

A NOTE ON HOUSING AND IMPUTED RENT

August 3, 2005

Summary

In most developing countries including Malawi, a small proportion of households explicitly pay for their dwelling with rent. The thin rental markets make it difficult to assess a value of owner-occupied dwellings. There are three ways to consider housing valuation for home-owners in the consumption aggregate: use self-reported estimated rents, use imputed rents based on sample of renters, or exclude from the consumption aggregate.

It should be noted that there is no consistency on whether and how housing valuation should be included in a consumption aggregate measure among countries in the region. The approaches that have been used by countries close to Malawi have been diverse. For example, Ethiopia, Kenya and Tanzania have excluded housing valuation; Mozambique have included actual rents and self-reported estimates; while Uganda and Zambia have included actual rental values estimated from regression models of actual renters.

The key findings from the analyses undertaken in this note are:

- *The lack of reliability of the estimated rental values suggests that the values should not be used for housing valuation or to extrapolate rental values for households with missing rental values.*
- *The small number of households renting in rural areas, combined with the notable differences in the dwelling characteristics of rented and owned houses in rural areas, suggests that the sample of renting households is too small to impute rental values for owned dwellings from actual rental values.*
- *Given the above conclusions, excluding housing valuation from the consumption aggregate should be given serious consideration.*

1.0 Introduction

Consumption aggregates are designed to estimate the total value of goods and services consumed by members of the household. For many items, this is fairly straightforward, for example, valuing the amount of food consumed in the past one week. For other areas, where the item “consumed” is a durable good or owner-occupied housing, this valuation is much more complicated. This note pertains to the issue of valuing the “consumption” of housing in Malawi. In the case of Malawi, where the vast majority of housing is owner-occupied, this issue becomes much more complicated. In areas where rental markets are thin (or non-existent, as in some rural areas), there is doubt as to whether the concept of renting (and hence estimated rental values) is understood by households. Most households in Malawi are either owners of their dwelling or otherwise do not pay rent. Overall, 91 percent of households in Malawi pay nothing for their dwelling. In rural areas, this figure is 97 percent.

This note evaluates the options with respect to valuing housing consumption for Malawi. The options available are:

1. Include the valuation of housing in the consumption aggregate, by one of the following methods:
 - a. Use actual rent and self-reported estimated rent for dwelling owners. For missing values of estimated rent (108 households), use regression estimates to predict estimated rent.
 - b. Use actual rent and imputed rental values for owner-occupied dwellings. The imputed values would be predicted based on regressions of actual rent values reported by renters.
2. Exclude the valuation of housing consumption from the consumption aggregate

In order to assess these options, this note evaluates several features of the data. It reviews the data available, including the housing characteristics of the renters and home-owners in rural and urban areas. It then reviews the correlates of rental values (actual and estimated). As an extension, the exercise tests whether the estimated rental values provided by non-renters were close to the predicted values based on actual rents paid by households living in houses of similar characteristics and in the same location. Based on the results from this analysis, a set of key findings are highlighted in the conclusions.

2.0 Data Available

In the Second Integrated Household Survey (IHS2), information on housing was collected in Section G of the questionnaire (Module G: Housing). The main parts that were used for this exercise were questions on estimated rental value (g03) and actual rent paid (g04). Other variables that were also used included housing characteristics. In addition, using the information on district and tradition authority, two dummy variables, regional and rural or urban, were also created. The next section gives a brief description of the main variables used in the regression models.

G03 Estimated rent values: For households who reside in owner-occupied dwellings (or those residing in free housing that is not technically owned by the household)¹, the household head was asked to estimate the value of rent for that dwelling.

G04 Actual rent paid: For households that paid rent, the household head provided information on the actual rent values they were paying.

Overall, 9 percent of households were paying rent for their housing (this corresponds to 1,068 households, unweighted).² There is a large difference between the prevalence of renting across urban and rural areas. Fifty percent of urban households were renting (720 out of 1,403 households, unweighted) compared to 3.3 percent of rural households (348 out of 9,667 households, unweighted).

2.1 Descriptive Statistics

Before moving to multivariate analysis, we first present descriptive statistics on housing characteristics for renters and non-renters. The variables included the following: roofing material (g07), type of roofing material (g09), floor material (g10), number of rooms occupied (g11), access to electricity (g20), main source of drinking water (g35) and type of toilet facility (refer to the table in Appendix 1). Since urban and rural differences are expected to be very salient, households are divided into four categories: urban and rural by renter and non-renter status.

The descriptive statistics show that there were significant differences between dwellings in urban and rural areas, and between rented and owned dwellings. Most of the houses in urban areas, both rented and owned, were constructed with better quality materials such as permanent materials, iron sheets, and smooth cement floors as compared to dwellings in rural areas. However, even so, there was also marked differences *within* the urban setting between those that were rented and those that were owned. Most of the rented dwellings in urban areas were of relatively better quality than the owned dwellings. The descriptive statistics also highlighted that in rural areas, the characteristics of rented houses were significantly different from owned houses. Renters in rural areas reside in dwellings of higher quality relative to home owners, as do their urban counterparts.

The differences in the characteristics of rented and owned dwellings may complicate the ability to impute rental values for owners from the sample of renters. This is because there may not be a sufficient number of comparable rented houses from which to estimate a model of rental values. This will be particularly problematic for rural areas. For example, in rural areas, there is a sample of 348 households who rent. Only about one-third of these dwellings are constructed of smoothed mud. However, the majority of owned dwellings in rural areas are constructed of smoothed mud. Thus, there are only approximately 100 rural renting households nationwide

¹ We use the term “owner-occupied” to refer to households that are not renting their dwelling. The majority of these households are, in fact, owner-occupied but not all are. In the IHS2 sample of *non-renting* households, 89 % were owners (or in the process of purchasing), 3.8 % live in employer provided housing, 7 % live in free, authorized housing, and 0.2 % are in free, unauthorized housing.

² All statistics are weighted using the preliminary household weights.

with which to estimate rental values for the sample of approximately 9300 rural home-owning households.

To understand the magnitude of the differences between characteristics of houses in different locations, houses were ranked according to the quality and facilities. A summary of the ranking of the 4 regions for each of the dwelling characteristics is presented in Table 1 below. Each housing characteristics is ranked from 1 to 4 with 1 indicated the area with the highest level in terms of quality of that characteristics, and 4 the lowest/poorest.

Table 1: Summary of Ranking of the Quality of Building Materials and Facilities

Variable	Urban		Rural	
	<i>Rented</i>	<i>Owned</i>	<i>Rented</i>	<i>Owned</i>
Type of construction material	2	3	1	4
Roofing material	1	3	2	4
Floor material	1	3	2	4
Access to electricity	1	2	3	4
Source of drinking water	1	2	3	4
Type of toilet facility	2	1	3	4

Note: 1 indicates the best/highest score/quality and 4 is the lowest/poorest score/quality for each indicator among the four types of the households (urban renter, urban owner, rural renter, and rural owner). See Appendix 1 for the underlying statistics for each of these indicators of housing quality.

The results of the ranking supports earlier findings that rented urban houses were better than urban owned, rural rented and rural owned houses. The rented urban houses scored higher than the rest in roofing material, floor and access to facilities such as electricity and water. The underlying statistics in Appendix 1 suggest minimum differences between urban owned and rural rented houses. The rural rented houses were better than the rest of the houses in the type of construction material while urban owned had superior toilet facilities. On the other hand, rural owned houses ranked the worst in all the categories.

Using housing valuation based on actual rent for renters and self-reported estimated rent for dwelling owners, we can evaluate the share in total consumption and correlation of housing to total consumption. Housing valuation as a share of total household consumption ranges from 11% to 6% (Table 2). Interestingly, the relationship is somewhat U-shaped, implying larger shares of housing in total consumption for the richest and poorest households, relative to households in the middle of the distribution.

Housing valuation is highly correlated with total consumption. The correlation coefficient between total consumption and housing valuation is 0.73. For most of the 4 groups (urban/rural, renters/non-renters), the correlations are similar: renting rural households .68, non-renting rural households .68, and renting urban households .64. For urban non-renters, the correlation is much higher, .88

Table 2: Housing as a Share of Total Consumption

Consumption Decile	Housing as a share of total consumption
1	11.1%
2	9.1%
3	8.1%
4	8.2%
5	7.2%
6	6.2%
7	6.0%
8	6.6%
9	5.9%
10	7.8%

Note: Housing valuation is based on actual rent for renters and self-reported estimated rent for dwelling owners. Total Consumption aggregate is based on preliminary version.

3.0 Results

3.1 Regression Models

In order to evaluate the self-reported estimated rental value of owner-occupied dwellings, this section imputes rental value using regression analysis. In the following section, the imputed rent values from the regressions will then be compared with the estimated rent values reported by households that do not pay rent.

Six regression models were estimated. Three were estimated on logged estimated rent paid by households in owner-occupied dwellings (all and by urban/rural). Three were estimated on logged actual rent values reported by renting households (all and by urban/rural). These logged dependent variables were regressed on the various house and geographic characteristics noted above (type of structure, type of roofing material, type of floor material, number of rooms occupied, access to electricity, main source of drinking water, type of toilet facility). Region dummy variables were also included. Results are presented in Appendix 2.

The first three models estimated regression equations using estimated rent for all households that reported estimated rent, as well as estimated rent for urban and rural rent only respectively. Comparing the models for the urban and rural samples (Appendix 2), the results from urban only model was able to explain about 72 percent of the variation in rental prices based on the household characteristics. On the other hand, the rural only model explained only 32 percent of the variations. These results suggest that households in urban areas have a better understanding of the concept of rental value than rural areas. This is presumably due to the fact that rental markets are more developed in these areas, which could lead to observable characteristics being more closely lined to self-reported rents.

Further, there was remarkable contradiction on the movement of the signs of the coefficients between the models of urban estimated rent and rural estimated rent. For example, main source of water (g30) and type of toilet facility (g36) had negative and significant coefficients for urban households, where lower quality was associated with lower rent. However, we do not see this pattern for rural households, implying that a lack of correlation between water quality and estimated rental values in rural areas. This further vindicated our earlier concerns that estimated rent was poorly correlated with characteristics of the house especially in rural areas.

Using the three actual rent paid models, the explanatory power of our models improved significantly as compared to the ones using estimated rent. The reported R^2 for the three actual rent models (columns 4-6 in Appendix 2) were 0.65, 0.68 and 0.50 for all, urban and rural households respectively. However, it was noted that the R^2 for the rural actual rent model was still low when compared with the other two models (column 6). There was a weak correlation between the characteristics of the house and the actual rent paid in rural areas. Although there were also some inconsistencies in the movement of the signs of the coefficients between actual rent model for all households paying rent and urban actual rent models with rural actual rent model, these were moderate. As such, the actual rent models might be preferred to impute rental values than using the self-estimated rental values reported by households.

4.0 Predictions

In order to compare the reliability of the estimated rent reported by home owners, the three specifications of actual rent (columns 4-6) were used to predict (impute) rental values.³

The predicted rental values from the regressions were then compared with the estimated rental values self-reported by home owners. Table 3 below shows the results of this comparison.

Table 3: Correlation Coefficients of the predicted and the estimated values

<i>Self-reported Rental Values</i>	<i>Predicted (imputed) rental values</i>		
	Specification 4 (All households)	Specification 5 (Urban)	Specification 6 (Rural)
All households (10,013)	0.58	0.57	0.55
Urban (694)	0.72	0.69	0.71
Rural (9,319)	0.46	0.46	0.41

Note: i) Figures in parenthesis are the number of observations. ii) All correlation coefficients are significantly different from zero.

The correlation coefficient between self-reported estimated rent and imputed rent is 0.58 when using specification 4 in Appendix 2. This was significantly different from zero. The predicted rental values for all households were also used to test its correlation with the estimated rent

³ Since the regression models above used a log-transformed dependent variable (estimated rent/actual rent), making predictions using a simple exponential approach will produce incorrect estimators. To prevent such misrepresentation of the results, we used the *predlog* command in Stata (See Stata manual for details).

values from the urban and rural households. The results yielded correlation coefficients of 0.72 and 0.46 for urban and rural respectively; these were also significantly different from zero.

The predicted figures from the sample of urban estimated rental values yielded correlation coefficients of 0.69 and 0.46 for urban and rural households respectively. The sample from rural households yielded 0.71 and 0.41 for urban and rural, respectively. These correlation coefficients are significantly different from zero.

As it can be seen from the results, there was a higher correlation between the predicted values for urban than rural areas, for all three model specifications. On the contrary, there was a weak correlation between the predicted values from the three models with the estimated rent for rural households. This further supported the hypothesis that the concept of renting is better developed in the urban areas than in the rural areas.

5.0 Conclusion

The note evaluated the options for valuating housing consumption for Malawi using several features of the IHS2 data. The key findings of this evaluation are:

- The basic statistics on housing characteristics revealed that there were significant differences in the dwelling characteristics of rented and owned dwellings. Further, there were also significant differences between the dwellings in urban and rural areas.
- The results of the estimated rent models in Appendix 2 showed that there were inconsistencies in the results of the correlates of estimated rent especially for rural households. This suggests that households had difficulties in estimating rental values that were consistent with the quality of their dwellings. This was especially apparent for rural households. Therefore, it can be concluded that the estimated rental values reported by households especially in the rural areas are unreliable.
 - *The lack of reliability of the estimated rental values suggests that the values should not be used for housing valuation or to extrapolate rental values for households with missing rental values.*
- A second option for including housing consumption in the consumption aggregate is to use the actual rental values reported to predict the rental values for owner-occupied dwellings. Out of the total 11,280 households interviewed, only 1,068 households are renting but most of these are urban households. The number of households paying rent is much lower for rural areas (348 renting households).
 - *The small number of households renting in rural areas, combined with the notable differences in the dwelling characteristics of rented and owned houses in rural areas, suggests that the sample of renting households is too small to impute rental values for owned dwellings from actual rental values.*

Appendix 1: Housing Characteristics

Housing Characteristics	Urban			Rural		
	All	Rented	Owned	All	Rented	Owned
<i>G07. (Type of Construction Material (%)</i>						
1. Permanent	42.6	47.1	37.9	12.1	53.3	10.6
2. Semi-Permanent	37.9	44.0	31.5	16.0	21.6	15.8
3. Traditional	19.5	8.9	30.6	71.9	25.2	73.6
<i>G09. Roofing Material (%)</i>						
1. Grass/Plastic Sheeting	23.6	12.0	35.8	81.5	32.9	83.2
2. Iron sheets/clay tiles/concrete	76.4	88.0	65.2	18.5	67.1	16.8
<i>G10. Floor Material (%)</i>						
1. Smoothed Mud /Sand	35.2	23.2	47.7	86.7	39.0	88.4
2. Wood/tile	2.4	1.8	3.1	0.0	0.0	0.0
3. Smooth Cement	62.8	75.0	49.2	13.3	61.0	11.6
<i>G11. Number of Separate Rooms Occupied (Mean)</i>	2.5	2.1	2.8	2.5	2.3	2.5
<i>G20. Has Access to Electricity (%)</i>	33.0	36.8	29.0	2.0	15.4	1.5
<i>G30. Main Source of Drinking Water (%)</i>						
1. Piped in house	12.6	11.9	13.4	0.8	2.7	0.7
2. Piped outside/communal pipe	64.4	79.5	48.6	11.9	31.8	11.2
3. Pump/protected spring	10.9	4.7	17.3	51.5	52.5	51.5
4. Unprotected water source	12.1	3.9	20.7	35.8	13.0	36.6
<i>G36. Type of Toilet Facility (%)</i>						
1. Flush toilet	14.0	12.4	15.7	1.2	6.0	1.1
2. Pit Latrine	83.0	87.2	78.7	79.8	91.4	79.4
3. None	3.0	0.5	5.6	19.0	2.6	19.5
Number of Observations	1,403	709	694	9,667	348	9,319

Note: Statistics are weighted using preliminary household weights. Total sample in IHS2 consists of 11,280 households. This table contains 11,070 households: 108 households missing actual or estimated rent and 101 households missing one or more of the dwelling characteristics in this table.

Appendix 2: Results of the Regression Models

	Estimated Rent			Actual Rent		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Urban	Rural	All	Urban	Rural
Type of Construction Material						
Semi-permanent	-0.180 (4.56)**	-0.372 (4.45)**	-0.121 (2.71)**	-0.120 (2.62)**	-0.095 (1.97)*	-0.194 (1.70)
Traditional	-0.385 (8.44)**	-0.339 (2.46)*	-0.341 (6.76)**	-0.208 (2.28)*	-0.165 (1.35)	-0.336 (2.04)*
Roofing Material						
Iron sheets/tiles	0.382 (11.33)**	0.351 (3.04)**	0.404 (11.33)**	0.169 (2.40)*	0.270 (2.83)**	0.050 (0.44)
Floor Material						
Wood/tiles	1.287 (7.35)**	0.947 (4.71)**	0.190 (0.42)	0.046 (0.27)	-0.003 (0.02)	
Smooth Cement	0.225 (6.08)**	0.061 (0.72)	0.261 (6.37)**	0.285 (5.64)**	0.277 (5.02)**	0.237 (2.17)*
# of separate rooms occupied	0.198 (31.39)**	0.323 (13.65)**	0.191 (29.16)**	0.203 (11.82)**	0.216 (10.23)**	0.187 (6.02)**
Electricity	0.876 (14.97)**	0.777 (8.71)**	0.870 (11.18)**	0.723 (15.43)**	0.685 (13.95)**	0.841 (7.46)**
Main Source of Drinking Water						
Piped water outside/communal	-0.540 (5.69)**	-0.585 (3.53)**	-0.293 (2.47)*	-0.180 (1.56)	0.113 (0.89)	-0.596 (2.09)*
Pump/protected spring	-0.552 (5.73)**	-0.832 (4.45)**	-0.297 (2.51)*	-0.208 (1.62)	0.094 (0.61)	-0.659 (2.26)*
Unprotected source	-0.593 (6.13)**	-0.754 (4.07)**	-0.344 (2.89)**	-0.199 (1.44)	0.062 (0.39)	-0.643 (2.10)*
Type of Toilet Facility						
Latrine	0.121 (1.51)	-0.116 (0.74)	0.244 (2.60)**	-0.361 (3.34)**	-0.715 (5.71)**	0.185 (0.92)
No toilet	-0.031 (0.38)	-0.303 (1.53)	0.091 (0.95)	-0.455 (2.37)*	-0.640 (2.12)*	-0.048 (0.16)
Region						
Northern region	-0.269 (10.15)**	-0.003 (0.03)	-0.289 (10.54)**	-0.019 (0.27)	0.178 (2.20)*	-0.306 (2.49)*
Central region	0.141 (8.45)**	-0.149 (2.12)*	0.155 (8.98)**	-0.012 (0.32)	-0.021 (0.46)	-0.054 (0.73)
Rural Residence	-0.258 (7.17)**			-0.372 (7.89)**		
Constant	8.632 (85.28)**	8.893 (47.30)**	7.968 (62.70)**	8.581 (71.34)**	8.513 (62.04)**	8.345 (27.89)**
Number of Observations	10,013	694	9,319	1,057	709	348
R-squared	0.41	0.72	0.32	0.65	0.68	0.50

Notes: Absolute value of t-statistics in parentheses. Regressions are weighted using preliminary household weights. * indicates significance at 5% level; ** significance at 1% level. Omitted categories are: permanent construction material, grass/plastic sheeting roofing material, smoothed mud/sand floor material, piped water in house, and flush toilet.

