, the findings point to the need for more concerted efforts to integrate market surveys into future HCESs. Creating best practice guidelines that can help scale up such a vision would promote cross-country harmonization and comparability, and consequently the quality of resulting price data. This will permit similar analyses in alternative contexts and can inform methodological decisions underlying official statistics pertaining to poverty and inequality.

References

Alchian, A. A., & Allen, W. R. (1967). *University Economics*. Wadsworth Publishing Co.

Beegle, K., De Weerdt, J., Friedman, J., & Gibson, J. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics*, *98*(1), 3–18.

Borcherding, T. E., & Silberberg, E. (1978). Shipping the good apples out: the Alchian and Allen theorem reconsidered. *Journal of Political Economy*, *86*(1), 131–138.

Chen, X., & Nordhaus, W. D. (2019). VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP. *Remote Sensing*, *11*(9), 1057.

Chitete, M., Mgomezulu, W., Bwanaisa, M., & Dzanja, J. (2021). Analysis of intra-region market integration and spatial price transmission in groundnut markets in Malawi. *Agrekon*, *60*(3), 280–296.

Deaton, A. (1988). Quality, quantity, and spatial variation of price. *The American Economic Review*, *78*(3), 418-430.

Deaton, A. (1990). Price elasticities from survey data: extensions and Indonesian results. *Journal of Econometrics*, *44*(3), 281-309.

Fafchamps, M., & Gabre-Madhin, E. Z. (2006). Agricultural markets in Benin and Malawi. *African Journal of Agricultural and Resource Economics*, *1*(1), 67–94.

FEWS NET (2018). *Malawi Enhanced Market Analysis*. Washington, DC: FEWS NET.

Foster, J., Greer, J., & Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica, 52*(3), 761–766.

Gibson, J. (2007). *A guide to using prices in poverty analysis*. World Bank: Washington, DC.

Gibson, J. (2013). The crisis in food price data. *Global Food Security*, *2*(2), 97–103.

Gibson, J., & Kim, B. (2013). Quality, quantity, and nutritional impacts of rice price changes in Vietnam. *World Development*, *43*, 329–340.

Gibson, J., & Kim, B. (2015). Hicksian separability does not hold over space: Implications for the design of household surveys and price questionnaires. *Journal of Development Economics*, *114*, 34–40.

Gibson, J., & Kim, B. (2019). Quality, quantity, and spatial variation of price: Back to the bog. *Journal of Development Economics*, *137*, 66–77.

Gibson, J., & Rozelle, S. (2005). Prices and unit values in poverty measurement and tax reform analysis. *The World Bank Economic Review*, *19*(1), 69-97.

Gibson, J. (2021). Better night lights data, for longer. *Oxford Bulletin of Economics and Statistics*, *83*(3), 770-791.

Gibson, J., & Boe-Gibson, G. (2021). Nighttime lights and county-level economic activity in the United States: 2001 to 2019. *Remote Sensing*, *13*(14), 2741.

Gould, J. P., & Segall, J. (1969). The substitution effects of transportation costs. *Journal of Political Economy*, *77*(1), 130–137.

Graf, B. (2020). *Consumer Price Index Manual, 2020: Concepts and Methods*. International Monetary Fund: Washington, DC.

Haughton, J., & Khandker, S. (2009). Handbook on poverty and inequality. World Bank Publications

Mancini, G., & Vecchi, G. (2022). *On the construction of a consumption aggregate for inequality and poverty analysis*. World Bank Group: Washington, DC.

Minten, B., & Kyle, S. (1999). The effect of distance and road quality on food collection, marketing margins, and traders' wages: evidence from the former Zaire. *Journal of Development Economics*, *60*(2), 467–495.

National Statistical Office and The World Bank, Poverty and Equity Global Practice. (2018). *Methodology for Poverty Measurement in Malawi (2016/17)*. National Statistical Office: Zomba

National Statistical Office. (2020). *Fifth Integrated Household Survey (IHS5) 2019-2020 Basic Information Document.* National Statistical Office: Zomba

National Statistical Office. (2021). Malawi Poverty Report 2020. National Statistical Office: Zomba

Prais, S. J., & Houthakker, H. S. (1971). *The analysis of family budgets*. Cambridge University Press.

Ravallion, M. (1992). *Poverty comparisons: A guide to concepts and methods* (LSMS Working Paper No. 88). Living Standards Measurement Study, World Bank, Washington DC.

Sutton, P. C., Elvidge, C. D., & Ghosh, T. (2007). Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics & Statistics*, *8*(S07), 5-21.

Appendix 1: Details on Market Survey Data Collection and Processing

The Market Survey aimed to collect measurement unit-specific prices and weights for a total of 163 food consumption and agricultural produce items. For every item-unit combination, up to three different vendors were approached in each market to record the kilogram weight and price. Table A1 provides the distribution of samples collected by month.

Table A1

Distribution of Number of Samples Collected by Month

Month	Samples
April 2019	21,894
May 2019	47,526
June 2019	41,118
July 2019	39,516
August 2019	54,468
September 2019	44,856
October 2019	43,254
November 2019	51,264
December 2019	27,768
January 2020	58,206
February 2020	39,516
March 2020	48,060
April 2020	18,156
Total	535,602

Planned versus Actual Months of Market Visits

The timing of the visits to the markets was planned to be in parallel with IHS5; however, due to COVID-19-related disruptions and administrative/logistical delays in different points of the fieldwork, there is a slight difference between the planned versus actual schedule of market visits. Table A2 compares the planned versus actual months of market visits, overall and by region.

Table A2

Distribution of Market Visits: Planned versus Actual Month of Data Collection

	Planned Data Collection				Actual Data Collection				
	North	Central	South	Total	North	Central	South	Total	Diff. Between Actual & Planned
April 2019	15	29	42	86	9	14	19	42	-44
May 2019	15	34	42	91	19	30	47	96	5
June 2019	9	33	42	84	10	33	37	80	-4
July 2019	14	34	42	90	8	26	45	79	-11
August 2019	15	34	41	90	17	48	52	117	27
September 2019	15	34	41	90	13	36	37	86	-4
October 2019	15	33	42	90	14	30	41	85	-5
November 2019	15	32	43	90	20	38	46	104	14
December 2019	15	33	42	90	5	19	29	53	-37
January 2020	15	32	42	89	23	41	55	119	30
February 2020	16	33	42	91	13	24	40	77	-14
March 2020	15	33	42	90	16	36	46	98	8
April 2020					7	19	9	35	35
Total	174	394	503	1,071	174	394	503	1,071	0

Data Cleaning

We implement several data cleaning techniques to make the market data suitable for analysis:

- We winsorized price and weight data at the top and bottom of the distribution i.e., if prices and weights were higher (lower) than the 95^{th} (5^{th}) percentile of their distribution, they are replaced by the 95^{th} (5^{th}) percentile. Each item-unit combination price was winsorized with cuts at the 5th and 95th percentile distribution of prices for that month while each item-unit combination weight was winsorized with cuts at the 5th and 95th percentile distribution of the region. This affected 13 percent of raw price and weight observations collected.
- In some cases, a single market may have been visited several times in a particular month; in cases like this, we combine these visits and define the item-unit prices for this market in this month as the median prices across the multiple visits for that month.
- We use the survey-gathered weights for item-unit combinations to convert prices of non-standard unit measures to KG-equivalent terms.

Handling Missing Data

As expected in a high-frequency survey of hundreds of food item-unit combinations sought after in a national sample of markets, gathering price data at the item-unit level can lead to a high proportion of missing entries, due to (a) items not being sold in specific units of measurement in specific markets, and (b) items not being available in a given month due to seasonality. Table A3 presents the proportion of missing price (and weight) data at the item-unit-month level across all markets and months.

Table A3

Frequency and Proportion of Missing Data at the Item-Unit-Month Level

Status at Item-Unit-Month Level	Frequency	Percent
Non-Missing	133,566	25
Missing	402,036	75
Total	535,602	100

Table A3 is at the item-unit level. There will be a high proportion of missing data at this level as we do not expect to find items being sold in all unit of measurement across all markets in all periods. For example, it is highly likely that some unit of measurements will be more prevalent in certain parts of the country than others. If we shift focus to the item level i.e., those missing all prices in all item-unit combinations in the market in a particular month, we can examine the proportion of missing data at the item-month level. At the item-month level, missing data represents the non-availability of this item in any unit in the market in that month. Table A4 presents the proportion of missing data at the item-month level.

Table A4

Total Frequency and Proportion of Missing Data at the Item-Month Level

Status at Item-Month Level	Frequency	Percent
Non-Missing	49,228	36
Missing	89,186	64
Total	138,414	100

Around 64 percent of data from the market survey prices are missing at the item-month level. Some of the missing data here will reflect seasonality and the non-availability of these items in particular times when they were visited. Tables A5 and A6 present the temporal and spatial distribution of the item-month-level availability. Table A5 shows that the highest proportion of missing data comes from the southern region while the northern region has the lowest proportion of missing data.

Table A5

Frequency and Proportion of Missing Data at the Item-Region-Month Level

Status at Item-Region-Month Level	North	Central	South	Total
Non-missing	9,280	17,799	22,149	49,228
Missing	12,248	32,709	44,229	89,186
Proportion Missing	57%	65%	67%	64%
Total	21,528	50,508	66,378	138,414

Note. Frequency and proportion of missing data at the item-region-month level over all data collection periods

Table A6 shows that the lowest proportions of missing data were at the initial months of the survey from April-September. This period coincides with the harvest season in Malawi, and it is not surprising

to find fewer missing data over this period. The final months of the actual data collection (from February 2020 to April 2020) have the highest proportion of missing data, with April 2020 (the last month of actual data collection) having the highest proportion of missing data at 71 percent.

Table A6

Monthly Frequency and Proportion of Missing Data at the Item-Month Level

Status at Item-Month Level	Non-Missing	Missing	Total	Proportion Missing
April 2019	2,492	3,166	5,658	56%
May 2019	4,667	7,615	12,282	62%
June 2019	3,811	6,815	10,626	64%
July 2019	3,808	6,404	10,212	63%
August 2019	5,363	8,713	14,076	62%
September 2019	4,122	7,470	11,592	64%
October 2019	3,883	7,295	11,178	65%
November 2019	4,551	8,697	13,248	66%
December 2019	2,339	4,837	7,176	67%
January 2020	5,293	9,749	15,042	65%
February 2020	3,397	6,815	10,212	67%
March 2020	4,154	8,266	12,420	67%
April 2020	1,348	3,344	4,692	71%
Total	49,228	89,186	138,414	64%

Imputation of Missing Data

Given the proportion of missing data, we take a cascading median approach to imputing missing prices. We perform the imputation at the item-unit-month level using the following approach: If the price of an item-unit (i) is missing in market (m) at month ($mnth$), we replace it with the median price of the item-unit combination in the associated traditional authority (*TA*), if the item-unit combination is missing at the TA-level, we replace it with the median of the item-unit combination in the district. This approach is illustrated below:

$$Price_i^m = \begin{bmatrix} Price_i^m|Mnth & \text{if } Price_i^m|Mnth \text{ is not missing} \\ Median[Price_i^m|\, TA, Mnth] & \text{if } Price_i^m|\, Mnth \text{ is missing} \\ Median[Price_i^m|District, Mnth] & \text{if } Median[Price_i^m|TA, Mnth] \text{ is missing} \\ Missing & if\ Median[Price_i^m|District, Mnth]\ is\ missing \end{bmatrix} \quad (10)$$

Tables A7 and A8 present the spatial and temporal distribution of observations at the market level after the imputation process.

Table A7

Frequency and Proportion of Missing Data at Item-Region-Month Level after Imputation

Status at Item-Region-Month Level After Imputation	North	Central	South	Total	Prop. of Total
Actual Survey (Raw)	9,280	17,799	22,149	49,228	36%
Imputed	3,153	9,587	9,296	22,036	16%
Missing	9,095	23,122	34,933	67,150	49%
Total	21,528	50,508	66,378	138,414	
Prop. missing post-imputation	42%	46%	53%	49%	

The imputation provides 16 percent more observations, thus reducing the proportion of missing observations from 64 percent of total observations to 49 percent at the item level.

Table A8

Frequency and Proportion of Missing Data at Item-Month Level after Imputation

Status at item-month level after imputation	Actual (Raw)	Imputed	Missing	Total	Prop. missing post-imputation
April 2019	2,492	448	2,718	5,658	48%
May 2019	4,667	2,334	5,281	12,282	43%
June 2019	3,811	1,676	5,139	10,626	48%
July 2019	3,808	1,614	4,790	10,212	47%
August 2019	5,363	2,510	6,203	14,076	44%
September 2019	4,122	1,831	5,639	11,592	49%
October 2019	3,883	1,531	5,764	11,178	52%
November 2019	4,551	2,171	6,526	13,248	49%
December 2019	2,339	931	3,906	7,176	54%
January 2020	5,293	2,953	6,796	15,042	45%
February 2020	3,397	1,403	5,412	10,212	53%
March 2020	4,154	2,013	6,253	12,420	50%
April 2020	1,348	621	2,723	4,692	58%
Total	49,228	22,036	67,150	138,414	49%

This imputed data set is the base price dataset for all our analysis. We calculate district prices for all items in each month from this base dataset. This district prices are the median prices across all markets in a particular district in each month.

For the poverty re-estimation presented in Section 4, we harmonize the IHS5 data with the market price data. The prices of items gathered from the market survey is allocated to each household at the district-month level, i.e., if household A in District 101 is visited in April, then the household is allocated the market prices for District 101 in April from the market survey. Our poverty re-estimation thus compares the sensitivity of poverty and inequality estimates based on unit value (as a proxy for prices) with an alternative based on measures of price coming from the market survey.

Please note that in cases where there is no market price for a particular item in a district for a particular month, we adopt a second-round imputation process such that: If the median price of an item (i) is missing in district (d) at month ($mnth$), we replace it with the median price of the item at the region level for that month, and if missing at the region level for that month, we replace with the median price for all that region level in all periods. If missing at the region level in all periods, we replace with the national median price for that item. The second-round imputation process is summarised in Equation 11 below. The price of item (i) for household (h) in IHS5 living in district (d) is:

$$Price_i^{h,d} = \begin{bmatrix} Median[Price_i^d|Mnth] & \text{if } Price_i^d|Mnth \text{ is not missing} \\ Median[Price_i|\, region, Mnth] & \text{if } Price_i^d|\, Mnth \text{ is missing} \\ Median[Price_i|\, region] & \text{if } Median[Price_i|\, region, Mnth] \text{ is missing} \\ [Median[Price_i] & if\ Median[Price_i|\, region]\ is\ missing \end{bmatrix} \quad (11)$$

This second-round imputation approach is the same as the approach used to deal with missing values in the traditional poverty estimation method based on unit value as a proxy for price. Once our price data has been harmonized with the IHS5 data, we follow all the steps in the traditional unit value-based analysis to estimate welfare measures (poverty and inequality). This ensures that the only difference between our market price approach and the unit value approach is our price data.

Appendix 2: Food Groups and Elementary Specifications in the Market Survey

Table A9

Breakdown of Food Groups and Elementary Specifications in the Market Survey

Food Group in the IH5 Food Consumption Module	Corresponding Elementary Goods Included in the Market Survey	1:1 Correspondence to an Item in the IHS5 Food Consumption Module?
Maize (Ufa)	Normal Flour	Yes
	Refined (Fine Flour)	Yes
	Bran Flour, Processed	No
	Bran Flour, Unprocessed	No
	Maize, Yellow Flour	No
Rice	Rice, 25% Broken, White	No
	Medium Grain Rice	No
	Short Grain Rice	No
Bread	Brown Bread	No
	White Bread	No
Buns, scones	Buns, Scones, White	No
	Buns, Scones, Yellow	No
Cassava tubers	Cassava Tubers	Yes
	Dried Cassava Makaka	No
Irish Potato	Irish Potato (Local)	No
	Irish Potato (Hybrid)	No
Pigeon pea (Nandolo)	Pigeon pea (Nandolo)	Yes
	Pigeon pea (Nandolo), Fresh	No
Beans	Beans, Demetta	No
	Beans, Kenyata	No
	Bean, Spotted	No
Onion	Onion (White)	No
	Onion (Red)	No
Chinese cabbage	Chinese Cabbage, Fresh	No
	Chinese Cabbage, Wilt	No
Eggs	Eggs Local	No
	Eggs Hybrid	No
Beef	Beef With Bones	No
	Beef Without Bones	No
	Beef, Ham	No
	Beef, Liver	No
	Beef, Minced	No
	Beef, Prepacked	No
	Beef, Rump Steak	No
	Beef, Fillet	No
Pork	Pork, Fillet	No
	Pork, Shoulder	No
Chicken	Chicken (Frozen), Whole	No
	Chicken (Traditionally), Bred Live	No
	Chicken Breast Without Skin	No
	Chicken Legs	No
	Chicken Wings	No
Banana	Banana, Short	No
	Banana, Long	No
Fresh milk	Fresh Milk, Unskimmed	No
	Fresh Milk, Low Fat	No
Sugar	Sugar, White	No
	Sugar, Brown	No

Appendix 3: Number and proportion of individuals in each quintile: Unit value vs Market Price Approach

Table A10

Number of Individuals in Aggregate per Capita Expenditure Quintile Group: Unit Value vs Market Price Approach (Unit Value Approach classification is used as the base)

	Unit Value Quintile					
Market Price Quintile	1	2	3	4	5	Total
1	3,096,647	522,509	7,355	1,528	-	3,628,039
2	496,042	2,554,360	568,236	7,410	-	3,626,049
3	34,348	522,693	2,589,858	482,503	-	3,629,402
4	-	27,332	458,562	2,858,965	279,306	3,624,164
5	-	-	4,322	275,353	3,346,773	3,626,449
Total	3,627,038	3,626,894	3,628,334	3,625,759	3,626,079	18,134,103

Table A11

Number of Individuals in Aggregate per Capita Expenditure Quintile Group: Unit Value vs Market Price Approach (Market Price Approach Classification is used as the base)

	Market Price Quintile					
Unit Value Quintile	1	2	3	4	5	Total
1	3,096,647	496,042	34,348	-	-	3,627,038
2	522,509	2,554,360	522,693	27,332	-	3,626,894
3	7,355	568,236	2,589,858	458,562	4,322	3,628,334
4	1,528	7,410	482,503	2,858,965	275,353	3,625,759
5	-	-	-	279,306	3,346,773	3,626,079
Total	3,628,039	3,626,049	3,629,402	3,624,164	3,626,449	18,134,103

Table A12

Proportion of Individuals in Aggregate per Capita Real Expenditure Quintile Group: Unit Value vs Market Price (Unit Value Approach Classification is used as the base)

	Unit Value Quintile					
Market Price Quintile	1	2	3	4	5	Total
1	85.4%	14.4%	0.2%	-	-	3,628,039
2	13.7%	70.4%	15.7%	0.2%	-	3,626,049
3	0.9%	14.4%	71.4%	13.3%	-	3,629,402
4	-	0.8%	12.6%	78.9%	7.7%	3,624,164
5	-	-	0.1%	7.6%	92.3%	3,626,449
Total	3,627,038	3,626,894	3,628,334	3,625,759	3,626,079	18,134,103
	100%	100%	100%	100%	100%	
% Ranked Within One Quintile	99.1%	99.2%	99.7%	99.8%	100.0%	

Table A13

Proportion of Individuals in Aggregate per Capita Real Expenditure Quintile Group: Unit Value vs Market Price (Market Price Approach Classification is used as the base)

	Market Price Quintile					
Unit Value Quintile	1	2	3	4	5	Total
1	85.4%	13.7%	0.9%	-	-	3,627,038
2	14.4%	70.4%	14.4%	0.8%	-	3,626,894
3	0.2%	15.7%	71.4%	12.7%	0.1%	3,628,334
4	-	0.2%	13.3%	78.9%	7.6%	3,625,759
5	-	-	-	7.7%	92.3%	3,626,079
Total	3,628,039	3,626,049	3,629,402	3,624,164	3,626,449	18,134,103
	100%	100%	100%	100%	100%	
% Ranked Within One Quintile	99.8%	99.8%	99.1%	99.2%	99.9%	

Appendix 4: Sensitivity of Poverty Lines to the Choice of Calorie Requirement and Reference Group

Table A14

Sensitivity of Market Price Derived Food and Total Poverty Lines (in Malawian Kwacha) to Calorie Requirements

Calorie Req.	**Food Poverty Lines**	**Total Poverty Lines**	Reference Group
2215	413.0	679.2	Deciles 5 and 6
2315	431.6	708.9	Deciles 5 and 6
2415	450.3	738.6	Deciles 5 and 6
2515	468.9	768.2	Deciles 5 and 6
2615	487.5	797.8	Deciles 5 and 6
2715	506.2	827.4	Deciles 5 and 6
2815	524.8	856.9	Deciles 5 and 6

Table A15

Sensitivity of Market Price Derived Food and Total Poverty Lines (in Malawian Kwacha) to Reference Group

Food Poverty line	**Total Poverty line**	**Reference Group**
393.9	585.2	Deciles 4 and 5
413.0	679.2	Deciles 5 and 6
420.2	563.0	Deciles 6 and 7
428.8	417.6	Deciles 7 and 8

Appendix 5: Sensitivity of Poverty Lines to Item Coverage in Market Survey

Table A16

Sensitivity of Poverty Lines to Item Coverage in Market Survey

	Calorie Req.	Reference-group Deciles	Unit Value			Complete Market Price Survey		
Item Coverage in Market Price Survey			**Food Poverty Line**	**Total Poverty Line**	**Non-food share**	**Food Poverty Line**	**Total Poverty Line**	**Non-food Share**
Full item Coverage	2215	Deciles 5 and 6	277.5	453.4	0.63	413.0	679.2	0.64
90	2215	Deciles 5 and 6				412.2	641.2	0.56
80	2215	Deciles 5 and 6				412.2	613.3	0.49
70	2215	Deciles 5 and 6				409.4	644.1	0.57
60	2215	Deciles 5 and 6				410.3	609.1	0.48
50	2215	Deciles 5 and 6				406.7	622.8	0.53

Note. These estimates are based on calorie requirement of 2215 Kcal and reference group of 5th and 6th deciles of aggregate real household consumption.

Appendix 6: Validation of Poverty and Consumption Estimates with Night Lights at Territorial Area and Enumeration Area Levels.

Results in Appendix Table A17 and A18 confirms a stronger association between poverty rates derived through the market price approach and the night lights measured at the traditional authority- and enumeration area-levels.

Table A17: Relationship between Poverty rates and Nightlights at Territorial Area (TA) level: Comparison of unit value and market price approaches

	Log TA CBN Poverty Rates (Unit Value)	Log TA CBN Poverty Rates (Market Price)	Log TA Food Poverty Rates (Unit Value)	Log TA Food Poverty Rates (Market Price)	TA CBN Poverty Rates (Unit Value)	TA CBN Poverty Rates (Market Price)	TA Food Poverty Rates (Unit Value)	TA Food Poverty Rates (Market Price)
Log TA Average Night Lights	-0.338***	-0.316***	-0.350***	-0.384***				
	(0.0221)	(0.0174)	(0.0388)	(0.0322)				
TA Average Night Lights					-6.884***	-7.955***	-3.298***	-4.915***
					(0.466)	(0.457)	(0.349)	(0.451)
Constant	3.384***	3.697***	2.455***	2.831***	54.74***	69.27***	21.70***	34.45***
	(0.0352)	(0.0277)	(0.0622)	(0.0511)	(1.178)	(1.154)	(0.883)	(1.139)
Observations	317	322	256	287	328	328	328	328
R-squared	0.428	0.508	0.243	0.334	0.401	0.482	0.215	0.267
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1								

Table A18: Relationship between Poverty rates and Nightlights at Enumeration Area (EA) level: Comparison of unit value and market price approaches

	Log EA CBN Poverty Rates (Unit Value)	Log EA CBN Poverty Rates (Market Price)	Log EA Food Poverty Rates (Unit Value)	Log EA Food Poverty Rates (Market Price)	EA CBN Poverty Rates (Unit Value)	EA CBN Poverty Rates (Market Price)	EA Food Poverty Rates (Unit Value)	EA Food Poverty Rates (Market Price)
Log EA Average Night Lights	-0.243***	-0.254***	-0.154***	-0.169***				
	(0.0152)	(0.0124)	(0.0266)	(0.0226)				
EA Average Night Lights					-6.459***	-7.488***	-3.047***	-4.528***
					(0.377)	(0.365)	(0.291)	(0.365)
Constant	3.436***	3.696***	2.779***	3.094***	53.22***	67.94***	20.87***	33.22***
	(0.0278)	(0.0226)	(0.0500)	(0.0422)	(0.875)	(0.848)	(0.675)	(0.847)
Observations	672	693	499	586	717	717	717	717
R-squared	0.277	0.379	0.063	0.088	0.291	0.370	0.133	0.177
Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1								

Appendix 7: Distribution of Markets at Reduced Market Coverage Scenario.

Appendix Table A19 presents the distribution of the number of markets in each level of market coverage scenario in the reduced market coverage scenarios. Please note that due to differences in the number of markets visited per district with some districts having fewer than 10 markets visited, the differences in the number of markets visited in each sensitivity analysis scenario will not exactly be different by 10 percent i.e., the 90 percent market coverage will not have exactly 10 percent less markets than the complete market coverage. In addition, one district (Likoma Island) had only one market visited in the complete survey. We have included this single market in all scenarios to prevent a situation where there is no information to represent market prices in this area in the sensitivity scenarios.

Compared to complete coverage where prices were collected from 377 markets across Malawi. Market coverage for the sensitivity scenarios ranged from 184 markets in the top 50 percent market coverage scenario to 325 markets in the top 90 percent coverage scenario.

Appendix Table A19

Number of markets in each level of market coverage

	Top 50% Market Coverage	Top 60% Market Coverage	Top 70% Market Coverage	Top 80% Market Coverage	Top 90% Market Coverage	Complete Market Coverage
Urban	52	58	59	66	70	75
Rural North	20	25	30	35	40	47
Rural Center	49	61	71	83	95	113
Rural South	63	74	94	105	120	142
National	184	218	254	289	325	377

Appendix 8: Additional inequality Statistics: Comparison of Unit Value and Market Price Approach

Appendix Table A20

Comparison of Inequality Statistics According to Unit Value versus Market Price Approach

					Decile Ratios							
	MLD		CoV		90-10		80-20		90-50		50-10	
	Unit Value	Market Price	Unit Value	Market Price	Unit Value	Market Price	Unit Value	Market Price	Unit Value	Market Price	Unit Value	Market Price
Urban	0.2528	0.2350	1.13	1.07	5.55	5.14	3.17	3.02	2.52	2.38	2.21	2.15
Rural-North	0.1680	0.1457	0.70	0.63	4.03	3.83	2.44	2.31	2.10	2.02	1.91	1.89
Rural-Center	0.1741	0.1778	0.71	0.70	4.25	4.42	2.55	2.58	2.11	2.12	2.01	2.09
Rural-South	0.1742	0.1716	0.76	0.73	3.97	4.08	2.46	2.46	2.10	2.10	1.89	1.94
National	0.2367	0.2264	1.06	0.99	5.02	4.93	2.82	2.79	2.43	2.35	2.06	2.10