Asian Journal of Empirical Research, 3(4)2013: 433-446

Asian Journal of Empirical Research

journal homepage: http://aessweb.com/journal-detail.php?id=5004

A REDUCED WEALTH INDEX FOR PHONE POLLS

Alyaa R. ZAHRAN[1]
Noha A. Abd El-SAMEA[2]
Rojan A. NAFEA[3]

ABSTRACT

The Reduced Wealth Index (RWI) used at the Public Opinion Poll Center (POPC) of the Information and Decision Support Center (IDSC) - Egyptian Cabinet, is evaluated for possible change in the importance pattern of the assets used to create the wealth index between the years 2005 and 2008. The predictive power of the reduced set of assets is studied and reasons behind its inability to represent the middle class are investigated. Two approaches of choosing/selecting the reduced set are utilized to improve predictive power of the reduced set of assets. It turns out that Best Reduced Wealth Index (BRWI) is the best reduced index.

Keywords: Assets index, Regression, Best subset selection technique, Socio-economic status, Principal component

INTRODUCTION

The Wealth Index of Filmer and Pritchett, (1998, 2001) is a creative way to get rid of the possible problems associated with asking about income in a survey such as: recall/ measurement errors, income seasonality nature, shortcoming of measuring monetary resources, sensitivity issues, and elderly/retired ones who do not earn income or rely on family members for their material supply. In addition, assets may provide a better picture of long term living standards than income (Moser and Felton 2007; Zimmer 2006). As a proxy for household wealth, the Wealth Index is found to be robust to the set of assets included as well as having internal and external coherence (Filmer and Pritchett, 2001). As a consequence, the use of the Wealth Index spread out and became a standard output in DHS and World Bank reports.

[1] Associate Prof.; STAT dept. FEPS, Cairo University, Egypt Emails: azahran@vt.edu, azahran@feps.edu.eg

[2] Senior Statistician POPC, IDSC

[3] Statistician POPC, IDSC

Before 2006, the Public Opinion Poll Center (POPC) of the Information and Decision Support Center (IDSC) - Egyptian Cabinet has relied on respondent telephone bill to categorize the respondent household in one of three socio-economic statuses (SES): high, medium, low. Following the Egypt-DHS (EDHS) office, POPC-IDSC decided to use the Wealth Index to proxy household SES. However, the EDHS uses more than 60 asset-indicators (household ownership of consumer durables, characteristics of the household's dwelling, and household landownership) to create the wealth index; a huge number of indicators to fit in a telephone interview. In addition, many of the dwelling characteristics do not fit a telephone interview. In 2006, Al-kasabi & El-Tawila at POPC proposed reducing this number to twelve asset-indicators. They believed this reduced set could capture most information of household wealth score. Currently POPC is using a version of their reduced wealth index, where one of their asset indicators (refrigerator) is dropped. It has been noticed, however, that this Reduced Wealth Index (RWI) suffers from under-representing the middle class. We evaluate RWI and investigate the reasons behind its limitation to represent the whole society correctly. Two approaches are used to better select a reduced subset of asset-indicators and the resulted reduced indices are evaluated as a proxy to the EDHS Wealth Index (EDHS-WI).

The remainder of this paper is organized as follows: Section 2 briefly reviews the asset index and the RWI. Section 3 compares the relative importance of the assets in the Wealth Index in the two rounds EDHS05 and EDHS08. In addition, it highlights the RWI shortening in estimating the EDHS-WI and postulates its reasons. In Section 4, a more objective strategy is proposed to select a subset of the assets that would better represent the wealth level in phone polls. In addition, the regression best subset technique is used as another approach to choose the best subset. Finally, Section 5 concludes the paper.

The reduced and EDHS wealth scores

The basic idea of the wealth index stems from the fact that assets ownership reflects household welfare in a non-monetary way. They considered a set of 21 asset indicators that reflect household ownership of consumer durables, characteristics of the household's dwelling, and household landownership. The principal component analysis is used to extract the first principal component which by definition would capture the largest amount of information in all the assets. This first principal component is a weighted average of the assets and is defined to be the wealth score. It is calculated for each household, transformed to population wealth score and then its range is divided into classes of Wealth Index that reflect population SES. Finally, each household is being classified into one class according to its wealth score and each household member is given his/her household wealth score. The DHS offices began to adopt this wealth index, but with a slight modification where the Factor analysis with PRINCIPAL COMPONENT extraction method is used. The assumed distribution of some of these assets over the range of the wealth score is provided in (Rutstein and Johnson, 2004).

The assets used in the wealth index should stem from the society under consideration (Filmer and Pritchett, 2001). The EDHS office decided to use more than 60 asset-indicators in the construction of its Wealth Index (different consumer durables and some dwelling characteristics such as flooring material; type of drinking water source; toilet facilities) (El-Zanaty and Way, 2005). However, in a telephone interview one should ask as few questions as necessary for the sake of response rate and time. In addition, it is somehow sensitive to ask about detailed dwelling characteristics over the telephone. Al-kasabi and El-Tawila chose a set of 11 durable assets from the set of consumer durables used by EDHS; and the dwelling type indicator (apartment or not) as the sole dwelling characteristic in their reduced index. They believed this 12-assets set does capture most information of household wealth score. POPC used their reduced set to create a wealth index for a while till it decided to drop the refrigerator from this 12-assets-set. The resulted index contains 11-assets (10 durables and the type of dwelling indicator). We refer to this index by Reduced Wealth Index (RWI). In addition, we refer to the 30 consumer durables used by EDHS and the dwelling type by full set of assets. Note that the telephone indicator is dropped from the set of consumer durables since it is redundant in a telephone interview. In addition, all the remaining dwelling indicators are left out form this definition of the full set, because it is inappropriate to ask about them in a telephone survey. It is like we create first a suitable and feasible pool of assets and then begin to choose from it the suitable subset for the telephone interview. Table 1 lists the full set of assets and the RWI subset. Other information in this table will be clarified subsequently in the paper.

Using EDHS05 data, Al-kasabi and El-Tawila created the 12-assets wealth scores and divided its range according to the EDHS wealth quintiles. Cross-classifying people by EDHS and reduced wealth indices, they found that the correct-classification percentages ranged between 42.5% and 93.3%, while the false-classification percentages were relatively small. They proposed the use of their 12-assets reduced wealth index in any poll performed by the POPC-IDSC with the recommendation to update its quintiles cutoff points according to the most recent released EDHS data set. To get a clear picture of these cross-classifications we depict them in FIG 1. The classification performance of the reduced index is good at the most outer quintiles (richest and poorest), but it loses gradually its performance as it heads toward the middle quintile from both sides with a slightly better performance at higher quintile compared to poorer quintile. This poses three questions: 1) why does the index behave this way? 2) how far is the reduced subset able to reflect EDHS wealth index? and 3) how could we better choose the reduced subset? Questions 1 & 2 are discussed in the Section 3, while question 3 is tackled in Section 4.

Evaluating the RWI used at POPC

Before evaluating the reduced set of assets, we investigate the assets relative importance in the full set of assets for the two consecutive rounds EDHS05 and EDHS08. The importance of each asset in the wealth index could be read from two quantities: the factor scoring and asset effect. Factor scoring is the weight of the standardized asset indicator in the wealth score equation, whereas the

asset effect is the factor scoring over the standard deviation of the asset indicator. FIG 2 compares the factor scoring in the two consecutive rounds: EDHS05 and EDHS08. A similar factor scoring pattern is present in the two rounds. The set of assets (stone drinking-water container "Kola/Zeer"; livestock: herds/farm animals; small table "tablia"; non-automatic washing machine; land for agriculture; hanging lamp; animal-drawn cart; bicycle) that was negatively associated to the wealth level in EDHS05 remained negative in EDHS08. Negative sign of any asset means that owning it will reduce the wealth score. It is worth noting that all of these negative scored assets are related to the rural areas in Egypt. Apart from electric fan, freezer, mobile, refrigerator, satellite, sewing machine, sofa, and video/DVD, all the 2008-wealth scorings are higher than the 2005-wealth scorings. FIG 3 shows how the asset-ownership is related to wealth scale. Some of the assets are sharply related to high levels of wealth (dotted line), some are more related to high levels than to low levels (dashed line), some are highly related to low levels (long dashed line), while some others are fairly associated with all levels of wealth (solid line).

FIG 4 shows the asset effect, which quantifies the effect of owning an asset on the household wealth score, keeping other assets unchanged. For the practitioner, this is usually more attractive to know and easier to understand. Comparing year 2008 to 2005, the wealth score is more affected by most of the assets regardless of the sign (direction) of the effect size. FIG 2 and Fig 4 indicate that the relative position of the full-set of assets on the wealth barometer is dynamic; some assets lose/gain some of its indicative power to/from some other assets over the time period 2005-2008. This highlights the need to review the choice of the reduced subset that is used in the POPC-IDSC with each new release of EDHS data. FIG 2 clearly shows that assets in the Al-kasabi and El-Tawila reduced set have factor scoring of no less than 0.059. The ten durable assets of the currently used reduced set of assets at POPC (leaving out the type of dwelling) are among the assets that we classified in FIG 3 as "sharply related to high levels of wealth". FIG 4 reveals that these durable assets have an effect of at least 0.22 in 2005. In DHS08, the picture did change a little bit, where two assets fall below that value. However, with all the assets having such positive effect on WI, the ability of RWI to measure a three-level SES, or in general a more than two-level SES, is limited. This indicates that RWI does have a strong identifiable power of the rich people, but only the rich vs others (middle and poor together), it is a binomial measure in nature.

FIG 5 & FIG 6 show the histogram of the EDHS-WI and RWI, respectively, in 2008. The EDHS08 wealth histogram is bimodal, smooth and almost symmetric, whereas the RWI distribution show clumping pattern. This clumping pattern reflects the inability of the reduced set to discriminate among different levels of wealth. In addition, RWI histogram is skewed to the right and does not follow the EDHS histogram. While McKenzie, (2005) postulated that the clumping property is related to the use of insufficient number of assets, we believe that this clumping property could be lessened by a better choice of assets inside the reduced subset. A bad choice of members of this subset will result in less predictive power; hence false modes in the histograms, where many

household will be clumped in few groups. On the other hand, FIG 6 indicates that the distribution suffers from a truncation problem, where the problem is more severe at the bottom of the distribution. McKenzie, (2005) claimed that the truncation property represents the fact that no assets are used in the index to differentiate between two neighboring socio-economic levels (commonly the poor vs the poorest or the richest vs the rich). From all the above, we conjecture that the use of the current reduced set of assets at POPC-IDSC lowers the predictive power of their wealth index and results in under-representing the middle class. In the next section we discuss how to improve the choice of the subset to better approximate the EDHS-WI.

Objective approaches to choose the reduced subset

As we have shown in the previous section, the current choice of the reduced set at POPC makes the RWI a binomial index in the sense that it could differentiate rich from not rich people, but could not, however, differentiate middle class people from the poor ones. From our point of view, the index could have been able to differentiate among the three socio-economic classes, if it contained not only high-effect assets but also assets with moderate- or low-effect. Another possible approach to better choose the reduced set is to utilize the regression best subset approach. Below both approaches are applied using EDHS08 data. The resulted indices are assessed in terms of their closeness to the EDHS08 wealth index. First we modify the choice of the reduced subset by including assets with moderate- and high-effect values. In this paper, a high effect is of value no less than 0.22 while a low effect is of a negative value. In the light of the EDHS08 data (FIG 4), the chosen high-effect assets are: water heater, clock, dishwasher, freezer, car/truck, and personal computer; while the moderate-effect assets are: apartment, sewing machine, satellite, radio, and TV. This set of assets is meant to be just an example of how one can choose the subset based on its effect classes: high/moderate, high/low, or moderate/low. FIG 7 shows the histogram of this Modified RWI (MRWI). The MRWI performs better than RWI, where the truncation problem is not severe but the clumping problem still exists. Compared to the EDHS08-WI histogram, the MRWI histogram is approaching the symmetric shape of the EDHS-WI. We conjecture that MRWI could be improved by changing some of the chosen assets.

Although the previous approach improved the resulted index, the selection remains subjective. An objective selection could be obtained if we applied one of the statistical selection procedures. The best subset of the regression analysis is used here to get the best subset of assets conditioned on a choice of eleven assets (the number of assets currently used at POPC assuming this is the optimal size of a subset). The first two eleven-asset models (Models 1 & 2) from the best subset regression technique are considered. The last two columns of Table 1 show the assets of these two models. Some of the assets were not included in the best subset technique (highlighted gray in Table 1) because of its sensitivity to be asked about over the phone in the Egyptian Society. FIG 8 & FIG 9 show the histograms of Best-subset RWI (BRWI) using Model 1 and Model 2, respectively. In terms of the histograms, both BRWIs perform better than RWI. The truncation problem is smaller

and no evidence of severe clumping. Model 1 BRWI histogram is smooth and approaches the shape of the EDHS-WI, but Model 2 histogram is skewed. There is no evidence, though, that the BRWIs outperform the MRWI. FIG 10 depicts the cumulative distributions of EDHS08, RWI, MRWI and BRWIs wealth scores. BRWI 1 behaves the best, whereas RWI is the worst. Though, BRWI 2 beats the BRWI 1 at the left tail. But in general, quintiles of both BRWIs almost coincides the EDHS08 quintiles. Although MRWI performs better than RWI, both indices fail to match the quintiles of the EDHS08. Table 1 shows also the performance of the four indices in terms of the coefficient of determination. MRWI (76.7%) slightly performs better than RWI (74.6%), while the BRWIs could capture 82.7% of the total variation in EDHS08-WI.

CONCLUSION

Reducing the number of assets used in wealth index is mandatory in phone polls. However, the choice of the reduced subset is critical and heavily affects the predictive power of the resulted index. The Reduced Wealth Index (RWI), however, suffers from under-representing the middle class. We show that the inability of RWI stems from the bad choice of the chosen set of assets in its reduced set, which contains only assets that are heavily owned by rich people. As a consequence RWI could only differentiate rich from not rich people, but could not, differentiate middle class people from the poor ones. Two approaches are offered to enhance the selection procedure of the reduced set. The Modified Reduced Wealth Index (MRWI) depends on choosing assets from two types (classes) of assets: high- & low-effect assets, high- & moderate-effect assets or low- & moderate-effect assets. The other approach utilizes the regression best subset approach. Using EDHS08 data, two best subsets are identified. These two indices outperform both the MRWI and the RWI. For phone surveys and also not-a-long face to face surveys, the Best Subset Reduced Wealth index offers a balance between cost (in terms of time and effort) and the identifiable power of correct classifying households in their wealth levels that the DHS WI would have declared. Any analysis at POPC that does involve the wealth score will be of more precision if this new version of the reduced wealth index is used.

REFERENCES

El-Zanaty, F and A. Way (2005). Egypt demographic and health survey 2005 http://www.measuredhs.com/pubs/pdf/FR176/FR176.pdf Accessed 32 Oct 2011

Filmer, D. and L. Pritchett (1998). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India. World Bank Policy Research working Paper number 1994, Washington, D. C. DECRG, The World Bank.

Filmer, D. and L. Pritchett (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India Demography Vol. 38, pp. 115-132.

McKenzie, D. J. (2005). Measuring inequality with assets indicators, *Journal of Population Economics*, Vol. 18, pp. 229-260.

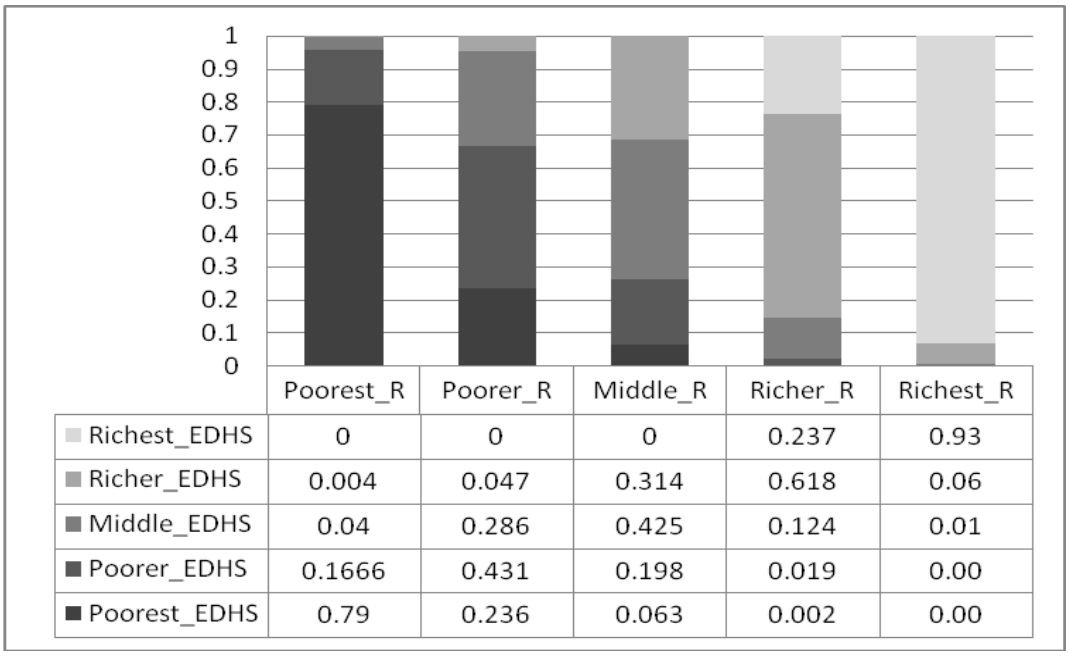
Moser, C. and A. Felton (2007). The construction of an asset index measuring asset accumulation in Ecuador *CPRC Working Paper number* 87, *Chronic Poverty Research Centre*, ISSN 1-904049-86-9, The Brookings Institution, Washington D. C. USA.

Rutstein, S. O. and K. Johnson (2004). The DHS wealth index *DHS Comparative Reports 6* ORC Macro Calverton, Maryland USA.

Zimmer, Z. (2006). Poverty, wealth inequality, and health among older adults in rural Cambodia *Policy Research Division, Population Council*, Working paper number 217, NY, USA.

Table 1: Assets considered in full set and reduced sets

	Indicator of Assets	Full	Reduced R^2=74.6%	Modified R^2=76.7%	Best Subset1 R^2=82.7%	Best Subset2 R^2=82.7%
1	Large/small stone water container (Kola/Zeer)	√				
2	Livestock, herds, or farm animals	√			√	√
3	Small table (tablia)	√				
4	Non automatic washing machine	√				
5	Land for agriculture	√				
6	Hanging lamp (green lamp)	√				
7	Animal-drawn cart	√				
8	Has bicycle	√				
9	Has motorcycle/scooter	√				
10	Sewing machine	√		√		
11	Bed	√				
12	Sofa	√				
13	TV	√		√	√	√
14	Electric fan	√				√
15	Table	√				
16	Dishwasher	√	√	√		
17	Has radio	√		√	√	√
18	Clock (watch)	√		√	√	√
19	Chair	√				
20	Has refrigerator	√			√	√
21	Bank account	√				
22	Air condition	√	√			
23	Has car/truck	√	√	√		
24	Freezer	√	√	√		
25	dwelling type (apartment not)	√	√	√	√	√
26	Satellite dish	√	√	√	√	√
27	Video/DVD	√	√			
28	Personal home computer	√	√	√	√	
29	Mobile telephone	√	√		√	√
30	Water heater	√	√	√	√	√
31	Automatic washing machine	√	√		√	√



	Poorest_R	Poorer_R	Middle_R	Richer_R	Richest_R
Richest_EDHS	0	0	0	0.237	0.93
Richer_EDHS	0.004	0.047	0.314	0.618	0.06
Middle_EDHS	0.04	0.286	0.425	0.124	0.01
Poorer_EDHS	0.1666	0.431	0.198	0.019	0.00
Poorest_EDHS	0.79	0.236	0.063	0.002	0.00

FIG 1: Cross classification percentages of people by EDHS and RWI quintiles

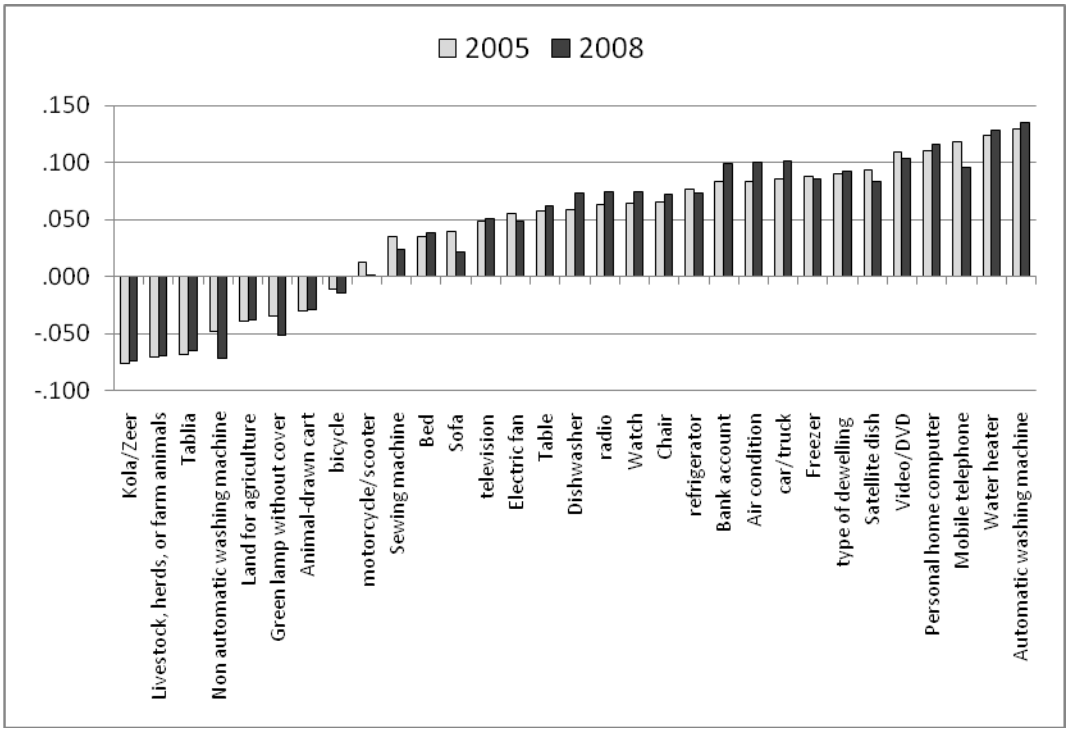


FIG 2: Wealth scoring on the full-set of assets (DATA: EDHS08 and EDHS05)

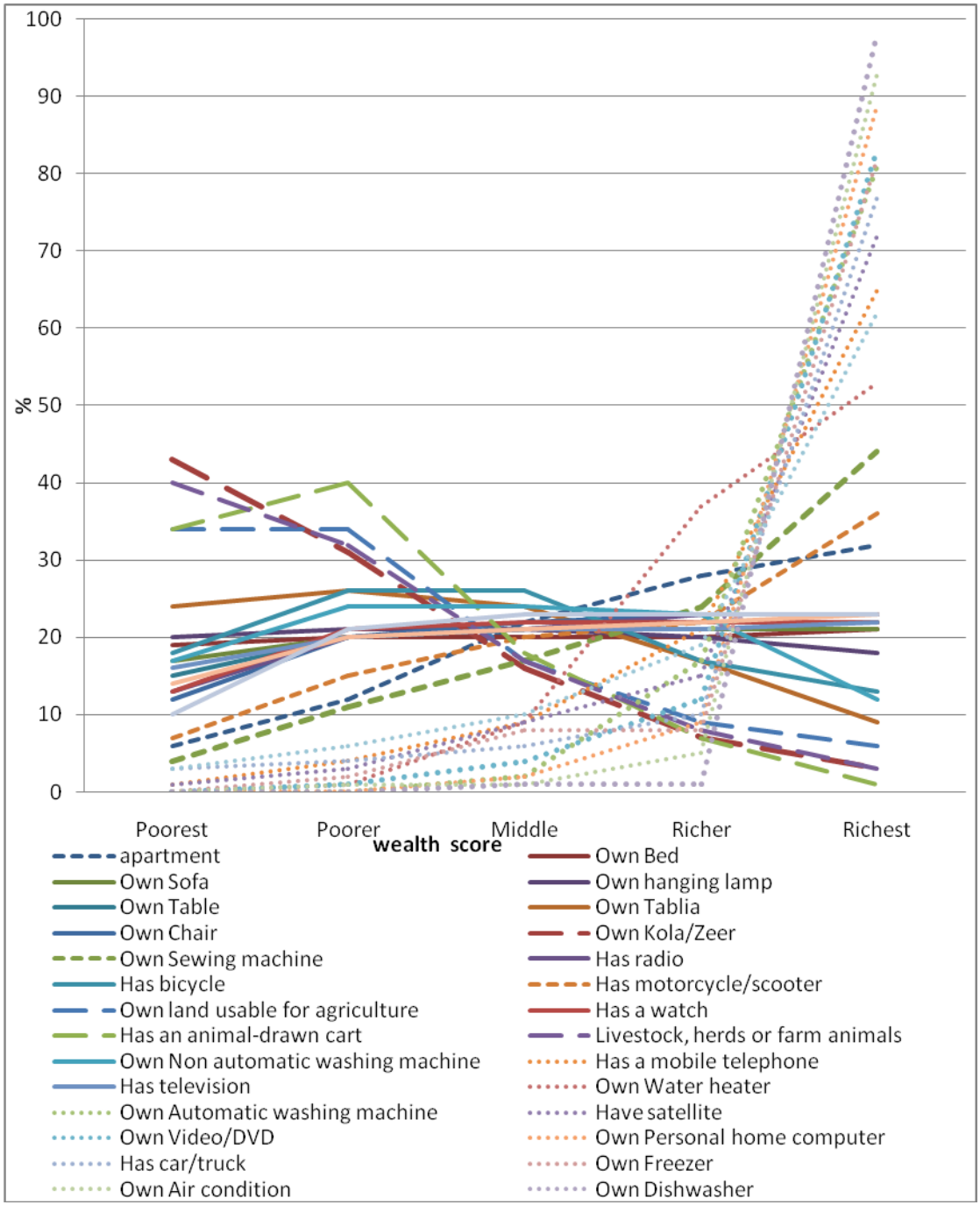


FIG 3: Relative ownership distributions of the full-set assets over the range of the wealth level

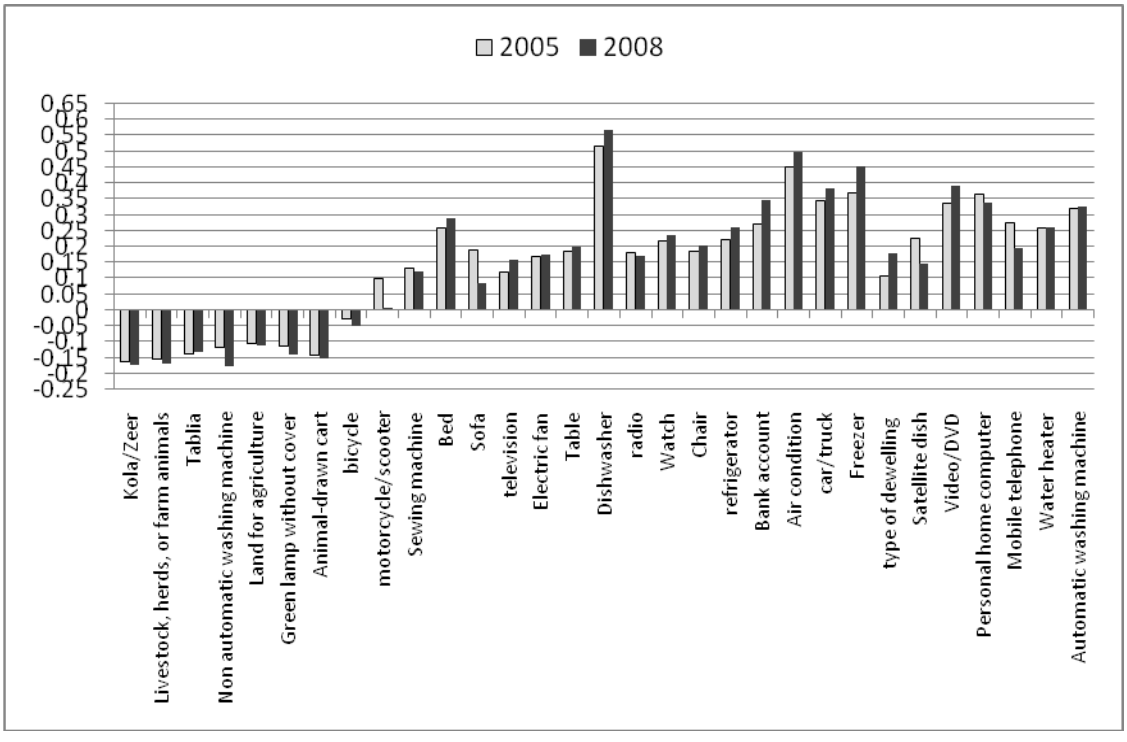


FIG 4: Full-set assets effect on the wealth score (DATA: EDHS05 and EDHS08)

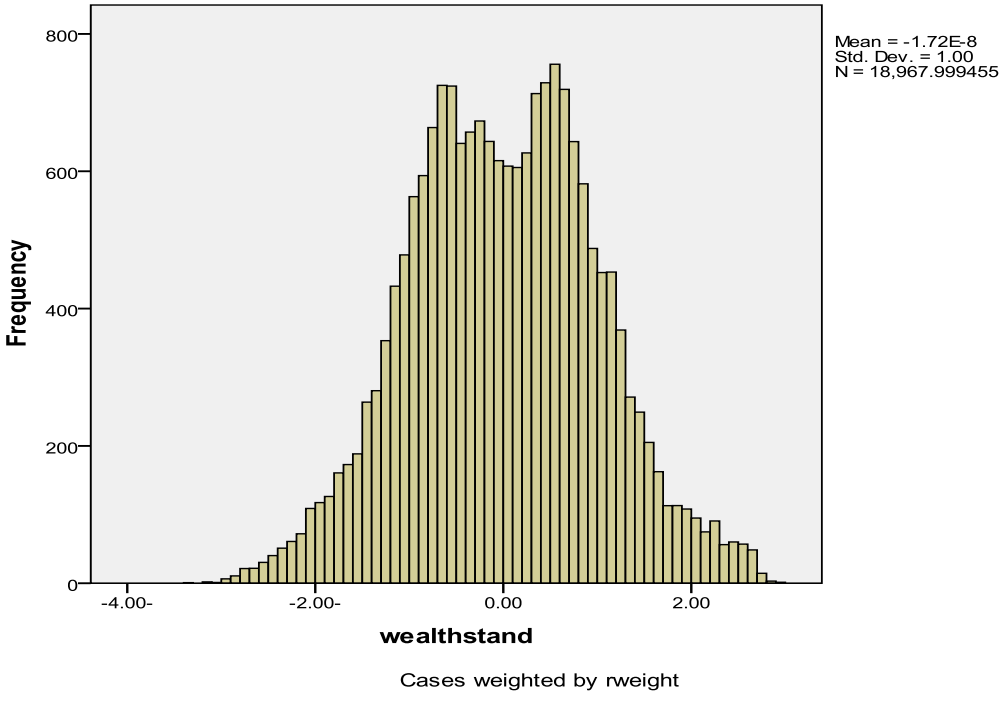


FIG 5: Histogram of Household EDHS standardized Wealth Score (DATA: EDHS08)

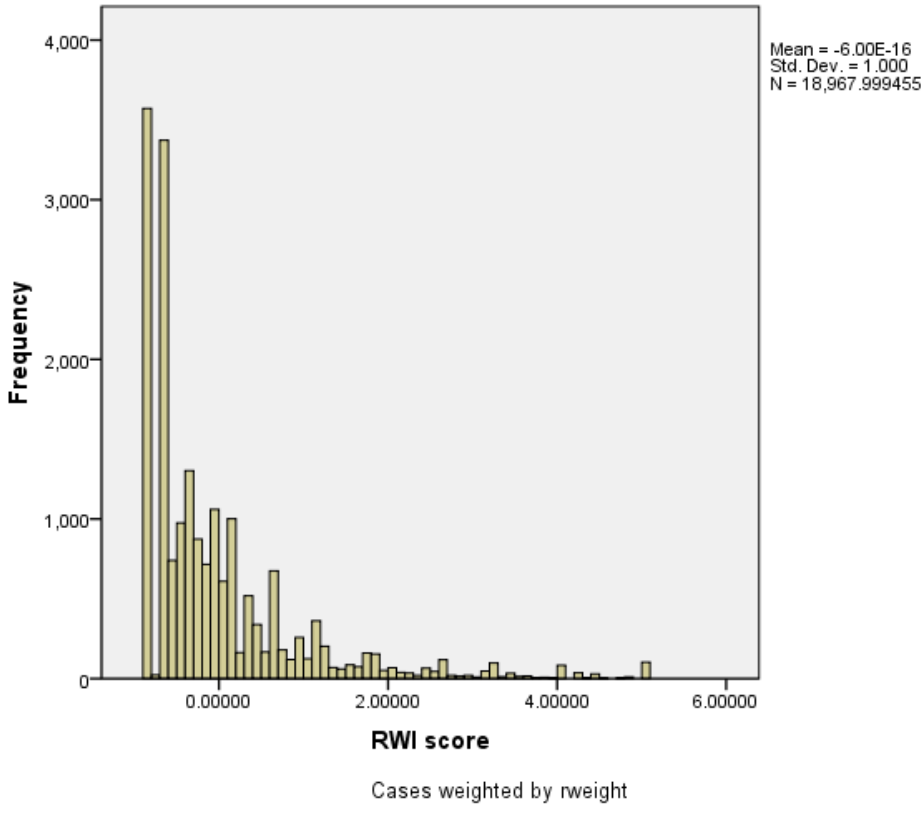


FIG 6: Histogram of household RWI score (DATA: EDHS08)

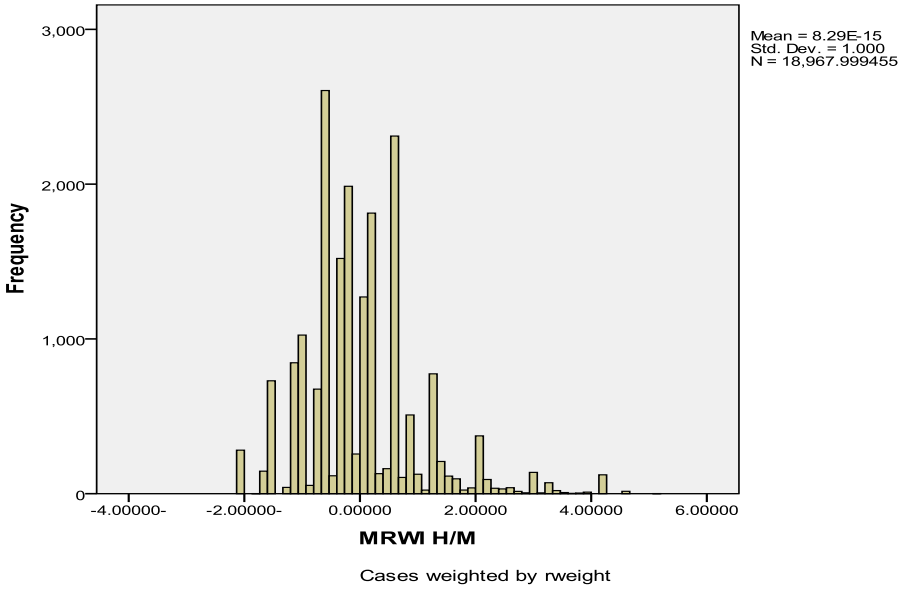


FIG 7: Histogram of MRWI Score (DATA: EDHS08)

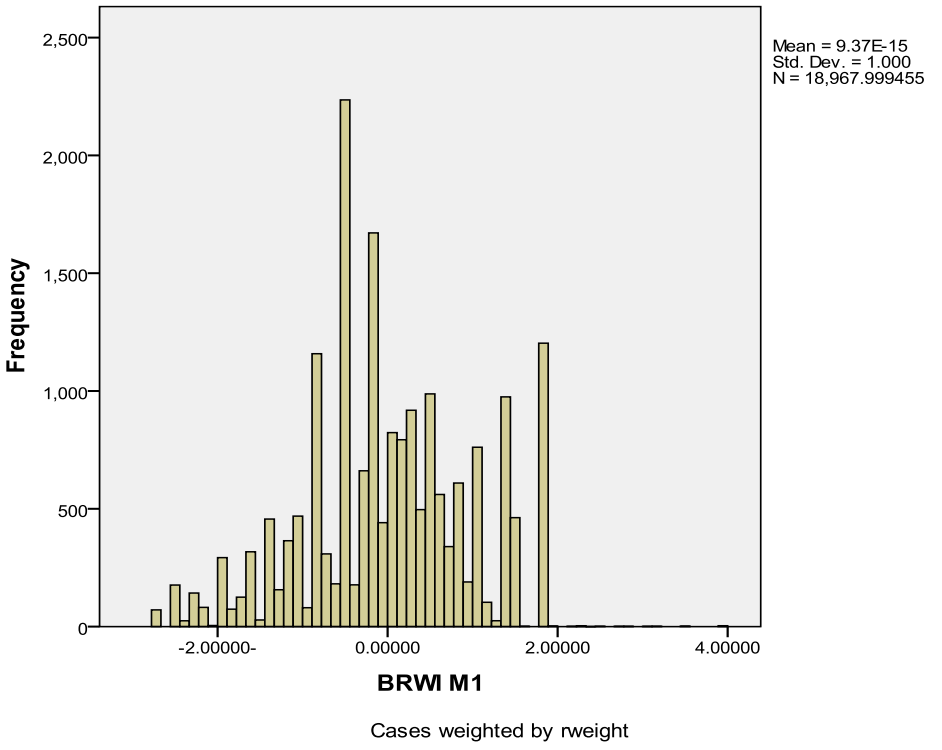


FIG 8: Histogram BRWI scores of model 1 (DATA: EDHS08)

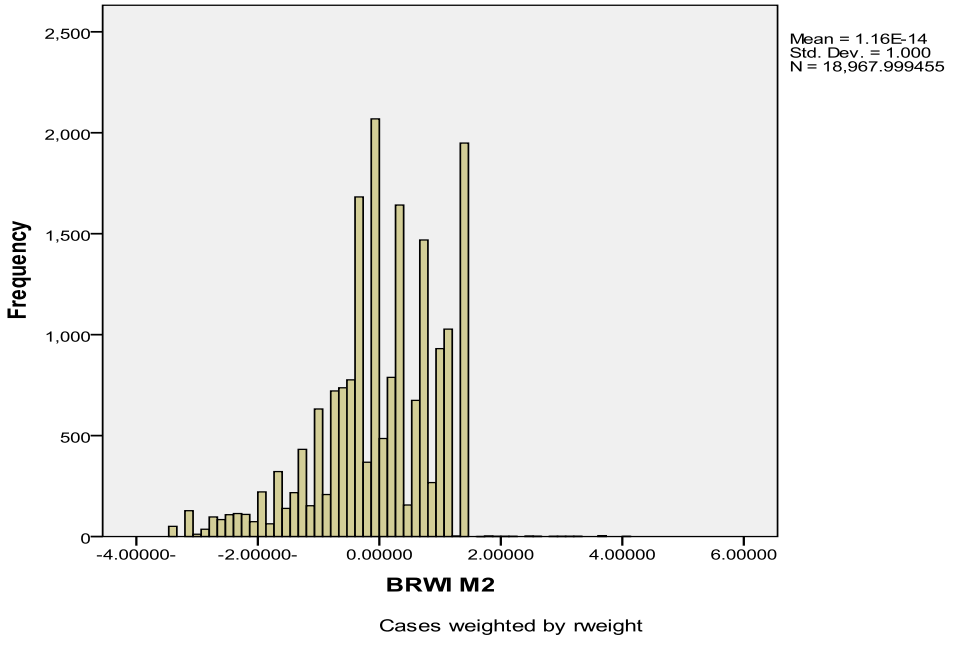


FIG 9: Histogram of BRWI score of model 2 (DATA: EDHS08)

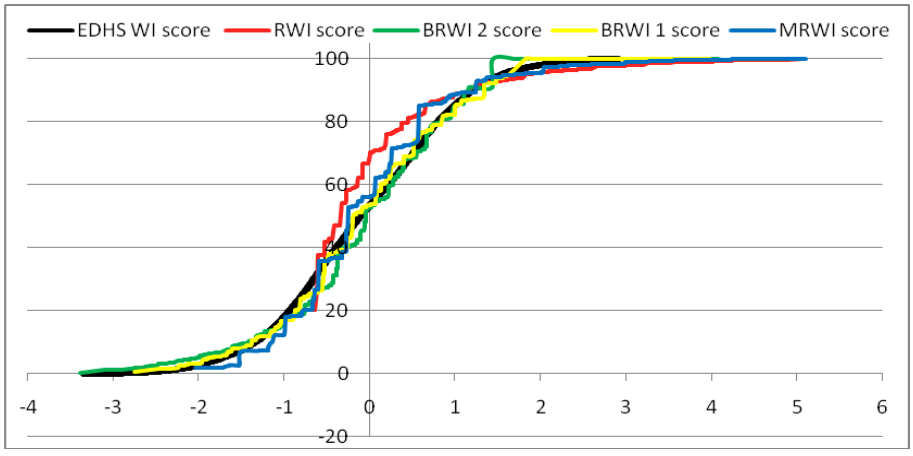


FIG 10: Cumulative distribution function of the different household wealth indices